Testing Slope Homogeneity in Large Panels

M. Hashem Pesaran and Takashi Yamagata

March 2005

CWPE 0513

Abstract

This paper proposes a modified version of Swamy's test of slope homogeneity for panel data models where the cross section dimension (N) could be large relative to the time series dimension (T). The proposed test exploits the cross section dispersion of individual slopes weighted by their relative precision. In the case of models with strictly exogenous regressors and normally distributed errors, the test is shown to have a standard normal distribution as $(N, T) \xrightarrow{j} \infty$. Under non-normal errors and in the case of stationary dynamic models, the condition on the relative expansion rates of N and T for the test to be valid is given by $\sqrt{N}/T \to 0$, as $(N, T) \xrightarrow{j} \infty$. Using Monte Carlo experiments, it is shown that the test has the correct size and satisfactory power in panels with strictly exogenous regressors for various combinations of N and T. For autoregressive (AR) models the proposed test performs well for moderate values of the root of the autoregressive process. But for AR models with roots near unity a bias-corrected bootstrapped version of the test is proposed which performs well even if N is large relative to T. The proposed cross section dispersion tests are applied to testing the homogeneity of slopes in autoregressive models of individual earnings using the PSID data. The results show statistically significant evidence of slope heterogeneity in the earnings dynamics, even when individuals with similar educational backgrounds are considered as sub-sets.

JEL-Classification: C12, C33.

Keywords: Testing Slope Homogeneity, Hausman Type Tests, Cross Section Dispersion Tests, Monte Carlo Results, PSID Earnings Dynamics

Testing Slope Homogeneity in Large Panels[1]

January 2005

M. Hashem Pesaran
Cambridge University & USC
Takashi Yamagata
Cambridge University

1. Introduction

In many empirical studies, it is assumed that the slope coefficients of interest in panel data models are homogeneous across individual units. When the cross section dimension (N) is relatively small and the time series dimension of the panel (T) large, the hypothesis of slope homogeneity can be tested using the SURE (seemingly unrelated regression equation) framework of Zellner (1962). This framework is particularly attractive as it also automatically deals with the possibility of cross section error correlations and dynamics when N is reasonably small (around 5-10) and T sufficiently large (around 80-100). However, in many empirical applications N is often (much) larger than T and the SURE approach would not be applicable.

In view of this Pesaran, Smith and Im (1996) proposed the application of the Hausman (1978) testing procedure where the standard fixed effects estimator is compared to the mean group estimator. However, as will be discussed below, such a procedure is not applicable in the case of panel data models that contain only strictly exogenous regressors and/or in the case of pure autoregressive models.

Recently Phillips and Sul (2003) have also proposed a 'Hausman type' test for slope homogeneity for stationary first-order autoregression (AR(1)) panel data models in presence of cross section dependence, with

[1] We would like to thank Ron Smith and Aman Ullah for helpful discussions. We are also grateful to Costas Meghir and Luigi Pistaferri for kindly providing us with their processed PSID data set. Also our thanks go to Donggyu Sul for the Gauss codes used for the implementation of Phillips and Sul's G test. Financial support from the ESRC (Grant No. RES-000-23-0135) is gratefully acknowledged.

N fixed as T goes to infinity. It will be shown below that their testing approach is not valid under cross section independence.

This paper proposes a modified version of the test proposed by Swamy (1970) that applies to panel data models where the cross section dimension could be large relative to the time series dimension. The proposed test is applicable to static as well as to stationary dynamic panel data models, possibly with heteroskedastic errors. In the case of models with strictly exogenous regressors and normally distributed errors, the proposed test is shown to have a standard normal distribution as $(N,T) \xrightarrow{j} \infty$, where $(N,T) \xrightarrow{j} \infty$ denotes N and $T \to \infty$ jointly. Under non-normal errors and in the case of stationary dynamic models, the condition on the relative expansion rates of N and T for the test to be valid is given by $\sqrt{N}/T \to 0$, as $(N,T) \xrightarrow{j} \infty$.

The small sample properties of the proposed test are investigated by means of Monte Carlo experiments. It is shown that the test has satisfactory size and power for T as small as 10 with N as large as 200 in panel data models containing only strictly exogenous regressors, even with non-normal errors. For autoregressive (AR) models the proposed test performs well for moderate values of the root of the AR process under various N and T combinations. But for AR panels with $T < N$, and roots near unity, a bias-corrected bootstrapped version of the test is proposed which is shown to perform well even if N is large relative to T.

The use of slope homogeneity tests in empirical contexts is illustrated by applying them to testing the homogeneity of slopes in autoregressive models of earnings using the Panel Study of Income Dynamics (PSID) data. The results show evidence of slope heterogeneity in the real earnings dynamics, even when individuals with similar educational backgrounds are considered as sub-sets.

The plan of the paper is as follows. Section 2 sets up the model and reviews existing tests of slope homogeneity. Section 3 considers the asymptotic distribution of alternative dispersion type tests of slope homogeneity and establishes their asymptotic distribution in the context of panel data models where N could be large relative to T. Section 4 considers the application of the proposed Δ test to stationary dynamic panel data models and develops the biased-corrected bootstrapped version of the test. Section 5 sets up the Monte Carlo design and summarizes the results. Section 6 discusses the empirical application, and Section 7 provides some concluding remarks.

2. The Model and Existing Tests of Slope Homogeneity

Consider the panel data model with fixed effects and heterogeneous slopes

$$y_{it} = \alpha_i + \beta_i' x_{it} + \varepsilon_{it}, \; i = 1, \dots, N, \; t = 1, \dots, T \tag{2.1}$$

where α_i is bounded on a compact set, x_{it} is a $k \times 1$ vector of regressors, β_i is a $k \times 1$ vector of unknown slope coefficients. Stacking the time series observations for i yields

$$y_i = \alpha_i \tau_T + X_i \beta_i + \varepsilon_i, \; i = 1, 2, \dots, N; \tag{2.2}$$

where $y_i = (y_{i1}, \dots, y_{iT})'$, τ_T is a $T \times 1$ vector of ones, $X_i = (x_{i1}, \dots, x_{iT})'$, and $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{iT})'$. Let

$$Q_{iT} = T^{-1}\left(X_i' M_\tau X_i\right); \; \xi_{iT} = T^{-1/2} X_i' M_\tau \varepsilon_i; \tag{2.3}$$

and

$$Q_N = (NT)^{-1}\left(\sum_{i=1}^{N} \sigma_i^{-2} X_i' M_\tau X_i\right); \tag{2.4}$$

where $M_\tau = I_T - \tau_T\left(\tau_T'\tau_T\right)^{-1}\tau_T'$, and I_T is an identity matrix of order T.

Consider now the following assumptions:

Assumption 1: $\varepsilon_{it} \sim IID(0, \sigma_i^2)$ with $0 < \sigma_i^2 < \infty$ for all i, and ε_{it} and ε_{js} are independently distributed for $i \neq j$ and/or $t \neq s$.

Assumption 2: The $k \times k$ matrices Q_{iT}, $i = 1, 2, \dots, N$, defined by (2.3) are positive definite, Q_{iT}^{-1} has finite second order moments for each i, and Q_{iT} tends to a non-stochastic positive definite matrix, Q_i, as $T \to \infty$.

Assumption 3: The $k \times 1$ vectors ξ_{iT}, $i = 1, 2, ..., N$ defined by (2.3) are independently distributed across i, and for each i, $\xi_{iT} \to_d N\left(0, \sigma_i^2 Q_i\right)$, as $T \to \infty$.

Assumption 4: The $k \times k$ pooled observation matrix Q_N defined by (2.4) is positive definite, and tends to a non-stochastic positive definite matrix, Q, as $(N, T) \xrightarrow{j} \infty$.[2]

Assumption 5: $(T-1)\left(\varepsilon_i' M_\tau \varepsilon_i\right)^{-1}$ has finite second order moments for each i.

The null hypothesis of interest is

$$H_0 : \beta_i = \beta \text{ for all } i; \ \|\beta\| < K < \infty, \tag{2.5}$$

against the alternatives

$$H_1 : \beta_i \neq \beta, \text{ for a non-zero fraction of slopes.}$$

Assumption 6: Under H_1, the fraction of slopes that are not the same does not tend to zero as $N \to \infty$.

Remark 1. In the case of randomly distributed slopes, where $\beta_i \sim IID(\beta, \Sigma_\beta)$, the null and the alternative hypothesis can be characterized by $H_0 : \Sigma_\beta = 0$, and $H_1 : \Sigma_\beta \neq 0$, respectively.

Remark 2. The above assumptions cover both cases of strictly exogenous regressors, as well as when x_{it} contains lagged values of y_{it}.

Remark 3. In the case where the errors, ε_{it}, are normally distributed Assumption 5 is met if $T > 5$. See, for example, Smith (1988) for a proof.[3]

2.1. The Standard F Test

There are a number of procedures that can be used to test H_0, the most familiar of which is the standard F test defined by

$$F = \left[\frac{N(T-k-1)}{k(N-1)}\right]\left(\frac{RSSR - USSR}{USSR}\right), \tag{2.6}$$

[2] $(N, T) \xrightarrow{j} \infty$ denotes joint asymptotics with N and $T \to \infty$ in no particular order.

[3] Under normality, the r^{th} moment of the inverse of $\varepsilon_i' A \varepsilon_i$ exits if $\text{rank}(A) > 2r$, where A is a $T \times T$ positive semi-definite symmetric matrix.

where RSSR and USSR are restricted and unrestricted residual sum of squares, respectively, obtained under the null ($\beta_i = \beta$) and the alternative hypotheses. This test is applicable when the regressors are strictly exogenous and the error variances homoskedastic, $\sigma_i^2 = \sigma^2$. But it is likely to perform rather poorly in cases where the regressors might contain lagged values of the dependent variable and/or if the error variances are cross sectionally heteroskedastic.

2.2. Hausman Type Test by Pesaran, Smith and Im

For cases where $N > T$, Pesaran, Smith and Im (1996) propose using the Hausman (1978) test where the standard fixed effects (FE) estimator of β,

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{N} X_i' M_\tau X_i \right)^{-1} \sum_{i=1}^{N} X_i' M_\tau y_i, \tag{2.7}$$

is compared to the mean group (MG) estimator defined by

$$\hat{\beta}_{MG} = N^{-1} \sum_{i=1}^{N} \hat{\beta}_i, \tag{2.8}$$

where

$$\hat{\beta}_i = \left[X_i' M_\tau X_i \right]^{-1} X_i' M_\tau y_i. \tag{2.9}$$

For the Hausman test to have the correct size and be consistent two conditions must be met, however.

(a) Under the null hypothesis, $\hat{\beta}_{FE}$ and $\hat{\beta}_{MG}$ must both be consistent, with $\hat{\beta}_{FE}$ being asymptotically more efficient such that

$$\mathrm{Avar}\left(\hat{\beta}_{MG} - \hat{\beta}_{FE}\right) = \mathrm{Avar}\left(\hat{\beta}_{MG}\right) - \mathrm{Avar}\left(\hat{\beta}_{FE}\right) > 0.$$

(b) Under the alternative hypothesis $\hat{\beta}_{MG} - \hat{\beta}_{FE}$ should tend to a non-zero vector.

In the context of dynamic panel data models with exogenous regressors both of these conditions are met, so long as the exogenous regressors are not drawn from the same distribution, and a Hausman type test based on the difference $\hat{\beta}_{FE} - \hat{\beta}_{MG}$ would be valid and is shown to have reasonable small sample properties. See Pesaran, Smith and Im (1996) and Hsiao and Pesaran (2004).

However, there are two major concerns with the routine use of the Hausman procedure as a test of slope homogeneity. It could lack power for certain parameter values, as its implicit null does not necessarily coincide with the null hypothesis of interest. Second, and more importantly, the Hausman test will not be applicable in the case of panel data models containing only strictly exogenous regressors, and/or in the case of pure autoregressive models. In the former case, both estimators, $\hat{\beta}_{FE}$ and $\hat{\beta}_{MG}$, are unbiased under the null and the alternative hypotheses and condition (b) will not be satisfied. Whilst, in the case of pure autoregressive panel data models $\sqrt{NT}\left(\hat{\beta}_{FE} - \beta\right)$ and $\sqrt{NT}\left(\hat{\beta}_{MG} - \beta\right)$ will be asymptotically equivalent and condition (a) will not be met.

2.3. G Test of Phillips and Sul

Phillips and Sul (2003) propose a different type of Hausman test where instead of comparing two different pooled estimators of the regression coefficients (as discussed above), they propose basing the test of slope homogeneity on the difference between the individual estimates and a suitably defined pooled estimator. In the context of the panel regression model (2.2), their test statistic can be written as

$$G = \left(\hat{\beta}_N - \tau_N \otimes \hat{\beta}_{pooled}\right)' \hat{\Sigma}_g^{-1} \left(\hat{\beta}_N - \tau_N \otimes \hat{\beta}_{pooled}\right),$$

where $\hat{\beta}_N = (\hat{\beta}_1', \hat{\beta}_2', ..., \hat{\beta}_N')'$ is an $Nk \times 1$ stacked vector of all the N individual estimates, $\hat{\beta}_{pooled}$ is a suitable pooled estimator of β $(= \beta_i)$, and $\hat{\Sigma}_g$ is a consistent estimator of Σ_g, the asymptotic variance matrix of $\hat{\beta}_N - \tau_N \otimes \hat{\beta}_{pooled}$, under H_0. Under Assumptions 1-4 and assuming H_0 holds and N is fixed, then $G \to_d \chi^2_{Nk}$ as $T \to \infty$, so long as the Σ_g is a non-stochastic positive definite matrix.

As compared to the Hausman test based on $\hat{\beta}_{MG} - \hat{\beta}_{FE}$, the G test is likely to be more powerful; but its use will be limited to panel data models where N is small relative to T. Also, the G test will not be valid in the case of pure dynamic models, very much for the same kind of reasons noted above in relation to the Hausman test based on $\hat{\beta}_{MG} - \hat{\beta}_{FE}$. This is easily established in the case of the stationary first order autoregressive panel data model considered by Phillips and Sul (2003) where

$$y_{it} = \alpha_i(1-\lambda_i) + \lambda_i y_{i,t-1} + \varepsilon_{it}; \quad |\lambda_i| < 1;$$

and the aim is to test $H_0 : \lambda_i = \lambda$. Phillips and Sul also consider the case where the errors, ε_{it}, are cross sectionally dependent through a single factor model. But, given the focus of our analysis, we shall abstract from this problem and continue to assume that ε_{it} are cross sectionally independent. Under this set up the appropriate form of the G statistic is given by

$$G = \left(\hat{\lambda}_N - \tau_N \hat{\lambda}_{pooled}\right)' \hat{\Sigma}_g^{-1} \left(\hat{\lambda}_N - \tau_N \hat{\lambda}_{pooled}\right);$$

where $\hat{\lambda}_N = \left(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_N\right)'$ is the $N \times 1$ vector of the individual estimates and $\hat{\lambda}_{pooled}$ is a pooled estimator, such that $(\hat{\lambda}_N', \hat{\lambda}_{pooled})' \to_p \lambda \tau_{N+1}$ under the null hypothesis. Phillips and Sul consider a number of different estimators, including Andrew's (1993) median unbiased estimator and its extension to panels. But, as they note, all such estimators yield the same asymptotic covariance matrix as $T \to \infty$. Using the fixed effects estimator for $\hat{\lambda}_{pooled}$, and the least squares estimators of λ_i for $\hat{\lambda}_i$, it is easily verified that under H_0

$$\begin{aligned}
Avar\left[\sqrt{T}\left(\hat{\lambda}_i - \hat{\lambda}_{FE}\right)\right] &= Avar\left[\sqrt{T}\left(\hat{\lambda}_i - \lambda\right) - \sqrt{T}\left(\hat{\lambda}_{FE} - \lambda\right)\right] \\
&= \left(1-\lambda^2\right) - \left(\frac{1-\lambda^2}{N}\right); \\
Acov\left[\sqrt{T}\left(\hat{\lambda}_i - \hat{\lambda}_{FE}\right), \sqrt{T}\left(\hat{\lambda}_j - \hat{\lambda}_{FE}\right)\right] &= -\left(\frac{1-\lambda^2}{N}\right).
\end{aligned}$$

Therefore

$$\Sigma_g = \left(\frac{1-\lambda^2}{T}\right)\left(I_N - N^{-1}\tau_N\tau_N'\right),$$

where $\text{Rank}(\Sigma_g) = N - 1$ and Σ_g is non-invertible.

2.4. Swamy's Test

Swamy (1970) bases his test of slope homogeneity on the dispersion of individual slope estimates from a suitable pooled estimator. Like the F test, Swamy's test is developed for panels where N is small relative to T, but allows for cross section heteroskedasticity. Swamy's statistic applied to the slope coefficients can be written as

$$\hat{S} = \sum_{i=1}^{N}\left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)'\frac{X_i'M_\tau X_i}{\hat{\sigma}_i^2}\left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right), \tag{2.10}$$

where

$$\hat{\sigma}_i^2 = \frac{\left(y_i - X_i\hat{\beta}_i\right)'M_\tau\left(y_i - X_i\hat{\beta}_i\right)}{(T-k-1)}, \tag{2.11}$$

and $\tilde{\beta}_{WFE}$ is the weighted FE (WFE) pooled estimator of slope coefficients defined by

$$\tilde{\beta}_{WFE} = \left(\sum_{i=1}^{N}\frac{X_i'M_\tau X_i}{\hat{\sigma}_i^2}\right)^{-1}\sum_{i=1}^{N}\frac{X_i'M_\tau y_i}{\hat{\sigma}_i^2}.$$

In the case where N is fixed and T tends to infinity, under H_0 the Swamy statistic, $\hat{S}$, is asymptotically chi-square-distributed with $k(N-1)$ degrees of freedom.[4]

3. Dispersion Type Tests for Large Panels

Our survey of the literature suggests that there are no satisfactory tests of slope homogeneity in panels where N is large relative to T. The standard F test and its extension by Swamy (1970) are appropriate for

[4] See also Hsiao (2003, p.149).

panels where N is small relative to T. Hausman type tests advanced by Pesaran, Smith and Im (1996) apply to large N panels, but are not generally applicable and can suffer from low power. In this paper we propose standardized dispersion statistics that are asymptotically normally distributed as $(N,T) \stackrel{j}{\to} \infty$.

In addition to Swamy's test statistic, $\hat{S}$, defined by (2.10), we also consider the following version

$$\tilde{S} = \sum_{i=1}^{N} \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)' \frac{X_i' M_\tau X_i}{\tilde{\sigma}_i^2} \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right) \tag{3.1}$$

where $\tilde{\sigma}_i^2$ is an estimator of σ_i^2 based on $\hat{\beta}_{FE}$, namely

$$\tilde{\sigma}_i^2 = \frac{\left(y_i - X_i \hat{\beta}_{FE}\right)' M_\tau \left(y_i - X_i \hat{\beta}_{FE}\right)}{T-1}, \tag{3.2}$$

and $\tilde{\beta}_{WFE}$ is the weighted FE estimator also computed using $\tilde{\sigma}_i^2$, namely

$$\tilde{\beta}_{WFE} = \left(\sum_{i=1}^{N} \frac{X_i' M_\tau X_i}{\tilde{\sigma}_i^2}\right)^{-1} \sum_{i=1}^{N} \frac{X_i' M_\tau y_i}{\tilde{\sigma}_i^2}. \tag{3.3}$$

Although the difference between $\hat{S}$ and $\tilde{S}$ might appear slight at first, the choice of the estimator of σ_i^2 has important implications for the properties of the two dispersion tests as N and T tends to infinity.

To establish the asymptotic results for the Swamy's version of the dispersion test we need the following more restrictive version of Assumption 5:

Assumption 5$'$: $\hat{\sigma}_i^2$ is a consistent estimator of σ_i^2 such that

$$\frac{\sigma_i^2}{\hat{\sigma}_i^2} = 1 + O_p\left(T^{-1}\right); \tag{3.4}$$

and $E\left(1/\hat{\sigma}_i^2\right)$ exists and is bounded.

We also note that under Assumptions 1-4

$$\xi_{iT}' Q_{iT}^{-1} \xi_{iT} = O_p(1); \quad N^{-1} \sum_{i=1}^{N} \sigma_i^{-2} Q_{iT} = Q_N = O_p(1); \tag{3.5}$$

and

$$\mathrm{Avar}\left(N^{-1/2}\sum_{i=1}^{N}\sigma_i^{-2}\xi_{iT}\right) = \lim_{N\to\infty}\left(N^{-1}\sum_{i=1}^{N}\sigma_i^{-2}Q_i\right) = Q = O(1). \tag{3.6}$$

Consider first the Swamy's version of the dispersion test. Under H_0 we have

$$\left(\hat{\beta}_i - \hat{\beta}_{WFE}\right) = T^{-1/2}Q_{iT}^{-1}\xi_{iT} - T^{-1/2}N^{-1/2}\left(N^{-1}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}Q_{iT}\right)^{-1}\left(N^{-1/2}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}\right), \tag{3.7}$$

where Q_{iT} and ξ_{iT} are given by (2.3). Using this result in (2.10) it is easily seen that

$$\frac{1}{\sqrt{N}}\hat{S} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\xi'_{iT}Q_{iT}^{-1}\xi_{iT}}{\tilde{\sigma}_i^{2}} - \frac{1}{\sqrt{N}}\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}}{\sqrt{N}}\right)'\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}Q_{iT}}{N}\right)^{-1}\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}}{\sqrt{N}}\right).$$

In view of (3.5) and (3.6) and using (3.4) we have,

$$N^{-1/2}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi'_{iT}Q_{iT}^{-1}\xi_{iT} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\xi'_{iT}Q_{iT}^{-1}\xi_{iT}}{\sigma_i^{2}}\right) + O_p\left(\frac{\sqrt{N}}{T}\right);$$

$$N^{-1/2}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT} = N^{-1/2}\sum_{i=1}^{N}\sigma_i^{-2}\xi_{iT} + O_p\left(\frac{\sqrt{N}}{T}\right);$$

and

$$N^{-1}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}Q_{iT} = N^{-1}\sum_{i=1}^{N}\sigma_i^{-2}Q_{iT} + O_p\left(\frac{1}{T}\right).$$

Hence (again using (3.5) and (3.6))

$$N^{-1/2}\hat{S} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\xi'_{iT}Q_{iT}^{-1}\xi_{iT}}{\sigma_i^{2}}\right) + O_p\left(\frac{\sqrt{N}}{T}\right) + O_p\left(\frac{1}{\sqrt{N}}\right);$$

or equivalently

$$N^{-1/2}\hat{S} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\varepsilon'_i P_i \varepsilon_i}{\sigma_i^{2}}\right) + O_p\left(\frac{\sqrt{N}}{T}\right) + O_p\left(\frac{1}{\sqrt{N}}\right); \tag{3.8}$$

where

$$P_i = M_\tau X_i \left[X_i' M_\tau X_i\right]^{-1} X_i' M_\tau. \tag{3.9}$$

Consider now our modified version of Swamy's statistic, $\tilde{S}$, which under H_0 can be similarly written as

$$\frac{1}{\sqrt{N}}\tilde{S} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\xi_{iT}' Q_{iT}^{-1}\xi_{iT}}{\tilde{\sigma}_i^2} - \frac{1}{\sqrt{N}}\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}}{\sqrt{N}}\right)'\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}Q_{iT}}{N}\right)^{-1}\left(\frac{\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}}{\sqrt{N}}\right). \tag{3.10}$$

Using (3.2) first note that after some algebra under H_0 we have

$$(T-1)\tilde{\sigma}_i^2 = \varepsilon_i' M_\tau \varepsilon_i - 2N^{-1/2}\xi_N^{*\prime} Q_N^{*-1}\xi_{iT} + N^{-1}\xi_N^{*\prime} Q_N^{*-1} Q_{iT} Q_N^{*-1}\xi_N^{*}.$$

where

$$Q_N^* = N^{-1}\sum_{i=1}^{N} Q_{iT}; \quad \xi_N^* = N^{-1/2}\sum_{i=1}^{N}\xi_{iT}.$$

But using results in (3.5) and (3.6) and recalling that $0 < \sigma_i^2 < \infty$, we also have[5]

$$Q_N^* = O_p(1); \quad \xi_N^* = O_p(1).$$

Therefore,

$$\tilde{\sigma}_i^2 = \frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1} + O_p\left(N^{-1/2}T^{-1}\right). \tag{3.11}$$

It is also clear from (3.11) that for N sufficiently large $\tilde{\sigma}_i^{-2}$ has second order moments for any T so long as Assumption 5 is satisfied. Therefore, under Assumptions 2 and 3 the second order moments of $\tilde{\sigma}_i^{-2}\xi_{iT}$ and $\tilde{\sigma}_i^{-2}Q_{iT}$ will exist for N large and we have

$$N^{-1/2}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT}' Q_{iT}^{-1}\xi_{iT} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1}\right)^{-1}\xi_{iT}' Q_{iT}^{-1}\xi_{iT} + O_p\left[T^{-1}\right];$$

$$N^{-1/2}\sum_{i=1}^{N}\tilde{\sigma}_i^{-2}\xi_{iT} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1}\right)^{-1}\xi_{iT} + O_p\left[T^{-1}\right],$$

[5] Note that by assumption Q_{iT} and ξ_{iT} have finite second order moments.

and

$$N^{-1}\sum_{i=1}^{N}\sigma_i^{-2}Q_{iT} = N^{-1}\sum_{i=1}^{N}\left(\frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1}\right)^{-1} Q_{iT} + O_p\left(N^{-1/2}T^{-1}\right):$$

Using these results in (3.10)

$$N^{-1/2}\tilde{S} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1}\right)^{-1}\xi_{iT}' Q_{iT}^{-1}\xi_{iT} + O_p\left(T^{-1}\right) + O_p\left(N^{-1/2}\right);$$

or equivalently, since $\xi_{iT}' Q_{iT}^{-1}\xi_{iT} = \varepsilon_i' P_i \varepsilon_i$,

$$N^{-1/2}\tilde{S} = N^{-1/2}\sum_{i=1}^{N} z_i + O_p\left(T^{-1}\right) + O_p\left(N^{-1/2}\right); \tag{3.12}$$

where

$$z_i = \left(\frac{\varepsilon_i' M_\tau \varepsilon_i}{T-1}\right)^{-1}\varepsilon_i' P_i \varepsilon_i. \tag{3.13}$$

A comparison of (3.8) and (3.12) clearly shows that for N and T large the $\tilde{S}$ version of the dispersion test could be preferable to the $\hat{S}$ version since the latter requires N and T to increase at the same rates whilst the former does not necessarily require this condition. In fact, as we shall see below, in the case of strictly exogenous regressors the slope homogeneity test based on $\tilde{S}$ would be valid for $(N, T) \xrightarrow{j} \infty$, whilst a test based on $\hat{S}$, in addition to $(N, T) \xrightarrow{j} \infty$ would also require that $\sqrt{N}/T \to 0$. In the case of dynamic panels both versions of the dispersion test require the additional condition $\sqrt{N}/T \to 0$, and a bias-corrected bootstrapped test will be considered.

Before proceeding further we summarize the above results in the following theorem.

Theorem 3.1. Consider the panel data model (2.1), and suppose that Assumptions 1-5 hold. Then the dispersion statistics $\hat{S}$ and $\tilde{S}$ defined by (2.10) and (3.1), respectively, can be written as

$$N^{-1/2}\hat{S} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\varepsilon_i' P_i \varepsilon_i}{\sigma_i^2}\right) + O_p\left(N^{1/2}T^{-1}\right) + O_p\left(N^{-1/2}\right); \tag{3.14}$$

$$N^{-1/2}\tilde{S} = N^{-1/2}\sum_{i=1}^{N} z_i + O_p\left(T^{-1}\right) + O_p\left(N^{-1/2}\right); \qquad (3.15)$$

where P_i and z_i are defined by (3.9) and (3.13), respectively.

This theorem is fairly general and applies irrespective of whether the regressors are strictly exogenous or contain lagged dependent variables, and holds for non-normal errors.

Consider now the case where the regressors are strictly exogenous and the errors are normally distributed, $\varepsilon_i \sim IIDN\left(0, \sigma_i^2 I_T\right)$. In this case $\sigma_i^{-2}\varepsilon_i' P_i \varepsilon_i$ has a chi-square distribution with k degrees of freedom and the following standardized version of $\hat{S}$ could be used when N and T are both large

$$\hat{\Delta} = \sqrt{N}\left(\frac{N^{-1}\hat{S} - k}{2k}\right). \qquad (3.16)$$

Using (3.14) it is easily seen that

$$\hat{\Delta} = N^{-1/2}\sum_{i=1}^{N}\left(\frac{\sigma_i^{-2}\varepsilon_i' P_i \varepsilon_i - k}{2k}\right) + O_p\left(N^{1/2}T^{-1}\right) + O_p\left(N^{-1/2}\right);$$

and under H_0, $\hat{\Delta} \to_d N(0,1)$ as $(N,T) \xrightarrow{j} \infty$ such that $\sqrt{N}/T \to 0$.

Turning to the $\tilde{S}$ version of the test, using well known results in von Neumann (1941) on moments of the ratio of quadratic forms in standard normal variates, we first note that

$$E(z_i) = \frac{E(\varepsilon_i' P_i \varepsilon_i)}{E\left(\varepsilon_i' M_\tau \varepsilon_i\right)/(T-1)} = \frac{(T-1)\operatorname{tr}(P_i)}{\operatorname{tr}(M_\tau)} = k,$$

and

$$E\left(z_i^2\right) = \frac{E(\varepsilon_i' P_i \varepsilon_i \varepsilon_i' P_i \varepsilon_i)}{E\left(\varepsilon_i' M_\tau \varepsilon_i \varepsilon_i' M_\tau \varepsilon_i\right)/(T-1)^2} = \left(\frac{T-1}{T+1}\right)\left(k^2 + 2k\right);$$

so that

$$\operatorname{var}(z_i) = v^2(T,k) = \frac{2k(T-1) - 2k^2}{T+1}. \qquad (3.17)$$

These results, therefore, motivate the following standardized version of the $\tilde{S}$ statistic

$$\tilde{\Delta} = \sqrt{N}\left(\frac{N^{-1}\tilde{S} - k}{v(T,k)}\right), \tag{3.18}$$

which in view of (3.15) can also be written as

$$\tilde{\Delta} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{z_i - k}{v(T,k)}\right) + O_p\left(T^{-1}\right) + O_p\left(N^{-1/2}\right).$$

Since $(z_i - k)/v(T,k) \sim IID(0,1)$, using standard central limit theorems it follows that under H_0, $\tilde{\Delta} \to_d N(0,1)$ as $(N,T) \xrightarrow{j} \infty$. The following theorem provides a formal statement of these results.

Theorem 3.2. Consider the panel data model (2.1), suppose that the $k \times 1$ regressors x_{it} are strictly exogenous, $\varepsilon_i \sim IIDN\left(0, \sigma_i^2 I_T\right)$, and Assumptions 1-5 hold. Then under H_0

$$\hat{\Delta} \to_d N(0,1), \text{ as } (N,T) \xrightarrow{j} \infty, \text{ such that } \sqrt{N}/T \to 0,$$

and

$$\tilde{\Delta} \to_d N(0,1), \text{ as } (N,T) \xrightarrow{j} \infty,$$

where the standardized dispersion statistics, $\hat{\Delta}$ and $\tilde{\Delta}$ are defined by (3.16) and (3.18), respectively.

Under strictly exogenous regressors and normal errors the null distribution of the $\tilde{\Delta}$ statistic does not depend on the relative expansion rates of N and T, whilst the same is not true of the Swamy version of the test. The differences between the two versions are, however, less clear cut as the exogeniety and the normality assumptions are relaxed. For example, if the normality assumption is relaxed, to eliminate the dependence of $\tilde{\Delta}$ on the higher order moments of ε_{it}, we also need $\sqrt{N}/T \to 0$, as $(N,T) \xrightarrow{j} \infty$. This result is summarized in the following corollary to Theorem 3.2.

Corollary 3.3. Suppose that the conditions of Theorem 3.2 are met, but the errors, ε_{it}, are not necessarily normally distributed. Instead assume that they are independently distributed over i and t and have finite fourth order moments. Then as $(N,T) \xrightarrow{j} \infty$, $\sqrt{N}\left(N^{-1}\tilde{S} - k\right) \to_d N(0, var(z_i))$; if it is also required that $\sqrt{N}/T \to 0$, as $(N,T) \xrightarrow{j} \infty$. The finite T expression for $var(z_i)$ in the case of non-normal errors would be rather complicated to obtain, but the result in (3.17) derived for the normal error case is likely to provide a reasonable approximation in practice.

See Appendix A.1 for a proof.

Remark 4. The proposed testing approach can be readily extended to testing the homogeneity of a sub-set of slope coefficients. Consider the following partitioned form of (2.1):

$$\underset{T\times 1}{y_i} = \alpha_i \tau_T + \underset{T\times k_1}{X_{i1}}\beta_{i1} + \underset{T\times k_2}{X_{i2}}\beta_{i2} + \varepsilon_i, \ i = 1, 2, \ldots, N;$$

or

$$\underset{T\times 1}{y_i} = \underset{T\times(k_1+1)}{X_{i1}^{*}}\theta_i + \underset{T\times k_2}{X_{i2}}\beta_{i2} + \varepsilon_i;$$

where $X_{i1}^{*} = (\tau_T, X_{i1})$ and $\theta_i = \left(\alpha_i, \beta_{i1}'\right)'$. Suppose the slope homogeneity hypothesis of interest is given by

$$H_0 : \beta_{i2} = \beta_2, \text{ for } i = 1, 2, \ldots, N. \tag{3.19}$$

Our version of the dispersion test statistic in this case is given by

$$\tilde{S}_2 = \sum_{i=1}^{N}\left(\hat{\beta}_{i2} - \tilde{\beta}_{2,WFE}\right)' \frac{X_{i2}'M_{i1}^{*}X_{i2}}{\tilde{\sigma}_i^2}\left(\hat{\beta}_{i2} - \tilde{\beta}_{2,WFE}\right),$$

where

$$\hat{\beta}_{i2} = \left(X_{i2}'M_{i1}^{*}X_{i2}\right)^{-1} X_{i2}'M_{i1}^{*}y_i.$$

$$\tilde{\beta}_{2,WFE} = \left(\sum_{i=1}^{N}\frac{X_{i2}'M_{i1}^{*}X_{i2}}{\tilde{\sigma}_i^2}\right)^{-1}\sum_{i=1}^{N}\frac{X_{i2}'M_{i1}^{*}y_i}{\tilde{\sigma}_i^2};$$

$$M^*_{i1} = I_T - X^*_{i1}\left(X^{*\prime}_{i1}X^*_{i1}\right)^{-1}X^{*\prime}_{i1},$$

$$\sigma_i^2 = \frac{\left(y_i - X_{i2}\hat{\beta}_{2,FE}\right)' M^*_{i1}\left(y_i - X_{i2}\hat{\beta}_{2,FE}\right)}{T_i - k_1 - 1};$$

and

$$\hat{\beta}_{2,FE} = \left(\sum_{i=1}^{N} X_{i2}' M^*_{i1} X_{i2}\right)^{-1} \sum_{i=1}^{N} X_{i2}' M^*_{i1} y_i.$$

Using a similar line of reasoning as above, it is now easily seen that under H_0 defined by (3.19)

$$\Delta_2 = \sqrt{N}\left(\frac{N^{-1}\tilde{S}_2 - k_2}{v(T, k_1, k_2)}\right) \to_d N(0,1), \text{ as } (N,T) \xrightarrow{j} \infty,$$

where

$$v^2(T, k_1, k_2) = \frac{2k_2(T - k_1 - 1) - 2k_2^2}{T - k_1 + 1}.$$

Remark 5. The proposed slope homogeneity tests can also be extended to unbalanced panels. Denoting the number of time series observations on the i^{th} cross section by T_i, our version of the standardized dispersion statistic is given by

$$\Delta = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\tilde{d}_i - k}{v(T_i, k)}\right), \tag{3.20}$$

where

$$v^2(T_i, k) = \frac{2k(T_i - 1) - 2k^2}{T_i + 1},$$

$$\tilde{d}_i = \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right)' \frac{X_i' M_{\tau i} X_i}{\sigma_i^2} \left(\hat{\beta}_i - \tilde{\beta}_{WFE}\right),$$

$X_i = (x_{i1}, x_{i2}, \ldots, x_{iT_i})$, $M_{\tau i} = I_{T_i} - \tau_{T_i}\left(\tau_{T_i}'\tau_{T_i}\right)^{-1}\tau_{T_i}'$ with τ_{T_i} being a $T_i \times 1$ vector of unity,

$$\hat{\beta}_i = \left(X_i' M_{\tau i} X_i\right)^{-1} X_i' M_{\tau i} y_i; \tag{3.21}$$

$$\tilde{\beta}_{WFE} = \left(\sum_{i=1}^{N} \frac{X_i' M_{\tau_i} X_i}{\tilde{\sigma}_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{X_i' M_{\tau_i} y_i}{\tilde{\sigma}_i^2}, \tag{3.22}$$

$y_i = (y_{i1}, y_{i2}, \ldots, y_{iT_i})'$

$$\tilde{\sigma}_i^2 = \frac{\left(y_i - X_i \hat{\beta}_{FE} \right)' M_{\tau_i} \left(y_i - X_i \hat{\beta}_{FE} \right)}{T_i - 1},$$

and

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^{N} X_i' M_{\tau_i} X_i \right)^{-1} \sum_{i=1}^{N} X_i' M_{\tau_i} y_i. \tag{3.23}$$

An extension to testing the homogeneity of a sub-set of slope coefficients in the case of the unbalanced panels is straightforward and is easily derived using the result in Remark 4.

3.1. Asymptotic Local Power of the Δ Test

For the analysis of the asymptotic power of the Δ test, we adopt the following local alternatives[6]

$$H_{1,NT}: \beta_i = \beta + \frac{\delta_i}{N^{1/4} T^{1/2}}, \quad i = 1, 2, \ldots, N, \tag{3.24}$$

where δ_i, $i = 1, 2, \ldots, N$ are $k \times 1$ vectors of fixed constants. As we shall with $N \to \infty$, it is not necessary that δ_i are non-zero for all i.

Under the above local alternatives and assuming that the regressors are strictly exogenous we have[7]

$$\Delta = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{z_i - k}{v(T,k)} \right) + \frac{\psi_{NT}}{v(T,k)} + O_p\left(N^{-1/4} \right) + O_p\left(T^{-1} \right),$$

where

$$\psi_{NT} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2} \delta_i' Q_{iT} \delta_i - \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2} \delta_i' Q_{iT} \right) \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2} Q_{iT} \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^{-2} Q_{iT} \delta_i \right).$$

[6] Similar results also hold for the $\tilde{\Delta}$ version of the test.

[7] For a proof see Appendix A.2.

Hence, it readily follows that under $H_{1,NT}$

$$\Delta \to_d N\left(\frac{\psi}{\sqrt{2k}}, 1\right), \text{ as } (N, T) \xrightarrow{j} \infty,$$

where

$$\psi = \lim_{N\to\infty} \left\{ \frac{1}{N}\sum_{i=1}^{N} \sigma_i^{-2}\delta_i' Q_i \delta_i - \left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^{-2}\delta_i' Q_i\right)\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^{-2} Q_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^{-2} Q_i \delta_i\right)\right\}.$$

Recall that $Q_i = \text{plim}_{T\to\infty}\left(T^{-1}X_i' M_\tau X_i\right)$. The Δ test has power against local alternatives if $\psi > 0$. Since Q_i is a symmetric positive definite matrix, using the the Cholesky decomposition, $Q_i = C_i' C_i$, and setting $\eta_i = C_i \delta_i / \sigma_i$, and $W_i = \sigma_i^{-1} C_i$ we have

$$\begin{aligned}\psi &= \lim_{N\to\infty}\left\{\frac{1}{N}\sum_{i=1}^{N}\eta_i'\eta_i - \left(\frac{1}{N}\sum_{i=1}^{N}\eta_i' W_i\right)\left(\frac{1}{N}\sum_{i=1}^{N} W_i' W_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} W_i'\eta_i\right)\right\} \\ &= \lim_{N\to\infty}\frac{1}{N}\left\{\sum_{i=1}^{N}\eta_i'\eta_i - \left(\sum_{i=1}^{N}\eta_i' W_i\right)\left(\sum_{i=1}^{N} W_i' W_i\right)^{-1}\left(\sum_{i=1}^{N} W_i'\eta_i\right)\right\}.\end{aligned}$$

Let $\eta = \left(\eta_1', \eta_2', \ldots, \eta_N'\right)'$ and $W = \left(W_1', W_2', \ldots, W_N'\right)'$, and write ψ as

$$\psi = \lim_{N\to\infty}\left(\frac{\eta' M_W \eta}{N}\right),$$

where $M_W = I_T - W(W'W)^{-1}W$. Hence, $\psi \geq 0$, and in general the Δ test is asymptotically powerful if $\delta_i \neq 0$ for a non-zero fraction of the cross section units in the limit, as specified under Assumption 6. Such an alternative, for example, allows a sub-set of the slope coefficients and/or a sub-set of cross section units to be homogeneous.

The above result also suggests that the power of the Δ test is likely to increase faster with T than with N.

4. Testing Slope Homogeneity in Autoregressive Models

Consider the stationary p^{th} order autoregressive (AR(p)) processes

$$y_{it} = \alpha_i + \sum_{j=1}^{p} \lambda_{ij} y_{i,t-j} + \varepsilon_{it}, \tag{4.1}$$

$i = 1, 2, \ldots, N$, where the roots of the characteristic equation $1 = \sum_{j=1}^{p} \lambda_{ij} x^j$, fall outside the unit circle, and assume that $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$. Testing the homogeneity of the slopes

$$H_0 : \lambda_{ij} = \lambda_j \text{ for all } i = 1, 2, \ldots, N \text{ and } j = 1, 2, \ldots, p; \tag{4.2}$$

can be carried out as computing the dispersion statistic, (3.1), with

$$\begin{aligned} X_i &= (y_{i,-1}, y_{i,-2}, \ldots, y_{i,-p}); \\ y_{i,-j} &= (y_{i,-j+1}, y_{i,-j+2}, \ldots, y_{i,T-j})'; \; j = 1, 2, \ldots, p. \end{aligned}$$

Using standard results from the literature of stationary autoregressive processes, it is easily established that Assumptions 1-5 are satisfied in the case of stationary autoregressive processes, and as a result Theorem 3.1 continues to hold in this case as well. In particular we have

$$N^{-1/2} \tilde{S} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} z_i + O_p\left(T^{-1}\right) + O_p\left(N^{-1/2}\right); \tag{4.3}$$

where z_i is defined by (3.13), with $X_i = (y_{i,-1}, y_{i,-2}, \ldots, y_{i,-p})$. However, in the case of AR processes exact expressions for the mean and variance of z_i are not easy to derive, and more importantly such exact results would in general depend on the unknown autoregressive coefficients, λ_{ij}, which further complicates any test that is directly based on the Swamy statistic, $\tilde{S}$. To deal with this problem we explore two alternative approaches. (i) An asymptotic procedure where $E(z_i)$ and $Var(z_i)$ are approximated by terms of up to order T^{-1}. (ii) A bootstrap approach where the small sample dependence of $E(z_i)$ and $Var(z_i)$ on $\lambda_i = (\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ip})'$ is taken into account using re-sampling techniques based on bias-corrected estimates of $\hat{\lambda}_i$.

In the case of dynamic models the two versions of the dispersion tests, $\hat{\Delta}$ and Δ, are asymptotically equivalent. Consider the Δ version of the test and using (3.9) in (3.13) first note that

$$z_i = \frac{\left(\frac{\varepsilon_i' M_\tau X_i}{\sqrt{T}}\right)\left(\frac{X_i' M_\tau X_i}{T}\right)^{-1}\left(\frac{X_i' M_\tau \varepsilon_i}{\sqrt{T}}\right)}{(T-1)^{-1}\varepsilon_i' M_\tau \varepsilon_i}. \tag{4.4}$$

Since (4.1) is a stationary process it then readily follows that under H_0

$$z_i \to_d \chi^2_p, \text{ as } T \to \infty.$$

Therefore, it is reasonable to conjecture that up to order T^{-1}, $E(z_i)$ and $Var(z_i)$ are given by p and $2p$, respectively. The proof of this conjecture turns out to be quite complicated. A rigorous proof is given in Appendix A.3 for the AR(1) case where it is established that indeed

$$E(z_i) = 1 + O\left(T^{-1}\right).$$

Supposing now that this result holds more generally, namely

$$E(z_i) = p + O\left(T^{-1}\right), \tag{4.5}$$

and write (4.3) as

$$\sqrt{N}\left(\frac{N^{-1}\tilde{S} - p}{v_z}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{z_i - E(z_i)}{v_z}\right) + \frac{1}{N}\sum_{i=1}^{N}\frac{\sqrt{N}\left[E(z_i) - p\right]}{v_z} + O_p\left(T^{-1}\right);$$

where $Var(z_i) = v_z^2$. Hence, using (4.5) we have

$$\sqrt{N}\left(\frac{N^{-1}\tilde{S} - p}{v_z}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{z_i - E(z_i)}{v_z}\right) + O\left(\frac{\sqrt{N}}{T}\right) + O_p\left(T^{-1}\right).$$

Under H_0, the first term in this expression is scaled sums of i.i.d. random variables and tends to $N(0,1)$ as $N \to \infty$. Therefore, under Assumptions 1-4, and assuming that (4.5) holds, we have (under H_0):

$$\Delta = \sqrt{N}\left(\frac{N^{-1}\bar{S} - p}{\sqrt{2p}}\right) \to_d N(0,1) \text{ as } (N,T) \stackrel{j}{\to} \infty\text{, such that } \frac{\sqrt{N}}{T} \to 0. \tag{4.6}$$

One important implication of this result is that the test is valid even when N increases faster than T, so long as $\sqrt{N}/T \to 0$. The Δ test is clearly more restrictive when applied to dynamic models and requires T to be sufficiently large so that the small sample bias of $E(z_i)$ and $Var(z_i)$ become negligible relative to $\sqrt{N}$.

4.2. Bias-Corrected Bootstrap Tests of Slope Homogeneity for the AR(1) Model

One possible way of improving over the asymptotic test developed for the AR models would be to follow the recent literature and use bootstrap techniques.[8] Here we make use of a bias-corrected version of the recursive bootstrap procedure.[9]

One of the main problems in application of bootstrap techniques to dynamic models in small T samples is the fact that the OLS estimates of the individual coefficients, λ_i, or their FE (or WFE) counterparts are biased when T is small; a bias that persists with $N \to \infty$. To deal with this problem we focus on the AR(1) case and use the bias-corrected version of $\tilde{\lambda}_{WFE}$ as proposed by Hahn and Kuersteiner (2002).[10] Denoting the bias-corrected version of $\tilde{\lambda}_{WFE}$ by $\mathring{\lambda}_{WFE}$, we have

$$\mathring{\lambda}_{WFE} = \tilde{\lambda}_{WFE} + \frac{1}{T}\left(1 + \tilde{\lambda}_{WFE}\right); \tag{4.7}$$

and estimate the associated intercepts as

$$\alpha_{i,WFE} = \bar{y}_i - \mathring{\lambda}_{WFE}\,\bar{y}_{i,-1};$$

[8] For example, see Beran (1988), Horowitz (1994), Li and Maddala (1996) and Bun (2004), although none of these authors make any bias corrections in their bootstrapping procedures.

[9] Bias-corrected estimates are also used in the literature on the derivation of the bootstrap confidence intervals to generate the bootstrap samples in dynamic AR(p) models. See Kilian (1998), among others.

[10] Bias corrections for the OLS estimates of individual λ_i are provided by Marriott and Pope (1954), and further elaborated by Kendall (1954) and Orcutt and Winokur (1969). Bias corrections for the OLS estimates in the case of higher order AR processes are provided in Shaman and Stine (1988). No bias corrections seem to be available for FE or WFE estimates of AR(p) panel data models in the case of $p \geq 2$.

where $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$, and $\bar{y}_{i,-1} = T^{-1}\sum_{t=1}^{T} y_{i,t-1}$. The residuals are given by

$$\tilde{e}_{it} = y_{it} - \alpha_{i,WFE} - \tilde{\lambda}_{WFE}\, y_{i,t-1},$$

with the associated bias-corrected estimator of σ_i^2 given by $\tilde{\sigma}_i^2 = (T-1)^{-1}\sum_{t=1}^{T}(\tilde{e}_{it})^2$. The b^{th} bootstrap sample, $y_{it}^{(b)}$ for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ can now be generated as

$$y_{it}^{(b)} = \alpha_{i,WFE} + \tilde{\lambda}_{WFE}\, y_{i,t-1}^{(b)} + \tilde{\sigma}_i \varepsilon_{it}^{(b)}, \text{ for } t = 1, 2, ..., T,$$

where $y_{i0}^{(b)} = y_{i0}$, and $\varepsilon_{it}^{(b)}$ are random draws with replacements from the set of pooled standardized residuals, $\tilde{e}_{it}/\tilde{\sigma}_i$, $i = 1, 2, ..., N$, and $t = 1, 2, ..., T$. With $y_{it}^{(b)}$, for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$ the bootstrap statistics

$$\Delta^{(b)} = \sqrt{N}\left(\frac{N^{-1}\tilde{S}^{(b)} - 1}{\sqrt{2}}\right), \; b = 1, 2, ..., B,$$

can be computed using (3.1) to obtain the bootstrap p-values

$$p_B = \frac{1}{B}\sum_{b=1}^{B} I\left(\Delta^{(b)} - \Delta\right),$$

where B is the number of bootstrap sample, $I(A)$ takes the value of unity if $A > 0$ or zero otherwise, and Δ is the standardized dispersion statistic applied to the actual observations. If $p_B < 0.05$, say, we reject the null hypothesis of slope homogeneity.

5. Finite Sample Properties of Slope Homogeneity Tests

In this section we shall use Monte Carlo techniques to evaluate the finite sample properties of the alternative tests of slope homogeneity. We shall focus on our proposed test, Δ defined by (3.18) and compare its performance to the Swamy and Hausman tests of slope homogeneity. We also considered the G test of Phillips and Sul (2003), but the G statistic could not be computed due to the singularity problem discussed in Section 2.3.[11] The Swamy's $\hat{S}$ statistic is defined by (2.10) which we consider to be distributed as $\chi^2_{k(N-1)}$ under H_0. For the Hausman test (called H test) we make use of the following statistic[12]

$$H = \left(\hat{\beta}_{MG} - \tilde{\beta}_{WFE}\right)' \hat{V}_H^{-1} \left(\hat{\beta}_{MG} - \tilde{\beta}_{WFE}\right) \overset{a}{\sim} \chi^2_k, \qquad (5.1)$$

where $\hat{\beta}_{MG}$ and $\tilde{\beta}_{WFE}$ are given by (2.8) and (3.3), respectively, and

$$\hat{V}_H = \frac{1}{N^2} \sum_{i=1}^{N} \hat{\sigma}_i^2 \left(X_i' M_\tau X_i\right)^{-1} - \left(\sum_{i=1}^{N} \frac{X_i' M_\tau X_i}{\tilde{\sigma}_i^2} \right)^{-1}, \qquad (5.2)$$

with $\hat{\sigma}_i^2$ and $\tilde{\sigma}_i^2$ being defined by (2.11) and (3.2), respectively. We report empirical size and power of these tests at 5% nominal level, for various pairs of N and T, including cases where N is much larger than T which is often encountered with micro data sets, as well as when $T > N$ which is more prevalent in the case of macro data sets. We consider panels with strictly exogenous regressors, as well as simple dynamic panels.

Initially, we consider the following simple data generating process (DGP):

$$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it},\ t = 1, 2, \ldots, T,\ i = 1, 2, \ldots, N;$$

where $\alpha_i \sim N(1, 1)$, with x_{it} generated as

$$x_{it} = \alpha_i(1 - \rho_i) + \rho_i x_{i,t-1} + \left(1 - \rho_i^2\right)^{1/2} v_{it};\ t = -49, \ldots, 0, \ldots, T,\ i = 1, 2, \ldots, N; \quad (5.3)$$

where $\rho_i \sim IIDU(0.05, 0.95)$, $v_{it} \sim IIDN\left(0, \sigma_{ix}^2\right)$ with $\sigma_{ix}^2 \sim IID\chi^2(1)$. ρ_i and σ_{ix}^2 are fixed across replications. The first 50 observations are discarded to reduce the effect of initial value on the generated values of x_{it}. $\varepsilon_{it} \sim IID\left(0, \sigma_i^2\right)$ is drawn from (i) standard normal distribution and (ii) $\left(\chi^2(2) - 2\right)/2$, and $\sigma_i^2 \sim IID\chi^2(2)/2$.

[11] In e-mail correspondences Dr. Sul has confirmed to us that there is an error in equation (27) in Phillips and Sul (2003) that defines the G statistic.

[12] We also tried a number of other variants of the Hausman test. But they all performed very similarly.

Under the null hypothesis, $\beta_i = 1$ for all i, and under the alternative hypothesis, $\beta_i = 1$ for $i = 1, \ldots, [2N/3]$, and $\beta_j \sim N(1, 0.04)$, for $j = [2N/3] + 1, \ldots, N$, where $[2N/3]$ is the nearest integer value. α_i, β_i, and σ_i^2 are fixed across replications. All combinations of $T = 10, 20, 30, 50, 100, 200$ and $N = 20, 30, 50, 100, 200$ are used as sample sizes.

For examining empirical size and power of the tests in the case of regression models with different numbers of covariates, the following DGP is used:

$$y_{it} = \alpha_i + \sum_{\ell=1}^{k} x_{i\ell t}\beta_{i\ell} + \varepsilon_{it}, \quad i = 1, 2, \ldots, N; \; t = 1, 2, \ldots, T;$$

where, as before, $\alpha_i \sim IIDN(1, 1)$, $x_{i\ell t}$ is generated as specified in (5.3), $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$, $\sigma_i^2 \sim IID\left(k\chi^2(2)/2\right)$, $k = 1, \ldots, 4$, so that the population R^2 of individual equations in the panel are invariant to the number of included regressors. Under the null hypothesis $\beta_{i\ell} = 1$ for all i and ℓ, and under the alternative hypothesis we set $\beta_{i1} \sim IIDN(1, 0.04)$ and $\beta_{i\ell} = \beta_{i1}$ for $\ell = 2, 3, 4$. α_i, $x_{i\ell t}$, $\beta_{i\ell}$, and σ_i^2 are fixed across replications. For these experiments the sample sizes being considered are the combinations of $T = 20$, 30 and $N = 20, 30, 50, 100, 200$.

In the case of dynamic models, two specifications are considered. The first is the AR(1) specification

$$y_{it} = (1 - \lambda_i)\alpha_i + \lambda_i y_{it-1} + \varepsilon_{it}, \; t = -49, \ldots, 0, \ldots, T, \; i = 1, 2, \ldots, N;$$

where $\alpha_i \sim N(1, 1)$, λ_i is specified as (i) $\lambda_i = \lambda = 0.2, 0.4, 0.6, 0.8, 0.9$ under the null hypothesis, and (ii) $\lambda_i \sim IIDU(\lambda - 0.2, \lambda + 0.2)$ for $\lambda = 0.2, 0.4, 0.6, 0.8$ and $\lambda_i \sim IIDU(0.0, 1.0)$, under the alternative hypothesis. $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$ with $\sigma_i^2 \sim IID\chi^2(2)/2$. α_i, λ_i, and σ_i^2 are fixed across replications. The first 49 observations are discarded. For these experiments, we consider the combinations of sample sizes N and $T = 20$, 30, 50, 100, 200. For bootstrap, 499 bootstrap samples are generated and the combinations of the sample sizes $T = 20$, 30, 50 and $N = 20$, 30, 50, 100, 200 are considered.

The second dynamic DGP is:

$$y_{it} = (1 - \lambda_{i1} - \lambda_2)\alpha_i + \lambda_{i1} y_{i,t-1} + \lambda_2 y_{i,t-2} + \varepsilon_{it}, \quad \begin{array}{rcl} t & = & -49, \ldots, 0, \ldots, T, \\ i & = & 1, \ldots, N; \end{array}$$

where $\alpha_i \sim N(1, 1)$, $\lambda_2 = 0.2$, and (i) $\lambda_{1i} = 0.6$ for all i under the null hypothesis, and (ii) $\lambda_{1i} \sim IIDU(0.4, 0.8)$ under the alternative. $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$ with $\sigma_i^2 \sim IID\chi^2(2)/2$. α_i, λ_{1i}, and σ_i^2 are fixed across replications. The first 48 observations are discarded. For these experiments, we consider the combinations of sample sizes N and T = 20, 30, 50, 100, 200.

For all experiments 2,000 replications are used.

5.1. Results

Tables 1 to 3 summarize the results for the DGP with strictly exogenous regressors. First, as predicted by the asymptotic theory, Swamy's $\hat{S}$ test tends to over-reject when N is small relative to T, with the extent of over-rejection diminishing as T is increased relative to T. In the case of $T = 20$ and $N = 200$, more typical of micro data sets, the empirical size of the $\hat{S}$ test is as much as 34%, and only approaches its nominal size of 5% when T is increased to 200. The standardized dispersion test, Δ; and the Hausman test, H, both have correct sizes. The power of the Δ test also seems to be satisfactory. However, as our theory predicts, the H test has no power in the case of these experiments. Table 2 suggests that the effect of non-normal errors might not be very important for the Δ test. Size and power estimates in Tables 1 and 2 are very similar. Even when $N = 200$ and $T = 10$, where Corollary 3.3 predicts that the effects of error non-normality can be most serious for the Δ test, the empirical size of the Δ test is 4.50%. Table 3 reports the size and power of the tests in the case of regression models with different numbers of covariates, $k = 1, 2, 3, 4$. The results are similar to those provided in Table 1, although, considering that we have controlled for the overall fit of the regressions, the power of the Δ test decreases as k increases.

The results for the dynamic DGPs are given in Tables 4 and 5. In the case of these experiments the H test is not valid, and the $\hat{S}$ and $\tilde{\Delta}$ tests are asymptotically equivalent and their validity requires that $\sqrt{N}/T \to 0$ as $(N,T) \xrightarrow{j} \infty$. The results of the Monte Carlo experiments are in line with our theoretical findings. The H statistic is often negative, particularly for values of λ below 0.4, and in cases where it is positive (and hence applicable), the H test exhibits serious over-rejections. The dispersion tests have satisfactory sizes for most combinations of N and T, so long as λ is relatively small, namely $\lambda \leq 0.4$. For these values of λ the $\hat{S}$ test tends to be more powerful than the $\tilde{\Delta}$ test. The $\hat{S}$ test starts to over-reject as λ is increased to 0.6 and beyond. By comparison, the $\tilde{\Delta}$ test only shows evidence of significant over-rejection when λ is increased to 0.9 and only for values of N that are considerably larger than T. For the value of λ in the range of 0.6 to 0.8, the size of the $\tilde{\Delta}$ test continues to be close to its nominal value for all N and T. The same table also illustrates that the $\tilde{\Delta}$ test has reasonable power. Under the alternatives of $\lambda_i \sim IIDU(\lambda - 0.2, \lambda + 0.2)$, the power increases as λ increases, purely because the explanatory power of the estimated model increases. A power comparison of the $\hat{S}$ and $\tilde{\Delta}$ tests for values of $\lambda \geq 0.6$ is complicated by the over-rejection tendency of the former test. Table 5 reports the performance of the tests for the heteroskedastic AR(2) case. Basically the results are similar to those summarized in Table 4 for the AR(1) case.

Table 6 compares the standard normal approximation, (conventional) bootstrap approximation, and Hahn and Kuersteiner (2002) bias-corrected bootstrap approximation of the $\tilde{\Delta}$ test.[13] The bias-corrected bootstrap procedure controls the size remarkably well, even when the value of λ is above 0.8. On the other hand, the conventional bootstrap (non-bias-corrected version) fails to reduce the size distortion of the test. Except when $\lambda = 0.2$ and $T = 20$, the bias-corrected bootstrap method yields reasonable power.

Therefore, in practice, when $N \leq T$ and it is believed that λ is

[13] A bias-corrected bootstrapped test based on $\hat{S}$ could also be considered, but was not pursued as we expected it to perform very similarly to the bias-corrected bootstrapped test based on $\tilde{\Delta}$.

not close to unity (say the value of λ is below 0.8), the asymptotic version of the Δ test is recommended. For all N and T, and with the value of λ around 0.9, the Hahn and Kuersteiner (2002) bias-corrected bootstrapped Δ test seems to be more appropriate.

6. Application: Testing Slope Homogeneity in Earnings Dynamics

In this section we examine the slope homogeneity of the dynamic earnings equations using the Panel Study of Income Dynamics (PSID) data set used in Meghir and Pistaferri (2004). Briefly, these authors select male heads aged 25 to 55 with at least nine years of usable earnings data. The selection process leads to a sample of $2,069$ individuals and $31,631$ individual-year observations. We further select the individuals who have at least 15 observations, and this leaves us with $1,031$ individuals and $19,992$ individual-year observations. Following Meghir and Pistaferri (2004), we also categorize the individuals into three education groups: High School Dropouts (HSD, those with less than 12 grades of schooling), High School Graduates (HSG, those with at least a high school diploma, but no college degree), and College Graduates (CLG, those with a college degree or more). In what follows the earning equations for the different educational backgrounds; HSD, HSG, and CLG are denoted by the superscripts $e = 1, 2,$ and 3, and for the pooled sample by 0. The number of individuals in the three categories are $N^{(1)} = 249$, $N^{(2)} = 531$, and $N^{(3)} = 251$. The panel is unbalanced with $t = 1, \ldots T_i^{(e)}$ and $i = 1, \ldots, N^{(e)}$, and an average time period of around 18 years.

In the research on earnings dynamics, it is standard to adopt a two-step procedure where in the first stage log of real earnings is regressed on a number of control variables such as age, race and year dummies. The dynamics are then modelled based on the residuals from this first stage regression. The use of the control variables and the grouping of the individuals by educational backgrounds is aimed at eliminating (minimizing) the effects of individual heterogeneities at the second stage.

It is, therefore, of interest to examine the extent to which the two-step strategy has been successful in dealing with the heterogeneity problem. With this in mind we follow closely the two-step procedure adopted by Meghir and Pistaferri (2004) and first run regressions of log real earnings, $w_{it}^{(e)}$, on the control variables: a square of "age" ($AGE_{it}^{(e)2}$), race ($WHITE_i^{(e)}$), year dummies ($YEAR(t)$), region of residence ($NE_{it}^{(e)}; CE_{it}^{(e)}; STH_{it}^{(e)}$), and residence in a Standard Metropolitan Statistical Area, ($SMSA_{it}^{(e)}$), for each education group $e = 0; 1; 2; 3$, separately.[14] The residuals from these regressions, which we denote by $y_{it}^{(e)}$, are then used in the second stage to estimate dynamics of the earnings process.

Specifically,

$$y_{it}^{(e)} = \alpha_i^{(e)} + \lambda^{(e)} y_{i,t-1}^{(e)} + \sigma_i^{(e)} \varepsilon_{it}^{(e)}, \quad e = 0; 1; 2; 3,$$

where within each education group $\lambda^{(e)}$ is assumed to be homogeneous across the different individuals. Our interest is to test the hypothesis that $\lambda^{(e)} = \lambda_i^{(e)}$ for all i in e.

The test results are given in the first panel of Table 7. The Δ statistics and the associated bootstrapped p values by education groups all lead to strong rejections of the homogeneity hypothesis. Judging by the size of the Δ statistics, the rejection is stronger for the pooled sample as compared to the sub-samples, confirming the importance of education as a discriminatory factor in the characterizations of heterogeneity of earnings dynamics across individuals. The test results also indicate the possibility of other statistically significant sources of heterogeneity within each of the education groups, and casts some doubt on the two-step estimation procedure adopted in the literature for dealing with heterogeneity; a point recently emphasized by Alvarez, Browning and Ejrnæs (2002).

In Table 7 we also provide a number of different FE estimates of $\lambda^{(e)}$;

[14] Log real earnings are computed as $w_{it}^{(e)} = \ln\left(LABY_{it}^{(e)}/PCED_t\right)$, where $LABY_{it}^{(e)}$ is earnings in current US dollar, and $PCED_t$ is the personal consumption expenditure deflator, base year 1992.

$e = 0, 1, 2, 3$, on the assumption of within group slope homogeneity. Given the relatively small number of time series observations available (on average 18), the bias corrections to the FE estimates are quite large. The cross section error variance heterogeneity also plays an important role in this application, as can be seen from a comparison of FE and WFE estimates with the latter being larger. Focussing on the bias-corrected WFE estimates, we also observe that the persistence of earnings dynamics rises systematically from 0.52 in the case of the school drop outs to 0.72 for the college graduates. This seems sensible, and partly reflects the more reliable job prospects that are usually open to individuals with a higher level of education.

The homogeneity test results suggest that further efforts are needed also to take account of within group heterogeneity. One possibility would be to adopt a Bayesian approach, assuming that $\lambda_i^{(e)}$, $i = 1, 2, ..., N^{(e)}$ are draws from a common probability distribution and focus attention on the whole posterior density function of the persistent coefficients, rather than the average estimates that tend to divert attention from the heterogeneity problem. Another possibility would be to follow Alvarez, Browning and Ejrnæs (2002) and consider particular parametric functions, relating $\lambda_i^{(e)}$ to individual characteristics as a way of capturing within group heterogeneity. Finally, one could consider a finer categorization of the individuals in the panel; say by further splitting of the education groups or by introducing new categories such as occupational classifications. The slope homogeneity tests provide an indication of the statistical importance of the heterogeneity problem, but are silent as how best to deal with the problem.

7. Concluding Remarks

In this paper we have developed simple tests of slope homogeneity in linear panel data models where N could be much larger than T. The proposed tests are based on modifications of Swamy's dispersion statistic and examine the cross section "dispersion" of individual slopes weighted by their relative precisions. It is shown that this test is valid when earlier tests based on Hausman (1978) procedure fail to be applicable. The Monte Carlo evidence shows that the proposed ¢ test has good small sample properties in the case of panel data models with strictly exogenous regressors even if N is much larger than T. The ¢ test has satisfactory performance for moderately large T and N of similar orders of magnitude in the case of stationary dynamic models, when the dominant root of the process is not close to unity. In cases where N is much larger than T and/or the dominant root of the dynamic process is near unity, a bias-corrected bootstrap procedure is proposed which seems to perform reasonably well based on Monte Carlo experiments.

The proposed tests are applied to testing the slope homogeneity of the dynamic earnings equations using PSID data, and the results show evidence of slope heterogeneity, even if attention is confined to the individuals with similar educational backgrounds.

An important further extension of the tests developed in this paper is to consider testing slope homogeneity in panel data models with multi-factor error structures recently examined in Pesaran (2004). This is, however, beyond the scope of the present paper.

. Appendix A: Mathematical Proofs

A.1. Proof of Corollary 3.3

We first note, suppressing the subscript i to simplify the notations, that z_i defined by (3.13) can be write as

$$
\begin{aligned}
z &= \frac{\varepsilon' P \varepsilon}{\varepsilon' M_\tau \varepsilon/(T-1)} = \frac{\varepsilon' P \varepsilon}{(1+W_T)} \\
&= \varepsilon' P \varepsilon \left(1 - W_T + \frac{W_T^2}{1+W_T}\right) \qquad \text{(A.1)}
\end{aligned}
$$

where

$$W_T = \frac{\varepsilon' M_\tau \varepsilon}{(T-1)} - 1;$$

$\varepsilon \sim IID(0, I_T)$ and P is defined by (3.9). Note also that in this case P is a function of strictly exogenous regressors and by Assumption 5 $E\left[1/(1+W_T)\right]$ is bounded.

By using the moments of the quadratic forms in i.i.d. random variables[15]

$$E\left(\varepsilon' P \varepsilon\right) = k,$$

and

$$E\left[\left(\varepsilon' P \varepsilon\right)\left(\varepsilon' M_\tau \varepsilon\right)\right] = \gamma_2 tr\left(P \odot M_\tau\right) + tr(P)\, tr\left(M_\tau\right) + 2tr\left(P M_\tau\right),$$

where γ_2 is the Pearson's measures of kurtosis, which is zero for normal distributions, and $\odot$ signifies Hadamard product. Since $tr\left(P \odot M_\tau\right) = tr\left(P \odot I_T - P \odot T^{-1}\tau_T\tau_T'\right) = T^{-1}(T-1)^{-1}k$, $tr\left(M_\tau\right) = T-1$, $PM_\tau = P$;

$$E\left[\left(\varepsilon' P \varepsilon\right)\left(\varepsilon' M_\tau \varepsilon\right)\right] = \gamma_2 \frac{T-1}{T} k + k(T-1) + 2k,$$

so that the expectation of the second term of (A.1) is

$$E\left[\left(\varepsilon' P \varepsilon\right) W_T\right] = \frac{\gamma_2 k}{T} + \frac{2k}{T-1},$$

which is $O\left(T^{-1}\right)$. Also,

$$
\begin{aligned}
E\left\{\varepsilon' P \varepsilon \left(\frac{W_T^2}{1+W_T}\right)\right\} &\le E\left|\varepsilon' P \varepsilon\right| E\left|W_T^2\right| E\left|\frac{1}{1+W_T}\right| \\
&= O\left(T^{-1}\right)
\end{aligned}
$$

since $E|\varepsilon' P \varepsilon| = O(1)$ and $E\left|1/(1+W_T)\right| = O(1)$, and $E\left(W_T^2\right) = \gamma_2/T + 2/(T-1) = O\left(T^{-1}\right)$; using results in Appendix A.5 of Ullah (2004). Hence,

$$E(z_i) = k + O\left(T^{-1}\right). \qquad \text{(A.2)}$$

[15] For example, see Appendix A.5 in Ullah (2004).

Using (3.15) note that

$$\sqrt{N}\left(N^{-1}\tilde{S}-k\right)=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}[z_i-E(z_i)]+\frac{\sqrt{N}}{T}\left\{\frac{1}{N}\sum_{i=1}^{N}T[E(z_i)-k]\right\}+O_p\left(T^{-1}\right):$$

However, in the light of (A.2) it is clear that

$$\frac{1}{N}\sum_{i=1}^{N}T[k-E(z_i)]=O(1);$$

and if $\sqrt{N}/T\to 0$ as $(N,T)\stackrel{j}{\to}\infty$ it will also follows that

$$\sqrt{N}\left(N^{-1}\tilde{S}-k\right)\to_d N(0,var(z_i)) \text{ as } (N,T)\stackrel{j}{\to}\infty \text{ such that } \sqrt{N}/T\to 0.$$

■

A.2. Proof of Asymptotic Power

Under the local alternatives (defined by (3.24))

$$\beta_i=\beta+\frac{\delta_i}{N^{1/4}T^{1/2}},$$

we first note that[16]

$$\sqrt{T}\left(\hat{\beta}_i-\tilde{\beta}_{WFE}\right)=\kappa_{iNT}+\zeta_{iNT},$$

where

$$\kappa_{iNT}=\bar{Q}_{iT}^{-1}\bar{\xi}_{iT}-N^{-1/2}\bar{Q}_N^{-1}\bar{\xi}_N;$$

and

$$\zeta_{iNT}=\frac{\delta_i}{N^{1/4}}-\frac{1}{N^{1/4}}\bar{Q}_N^{-1}\left(\frac{\sum_{i=1}^{N}\bar{Q}_{iT}\delta_i}{N}\right);$$

with

$$\bar{Q}_{iT}=\sigma_i^{-2}Q_{iT};\quad \bar{\xi}_{iT}=\sigma_i^{-2}\xi_{iT};\tag{A.3}$$

and

$$\bar{Q}_N=N^{-1}\sum_{i=1}^{N}\bar{Q}_{iT};\quad \bar{\xi}_N=N^{-1/2}\sum_{i=1}^{N}\bar{\xi}_{iT}.\tag{A.4}$$

Hence

$$\begin{aligned}N^{-1/2}\tilde{S} &= \frac{T}{\sqrt{N}}\sum_{i=1}^{N}\left(\hat{\beta}_i-\tilde{\beta}_{WFE}\right)'\bar{Q}_{iT}\left(\hat{\beta}_i-\tilde{\beta}_{WFE}\right)\\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\kappa_{iNT}'\bar{Q}_{iT}\kappa_{iNT}+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta_{iNT}'\bar{Q}_{iT}\zeta_{iNT}\\ &\quad+\frac{2}{\sqrt{N}}\sum_{i=1}^{N}\kappa_{iNT}'\bar{Q}_{iT}\zeta_{iNT}.\end{aligned}$$

[16] This relation generalizes (3.7).

The first term is the component of the test statistic that remains under the null hypothesis and is already shown to be given by

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\kappa_{iNT}'Q_{iT}\kappa_{iNT}=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}z_i+O_p\left(T^{-1}\right)+O_p\left(N^{-1/2}\right):$$

Similarly,

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\kappa_{iNT}'Q_{iT}\imath_{iNT}=N^{-1/4}\left[\left(\frac{\sum_{i=1}^{N}\xi_{iT}'\delta_i}{\sqrt{N}}\right)-\xi_N'Q_N^{-1}\left(\frac{\sum_{i=1}^{N}Q_{iT}\delta_i}{N}\right)\right]=O_p\left(N^{-1/4}\right);$$

and

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\imath_{iNT}'Q_{iT}\imath_{iNT}=\psi_{NT};$$

where

$$\psi_{NT}=\frac{1}{N}\sum_{i=1}^{N}\delta_i'Q_{iT}\delta_i-\left(\frac{1}{N}\sum_{i=1}^{N}\delta_i'Q_{iT}\right)\left(\frac{1}{N}\sum_{i=1}^{N}Q_{iT}\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}Q_{iT}\delta_i\right):$$

Therefore

$$N^{-1/2}S=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}z_i+\psi_{NT}+O_p\left(N^{-1/4}\right)+O_p\left(T^{-1}\right):$$

Using this result in (3.18) we have

$$\Delta=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\frac{z_i-k}{v(T;k)}\right)+\frac{\psi_{NT}}{v(T;k)}+O_p\left(N^{-1/4}\right)+O_p\left(T^{-1}\right),$$

as required. ■

A.3. Derivation of $E(z_i)$ in the Case of AR(1) Models with Normal Errors

Suppressing the subscript i to simplify the notations, the AR(1) model is given by

$$y_t=\alpha(1-\lambda)+\lambda y_{t-1}+\varepsilon_t,\ t=1,2,\ldots,T; \qquad \text{(A.5)}$$

where α is bounded on a compact set, $|\lambda|<1$, $\varepsilon_t\sim IIDN\left(0,\sigma^2\right)$ with $0<\sigma^2<\infty$, and it is assumed that the process is initialized with $y_0=\alpha+\varepsilon_0$, and $\varepsilon_0\sim IIDN\left(0,\delta^2\right)$. The choice of δ depends on the initialization of the process and will be given by $\delta=\sigma\left(1-\lambda^2\right)^{-1/2}$ if the process has started at $t=-M$, with $M\to\infty$. For this model specification z defined in (3.13) can be written as

$$z=\frac{\left(T^{-1/2}\varepsilon'M_\tau y_{-1}\right)^2}{\left[(T-1)^{-1}\varepsilon'M_\tau\varepsilon\right]\left(T^{-1}y_{-1}'M_\tau y_{-1}\right)};$$

where $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_T)'$, $y_{-1}=(y_0,y_1,\ldots,y_{T-1})'$; and as before $M_\tau=I_T-\tau_T\left(\tau_T'\tau_T\right)^{-1}\tau_T'$, with τ_T being a $T\times 1$ vector of unity.

Rewrite the AR(1) processes in matrix notations as

$$y^* = \alpha\tau_{T+1} + B^{-1}D\upsilon; \qquad (A.6)$$

where $y^* = (y_0; y_1; \ldots; y_T)'$, $\upsilon = (\varepsilon_0/\delta; \varepsilon_1/\sigma; \ldots; \varepsilon_T/\sigma)'$ so that $\upsilon \sim N(0_{T+1\times1}; I_{T+1})$, $0_{T+1\times1}$ is a $(T+1)\times1$ vector of zeros, I_{T+1} is an identity matrix of order $T+1$, τ_{T+1} is a $(T+1)\times1$ vector of ones, D is a $(T+1)\times(T+1)$ diagonal matrix with its first element equal to δ and the remaining elements equal to σ, and

$$B = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\lambda & 1 \end{bmatrix}.$$

Also $y = G_0y^*$, $y_{-1} = G_1y^*$, where $G_0 = (0_{T\times1}; I_T)$ and $G_1 = (I_T; 0_{T\times1})$. Hence, noting that $M_\tau G_1\tau_{T+1} = 0$ we have

$$z = \frac{(\upsilon'A\upsilon)^2}{(\upsilon'B\upsilon)(\upsilon'C\upsilon)},$$

where

$$A = \frac{G_0'M_\tau G_1B^{-1}D}{\sqrt{T}}; \qquad (A.7)$$

$$B = \frac{G_0'M_\tau G_0}{T-1}; \qquad (A.8)$$

$$C = \frac{DB^{-1\prime}G_1'M_\tau G_1B^{-1}D}{T}. \qquad (A.9)$$

Proposition A.1. Under the stationary AR(1) specification with normal errors given by (A.5), we have

$$\begin{aligned} E(z) &= E\left(\frac{4(\upsilon'A\upsilon)^2}{bc} - \frac{2(\upsilon'A\upsilon)^2(\upsilon'B\upsilon)}{b^2c}\right. \\ &\qquad \left. - \frac{2(\upsilon'A\upsilon)^2(\upsilon'C\upsilon)}{bc^2} + \frac{(\upsilon'A\upsilon)^2(\upsilon'B\upsilon)(\upsilon'C\upsilon)}{b^2c^2}\right) + O\left(T^{-1}\right) \qquad (A.10) \\ &= 1 + O\left(T^{-1}\right), \qquad (A.11) \end{aligned}$$

where υ, A, B, C are defined in (A.6), (A.7), (A.8), (A.9), respectively, and $tr(B) = b$ $(= 1)$; and $tr(C) = c > 0$.

Proof. Firstly we show (A.10), then (A.11). Define[17]

$$\begin{aligned} \upsilon'B\upsilon &= b(1 + X_T); \\ \upsilon'C\upsilon &= c(1 + Y_T); \end{aligned}$$

[17] Note that $b = 1$ and c is $O(1)$. See Appendix B.

where $X_T = b^{-1}(\varepsilon' B \varepsilon - b)$, $Y_T = c^{-1}(\varepsilon' C \varepsilon - c)$. We also note that since by the assumption $\varepsilon \sim N\left(0_{(T+1)\times 1}, I_{T+1}\right)$, and B and C are symmetric positive semi-definite matrices with rank $T-1$, then

$$E\left(\frac{b}{\varepsilon' B \varepsilon}\right) = O(1);$$

and

$$E\left(\frac{c}{\varepsilon' C \varepsilon}\right) = O(1),$$

so long as $T > 3$ (Smith 1988).

Also

$$\begin{aligned}
z &= \frac{(\varepsilon' A \varepsilon)^2}{bc}\left(1 - X_T + \frac{X_T^2}{1+X_T}\right)\left(1 - Y_T + \frac{Y_T^2}{1+Y_T}\right) \\
&= \frac{(\varepsilon' A \varepsilon)^2}{bc}\left[(1 - X_T)(1 - Y_T) + \frac{Y_T^2}{1+Y_T} - \frac{X_T Y_T^2}{1+Y_T} + \frac{X_T^2}{1+X_T} - \frac{Y_T X_T^2}{1+X_T} + \frac{Y_T^2 X_T^2}{(1+X_T)(1+Y_T)}\right].
\end{aligned}$$

As B and C are symmetric positive semi-definite matrices, by Lemma B.1 in Appendix B

$$E\left[X_T^2\right] = \frac{tr\left(B^2\right)}{[tr(B)]^2} = O\left(T^{-1}\right), \quad E\left[Y_T^2\right] = \frac{tr\left(C^2\right)}{[tr(C)]^2} = O\left(T^{-1}\right),$$

so that

$$\begin{aligned}
E[z] &= E\left[\frac{(\varepsilon' A \varepsilon)^2}{bc}(1 - X_T)(1 - Y_T)\right] + O\left(T^{-1}\right) \\
&= E\left[\frac{4(\varepsilon' A \varepsilon)^2}{bc} - \frac{2(\varepsilon' A \varepsilon)^2(\varepsilon' B \varepsilon)}{b^2 c} - \frac{2(\varepsilon' A \varepsilon)^2(\varepsilon' C \varepsilon)}{bc^2} + \frac{(\varepsilon' A \varepsilon)^2(\varepsilon' B \varepsilon)(\varepsilon' C \varepsilon)}{b^2 c^2}\right] \\
&\quad + O\left(T^{-1}\right),
\end{aligned}$$

since

$$\begin{aligned}
E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\frac{Y_T^2}{1+Y_T}\right| &< E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\right| E\left|\frac{1}{1+Y_T}\right| E\left|Y_T^2\right| \\
&= O\left(E\left|Y_T^2\right|\right) = O\left(T^{-1}\right),
\end{aligned}$$

and

$$\begin{aligned}
E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\frac{X_T Y_T^2}{1+X_T}\right| &< E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\right| E|X_T| E\left|\frac{1}{1+X_T}\right| E\left|Y_T^2\right| \\
&= O\left(T^{-1}\right).
\end{aligned}$$

Similarly $E\left|\left[b^{-1}c^{-1}(\varepsilon' A \varepsilon)^2\right]\left[X_T^2(1+X_T)^{-1}\right]\right|$ and $E\left|\left[b^{-1}c^{-1}(\varepsilon' A \varepsilon)^2\right]\left[Y_T X_T^2(1+Y_T)^{-1}\right]\right|$ are at most $O\left(T^{-1}\right)$, and

$$\begin{aligned}
E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\frac{Y_T^2 X_T^2}{(1+X_T)(1+Y_T)}\right| &< E\left|\frac{(\varepsilon' A \varepsilon)^2}{bc}\right| E\left|\frac{1}{1+Y_T}\right| E\left|\frac{1}{1+X_T}\right| E\left|Y_T^2\right| E\left|X_T^2\right| \\
&= O\left(E\left|Y_T^2\right| E\left|X_T^2\right|\right) = O\left(T^{-2}\right).
\end{aligned}$$

Consider now (A.11). By using the moments of the quadratic forms in i.i.d. standard normal random variables[18]

$$E\left[\left(\varepsilon'A\varepsilon\right)^2\right] = [\mathrm{tr}\,(A)]^2 + \mathrm{tr}\left(A^2+A'A\right).$$

Using (B.2) in Appendix B

$$E\left[\left(\varepsilon'A\varepsilon\right)^2\right] = c + O\left(T^{-1}\right).$$

Also, using results in Ullah (2004, Appendix A.4), together with (B.2) and (B.3) in Appendix B, and noting that $\mathrm{tr}\,(AC) = \mathrm{tr}\,(A'C)$,

$$\begin{aligned}
E\left[\left(\varepsilon'A\varepsilon\right)^2\left(\varepsilon'C\varepsilon\right)\right] &= [\mathrm{tr}\,(A)]^2\,\mathrm{tr}\,(C) + 4\mathrm{tr}\left(A^2C\right) + 2\mathrm{tr}\left(A'AC\right) + 2\mathrm{tr}\left(AA'C\right) \\
&\quad +4\mathrm{tr}\,(A)\,\mathrm{tr}\,(AC) + \mathrm{tr}\,(C)\,\mathrm{tr}\left(A^2 + A'A\right) \\
&= \mathrm{tr}\,(C)\,\mathrm{tr}\left(A'A\right) + O(T^{-1}) \\
&= c^2 + O(T^{-1}).
\end{aligned}$$

Next, again using results in Ullah (2004, Appendix A.4), together with (B.2) and (B.4) in Appendix B

$$\begin{aligned}
E\left[\left(\varepsilon'A\varepsilon\right)^2\left(\varepsilon'B\varepsilon\right)\right] &= [\mathrm{tr}\,(A)]^2\,\mathrm{tr}\,(B) + 4\mathrm{tr}\left(A^2B\right) + 2\mathrm{tr}\left(A'AB\right) + 2\mathrm{tr}\left(AA'B\right) \\
&\quad +4\mathrm{tr}\,(A)\,\mathrm{tr}\,(AB) + \mathrm{tr}\,(B)\,\mathrm{tr}\left(A^2 + A'A\right) \\
&= \mathrm{tr}\,(B)\,\mathrm{tr}\left(A'A\right) + O(T^{-1}) \\
&= bc + O(T^{-1}).
\end{aligned}$$

Finally, using results in Ullah (2004, Appendix A.4), together with (B.2) - (B.6) in Appendix B,

$$\begin{aligned}
&E\left[\left(\varepsilon'A\varepsilon\right)^2\left(\varepsilon'B\varepsilon\right)\left(\varepsilon'C\varepsilon\right)\right] \\
&= [\mathrm{tr}\,(A)]^2\,\mathrm{tr}\,(B)\,\mathrm{tr}\,(C) \\
&+ 8\mathrm{tr}\,(A)\,\mathrm{tr}\left(ABC + A'BC\right) \\
&+ \mathrm{tr}\,(B)\left[4\mathrm{tr}\left(A^2C\right) + 2\mathrm{tr}\left(A'AC\right) + 2\mathrm{tr}\left(AA'C\right)\right] \\
&+ \mathrm{tr}\,(C)\left[4\mathrm{tr}\left(A^2B\right) + 2\mathrm{tr}\left(A'AB\right) + 2\mathrm{tr}\left(AA'B\right)\right] \\
&+ 2\mathrm{tr}\left(A^2\right)\mathrm{tr}\,(BC) + 2\mathrm{tr}\left(A'A\right)\mathrm{tr}\,(BC) + 8\mathrm{tr}\,(AB)\,\mathrm{tr}\,(AC) \\
&+ 2\,[\mathrm{tr}\,(A)]^2\,\mathrm{tr}\,(BC) + 4\mathrm{tr}\,(A)\,\mathrm{tr}\,(B)\,\mathrm{tr}\,(AC) + 4\mathrm{tr}\,(A)\,\mathrm{tr}\,(C)\,\mathrm{tr}\,(AB) \\
&+ \mathrm{tr}\,(B)\,\mathrm{tr}\,(C)\,\mathrm{tr}\left(A^2\right) + \mathrm{tr}\,(B)\,\mathrm{tr}\,(C)\,\mathrm{tr}\left(A'A\right) \\
&+ 8\mathrm{tr}\left(A^2BC\right) + 8\mathrm{tr}\left(A'ABC\right) + 8\mathrm{tr}\left(AA'BC\right) + 8\mathrm{tr}\left(A^2CB\right) \\
&+ 8\mathrm{tr}\,(ABAC) + 4\mathrm{tr}\left(A'BAC\right) + 4\mathrm{tr}\left(ABA'C\right) \\
&= \mathrm{tr}\,(B)\,\mathrm{tr}\,(C)\,\mathrm{tr}\left(A'A\right) + O(T^{-1}) \\
&= bc^2 + O(T^{-1}).
\end{aligned}$$

Therefore, we can conclude

$$E\,(z) = 1 + O(T^{-1});$$

as required. ■

[18] For example, see Appendix A.4 Ullah (2004).

. Appendix B: Lemmas

Lemma B.1 Suppose H is a $(T \times T)$ symmetric positive semi-definite matrix with bounded eigenvalues where $\lambda_t(H) \geq 0$ for $t = 0; 1; \ldots; T$, where $\lambda_t(H) = O(1)$. Then,

$$\frac{\operatorname{tr}\left(H^2\right)}{[\operatorname{tr}(H)]^2} = O\left(T^{-1}\right). \tag{B.1}$$

Proof. We first note that

$$\frac{\operatorname{tr}\left(H^2\right)}{[\operatorname{tr}(H)]^2} = \frac{\sum_{t=1}^{T} \lambda_t^2(H)}{\left(\sum_{t=1}^{T} \lambda_t(H)\right)^2} = \frac{T^{-1}\sum_{t=1}^{T} \lambda_t^2(H)}{T\left(T^{-1}\sum_{t=1}^{T} \lambda_t(H)\right)^2}.$$

But

$$\left(\sum_{t=1} \lambda_t(H)\right)^2 = \sum_{t=1} \lambda_t^2(H) + 2\sum_{t>t'} \lambda_t(H)\lambda_{t'}(H);$$

so

$$\left(\sum_{t=1} \lambda_t(H)\right)^2 \geq \sum_{t=1} \lambda_t^2(H).$$

Hence (B.1) allows considering that $\sum_{t=1}^{T} \lambda_t(H) = O(T)$ and $\sum_{t=1}^{T} \lambda_t^2(H) = O(T)$. ■

Lemma B.2 Consider the non-stochastic matrices A, B, and C defined by (A.7), (A.8), and (A.9) in Appendix A.3, respectively. Then,

$$\operatorname{tr}(B) = 1; \operatorname{tr}\left(A'A\right) = \operatorname{tr}(C) = O(1); \operatorname{tr}(A) = O\left(T^{-1/2}\right); \operatorname{tr}\left(A^2\right) = O\left(T^{-1}\right); \tag{B.2}$$

$$\operatorname{tr}\left(A'C\right) = O\left(T^{-1/2}\right); \operatorname{tr}\left(A'AC\right) = O\left(T^{-1}\right); \operatorname{tr}\left(AA'C\right) = O\left(T^{-1}\right); \operatorname{tr}\left(A^2C\right) = O\left(T^{-1}\right); \tag{B.3}$$

$$\operatorname{tr}\left(A'AB\right) = O\left(T^{-1}\right); \operatorname{tr}\left(AA'B\right) = O\left(T^{-1}\right); \operatorname{tr}(AB) = O\left(T^{-3/2}\right); \operatorname{tr}\left(A^2B\right) = O\left(T^{-1}\right); \tag{B.4}$$

$$\operatorname{tr}(BC) = O\left(T^{-1}\right); \operatorname{tr}\left(A'BC\right) = O\left(T^{-3/2}\right); \operatorname{tr}(ABC) = O\left(T^{-3/2}\right); \tag{B.5}$$

and

$$\begin{aligned}&\operatorname{tr}\left(A^2BC\right); \operatorname{tr}\left(A'ABC\right); \operatorname{tr}\left(AA'BC\right); \operatorname{tr}\left(A^2CB\right);\\ &\operatorname{tr}(ABAC); \operatorname{tr}\left(A'BAC\right); \operatorname{tr}\left(ABA'C\right) \text{ are at most } O\left(T^{-2}\right):\end{aligned} \tag{B.6}$$

Proof.

We first note that

$$H_{01} = G_0' M_\tau G_1 = \begin{pmatrix} 0_{1\times T} & 0_{1\times 1} \\ M_\tau & 0_{T\times 1} \end{pmatrix};$$

and

$$G_0' M_\tau G_0 = \begin{pmatrix} 0_{1\times 1} & 0_{1\times T} \\ 0_{T\times 1} & M_\tau \end{pmatrix}; \ G_1' M_\tau G_1 = \begin{pmatrix} M_\tau & 0_{T\times 1} \\ 0_{1\times T} & 0_{1\times 1} \end{pmatrix}:$$

The matrices $G_0' M_\tau G_0$ and $G_1' M_\tau G_1$ are idempotent with two zero eigenvalues and $T - 1$ unit eigenvalues. Therefore, noting that B is a lower triangular matrix with unit diagonal elements and D is a diagonal matrix with $\sigma_{max} = \mathrm{Max}(\sigma;\delta) < K < 1$ we have, using (A.9),

$$0 \leq \nu_t(C) \leq \frac{\sigma_{max}}{T};$$

where $\nu_t(C)$ for $t = 0;1;...;T$ are the eigenvalues of C. Also it is easily verified that

$$G_0 G_0' = I_T\ ,\ A'A = C; \tag{B.7}$$

and

$$A'B = \frac{B'^{-1}G_1' M_\tau G_0 G_0' M_\tau G_0}{T^{1/2}(T-1)} = (T-1)^{-1}A'; \tag{B.8}$$

$$AA'B = (T-1)^{-1}AA'. \tag{B.9}$$

To prove the results in (B.2), we first note that

$$tr(B) = 1;\ tr(C) = \sum_{t=0}^{T} \nu_t(C) \leq \frac{(T+1)\sigma_{max}}{T} = O(1). \tag{B.10}$$

Since σ_{max} is bounded, to simplify the derivations and without loss of generality in what follows we set $\delta = \sigma = 1$; (so that $D = I_{T+1}$) and note that

$$B^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lambda^{T-1} & \lambda^{T-2} & \cdots & 1 & 0 \\ \lambda^{T} & \lambda^{T-1} & \cdots & \lambda & 1 \end{bmatrix};$$

$$\begin{aligned} A &= T^{-1/2} G_0' M_\tau G_1 B^{-1} \\ &= T^{-1/2}(E - F), \end{aligned} \tag{B.11}$$

$$E = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \\ \lambda & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \lambda^{T-2} & \lambda^{T-3} & \cdots & 1 & 0 & 0 \\ \lambda^{T-1} & \lambda^{T-2} & \cdots & \lambda & 1 & 0 \end{bmatrix},\ F = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ g_{T-1} & g_{T-2} & \cdots & g_0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{T-2} & g_{T-2} & \cdots & g_0 & 0 \\ g_{T-1} & g_{T-2} & \cdots & g_0 & 0 \end{bmatrix},$$

where

$$g_\ell = \frac{1}{T}\sum_{j=0}^{\ell} \lambda^j = \frac{1}{T}\left(\frac{1-\lambda^{\ell+1}}{1-\lambda}\right) = O\left(T^{-1}\right) \text{ (since } |\lambda| < 1\text{), for } \ell = 0;1;...;T-1.$$

Therefore,

$$tr(A) = \frac{-1}{\sqrt{T}}\sum_{\ell=0}^{T-2} g_\ell = \frac{-1}{T\sqrt{T}}\sum_{\ell=0}^{T-2}\left(\frac{1-\lambda^{\ell+1}}{1-\lambda}\right) = O(T^{-1/2}). \tag{B.12}$$

Consider now $\mathrm{tr}\left(A^2\right)$. Using (B.11)

$$\mathrm{tr}\left(A^2\right) = T^{-1}\left[\mathrm{tr}\left(E^2\right) + \mathrm{tr}\left(F^2\right) - 2\,\mathrm{tr}\,(EF)\right]. \tag{B.13}$$

But it is easily seen that

$$\begin{aligned}
\mathrm{tr}\left(E^2\right) &= 0;\\
\mathrm{tr}\left(F^2\right) &= \left(\sum_{\ell=0}^{T-1} g_\ell\right)\left(\sum_{\ell=0}^{T-2} g_\ell\right) = O(1);\\
\mathrm{tr}\,(EF) &= \sum_{\ell=0}^{T-3}\left(\frac{1-\lambda^{T-\ell-2}}{1-\lambda}\right) g_\ell = \frac{1}{T}\sum_{\ell=0}^{T-3}\left(\frac{1-\lambda^{T-\ell-2}}{1-\lambda}\right)\left(\frac{1-\lambda^{\ell+1}}{1-\lambda}\right) = O(1);
\end{aligned}$$

which together with (B.13) establishes that $\mathrm{tr}\left(A^2\right) = O\left(T^{-1}\right)$.

To prove the results in (B.3), we observe that[19]

$$\mathrm{tr}\left(A'AC\right) = \mathrm{tr}\left(C^2\right) = \sum_{t=0}^{T} o_t^2(C) \le \frac{\sigma_{\max}}{T} = O\left(T^{-1}\right).$$

By Cauchy-Schwarz inequality

$$\left[\mathrm{tr}\left(AA'C\right)\right]^2 \le \mathrm{tr}\left(AA'AA'\right)\mathrm{tr}\left(C'C\right) = \mathrm{tr}\left[\left(A'A\right)^2\right]\mathrm{tr}\left(C^2\right) = \left[\mathrm{tr}\left(C^2\right)\right]^2;$$

which establishes $|\mathrm{tr}\,(AA'C)| = O\left(T^{-1}\right)$. Similarly, again by Cauchy-Schwarz inequality and noting that $A'A = C$,

$$\left[\mathrm{tr}\left(A^2C\right)\right]^2 \le \mathrm{tr}\left(AAA'A'\right)\mathrm{tr}\left(C^2\right) = \mathrm{tr}\left(AA'C\right)\mathrm{tr}\left(C^2\right);$$

which establishes $\left|\mathrm{tr}\left(A^2C\right)\right| = O\left(T^{-1}\right)$. To derive the order of $\mathrm{tr}\,(A'C)$, again by Cauchy-Schwarz inequality

$$\left[\mathrm{tr}\left(A'C\right)\right]^2 \le \mathrm{tr}\left(A'A\right)\mathrm{tr}\left(C'C\right) = \mathrm{tr}(C)\mathrm{tr}(C^2).$$

Therefore, since $\mathrm{tr}(C) = O(1)$, it follows that $|\mathrm{tr}\,(A'C)| = O(T^{-1/2})$.

To establish the results in (B.4), by Cauchy-Schwarz inequality

$$\left[\mathrm{tr}\left(A^2B\right)\right]^2 \le \mathrm{tr}\left(AA'C\right)\mathrm{tr}\left(B^2\right).$$

But

$$\mathrm{tr}\left(B^2\right) = \frac{\mathrm{tr}\left[\left(G_0'M_\tau G_0\right)^2\right]}{(T-1)^2} = \frac{\mathrm{tr}\left[\left(G_0'M_\tau G_0\right)\right]}{(T-1)^2} = \frac{1}{T-1} = O\left(T^{-1}\right),$$

hence, $\left|\mathrm{tr}\left(A^2B\right)\right| = O\left(T^{-1}\right)$. Similarly,

$$\left[\mathrm{tr}\left(A'AB\right)\right]^2 = \left[\mathrm{tr}\,(CB)\right]^2 \le \mathrm{tr}\left(C^2\right)\mathrm{tr}\left(B^2\right) = O\left(T^{-2}\right);$$

[19] Recall that $C' = C$ and $A'A = C$.

which establishes $|\mathrm{tr}(A'AB)| = O\left(T^{-1}\right)$. Using (B.9)

$$\mathrm{tr}\left(AA'B\right) = (T-1)^{-1}\mathrm{tr}\left(A'A\right) = (T-1)^{-1}\mathrm{tr}(C) = O\left(T^{-1}\right).$$

Also

$$\begin{aligned}\mathrm{tr}(AB) &= T^{-1/2}(T-1)^{-1}\mathrm{tr}\left(G_0'M_\tau G_1 B'^{-1} G_0' M_\tau G_0\right)\\ &= T^{-1/2}(T-1)^{-1}\mathrm{tr}\left(G_0'M_\tau G_1 B'^{-1}\right) = \frac{1}{T-1}\mathrm{tr}(A) = O\left(T^{-3/2}\right).\end{aligned}$$

To prove the results in (B.5), a further application of the Cauchy-Schwarz inequality to A and BC now yields

$$\begin{aligned}\left[\mathrm{tr}\left(A'BC\right)\right]^2 &\leq \mathrm{tr}\left(A'A\right)\mathrm{tr}\left(C'B'BC\right) = \mathrm{tr}(C)\mathrm{tr}\left(B^2C^2\right);\\ [\mathrm{tr}(ABC)]^2 &\leq \mathrm{tr}\left(AA'\right)\mathrm{tr}\left(C'B'BC\right) = \mathrm{tr}(C)\mathrm{tr}\left(B^2C^2\right).\end{aligned}$$

But as easily seen
$$\left[\mathrm{tr}\left(B^2C^2\right)\right]^2 \leq \mathrm{tr}\left(B^4\right)\mathrm{tr}\left(C^4\right) \leq O\left(T^{-6}\right)$$

so that
$$\left|\mathrm{tr}\left(B^2C^2\right)\right| \leq O\left(T^{-3}\right),$$

and hence
$$\left|\mathrm{tr}\left(A'BC\right)\right| = O(T^{-3/2}); \text{ and } |\mathrm{tr}(ABC)| = O(T^{-3/2}).$$

Similarly,
$$[\mathrm{tr}(BC)]^2 \leq \mathrm{tr}\left(B^2\right)\mathrm{tr}\left(C^2\right) = O\left(T^{-2}\right);$$

and $|\mathrm{tr}(BC)| = O(T^{-1})$.

Finally, the various higher order terms in (B.6) can be established following similar lines. Firstly,
$$\mathrm{tr}\left(A'ABC\right) = \mathrm{tr}(BC^2) \leq \mathrm{tr}(B^2)\mathrm{tr}(C^4) = O(T^{-4});$$
so that $\left|\mathrm{tr}(BC^2)\right| = O(T^{-2})$; and

$$\begin{aligned}\left[\mathrm{tr}\left(A^2BC\right)\right]^2 &\leq \mathrm{tr}\left(AA'C\right)\mathrm{tr}\left(C^2B^2\right) = O\left(T^{-4}\right);\\ \left[\mathrm{tr}\left(A^2CB\right)\right]^2 &\leq \mathrm{tr}\left(AA'C\right)\mathrm{tr}\left(C^2B^2\right) = O\left(T^{-4}\right).\end{aligned}$$

Similarly,

$$[\mathrm{tr}(ABAC)]^2 \leq \mathrm{tr}\left(ABB'A'\right)\mathrm{tr}\left(C'A'AC\right) = \mathrm{tr}\left(B^2C\right)\mathrm{tr}\left(C^3\right) = O(T^{-4}).$$

Furthermore,

$$\left[\mathrm{tr}\left(AA'BC\right)\right]^2 = \left[\mathrm{tr}\left(A'BCA\right)\right]^2 \leq \mathrm{tr}\left(A'BB'A\right)\mathrm{tr}\left(A'C'CA\right) = \mathrm{tr}\left(B^2AA'\right)\mathrm{tr}\left(C^2AA'\right);$$

and

$$\begin{aligned}\left[\mathrm{tr}\left(A'BAC\right)\right]^2 &\leq \mathrm{tr}\left(A'BB'A\right)\mathrm{tr}\left(C'A'AC\right) = \mathrm{tr}\left(B^2AA'\right)\mathrm{tr}\left(C^3\right);\\ \left[\mathrm{tr}\left(ABA'C\right)\right]^2 &\leq \mathrm{tr}\left(ABB'A'\right)\mathrm{tr}\left(C'AA'C\right) = \mathrm{tr}\left(B^2C\right)\mathrm{tr}\left(C^2AA'\right).\end{aligned}$$

Also using (B.8) and (B.9) we have

$$\mathrm{tr}\left(AA'B^2\right) = \frac{1}{T-1}\mathrm{tr}\left(AA'B\right) = \frac{1}{(T-1)^2}\mathrm{tr}\left(AA'\right) = \frac{\mathrm{tr}(A'A)}{(T-1)^2} = O(T^{-2}).$$

$$\begin{aligned}\left[\mathrm{tr}\left(C^2AA'\right)\right]^2 &= \left[\mathrm{tr}\left(AA'C^2\right)\right]^2 \le \mathrm{tr}\left(AA'AA'\right)\mathrm{tr}\left(C^4\right) = \mathrm{tr}\left(A'AA'A\right)\mathrm{tr}\left(C^4\right)\\ &= \mathrm{tr}(C^2)\mathrm{tr}\left(C^4\right) = O(T^{-4}).\end{aligned}$$

Finally, it is easily established that

$$\mathrm{tr}\left(B^2C\right) = O(T^{-2});\ \mathrm{tr}\left(C^3\right) = O(T^{-2}).$$

Hence all the terms in (B.6) are of order $O(T^{-2})$. ■

References

Alvarez, J., Browning, M., Ejrnæs, M., (2002). Modelling income processes with lots of heterogeneity. Presented at: 10th International Conference on Panel Data, Berlin, July 5-6, 2002.

Andrews, D.W.K., (1993). Exactly median-unbiased estimation of …rst order autoregressive/unit root models. Econometrica 61, 139-165.

Beran, R., (1988). Prepivoting test statistics: A bootstrap view of asymptotic re…nements. Journal of the American Statistical Association 83, 687-697.

Bun, M.J.G., (2004). Testing poolability in a system of dynamic regressions with nonspherical disturbances. Empirical Economics 29, 89-106.

Hahn, J., Kuersteiner, G., (2002). Asymptotically unbiased inference for a dynamic panel model with …xed e¤ects when both n and T are large. Econometrica 70, 1639-1657.

Hausman, J.A., (1978). Speci…cation tests in econometrics. Econometrica 46, 1251-1271.

Horowitz, J.L., (1994). Bootstrap-based critical values for the information matrix test. Journal of Econometrics 61, 395-411.

Hsiao, C., (2003). Analysis of Panel Data, second edition, Cambridge: Cambridge University Press.

Hsiao, C., Pesaran, M.H., (2004). Random coe¢cient panel data models. CESifo Working Paper No.1229.

Kendall, M.G., (1954). Note on bias in the estimation of autocorrelation. Biometrika 41, 403-404.

Kilian, L., (1998). Con…dence intervals for impulse responses under departures from normality. Econometric Reviews 17, 1-29.

Li, H., Maddala, G.S., (1996). Bootstrapping time series models. Econometric Reviews 15, 115-158.

Marriott, F.H.C., Pope, J.A., (1954). Bias in the estimation of autocorrelations. Biometrika 41, 390-402.

Meghir, C., Pistaferri, L., (2004). Income variance dynamics and heterogeneity. Econometrica 72, 1-32.

Orcutt, G.H., Winokur, H.S., (1969). First order autoregression: Inference, estimation, and prediction. Econometrica 37, 1-14.

Pesaran, M.H., (2004). Estimation and inference in large heterogeneous panels with a multifactor error structure. CESifo Working Paper No.1331.

Pesaran, H., Smith, R., Im, K.S., (1996). Dynamic linear models for heterogenous panels. In: Matyas, L., Sevestre, P., eds, The Econometrics of Panel Data: A Handbook of the Theory with Applications, second revised edition, pp. 145-195.

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

Shaman, P., Stine, R.A., (1988). The bias of autoregressive coe¢cient estimators. Journal of the American Statistical Association 83, 842-848.

Smith, M.D., (1988). Convergent series expressions for inverse moments of quadratic forms in normal variables. The Australian Journal of Statistics 30, 235-246.

Swamy, P.A.V.B., (1970). E¢cient inference in a random coe¢cient regression model. Econometrica 38, 311-323.

Ullah, A., (2004). Finite Sample Econometrics. Oxford University Press.

von Neumann, J., (1941). Distribution of the ratio of the mean square successive di¤erence to the variance. The Annals of Mathematical Statistics 12, 367-395.

Zellner, A., (1962). An e¢cient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association 57, 348-368.

Table 1 : Size and Power of the Slope Homogeneity Tests, with Strictly Exogenous Regressors: Normal Errors

N\T	10	20	30	50	100	200
Size						
$\hat{S}$ test						
20	25.25	11.70	9.75	7.70	6.00	6.40
30	31.20	13.45	10.60	8.10	5.45	5.00
50	39.45	17.75	10.70	8.70	6.50	6.00
100	61.05	21.15	15.00	10.65	7.50	5.95
200	82.35	33.90	18.30	12.90	7.95	6.65
H test						
20	7.95	6.55	4.55	6.35	5.20	4.10
30	8.75	6.95	6.50	4.70	4.90	5.35
50	6.60	5.25	5.95	4.80	5.15	5.50
100	6.85	6.20	5.85	5.00	5.30	5.40
200	9.10	5.90	5.40	6.30	5.80	5.30
$\tilde{\Delta}$ test						
20	4.60	4.20	3.70	3.95	3.80	4.05
30	4.95	4.50	4.30	4.75	3.80	4.15
50	4.85	4.80	4.05	4.35	4.60	4.90
100	4.40	5.00	4.95	5.60	5.00	4.80
200	5.20	5.75	4.55	4.85	4.70	4.95
Power						
$\hat{S}$ test						
20	29.05	21.80	23.30	31.00	58.65	88.80
30	43.45	45.25	52.40	81.80	99.10	100.00
50	64.50	68.05	80.15	96.65	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00
200	100.00	100.00	100.00	100.00	100.00	100.00
H test						
20	6.65	5.15	5.60	5.70	5.40	3.55
30	7.15	5.70	5.55	5.20	5.30	5.90
50	6.70	5.75	5.20	5.10	5.90	5.05
100	6.00	6.30	5.90	5.20	5.85	5.45
200	6.35	6.30	4.65	5.00	6.10	5.70
$\tilde{\Delta}$ test						
20	4.65	6.90	9.45	16.60	43.65	82.55
30	7.10	15.65	26.45	65.65	98.15	100.00
50	8.55	26.85	46.80	85.95	99.90	100.00
100	26.80	71.35	99.25	100.00	100.00	100.00
200	44.95	99.15	100.00	100.00	100.00	100.00

Notes: $\hat{S}$ test, $\tilde{\Delta}$ test; and H test statistics are defined in (2.10), (3.18), and (5.1), respectively. The DGP is specified as $y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}$, $t = 1, 2, \dots, T$, $i = 1, 2, \dots, N$ where $\alpha_i \sim N(1, 1)$, with x_{it} generated as $x_{it} = \alpha_i(1 - \rho_i) + \rho_i x_{i,t-1} + \left(1 - \rho_i^2\right)^{1/2} v_{it}$; $t = -49, \dots, 0, \dots, T$, $i = 1, 2, \dots, N$ where $\rho_i \sim IIDU(0.05, 0.95)$, $v_{it} \sim IIDN\left(0, \sigma_{ix}^2\right)$ with $\sigma_{ix}^2 \sim IID\chi^2(1)$. ρ_i and σ_{ix}^2 are fixed across the replications. The first 50 observations are discarded to reduce the effect of initial value on the generated values of x_{it}. $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$, with $\sigma_i^2 \sim IID\chi^2(2)/2$. Under the null hypothesis, $\beta_i = 1$ for all i, and under the alternative hypothesis, $\beta_i = 1$ for $i = 1, \dots, [2N/3]$, and $\beta_j \sim N(1, 0.04)$, for $j = [2N/3] + 1, \dots, N$, where $[2N/3]$ is the nearest integer value. α_i, β_i, and σ_i^2 are fixed across replications. All tests are conducted at 5% nominal level. All the experiments are based on $2,000$ replications.

Table 2 : Size and Power of Slope Homogeneity Tests, with Strictly Exogenous Regressors: $\left(\chi^2(2) - 2\right)/2$ Errors

N\T	10	20	30	50	100	200
Size						
$\hat{S}$ test						
20	23:85	13:45	8:50	5:85	5:30	5:55
30	30:90	13:10	10:15	8:40	6:40	4:75
50	40:70	15:50	11:15	8:85	6:30	6:80
100	59:65	23:00	14:70	9:55	7:35	5:35
200	81:50	32:35	19:30	11:40	8:60	6:40
H test						
20	6:45	6:60	6:45	4:95	4:90	5:90
30	7:60	5:75	5:90	5:20	5:00	5:05
50	7:35	6:10	5:70	5:15	5:30	4:50
100	6:75	6:55	5:25	5:05	5:25	4:40
200	8:05	6:35	6:15	5:40	4:70	5:60
Δ test						
20	3:50	4:35	3:30	3:35	3:95	3:80
30	4:65	3:75	4:10	4:30	4:45	4:20
50	5:25	3:95	4:45	4:60	4:80	5:20
100	6:15	4:70	4:60	4:65	3:65	4:65
200	4:50	4:40	3:65	4:40	4:55	4:70
Power						
$\hat{S}$ test						
20	28:55	23:15	24:20	33:15	61:85	89:90
30	49:95	49:90	59:10	75:85	99:20	100:00
50	67:10	67:95	84:30	91:95	100:00	100:00
100	99:75	99:95	100:00	100:00	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00
H test						
20	6:20	5:40	5:05	4:30	4:75	6:25
30	7:30	6:00	5:20	5:45	5:25	6:30
50	7:40	6:20	5:10	5:65	5:50	5:70
100	6:80	5:40	4:80	5:30	6:20	5:55
200	6:40	5:20	5:35	5:15	5:30	5:70
Δ test						
20	4:40	7:10	9:95	18:65	48:30	84:25
30	7:00	19:15	33:15	60:40	98:10	100:00
50	9:95	26:15	54:40	79:35	99:95	100:00
100	26:55	77:40	99:00	100:00	100:00	100:00
200	62:25	99:60	100:00	100:00	100:00	100:00

Notes: See the notes on Table 1. The design is the same as that of Table 1 except $\varepsilon_{it} \sim IID\left(\left(\chi^2(2) - 2\right)/2\right)$.

Table 3 : Size and Power of the Slope Homogeneity Tests with Strictly Exogenous Regressors
with Di¤erent Numbers of Covariates (k)

	k = 1		k = 2		k = 3		k = 4	
NnT	20	30	20	30	20	30	20	30
SIZE								
Ŝ test								
20	11:40	8:35	16:55	10:80	20:40	13:70	26:50	16:30
30	13:40	11:20	21:25	13:10	25:70	15:70	31:90	17:05
50	15:80	11:60	24:85	15:90	36:45	20:15	43:00	22:45
100	22:45	15:25	35:90	20:90	48:90	26:85	60:90	33:95
200	34:50	20:10	54:55	30:05	72:10	41:95	83:10	50:30
H test								
20	5:60	5:40	6:20	6:45	6:50	5:40	5:45	5:90
30	5:50	6:00	6:45	5:30	6:55	5:95	6:75	6:40
50	5:45	5:75	6:85	6:30	6:75	6:30	7:65	8:15
100	6:90	5:80	7:10	5:60	6:50	5:85	6:10	6:05
200	5:60	5:70	6:20	5:15	5:75	6:70	7:05	6:05
¢ test								
20	3:80	3:20	5:00	4:75	5:20	5:35	5:60	6:05
30	3:55	4:75	5:50	4:70	5:10	5:05	5:00	5:20
50	4:60	4:15	5:55	4:75	5:40	6:05	4:70	5:05
100	4:80	4:35	5:25	5:65	5:15	5:05	5:60	5:20
200	5:05	5:25	4:80	5:50	5:30	5:25	4:70	4:75
Power								
Ŝ test								
20	75:15	94:15	75:65	90:25	71:40	87:10	62:70	69:60
30	94:25	98:95	74:00	93:85	81:55	92:00	94:50	98:70
50	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
100	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
H test								
20	6:90	8:40	6:40	7:05	6:25	6:90	6:75	5:90
30	5:95	4:55	7:35	5:95	6:45	6:15	6:20	5:80
50	6:30	5:40	6:70	5:85	6:90	6:65	6:35	5:75
100	4:85	5:65	6:00	4:95	6:55	6:75	6:80	6:30
200	6:10	5:10	6:90	5:25	6:00	5:80	5:95	5:65
¢ test								
20	46:35	84:50	38:10	67:30	21:70	52:80	10:40	25:05
30	70:15	94:25	28:50	68:35	26:55	58:15	32:05	73:05
50	100:00	100:00	99:05	100:00	94:85	100:00	75:70	99:90
100	99:65	100:00	98:70	100:00	98:25	100:00	98:45	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	99:95	100:00

Notes: The DGP is speci…ed as $y_{it} = ®_i + \sum_{`=1}^{k} x_{i`t}¯_{i`} + "_{it}$; $i = 1; 2; :::; N; t = 1; 2; :::; T$, where $®_i$ » IIDN (1; 1), $x_{i`t}$ is generated as speci…ed in the notes to Table 1, $"_{it}$ » IIDN $¡0; ¾_i^2¢$, where $¾_i^2$ » IID $¡kÂ^2(2)=2¢$, $k = 1; :::; 4$, so that the goodness of …t for each equation in the panel i is invariant to the number of regressors. Under the null hypothesis $¯_{i`} = 1$ for all $i$ and `, and under the alternative hypothesis we generates $¯_{i`}$ as $¯_{i1}$ » IIDN (1; 0:04) and $¯_{i`} = ¯_{i1}$ for ` = 2; 3; 4. $®_i$, $x_{i`t}$, $¯_{i`}$, and $¾_i^2$ are …xed across replications.

Table 4 : Size and Power of the Slope Homogeneity Tests for Heteroskedastic AR(1) Specifications

N\T	Size 20	30	50	100	200	Power 20	30	50	100	200
	$\lambda_i = 0.2$ for all i					$\lambda_i \sim IIDU(0.0, 0.4)$				
$\hat{S}$ test										
20	4.60	5.10	5.60	4.35	4.50	13.85	22.55	42.95	82.55	99.75
30	5.50	5.60	4.40	4.80	6.30	20.25	35.25	62.70	96.45	100.00
50	4.50	4.90	4.05	4.60	4.15	24.75	45.30	78.85	99.70	100.00
100	3.55	4.60	4.40	5.50	4.75	41.75	74.90	98.45	100.00	100.00
200	3.30	5.00	4.85	5.00	5.80	62.30	93.85	99.95	100.00	100.00
H test										
20	38.50	53.40	65.75	73.40	78.00	51.35	75.05	93.80	99.40	100.00
30	45.55	65.25	79.75	88.30	91.45	64.70	89.45	99.00	100.00	100.00
50	62.80	85.25	96.30	98.60	99.50	93.80	99.60	100.00	100.00	100.00
100	89.15	98.85	100.00	100.00	100.00	97.70	99.95	100.00	100.00	100.00
200	99.65	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
$\tilde{\Delta}$ test										
20	2.55	2.20	2.80	3.10	2.90	2.40	8.80	25.80	71.75	99.20
30	3.05	2.95	3.15	4.00	3.95	3.70	14.25	43.55	92.55	100.00
50	4.55	4.10	4.05	4.40	3.70	3.75	18.10	59.00	99.30	100.00
100	8.90	5.80	4.65	4.40	4.35	6.50	39.85	93.55	100.00	100.00
200	18.95	8.95	6.45	4.25	4.55	9.25	65.90	99.80	100.00	100.00
	$\lambda_i = 0.4$ for all i					$\lambda_i \sim IIDU(0.2, 0.6)$				
$\hat{S}$ test										
20	5.70	5.55	6.25	4.70	4.25	16.15	26.10	49.25	87.45	99.95
30	5.95	6.20	5.55	4.75	6.50	24.05	41.40	71.90	98.55	100.00
50	6.40	6.15	5.10	5.80	4.75	31.55	53.40	85.95	99.85	100.00
100	5.65	6.40	5.60	6.15	5.35	54.15	82.75	99.35	100.00	100.00
200	6.60	6.40	6.80	5.55	5.60	85.80	98.70	100.00	100.00	100.00
H test										
20	73.25	87.20	92.80	94.65	96.20	83.45	95.55	99.65	100.00	100.00
30	90.30	96.45	99.15	99.40	99.70	96.75	99.65	100.00	100.00	100.00
50	99.15	99.95	100.00	100.00	100.00	99.95	100.00	100.00	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
$\tilde{\Delta}$ test										
20	2.80	2.20	3.00	3.20	3.15	3.70	11.05	31.75	78.75	99.85
30	2.70	3.05	3.40	3.85	4.70	4.60	18.40	53.15	96.60	100.00
50	3.40	3.65	3.95	4.55	3.40	6.35	25.15	69.35	99.70	100.00
100	6.35	4.65	4.15	4.05	4.70	10.15	54.00	97.60	100.00	100.00
200	12.50	6.30	5.65	4.05	5.00	18.45	81.30	99.95	100.00	100.00

Notes: See notes to Table 1. $\hat{S}$ test, $\tilde{\Delta}$ test; and H test statistics are defined in (2.10), (4.6), and (5.1), respectively. The DGP is specified as $y_{it} = (1-\lambda_i)\alpha_i + \lambda_i y_{i,t-1} + \varepsilon_{it}$, $t = -49, \ldots, 0, \ldots, T$, $i = 1, 2, \ldots, N$, where; $\alpha_i \sim N(1,1)$; λ_i are as specified in the table; $\varepsilon_{it} \sim IIDN\left(0, \sigma_i^2\right)$ with $\sigma_i^2 \sim IID\chi^2(2)/2$. α_i, λ_i, and σ_i^2 are fixed for replications. $y_{i,-49} = \alpha_i$, and the first 49 observations are discarded. We reject the null hypothesis when we obtain negative (H test) statistics (due to negative variance estimates, $\hat{v}_H$).

(Continued)

	Size					Power				
NnT	20	30	50	100	200	20	30	50	100	200
	¸i = 0:6 for all i					¸i » IIDU(0:4; 0:8)				
Ŝ test										
20	7:50	7:60	7:10	5:40	4:85	21:90	35:50	62:00	95:10	99:95
30	9:15	8:40	6:65	5:45	6:45	34:10	55:40	85:55	99:90	100:00
50	9:55	9:55	7:65	6:15	5:15	45:70	69:60	94:70	100:00	100:00
100	11:00	10:50	8:00	6:70	6:50	73:65	95:10	100:00	100:00	100:00
200	15:75	13:20	10:60	7:10	7:60	93:10	99:65	100:00	100:00	100:00
H test										
20	91:45	97:00	98:50	98:95	99:35	96:45	99:50	100:00	100:00	100:00
30	98:95	99:70	100:00	100:00	100:00	99:90	100:00	100:00	100:00	100:00
50	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
100	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
¢ test										
20	2:40	2:35	4:15	3:05	3:05	5:95	16:00	44:60	90:50	99:95
30	2:35	3:10	3:60	3:40	5:10	8:75	29:40	72:90	99:40	100:00
50	2:65	3:75	3:95	4:15	4:40	12:15	40:30	87:10	100:00	100:00
100	4:35	3:65	3:80	4:40	4:90	25:30	77:30	99:90	100:00	100:00
200	6:15	4:00	4:80	3:90	4:70	44:35	96:20	100:00	100:00	100:00
	¸i = 0:8 for all i					¸i » IIDU(0:6; 1:0)				
Ŝ test										
20	13:40	11:45	10:05	6:35	5:50	34:55	54:20	84:45	99:95	100:00
30	16:10	14:25	11:35	7:35	8:20	54:75	78:65	98:10	100:00	100:00
50	20:75	17:00	12:40	9:25	7:05	73:15	94:05	100:00	100:00	100:00
100	29:55	24:40	17:90	10:50	9:20	93:80	99:90	100:00	100:00	100:00
200	46:30	35:30	25:45	13:70	11:45	99:85	100:00	100:00	100:00	100:00
H test										
20	96:30	99:15	99:40	99:85	99:90	98:50	100:00	100:00	100:00	100:00
30	99:70	100:00	100:00	100:00	100:00	99:95	100:00	100:00	100:00	100:00
50	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
100	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
¢ test										
20	3:05	3:70	4:95	3:90	3:60	11:70	32:55	72:35	99:85	100:00
30	3:45	4:45	4:45	4:00	4:90	21:55	56:90	95:20	100:00	100:00
50	2:95	5:35	4:35	4:75	4:85	36:35	78:95	99:40	100:00	100:00
100	4:35	5:70	5:55	4:75	5:80	64:95	97:90	100:00	100:00	100:00
200	4:70	7:05	8:30	5:95	6:25	89:90	100:00	100:00	100:00	100:00

(Continued)

$N \backslash T$	Size 20	Size 30	Size 50	Size 100	Size 200	Power 20	Power 30	Power 50	Power 100	Power 200
	$\lambda_i = 0.9$ for all i					$\lambda_i \sim IIDU(0.0, 1.0)$				
$\hat{S}$ test										
20	20:40	18:20	15:30	9:50	7:70	89:15	99:55	100:00	100:00	100:00
30	27:30	22:10	18:40	10:65	9:40	99:30	100:00	100:00	100:00	100:00
50	37:80	30:35	22:60	15:20	10:00	99:95	100:00	100:00	100:00	100:00
100	56:40	47:20	36:50	21:30	14:35	100:00	100:00	100:00	100:00	100:00
200	79:40	70:50	54:65	33:10	19:90	100:00	100:00	100:00	100:00	100:00
H test										
20	97:40	99:05	99:80	100:00	99:95	99:80	100:00	100:00	100:00	100:00
30	99:80	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
50	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
100	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00	100:00
Δ test										
20	4:85	6:10	6:75	5:20	4:60	69:00	97:55	100:00	100:00	100:00
30	6:00	7:55	8:00	5:95	6:25	94:70	100:00	100:00	100:00	100:00
50	8:15	9:70	10:05	8:10	5:95	99:65	100:00	100:00	100:00	100:00
100	13:85	15:90	14:95	11:30	8:95	100:00	100:00	100:00	100:00	100:00
200	24:00	26:35	25:05	16:25	10:80	100:00	100:00	100:00	100:00	100:00

Table 5 : Size and Power of the Slope Homogeneity Tests
for Heteroskedastic AR(2) Specifications

$N \backslash T$	Size 20	Size 30	Size 50	Size 100	Size 200	Power 20	Power 30	Power 50	Power 100	Power 200
	$\lambda_{1i} = 0.6$ for all i					$\lambda_{1i} \sim IIDU\,(0.4, 0.8)$				
$\hat{S}$ test										
20	14:60	13:30	10:40	6:80	6:50	30:45	39:60	63:70	96:90	100:00
30	18:65	16:20	11:05	7:00	6:50	46:15	59:25	88:80	100:00	100:00
50	25:75	19:85	13:25	8:75	6:15	60:50	78:50	97:75	100:00	100:00
100	37:05	26:95	19:10	12:05	9:25	85:05	97:20	99:95	100:00	100:00
200	56:50	42:45	26:90	15:50	10:05	97:60	100:00	100:00	100:00	100:00
H test										
20	90:65	96:25	98:35	98:45	97:60	95:00	98:60	96:80	98:05	100:00
30	98:45	99:55	99:50	99:35	99:00	99:60	99:25	97:90	99:75	100:00
50	100:00	100:00	99:90	99:80	99:75	99:95	99:55	99:20	100:00	100:00
100	100:00	100:00	100:00	99:95	100:00	100:00	99:80	99:45	100:00	100:00
200	100:00	100:00	100:00	100:00	100:00	100:00	100:00	99:80	100:00	100:00
Δ test										
20	3:10	3:05	4:10	4:00	4:80	4:70	12:50	38:00	92:30	100:00
30	2:95	3:55	3:95	3:65	4:45	7:25	24:05	71:30	99:80	100:00
50	2:20	3:65	3:90	5:05	5:15	13:30	38:70	88:90	100:00	100:00
100	2:65	3:80	5:10	5:40	5:90	25:20	74:90	99:85	100:00	100:00
200	3:40	4:55	5:75	5:95	4:65	43:55	96:50	100:00	100:00	100:00

Notes: See notes to Table 1 and 4. The DGP is $y_{it} = (1 - \lambda_{i1} - \lambda_2)\alpha_i + \lambda_{i1} y_{i,t-1} + \lambda_2 y_{i,t-2} + \varepsilon_{it}$, $t = -49, \ldots, 0, \ldots, T$, $i = 1, \ldots, N$, λ_{1i} is as specified in the table, $\lambda_2 = 0.2$. The first 48 observations are discarded.

Table 6 : Size and Power of the Bootstrap Test of Slope Homogeneity for Heteroskedastic AR(1) Specifications

	Size			Power		
N\T	20	30	50	20	30	50
	$\lambda_i = 0.2$ for all i			$\lambda_i \sim IIDU(0.0, 0.4)$		
Standard Normal						
20	1.60	1.65	2.60	3.50	8.45	26.85
30	3.30	3.40	2.75	4.05	12.90	42.85
50	5.20	3.60	4.35	3.95	17.35	58.35
100	9.20	4.90	4.65	7.00	40.25	92.30
200	19.25	10.75	6.75	9.55	67.75	99.80
Bootstrap						
20	4.30	4.45	4.80	5.40	11.10	30.60
30	4.90	5.05	4.45	5.60	16.25	47.40
50	5.05	5.05	5.45	4.45	18.70	61.95
100	5.20	4.40	5.00	4.55	38.35	92.60
200	6.15	5.35	5.30	2.85	57.20	99.80
Bias-Corrected Bootstrap						
20	4.25	4.55	4.70	5.20	11.30	30.75
30	5.00	5.20	4.70	5.90	16.35	47.35
50	5.40	5.30	5.75	4.85	18.75	61.55
100	5.75	4.25	5.10	5.00	38.90	92.50
200	7.30	5.50	5.45	2.80	58.00	99.80
	$\lambda_i = 0.4$ for all i			$\lambda_i \sim IIDU(0.2, 0.6)$		
Standard Normal						
20	2.45	2.65	2.75	3.60	10.40	29.05
30	2.35	3.05	3.15	4.85	18.75	52.85
50	4.50	3.90	3.70	6.10	23.15	69.55
100	7.90	5.85	4.15	11.60	52.50	97.60
200	12.75	6.95	5.90	18.85	82.00	99.95
Bootstrap						
20	5.40	5.50	4.65	6.35	13.95	33.90
30	4.25	5.45	4.95	6.50	22.50	56.70
50	5.25	5.35	5.05	7.15	25.50	72.25
100	5.65	5.65	5.15	9.60	51.90	97.80
200	5.20	4.80	5.35	9.90	78.30	99.95
Bias-Corrected Bootstrap						
20	5.80	5.55	4.90	6.15	13.85	34.10
30	4.70	5.35	4.75	6.75	22.25	56.40
50	5.40	5.80	4.65	7.15	25.70	72.15
100	6.35	6.10	5.05	10.25	53.20	98.00
200	6.00	5.30	5.80	11.85	79.20	99.95

Notes: See the notes to Table 4. 499 bootstrap samples are generated, and rejection frequencies are based on 2,000 replications. "Bootstrap" is based on the bootstrap samples generated using $\tilde{\lambda}_{WFE}$. The "Bias-Corrected Bootstrap" is based on the bootstrap samples generated using the bias-corrected estimator, $\mathring{\lambda}_{WFE}$. For further details see Section 4.2.

(Continued)

NnT	Size 20	Size 30	Size 50	Power 20	Power 30	Power 50
	¸ i = 0:6 for all i			¸ i » IIDU(0:4;0:8)		
Standard Normal						
20	1:40	2:75	3:85	5:40	15:55	43:75
30	2:60	3:80	2:55	7:50	30:50	73:65
50	2:45	3:55	3:35	11:30	40:45	87:50
100	4:25	3:55	4:15	28:00	78:50	99:80
200	5:75	3:65	4:30	45:45	96:75	100:00
Bootstrap						
20	4:25	6:05	5:20	8:70	19:85	48:15
30	4:85	5:70	3:70	10:40	34:65	76:40
50	4:30	5:40	4:95	14:25	45:80	89:15
100	4:50	4:60	4:90	29:75	80:50	99:85
200	4:30	4:30	4:95	40:80	96:80	100:00
Bias-Corrected Bootstrap						
20	4:35	5:50	5:35	8:80	19:60	48:05
30	4:90	5:20	3:85	10:75	35:35	76:10
50	4:40	5:35	4:85	14:85	45:80	89:55
100	5:50	5:25	4:95	31:85	80:80	99:85
200	5:55	4:45	5:20	45:40	97:05	100:00
	¸ i = 0:8 for all i			¸ i » IIDU(0:6;1:0)		
Standard Normal						
20	2:40	3:80	4:80	12:60	31:05	71:85
30	2:95	3:85	5:45	22:05	55:10	95:75
50	3:20	5:25	5:25	36:90	76:70	99:70
100	4:65	5:75	5:90	64:90	98:20	100:00
200	4:65	7:50	8:25	90:55	100:00	100:00
Bootstrap						
20	5:20	5:65	5:70	17:10	35:00	73:70
30	4:70	5:15	6:25	26:15	58:05	95:70
50	4:70	6:40	5:60	40:05	75:25	99:70
100	6:40	6:50	5:75	69:70	98:20	100:00
200	6:70	8:60	7:55	92:25	100:00	100:00
Bias-Corrected Bootstrap						
20	4:50	5:10	5:20	15:50	34:15	73:25
30	4:25	4:65	5:65	23:85	55:35	95:30
50	4:20	5:50	5:25	32:95	71:15	99:45
100	5:55	5:15	4:85	63:10	97:60	100:00
200	5:00	5:55	6:05	87:45	99:95	100:00

(Continued)

NnT	Size			Power		
	20	30	50	20	30	50
	¸ i = 0:9 for all i			¸ i » IIDU(0:0;1:0)		
Standard Normal						
20	5:30	6:20	7:60	68:35	97:95	100:00
30	7:25	7:35	9:35	94:85	99:95	100:00
50	7:90	9:25	8:95	99:35	100:00	100:00
100	12:55	15:80	15:45	100:00	100:00	100:00
200	21:30	27:65	25:40	100:00	100:00	100:00
Bootstrap						
20	6:00	6:10	7:50	74:15	98:50	100:00
30	7:95	6:95	7:85	96:05	99:95	100:00
50	8:45	7:30	6:55	99:60	100:00	100:00
100	11:95	9:95	8:50	100:00	100:00	100:00
200	18:35	16:25	10:55	100:00	100:00	100:00
Bias-Corrected Bootstrap						
20	4:45	4:70	6:25	74:20	98:60	100:00
30	5:20	4:50	6:05	96:30	99:95	100:00
50	4:45	3:60	5:35	99:60	100:00	100:00
100	5:05	4:80	4:95	100:00	100:00	100:00
200	4:50	5:75	5:65	100:00	100:00	100:00

Table 7: Slope Homogeneity Tests and Alternative Estimates of the Autoregressive Coefficient
of the Real Earnings Equations

		Pooled Sample $e=0$	High School Dropout $e=1$	High School Graduate $e=2$	College Graduate $e=3$
N		$1,031$	249	531	251
Average T_i		18.39	18.36	18.22	18.79
Total Observations		$18,961$	$4,572$	$9,673$	$4,716$
Tests for Slope Homogeneity					
Δ test	Statistic	25.59	7.20	13.65	18.32
	Normal approximation p-value	$[0.0000]$	$[0.0000]$	$[0.0000]$	$[0.0000]$
	Bias-corrected bootstrap p-value	$[0.0000]$	$[0.0000]$	$[0.0000]$	$[0.0000]$
Autoregressive Coefficient (λ)					
FE Estimates ($\hat{\lambda}_{FE}$)		0.4841 (0.0065)	0.4056 (0.0147)	0.4497 (0.0095)	0.5538 (0.0106)
WFE Estimates ($\tilde{\lambda}_{WFE}$)		0.5429 (0.0056)	0.4246 (0.0133)	0.5169 (0.0086)	0.6002 (0.0095)
Bias-Corrected WFE ($\mathring{\lambda}_{WFE}$)		0.6504 (0.0055)	0.5188 (0.0126)	0.6192 (0.0080)	0.7214 (0.0101)

Notes: Noting PSID data we used are unbalanced, FE estimator, and WFE estimator are defined by (3.23), and (3.22) in Remark 5, respectively, and their associated standard errors (shown in round brackets) are based on $\hat{V}\left(\hat{\lambda}_{FE}\right) = \hat{\sigma}^2\left(\sum_{i=1}^{N} y'_{i,-1}M_{\tau_i}y_{i,-1}\right)^{-1}$, where

$$\hat{\sigma}^2 = (T-N-1)^{-1}\sum_{i=1}^{N}\left(y_i - \hat{\lambda}_{FE}y_{i,-1}\right)' M_{\tau_i}\left(y_i - \hat{\lambda}_{FE}y_{i,-1}\right);$$

$T = \sum_{i=1}^{N} T_i$, and $\hat{V}\left(\tilde{\lambda}_{WFE}\right) = \left(\sum_{i=1}^{N} \hat{\sigma}_i^{-2} y'_{i,-1}M_{\tau_i}y_{i,-1}\right)^{-1}$.

Bias corrected estimates are based on $\mathring{\lambda}_{WFE} = \tilde{\lambda}_{WFE} + (T/N)\left(1+\tilde{\lambda}_{WFE}\right)$ and $\hat{V}\left(\mathring{\lambda}_{WFE}\right) = T^{-1}\left(1-\mathring{\lambda}_{WFE}^2\right)$. Bias-corrected bootstrapped tests also use $\mathring{\lambda}_{WFE}$ and the associated estimates to generate bootstrap samples (see Section 4.2 for further details).