ESSAYS IN TIME SERIES ECONOMETRICS

By

Nasreen Nawaz

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2015

ABSTRACT

ESSAYS IN TIME SERIES ECONOMETRICS

By

Nasreen Nawaz

In the first chapter, We focus on the estimation of the ratio of trend slopes between two time series where is it reasonable to assume that the trending behavior of each series can be well approximated by a simple linear time trend. We obtain results under the assumption that the stochastic parts of the two time series comprise a zero mean time series vector that has sufficient stationarity and has dependence that is weak enough so that scaled partial sums of the vector satisfy a functional central limit theorem (FCLT). We compare two obvious estimators of the trend slope ratio and propose a third bias-corrected estimator. We show how to use these three estimators to carry out inference about the trend slope ratio. When trend slopes are small in magnitude relative to the variation in the stochastic components (the trend slopes are small relative to the noise), we find that inference using any of the three estimators is compromised and potentially misleading. We propose an alternative inference procedure that remains valid when trend slopes are small or even zero. We carry out an extensive theoretical analysis of the estimators and inference procedures with positive findings. First, the theory points to one of the three estimators as being preferred in terms of bias. Second, the theory unambiguously suggests that our alternative inference procedure is superior both under the null and under the alternative with respect to the magnitudes of the trend slopes. Finite sample simulations indicate that the predictions made by the asymptotic theory are relevant in practice. We give concrete and specific advice to empirical practitioners on how to estimate a ratio of trend slopes and how to carry out tests of hypotheses about the ratio.

The second chapter is an extension of the first, where the stationarity assumption in the analysis is relaxed. It is assumed that the stochastic parts of the trending series follow an I(1) process. We consider the case of unit root in the noise term in the IV regression equation. We also consider the case of cointegration between the two series.

The theory explicitly captures the impact of the magnitude of the trend slopes on the estimation and inference about the trend slopes ratio. If the trend slopes are relatively large in magnitude, the IV estimator is consistent for both I(1) and I(0) regression errors. For medium and small trend slopes, the IV estimator is inconsistent for I(1) case, but consistent for I(0) regression error. For inference, the test based on IV estimator has been compared with the alternative testing approach. Asymptotic theory and finite sample simulations suggest that the alternative testing approach is superior both under the null and under the alternative with respect to the magnitudes of the trend slopes. Whether the noise term in the IV regression equation is I(0) or I(1) has an impact on the power performance of the test for the trend slopes ratio.

The third chapter is an empirical application of the methodology developed in the first and the second chapters. The empirical findings on convergence of per capita income across regions in convergence literature are mixed. There is evidence of convergence in a substantial number of cases, whereas evidence contrary to convergence has also been found. Where there is β-convergence found, it is interesting to come up with a measure of speed of convergence and estimate it. The speed of convergence has been shown to be proportional to a ratio of trend slopes, and using the methodology developed in the first and the second chapters, we estimate this ratio for all US regions which are converging. The higher the ratio, the greater is the speed of convergence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 ESTIMATION AND INFERENCE OF LINEAR TREND SLOPE RATIOS WITH I(0) ERRORS (with Timothy J. Vogelsang)

1.1 Introduction

A time trend refers to systematic behavior of a time series that can approximated by a function of time. Often when plotting macroeconomic or climate time series data, one notices a tendency for the series to increase (or decrease) over time. In some cases it is immediately apparent from the time series plot that the trend is approximately linear. In the econometrics literature there is a well developed literature on estimation and robust inference of deterministic trend functions with a focus on the case of the simple linear trend model. See for example, Canjels and Watson (1997), Vogelsang (1998), Bunzel and Vogelsang (2005), Harvey, Leybourne and Taylor (2009), and Perron and Yabu (2009).

When analyzing more than one time series with trending behavior, it may be interesting to compare the trending behavior across series as in Vogelsang and Franses (2005). Empirically, comparisons across trends are often made in the economic convergence literature where growth rates of gross domestic products (GDPs), i.e. trend slopes of DGPs, are compared across regions or countries. See for example, Fagerberg (1994), Evans (1997) and Tomljanovich and Vogelsang (2002). Often empirical work in the economic convergence literature seeks to determine whether countries or regions have growth rates that are consistent with convergence that has either occurred or is occurring. There is little, if any, focus on estimating and quantifying the relative speed by which convergence is occurring. In simple settings, quantifying the relative speed of economic convergence amounts to estimating the *ratio* of gross domestic product growth rates, i.e. estimating relative growth rates.

In the empirical climate literature there is a recent literature documenting the relative warming rates between surface temperatures and lower troposphere temperatures. See Santer, Wigley, Mears, Wentz, Klein, Seidel, Taylor, Thorne, Wehner, Gleckler et al. (2005), Klotzbach, Pielke, Christy and McNider (2009) and the references cited therein. In this literature there is an explicit interest in estimating trend slope ratios and reporting confidence intervals for them.

As far as we know, there are no formal statistical methodological papers in the econometrics or climate literatures focusing on estimation and inference of trend slope ratios. This paper fills that methodological hole in the literature. We focus on estimation of the ratio of trend slopes between two time series where is it reasonable to assume that the trending behavior of each series can be well approximated by a simple linear time trend. We obtain results under the assumption that the stochastic parts of the two time series comprise a zero mean time series vector that has sufficient stationarity and has dependence that is weak enough so that scaled partial sums of the vector satisfy a functional central limit theorem (FCLT). We compare two obvious estimators of the trend slope ratio and propose a third bias-corrected estimator. We show how to use these three estimators to carry out inference about the trend slope ratio. When trend slopes are small in magnitude relative to the variation in the stochastic components (the trend slopes are small relative to the noise), we find that inference using any of the three estimators is compromised and potentially misleading. We propose an alternative inference procedure that remains valid when trend slopes are small or even zero.

We carry out an extensive theoretical analysis of the estimators and inference procedures. Our theoretical framework explicitly captures the impact of the magnitude of the trend slopes on the estimation and inference about the trend slope ratio. Our theoretical results are constructive in two important ways. First, the theory points to one of the three estimators as being preferred in terms of bias. Second, the theory strongly suggests that our alternative inference procedure is superior both for robustness under the null and power under the alternative with respect to the magnitudes of the trend slopes. Finite sample simulations indicate that the predictions made by the asymptotic theory are relevant in practice. Therefore, we are able to give concrete and specific advice to empirical practitioners on how to estimate a ratio of trend slopes and how to carry out tests of hypotheses about the ratio.

The remainder of this chapter is organized as follows: Section 1.2 describes the model and analyzes the asymptotic properties of the three estimators of the trend slope ratio. Section 1.3 provides some finite sample evidence on the relative performance of the three estimators. Section 1.4 investigates inference regarding the trend slope ratio. We show how to construct heteroskedasticity autocorrelation (HAC) robust tests using each of the three estimators. We propose an alternative testing approach and show how to compute confidence intervals for this approach. We derive asymptotic results of the tests under the null and under local alternatives. The asymptotic theory clearly shows that our alternative testing approach is superior under both the null and local alternatives. Additional finite sample simulation results reported in Section 1.5 indicate that the predictions of the asymptotic theory are relevant in practice. In Section 1.6 we make some practical recommendations for empirical researchers and Section 1.7 concludes. All proofs are given in the Appendix.

1.2 The Model and Estimation

1.2.1 Model and Assumptions

Suppose the univariate time series y_{1t} and y_{2t} are given by

$$y_{1t} = \mu_1 + \beta_1 t + u_{1t}, \tag{1}$$

$$y_{2t} = \mu_2 + \beta_2 t + u_{2t}, \tag{2}$$

where u_{1t} and u_{2t} are mean zero covariance stationary processes. Assume that

$$T^{-1/2}\sum_{t=1}^{[rT]} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \Rightarrow \mathbf{\Lambda W}(r) \equiv \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}, \tag{3}$$

where $r \in [0,1]$, $[rT]$ is the integer part of rT and $W(r)$ is a 2×1 vector of independent standard Wiener processes. Λ is not necessarily diagonal allowing for correlation between u_{1t} and u_{2t}. In addition to (3), we assume that u_{1t} and u_{2t} are ergodic for the first and second moments.

Suppose that $\beta_2 \neq 0$ and we are interested in estimating the parameter

$$\theta = \frac{\beta_1}{\beta_2}$$

which is the ratio of trend slopes. Equation (1) can be rewritten so that y_{1t} depends on θ through y_{2t}. Rearranging (2) gives

$$t = \frac{1}{\beta_2}\left[y_{2t} - \mu_2 - u_{2t}\right], \tag{4}$$

and plugging this expression into Equation (1) and then rearranging, we obtain

$$y_{1t} = (\mu_1 - \frac{\beta_1}{\beta_2}\mu_2) + \frac{\beta_1}{\beta_2}y_{2t} + (u_{1t} - \frac{\beta_1}{\beta_2}u_{2t}) = (\mu_1 - \theta\mu_2) + \theta y_{2t} + (u_{1t} - \theta u_{2t}).$$

Defining $\delta = \mu_1 - \theta\mu_2$ and $\epsilon_t(\theta) = u_{1t} - \theta u_{2t}$ gives the regression model

$$y_{1t} = \delta + \theta y_{2t} + \epsilon_t(\theta). \tag{5}$$

Given the definition of $\epsilon_t(\theta)$, it immediately follows from (3) that

$$T^{-1/2}\sum_{t=1}^{[rT]}\epsilon_t(\theta) \Rightarrow \lambda_\theta w(r), \tag{6}$$

where $w(r)$ is a univariate standard Wiener process and

$$\lambda_\theta^2 = \begin{bmatrix} 1 & -\theta \end{bmatrix} \Lambda\Lambda' \begin{bmatrix} 1 & -\theta \end{bmatrix}'$$

is the long run variance of $\epsilon_t(\theta)$.

1.2.2 Estimation of the Trend Slope Ratio

Using regression (5), the natural estimator of θ is ordinary least squares (OLS) which is defined as

$$\widetilde{\theta} = \left(\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right)^{-1} \sum_{t=1}^{T}(y_{2t} - \overline{y}_2)(y_{1t} - \overline{y}_1), \tag{7}$$

where $\overline{y}_1 = T^{-1}\sum_{t=1}^{T} y_{1t}$ and $\overline{y}_2 = T^{-1}\sum_{t=1}^{T} y_{2t}$. Standard algebra gives the relationship

$$\widetilde{\theta} - \theta = \left(\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right)^{-1} \sum_{t=1}^{T}(y_{2t} - \overline{y}_2)\epsilon_t(\theta). \tag{8}$$

Alternatively, one could estimate θ by the analogy principle by simply replacing β_1 and β_2 with estimators. Let $\widehat{\beta}_1$ and $\widehat{\beta}_2$ be the OLS estimators of β_1 and β_2 based on regressions (1) and (2):

$$\widehat{\beta}_1 = \left(\sum_{t=1}^{T}(t - \overline{t})^2\right)^{-1} \sum_{t=1}^{T}(t - \overline{t})(y_{1t} - \overline{y}_1), \tag{9}$$

$$\widehat{\beta}_2 = \left(\sum_{t=1}^{T}(t - \overline{t})^2\right)^{-1} \sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2), \tag{10}$$

where $\overline{t} = T^{-1}\sum_{t=1}^{T} t$ is the sample average of time and define $\widehat{\theta} = \widehat{\beta}_1/\widehat{\beta}_2$. Simple algebra shows that

$$\widehat{\theta} = \frac{\widehat{\beta}_1}{\widehat{\beta}_2} = \left(\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2)\right)^{-1} \sum_{t=1}^{T}(t - \overline{t})(y_{1t} - \overline{y}_1) \tag{11}$$

which is the instrumental variable (IV) estimator of θ in (5) where t has been used as an instrument for y_{2t}. Standard algebra gives the relationship

$$\widehat{\theta} - \theta = \left(\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2)\right)^{-1} \sum_{t=1}^{T}(t - \overline{t})\epsilon_t(\theta). \tag{12}$$

1.2.3 Asymptotic Properties of OLS and IV

We now explore the asymptotic properties of the OLS and IV estimators of θ. The asymptotic behavior of the estimators depends on the magnitude of the trend slope parameters relative to the variation in the random components, u_{1t} and u_{2t}, i.e. the noise. The following theorem summarizes the asymptotic behavior of the estimators for fixed βs and for βs that are modeled as local to zero at rate $T^{-1/2}$.

Theorem 1 *Suppose that (6) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{3/2}\left(\widetilde{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \overline{\beta}_2 \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda^2}{\overline{\beta}_2^2}\right),$$

$$T^{3/2}\left(\widehat{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda^2}{\overline{\beta}_2^2}\right).$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$T\left(\widetilde{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \left[\overline{\beta}_2 \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) + E(u_{2t}\epsilon_t(\theta))\right]$$
$$\sim N\left(\frac{12}{\overline{\beta}_2^2} E(u_{2t}\epsilon_t), \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right),$$

$$T\left(\widehat{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

Theorem 1 makes some interesting predictions about the sampling properties of OLS and IV. When the trend slopes are fixed, i.e. when the trend slopes are large relative to the noise, OLS and IV converge to the true value of θ at the rate $T^{3/2}$ and are asymptotically normal with equivalent asymptotic variances. The precision of both estimators improves when there is less noise (λ_θ^2 is smaller) or when the magnitude of the trend slope parameter for y_{2t} increases (β_2 is larger).

When the trend slopes are modeled as local to zero at rate $T^{-1/2}$, i.e. when trend slopes are medium sized relative to the noise, asymptotic equivalence of OLS and IV no longer holds. The IV estimator essentially has the same asymptotic behavior as in the fixed slopes case because the implied approximations are the same:

$$\text{Case 1 (large trend slopes): } \widehat{\theta} \approx N\left(\theta, \frac{12\lambda_\theta^2}{T^3\overline{\beta}_2^2}\right) \equiv N\left(\theta, \frac{12\lambda_\theta^2}{T^3\beta_2^2}\right),$$
$$\text{Case 2 (medium trend slopes): } \widehat{\theta} \approx N\left(\theta, \frac{12\lambda_\theta^2}{T^2\overline{\beta}_2^2}\right) \equiv N\left(\theta, \frac{12\lambda_\theta^2}{T^3\beta_2^2}\right).$$

In contrast, the result for OLS is markedly different in Case 2. While OLS consistently estimates θ and the asymptotic variance is the same as in Case 1, OLS now has an asymptotic bias that could matter when trend slopes are medium sized.

The fact that OLS is asymptotically biased in Case 2 is not that surprising because $\epsilon_t(\theta)$ is correlated with y_{2t} through the correlation between $\epsilon_t(\theta)$ and u_{2t}. In Case 2, the trend slopes are small enough so that the covariance between u_{2t} and $\epsilon_t(\theta)$ asymptotically affects the OLS estimator. Because $E(u_{2t}\epsilon_t(\theta)) = E(u_{1t}u_{2t}) - \theta E(u_{2t}^2)$, the asymptotic bias will be non-zero unless $E(u_{1t}u_{2t}) = \theta E(u_{2t}^2)$ which only happens in very particular special cases. In general, OLS will have an asymptotic bias when trend slopes are medium sized.

In the next subsection we propose a bias correction for OLS estimator.

1.2.4 Bias Corrected OLS

According to Theorem 1, for the case of medium sized trend shifts OLS is asymptotically biased and this bias is driven by covariance between $\epsilon_t(\theta)$ and u_{2t}. The approximate bias of OLS suggested by Theorem 1 is given by the quantity

$$bias(\widetilde{\theta}) \approx T^{-1}\left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} E(u_{2t}\epsilon_t(\theta)).$$

We can estimate $\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds$ using $T^{-2}\sum_{t=1}^T (y_{2t} - \overline{y}_2)^2$, and we can estimate $E(u_{2t}\epsilon_t(\theta))$ using $T^{-1}\sum_{t=1}^T \widehat{u}_{2t}\widetilde{\epsilon}_t$ where $\widetilde{\epsilon}_t$ are the OLS residuals from regression (5) and $\widehat{u}_{2t}$ are the OLS residuals from regression (2). This leads to the bias corrected

OLS estimator of θ given by

$$\widetilde{\theta}^c = \widetilde{\theta} - T^{-1}\left(\frac{T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t}{T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2}\right) = \widetilde{\theta} - \frac{T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t}{T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2}. \tag{13}$$

The next theorem gives the asymptotic behavior of the bias corrected OLS estimator for the same cases covered by Theorem 1.

Theorem 2 *Suppose that (6) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{3/2}\left(\widetilde{\theta}^c - \theta\right) \Rightarrow \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \overline{\beta}_2 \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$T\left(\widetilde{\theta}^c - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right).$$

As Theorem 2 shows, the bias corrected OLS estimator is asymptotically equivalent to the IV estimator for both large and medium trend slopes.

1.2.5 Asymptotic Properties of Estimators for Small Trend Slopes

As shown by Theorem 1, the magnitudes of the trend slopes relative to the noise can affect the behavior of estimators of the trend slope ratio, θ. Intuitively, we know that as the trend slopes become very small in magnitude, we approach the case where the trend slopes are zero in which case θ is not well defined. While it is clear that OLS and possibly bias-corrected OLS will have problems when trends slopes are very small, IV is also expected to have problems in this case. If the trend slopes are very small, then the sample correlation between t and y_{2t} also becomes very small and t becomes a weak instrument for y_{2t}. It is well known in the literature that weak instruments have important implications for IV estimation (see Staiger and Stock 1997**?**) and estimation of θ is no exception.

The next two theorems provide asymptotic results for the estimators of θ for trend slopes that are local to zero at rates T^{-1} and $T^{-3/2}$ with the latter case corresponding to trend slopes that are very small relatively to the noise.

Theorem 3 *Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 3 (small trend slopes): For $\beta_1 = T^{-1}\overline{\beta}_1, \beta_2 = T^{-1}\overline{\beta}_2$,*

$$\widetilde{\theta} - \theta \xrightarrow{p} \Re, \qquad \widetilde{\theta}^c - \theta \xrightarrow{p} \Re_c,$$

$$T^{1/2}(\widehat{\theta} - \theta) \Rightarrow \left(\overline{\beta}_2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds\right)^{-1} \lambda_\theta \int_0^1 \left(s - \frac{1}{2}\right) dw(s) \sim N\left(0, \frac{12\lambda_\theta^2}{\overline{\beta}_2^2}\right),$$

where

$$\Re = \left(\overline{\beta}_2^2 \int_0^1 \left(s - \frac{1}{2}\right)^2 ds + E(u_{2t}^2)\right)^{-1} E(u_{2t}\epsilon_t(\theta)), \quad \Re_c = \frac{\Re^2 E(u_{2t}^2)}{E(u_{2t}\epsilon_t(\theta))}.$$

Case 4 (very small trend slopes): For $\beta_1 = T^{-3/2}\overline{\beta}_1, \beta_2 = T^{-3/2}\overline{\beta}_2$,

$$\widetilde{\theta} - \theta \xrightarrow{p} \frac{E(u_{2t}\epsilon_t(\theta))}{E(u_{2t}^2)}, \quad \widetilde{\theta}^c - \theta \xrightarrow{p} \frac{E(u_{2t}\epsilon_t(\theta))}{E(u_{2t}^2)},$$

$$\widehat{\theta} - \theta \Rightarrow \left(\overline{\beta}_2 \int_0^1 (s - \frac{1}{2})^2 ds + \int_0^1 (s - \frac{1}{2}) dB_2(s)\right)^{-1} \lambda_\theta \int_0^1 (s - \frac{1}{2}) dw(s).$$

Theorem 3 shows that for small trend slopes, OLS and bias-corrected OLS are biased and inconsistent. In contrast, IV has the same asymptotic behavior as for large and medium trend slopes. When trend slopes are very small, all three estimators are inconsistent and biased. The IV estimator converges to a ratio of normal random variables that are correlated with each other because $B_2(r)$ is correlated with $w(r)$ as long as u_{2t} is correlated with $\epsilon_t(\theta)$. Theorem 3 predicts that none of the estimators of θ will work well when trend slopes are very small. This is not surprising because it is very difficult to identify the ratio of trend slopes when the noise dominates the information in the sample regarding the trend slopes themselves.

1.2.6 Implications (Predictions) of Asymptotics for Finite Samples

Theorems 1-3 make clear predictions about the finite sample behavior of the three estimators of θ. For large βs the three estimators should have similar bias, variance and sampling properties given that they are asymptotically equivalent. When βs are of medium size, IV and bias-corrected OLS are asymptotically equivalent and should have similar behavior compared to the large βs case. In contrast, OLS should exhibit some finite sample bias in the medium βs case. With small or very small βs, OLS and bias-corrected OLS are inconsistent and biased. In contrast IV should continue to perform well even with small βs with the main implication of small βs being less precision given that the asymptotic variance of IV is inversely related to the magnitude of β_2. For very small βs IV is inconsistent and can exhibit substantial variability given that IV is approximately a ratio of two normal random variables.

For a given sample size and variability of the noise, as the trend slopes decrease from being large to becoming very small, we should see the performance of all three estimators deteriorating with OLS deteriorating quickest followed by bias-corrected OLS followed by IV.

1.3 Finite Sample Means and Standard Deviations of Estimators

In this section we illustrate the finite sample performance of the estimators via a Monte Carlo simulation study. For the data generating process (DGP) that we consider, the finite sample behavior of the three estimators closely follows the predictions suggested by Theorems 1-3.

The following DGP was used. The y_{1t} and y_{2t} variables were generated by models

(1) and (2) where the noise is given by

$$u_{1t} = 0.4u_{2t} + 0.3u_{1t-1} + \varepsilon_{1t},$$

$$u_{2t} = 0.5u_{2t-1} + \varepsilon_{2t},$$

$$[\varepsilon_{1t}, \varepsilon_{2t}]' \sim i.i.d.\ N(0, I_2),$$

$$u_{10} = u_{20} = 0.$$

Given that all three estimators are exactly invariant to the values of $\mu_1$and μ_2, we set $\mu_1 = 0, \mu_2 = 0$ without loss of generality. We report results for various magnitudes of β_1 and β_2 where it is almost always the case that $\theta = \beta_1/\beta_2 = 2$. The exception is when $\beta_1 = 0,\ \beta_2 = 0$ in which case θ is not defined. We report results for $T = 50, 100, 200$ with $10,000$ replications used in all cases.

Given that the bias-corrected OLS estimator uses a bias correction based on the OLS residuals from (5), we experimented with an iterative procedure for the bias-corrected OLS estimator that improved its finite sample performance. We first compute $\widetilde{\theta}^c$ as given by (13). Then we updated the OLS residuals using $\widetilde{\theta}^c$ in place of $\widetilde{\theta}$ and recalculated $\widetilde{\theta}^c$. We iterated between updated residuals and bias-correction 100 times.

Table 1 reports estimated means and standard deviations of the three estimators across the $10,000$ replications. Focusing on the $T = 50$ case we see that when the trend slopes are large ($\beta_2 = 10, 5$), the means and standard deviations of the three estimators are the same and none of the estimators shows any bias. For medium sized trend slopes ($\beta_2 = 2, .2, .15, .1$), bias-corrected OLS and IV have the same means and standard deviations with little bias being present. In contrast, OLS shows bias that increases substantially as β_2 decreases. For all three estimators we see that the standard deviations increase as β_2 decreases as expected. For small and very small trend slopes ($\beta_2 = .05, .02, .002$), OLS and bias-corrected OLS show substantial bias. It is difficult to determine whether IV is biased given the very large standard deviation of IV in this case. Overall, IV has the least bias but IV becomes very imprecise as the trend slopes approach zero.

Results for the cases of $T = 100, 200$ are similar to the $T = 50$ case. The only difference is that the bias of OLS and bias-corrected OLS kicks in more slowly as β_2 decreases. With $T = 200$, bias-corrected OLS and IV have the same means and standard deviations for β_2 as small as 0.02.

The results for $\beta_1 = 0$, $\beta_2 = 0$ at first may look surprising but make sense upon deeper inspection. The OLS estimator is no longer estimating θ which is not defined. Instead, OLS is estimating the population quantity $E(u_{2t}\epsilon_t(\theta))/E(u_{2t}^2)$ which is very close to 0.469 in our DGP. Bias-corrected OLS is attempting to correct the wrong bias and the IV estimator is based on an instrument that has zero correlation with y_{2t}. In fact, the estimators are behaving as expected when the trend slopes are zero.

Overall, the finite sample means and variances exhibit patterns as predicted by the asymptotic theory. IV and bias-corrected OLS work equally well for large, medium and somewhat small trend slopes. As trend slopes become very small, none of the estimators are very good and this is to be expected given that the data has relatively little information about the trend slope ratio when trend slopes are small relative to the noise and/or the sample size is small.

1.4 Inference

In this section we analyze test statistics for testing simple hypotheses about θ. Suppose we are interested in testing the null hypothesis

$$H_0 : \theta = \theta_0, \tag{14}$$

against the alternative hypothesis

$$H_1 : \theta = \theta_1 \neq \theta_0.$$

It is straightforward to construct HAC robust statistics using the three estimators of θ as

$$t_{OLS} = \frac{(\widetilde{\theta} - \theta_0)}{\sqrt{\widetilde{\lambda}_{\theta}^2 \left[\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}}, \tag{15}$$

$$t_{BC} = \frac{(\widetilde{\theta}^c - \theta_0)}{\sqrt{\widetilde{\lambda}_{\theta c}^2 \left[\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}}, \tag{16}$$

$$t_{IV} = \frac{(\widehat{\theta} - \theta_0)}{\sqrt{\widehat{\lambda}_{\theta}^2 \left[\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2)\right]^{-2} \sum_{t=1}^{T}(t - \overline{t})^2}}, \tag{17}$$

where the estimated long run variance estimators are given by

$$\widetilde{\lambda}_{\theta}^2 = \widetilde{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widetilde{\gamma}_j, \quad \widetilde{\gamma}_j = T^{-1}\sum_{t=j+1}^{T} \widetilde{\epsilon}_t \widetilde{\epsilon}_{t-j},$$

$$\widetilde{\lambda}_{\theta c}^2 = \widetilde{\gamma}_0^c + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widetilde{\gamma}_j^c, \quad \widetilde{\gamma}_j^c = T^{-1}\sum_{t=j+1}^{T} \left(\widetilde{\epsilon}_t^c - \overline{\widetilde{\epsilon}^c}\right)\left(\widetilde{\epsilon}_{t-j}^c - \overline{\widetilde{\epsilon}^c}\right),$$

$$\widehat{\lambda}_{\theta}^2 = \widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widehat{\gamma}_j, \quad \widehat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T} \widehat{\epsilon}_t \widehat{\epsilon}_{t-j},$$

with $\widetilde{\epsilon}_t = y_{1t} - \widetilde{\delta} - \widetilde{\theta} y_{2t}$ being the OLS residuals from (5), $\widetilde{\epsilon}_t^c = y_{1t} - \widetilde{\delta} - \widetilde{\theta}^c y_{2t}$ being the bias-corrected OLS residuals, and $\widehat{\epsilon}_t = y_{1t} - \widehat{\delta} - \widehat{\theta} y_{2t}$ being the IV residuals from (5). Because

$$\widetilde{\epsilon}_t^c = y_{1t} - \widetilde{\delta} - \widetilde{\theta} y_{2t} - \left(\widetilde{\theta}^c - \widetilde{\theta}\right) y_{2t} = \widetilde{\epsilon}_t - \left(\widetilde{\theta}^c - \widetilde{\theta}\right) y_{2t},$$

the bias-corrected OLS residuals do not sum to zero. Therefore, we construct $\widetilde{\lambda}_{\theta c}^2$ using $\widetilde{\epsilon}_t^c - \overline{\widetilde{\epsilon}^c}$ where $\overline{\widetilde{\epsilon}^c} = T^{-1}\sum_{t=1}^{T}\widetilde{\epsilon}_t^c$, i.e. we demean $\widetilde{\epsilon}_t^c$ before computing $\widetilde{\lambda}_{\theta c}^2$. The long run variance estimators are constructed using the kernel weighting function $k(x)$ and M is the bandwidth tuning parameter.

1.4.1 Linear in Slopes Approach

Because all three estimators of θ deteriorate as the trend slopes approach zero, we consider a fourth test statistic that is exactly invariant to the true values of the slope parameters under H_0. Given the null value of θ_0, H_0 and H_1 can be written in terms of the trend slopes as

$$H_0 : \frac{\beta_1}{\beta_2} = \theta_0, \quad H_1 : \frac{\beta_1}{\beta_2} = \theta_1$$

which is a nonlinear restriction on the trends slopes. Obviously, the restrictions implied by these hypotheses can be written as linear functions of β_1 and β_2 as

$$H_0 : \beta_1 - \beta_2\theta_0 = 0,$$

$$H_1 : \beta_1 - \beta_2\theta_0 = \beta_2\theta_1 - \beta_2\theta_0 = \beta_2\left(\theta_1 - \theta_0\right) \neq 0.$$

Given θ_0, define the univariate time series

$$z_t\left(\theta_0\right) = y_{1t} - \theta_0 y_{2t},$$

where it follows from (1) and (2) that

$$z_t\left(\theta_0\right) = \pi_0(\theta_0) + \pi_1(\theta_0)t + v_t\left(\theta_0\right), \tag{18}$$

where $\pi_0(\theta_0) = \mu_1 - \theta_0\mu_2$, $\pi_1(\theta_0) = \beta_1 - \theta_0\beta_2$ and $v_t\left(\theta_0\right) = u_{1t} - \theta_0 u_{2t}$. Notice that $v_t\left(\theta_0\right) = \epsilon_t(\theta_0)$ and therefore the long run variance of $v_t\left(\theta_0\right)$ is

$$\lambda_{\theta_0}^2 = \begin{bmatrix} 1 & -\theta_0 \end{bmatrix} \Lambda\Lambda' \begin{bmatrix} 1 & -\theta_0 \end{bmatrix}'.$$

Under H_0 it follows that $\pi_1(\theta_0) = 0$ whereas under H_1 it follows that $\pi_1(\theta_0) = \beta_2\left(\theta_1 - \theta_0\right) \neq 0$. We can test the original null hypothesis given by (14) by testing $H_0 : \pi_1(\theta_0) = 0$ in (18) against the alternative $H_1 : \pi_1(\theta_0) \neq 0$ using the following

t-statistic:

$$t_{\theta_0} = \frac{\widehat{\pi}_1(\theta_0)}{\sqrt{\widehat{\lambda}^2_{\theta_0}\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}}}, \tag{19}$$

where

$$\widehat{\pi}_1(\theta_0) = \left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}\sum_{t=1}^{T}(t-\bar{t})\left(z_t(\theta_0) - \bar{z}(\theta_0)\right),$$

$$\widehat{\lambda}^2_{\theta_0} = \widehat{\gamma}^{\theta_0}_0 + 2\sum_{j=1}^{T-1}k\left(\frac{j}{M}\right)\widehat{\gamma}^{\theta_0}_j, \quad \widehat{\gamma}^{\theta_0}_j = T^{-1}\sum_{t=j+1}^{T}\widehat{v}_t(\theta_0)\,\widehat{v}_{t-j}(\theta_0),$$

$$\widehat{v}_t(\theta_0) = z_t(\theta_0) - \bar{z}(\theta_0) - \widehat{\pi}_1(\theta_0)\left(t-\bar{t}\right).$$

Note that $\widehat{\pi}_1(\theta_0)$ is simply the OLS estimator of $\pi_1(\theta_0)$ from (18) and $\widehat{v}_t(\theta_0)$ are the corresponding OLS residuals.

1.4.2 Confidence Intervals Using t_{θ_0}

Confidence intervals for θ can be constructed by finding the values of θ_0 such that

$$|t_{\theta_0}| \leq cv_{\alpha/2} \tag{20}$$

where $cv_{\alpha/2}$ is the two-tail critical value for significance level α. Because both $\widehat{\pi}_1$ and $\widehat{\lambda}^2_{\theta_0}$ are functions of θ_0, finding the values of θ_0 that result in a non-rejection, i.e. satisfy (20), is equivalent to finding the roots of a particular second-order polynomial. Depending on whether the roots are real or complex, the confidence interval for θ_0 can be a closed interval on the real line, the complement of an open interval on the real line, or the entire real itself.

The form of the confidence interval depends on the magnitudes of the trend slopes relative to the noise as we now explain. Let

$$\widehat{\mathbf{\Omega}} = \widehat{\mathbf{\Gamma}}_0 + \sum_{j=1}^{T-1}k\left(\frac{j}{M}\right)(\widehat{\mathbf{\Gamma}}_j + \widehat{\mathbf{\Gamma}}'_j)$$

where

$$\widehat{\mathbf{\Gamma}}_j = T^{-1} \sum_{t=j+1}^{T} \widehat{\boldsymbol{u}}_t \widehat{\boldsymbol{u}}_{t-j}', \quad \widehat{\boldsymbol{u}}_t = [u_{1t},\ u_{2t}]'.$$

Partition $\widehat{\mathbf{\Omega}}$ as

$$\widehat{\mathbf{\Omega}} \equiv \begin{bmatrix} \widehat{\Omega}_{11} & \widehat{\Omega}_{12} \\ \widehat{\Omega}_{21} & \widehat{\Omega}_{22} \end{bmatrix}.$$

It is easy to show that

$$\widehat{\pi}_1(\theta_0) = \widehat{\beta}_1 - \theta_0 \widehat{\beta}_2,$$

and

$$\widehat{\lambda}_{\theta_0}^2 = \widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22},$$

allowing us to write (20) as

$$\frac{\left|\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2\right|}{\sqrt{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right) \left(\sum_{t=1}^{T} (t - \bar{t})^2\right)^{-1}}} \leq cv_{\alpha/2},$$

or equivalently

$$\frac{\left(\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2\right)^2}{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right) \left(\sum_{t=1}^{T} (t - \bar{t})^2\right)^{-1}} \leq cv_{\alpha/2}^2. \tag{21}$$

The inequality given by (21) can be rewritten as

$$c_2 \theta_0^2 + c_1 \theta_0 + c_0 \leq 0, \tag{22}$$

where

$$c_2 = \widehat{\beta}_2^2 - \Psi \widehat{\Omega}_{22}, \quad c_1 = \widehat{\beta}_1 \widehat{\beta}_2 - \Psi \widehat{\Omega}_{12}, \quad c_0 = \widehat{\beta}_1^2 - \Psi \widehat{\Omega}_{11}, \quad \Psi = cv_{\alpha/2}^2 \left(\sum_{t=1}^{T} (t - \bar{t})^2\right)^{-1}.$$

The values of θ_0 that solve (22) depend on the roots of the polynomial $p(\theta_0) =$

$c_2\theta_0^2 + c_1\theta_0 + c_0$. Let r_1 and r_2 denote the roots of $p(\theta_0)$ and when r_1 and r_2 are real, order the roots so that $r_1 \leq r_2$. There are three potential shapes for the confidence interval for θ_0:

Case 1: Suppose that $c_2 > 0$ and $c_1^2 - 4c_2c_0 \geq 0$. In this case, the roots are real and $p(\theta_0)$ opens upwards. The confidence interval is the values of θ_0 between the two roots, i.e. $\theta_0 \in [r_1, r_2]$. The inequality $c_2 > 0$ is equivalent to the inequality

$$\frac{\widehat{\beta}_2^2}{\widehat{\Omega}_{22}\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}} > cv_{\alpha/2}^2. \tag{23}$$

Notice that the left hand side of (23) is simply the square of the HAC robust t-statistic for testing that the trend slope of y_{2t} is zero. Inequality (23) holds if the trend slope of y_{2t} is statistically different from zero at the α level. This occurs when the trend slope for y_2 is large relative to the variation in u_{2t}. Mechanically, c_2 will be positive when the t-statistic for testing that $\beta_2 = 0$ is large in magnitude. Although not obvious, if $c_2 > 0$, it is impossible for $c_1^2 - 4c_2c_0 < 0$ to hold. In this case $p(\theta_0)$ opens upward and has its vertex above zero and the roots are complex; therefore, there are no solutions to (22) and the confidence interval would be empty. This case is impossible because the confidence interval cannot be empty because $\theta_0 = \widehat{\theta} = \widehat{\beta}_1/\widehat{\beta}_2$ is always contained in the interval given that $\widehat{\beta}_1 - \theta_0\widehat{\beta}_2 = 0$ in which case (21) must hold (equivalently (22) must hold).

Case 2: Suppose that $c_2 < 0$ and $c_1^2 - 4c_2c_0 > 0$. In this case, $p(\theta_0)$ has two real roots and opens downwards and the confidence interval is the values of θ_0 **not** between the two roots. In this case the confidence interval is the union of two disjoint sets and is given by $\theta_0 \in (-\infty, r_1] \cup [r_2, \infty)$.

Case 3: Suppose that $c_2 < 0$ and $c_1^2 - 4c_2c_0 \leq 0$. In this case, $p(\theta_0)$ opens downward and has a vertex at or below zero. The entire $p(\theta_0)$ function lies at or below zero and the confidence interval is the entire real line: $\theta_0 \in (-\infty, \infty)$.

Although it is a zero probability event, should $c_2 = 0$ then the confidence interval will take the form of either $\theta_0 \in (-\infty, -c_0/c_1]$ when $c_1 > 0$ or $\theta_0 \in [-c_0/c_1, \infty)$ when $c_1 < 0$.

While the confidence intervals constructed using t_{θ_0} can be wide when the trend slopes are small and there is no guarantee that these confidence intervals will contain the OLS or bias-corrected OLS estimators of θ, t_{θ_0} has a major advantage over the other t-statistics. Recall that we can write t_{θ_0} as

$$t_{\theta_0} = \frac{\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2}{\sqrt{\left(\widehat{\Omega}_{11} - 2\theta_0 \widehat{\Omega}_{12} + \theta_0^2 \widehat{\Omega}_{22}\right) \left(\sum_{t=1}^{T} (t - \bar{t})^2\right)^{-1}}}.$$

The denominator is a function $\widehat{u}_{1t}$ and $\widehat{u}_{2t}$ each of which are exactly invariant to the true values of β_1 and β_2. When H_0 is true, it follows that $\beta_1 - \beta_2\theta_0 = 0$ and we can write the numerator of t_{θ_0} as

$$\widehat{\beta}_1 - \theta_0 \widehat{\beta}_2 = \widehat{\beta}_1 - \theta_0 \widehat{\beta}_2 - (\beta_1 - \beta_2 \theta_0) = \left(\widehat{\beta}_1 - \beta_1\right) - \theta_0 \left(\widehat{\beta}_2 - \beta_2\right).$$

Because $\widehat{\beta}_1 - \beta_1$ and $\widehat{\beta}_2 - \beta_2$ are only functions of t and u_{1t}, u_{2t}, the numerator of t_{θ_0} is also exactly invariant to the true values of β_1 and β_2. Therefore, the null distribution of t_{θ_0} is exactly invariant to the true values of β_1 and β_2 including the case where both trend slopes are zero. In contrast, the other t-statistics have null distributions that depend on the magnitudes of β_1 and β_2. Because of its exact invariance to β_1 and β_2 under the null, t_{θ_0} will deliver much more robust inference (with respect to the magnitudes of β_1 and β_2) than the other t-statistics.

1.4.3 Asymptotic Results for t-statistics

In this section we provide asymptotic limits of the four t-statistics described in the previous sub-section. We derive asymptotic limits under alternatives that are local to the null given by (14). Suppose that $\beta_2 = T^{-\kappa}\overline{\beta}_2$. Then the alternative value of θ_1 is modeled local to θ_0 as

$$\theta_1 = \theta_0 + T^{-3/2+\kappa}\overline{\theta}_\Delta. \quad (24)$$

The parameter $\overline{\theta}_\Delta$ measures the magnitude of the departure from the null under the local alternative. In the results presented below, the asymptotic null distributions of the t-statistics are obtained by setting $\overline{\theta}_\Delta = 0$.

Recall that for the t_{θ_0} statistic, under the null that $\theta = \theta_0$ it follows that $\pi_1(\theta_0) = 0$. Under the local alternative (24), it follows that

$$\pi_1(\theta_0) = \beta_2(\theta_1 - \theta_0) = \beta_2 T^{-3/2+\kappa}\overline{\theta}_\Delta = T^{-\kappa}\overline{\beta}_2 T^{-3/2+\kappa}\overline{\theta}_\Delta = T^{-3/2}\overline{\beta}_2\overline{\theta}_\Delta, \quad (25)$$

regardless of the magnitude of the trend slopes. Therefore, the asymptotic limit of t_{θ_0} is invariant to the magnitude of the trend slopes under both the null and local alternative for θ.

We derive the limits of the various HAC estimators using fixed-b theory following Bunzel and Vogelsang (2005). The form of these limits depends on the type of kernel function used to compute the HAC estimator. We follow Bunzel and Vogelsang (2005) and use the following definitions.

Definition 1 *A kernel is labelled Type 1 if $k(x)$ is twice continuously differentiable everywhere and as a Type 2 kernel if $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$ and $k(x)$ is twice continuously differentiable everywhere except at $|x| = 1$.*

We also consider the Bartlett kernel (which is neither Type 1 or 2) separately. The fixed-b limiting distributions are expressed in terms of the following stochastic functions.

Definition 2 *Let $Q(r)$ be a generic stochastic process. Define the random variable*

$P_b(Q(r))$ *as*

$$P_b(Q(r)) = \begin{cases} \int_0^1 \int_0^1 -k^{*\prime\prime}(r-s)\, Q(r)Q(s)drds & \text{if } k(x) \text{ is Type 1} \\ \\ \int\int_{|r-s|<b} -k^{*\prime\prime}(r-s)\, Q(r)\, Q(s)\, drds & \\ \quad +2k_{-}^{*\prime}(b) \int_0^{1-b} Q(r+b)\, Q(r)\, dr & \text{if } k(x) \text{ is Type 2} \\ \\ \frac{2}{b}\int_0^1 Q(r)^2\, dr - \frac{2}{b}\int_0^{1-b} Q(r+b)\, Q(r)\, dr & \text{if } k(x) \text{ is Bartlett} \end{cases}$$

where $k^*(x) = k\left(\frac{x}{b}\right)$ *and* $k_{-}^{*\prime}$ *is the first derivative of* k^* *from below.*

The following theorems summarize the asymptotic limits of the t-statistics for testing (14) when the alternative is given by (24).

Theorem 4 *(Large Trend Slopes) Suppose that (6) holds. Let* $M = bT$ *where* $b \in (0,1]$ *is fixed. Let* $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$ *where* $\overline{\beta}_1, \overline{\beta}_2$ *are fixed with respect to* T, *and let* $\theta_1 = \theta_0 + T^{-3/2}\overline{\theta}_\Delta$. *Then as* $T \to \infty$,

$$t_{OLS}, t_{BC}, t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\beta_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\beta_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where $Z \sim N(0,1)$, $Q(r) = \widetilde{w}(r) - 12L(r)\int_0^1 \left(s - \frac{1}{2}\right) dw(s)$, $\widetilde{w}(r) = w(r) - rw(1)$, $L(r) = \int_0^r \left(s - \frac{1}{2}\right) ds$ *and* Z *and* $Q(r)$ *are independent.*

Theorem 5 *(Medium Trend Slopes): Suppose that (6) holds. Let* $M = bT$ *where* $b \in (0,1]$ *is fixed. Let* $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$ *where* $\overline{\beta}_1, \overline{\beta}_2$ *are fixed with respect* *to T, and let $\theta_1 = \theta_0 + T^{-1}\overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$t_{OLS} \Rightarrow \frac{Z}{\sqrt{P_b(H_1(r))}} + \frac{12\overline{\beta}_2^{-1}E(u_{2t}\epsilon_t(\theta))}{\sqrt{12\lambda_{\theta_1}^2 P_b(H_1(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(H_1(r))}},$$

$$t_{BC}, t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where $H_1(r) = Q(r) - 12\left(\lambda_{\theta_1}\overline{\beta}_2\right)^{-1} L(r) \cdot E(u_{2t}\epsilon_t(\theta))$.

Theorem 6 *(Small Trend Slopes): Suppose that (6) holds. Let $M = bT$ where $b \in (0,1]$ is fixed. Let $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + T^{-1/2}\overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$t_{OLS}, t_{BC} \xrightarrow{d} \frac{1}{\sqrt{\overline{\beta}_2^2 P_b\left(L(r)\right)\left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds + E\left(u_{2t}^2\right)\right)^{-1}}},$$

$$t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}}.$$

Theorem 7 *(Very Small Trend Slopes): Suppose that (3) and (6) hold. Let $M = bT$ where $b \in (0,1]$ is fixed. Let $\beta_1 = T^{-3/2}\overline{\beta}_1$, $\beta_2 = T^{-3/2}\overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + \overline{\theta}_\Delta$. Then as $T \to \infty$,*

$$T^{-1/2}t_{OLS}, T^{-1/2}t_{BC} = \frac{\left(E(u_{2t}^2)\right)^{-1}E(u_{2t}\epsilon_t(\theta)) + \overline{\theta}_\Delta}{\sqrt{P_b\left(H_2(r)\right)\left[E\left(u_{2t}^2\right)\right]^{-1}}},$$

$$t_{IV} \Rightarrow \frac{\int_0^1 (s-\frac{1}{2})dw(s) + \left(\overline{\beta}_2\int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)\lambda_\theta^{-1}\overline{\theta}_\Delta}{\sqrt{P_b\left(H_3(r)\right)\int_0^1 (s-\frac{1}{2})^2 ds}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_0}^2 P_b(Q(r))}},$$

where

$$H_2(r) = \widetilde{w}(r) - \left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2}) dB_2(s)\right)^{-1}$$
$$* \int_0^1 (s-\frac{1}{2}) dw(s) \left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right),$$
$$H_3(r) = \widetilde{w}(r) - 12\left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right)\left(\overline{\beta}_2 + 12\int_0^1 \left(s-\frac{1}{2}\right) dB_2(s)\right)^{-1}$$
$$* \int_0^1 \left(s-\frac{1}{2}\right) dw(s),$$

and $\widetilde{B}_2(r) = B_2(r) - rB_2(1)$.

Some interesting results and predictions about the finite sample behavior of the t-statistics are given by the Theorems 4-7. First examine the limiting null distributions that are obtained when $\overline{\theta}_\Delta = 0$. For large trend slopes, all four t-statistics have the same asymptotic null limit and the limiting random variable is the same fixed-b limit obtained by Bunzel and Vogelsang (2005) for inference regarding the trend slope in a simple linear trend model with stationary errors. Therefore, fixed-b critical values are available from Bunzel and Vogelsang (2005). As the trend slopes become smaller relative to the noise, differences among the t-statistics emerge. As anticipated, t_{θ_0} has the same limiting null distribution regardless of the magnitudes of the trend slopes. Except for very small trends slopes, t_{IV} has the same limiting null distribution as t_{θ_0}. The bias in OLS affects the null limit of t_{OLS} for medium, small and very small trend slopes. The bias correction helps for medium trend slopes in which case t_{BC} has the same null limit as t_{IV} and t_{θ_0}. For small and very small trend slopes the bias correction no longer works effectively and t_{BC} has the same limiting behavior as t_{OLS}. Both tests will tend to over-reject under the null when trends slopes are very small given that they diverge with the sample size.

In terms of finite sample null behavior of the t-statistics, the asymptotic theory predicts that t_{OLS} and t_{BC} will only work well when trend slopes are relatively large whereas t_{IV} should work well except when trend slopes are very small. The most reliable test in terms of robustness to magnitudes of trend slopes under the null should be t_{θ_0}.

When $\overline{\theta}_\Delta \neq 0$, in which case we are under the alternative, the t-statistics have additional terms in their limits which push the distributions away from the null distributions giving the tests power. When trend slopes are large, all four t-statistics have the same limiting distributions with the only minor difference being that t_{θ_0} depends on $\lambda^2_{\theta_0}$ rather than $\lambda^2_{\theta_1}$ as for the other t-statistics. In general we cannot rank $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$ as any difference depends on the joint serial correlation structure of u_{1t} and u_{2t}. Unless $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$ are nontrivially different, we would expect power of the tests to be similar in the large trend slope cases. As the trends slopes become smaller, power of t_{OLS} and t_{BC} becomes meaningless given that the statistics have poor behavior under the null. In Theorem 6, the limits of t_{OLS} and t_{BC} do not depend on $\overline{\theta}_\Delta$ which suggests that power will very low when trend slopes are small. In Theorem 7, t_{OLS} and t_{BC} diverge with the sample size which suggests large rejections with very small trend slopes. In constrast power of t_{IV} should be similar to t_{θ_0} except when trend slopes are very small. As in the case of null behavior, the asymptotic theory predicts that t_{θ_0} should perform the best in terms of power.

One thing to keep in mind regarding the limit of t_{θ_0} in Theorems 4-7 is that while the limit under the alternative is the same in each case, the relevant values of $\overline{\theta}_\Delta$ are farther away from the null in the case of smaller trends slopes compared to the case of larger trend slopes. Therefore, θ_1 needs to be much farther away from θ_0 in the case of small trend slopes than for the case of large trend slopes for power of t_{θ_0} to be the same in both cases. In other words, for a given value of θ_1, power of t_{θ_0} decreases as the trend slopes become smaller. This relationship between power and magnitudes of the trend slopes can be seen clearly in Theorem 4 where we can see that as $\beta_2 \to 0$ the limiting distribution under the local alternative for θ approaches the null distribution and power decreases.

1.5 Finite Sample Null Rejection Probabilities and Power

Using the same DGP as used in Section 3 we simulated finite sample null rejection probabilities and power of the four t-statistics. Table 2 reports null rejection probabilities for 5% nominal level tests for testing $H_0 : \theta = \theta_0 = 2$ against the two-sided alternative $H_1 : \theta \neq 2$. Results are reported for the same values of β_1, β_2 as used in Table 1 for $T = 50, 100, 200$ and $10,000$ replications are used in all cases. The HAC estimators are implemented using the Daniell kernel. Results for three bandwidth sample size ratios are provided: $b = 0.1, 0.5, 1.0$. For a given sample size, T, we use the bandwidth $M = bT$ for each of the three values of b. We compute empirical rejections using fixed-b asymptotic critical values using the critical value function

$$cv_{0.025}(b) = 1.9659 + 4.0603b + 11.6626b^2 + 34.8269b^3 - 13.9506b^4 + 3.2669b^5,$$

as given by Bunzel and Vogelsang (2005) for the Daniell kernel.

The patterns in the empirical null rejections closely match the predictions of the asymptotic results. When the trend slopes are large, $\beta_1 \geq 4$, $\beta_2 \geq 2$, null rejections are the essentially the same for all t-statistics and are close to 0.05 even when $T = 50$. This is true for all three bandwidth choices which illustrates the effectiveness of the fixed-b critical values. For medium sized trend slopes, $0.1 \leq \beta_1 \leq 0.4$, $0.05 \leq \beta_2 \leq 0.2$, t_{OLS} begins to show over-rejection problems that become very severe as the trend slopes decrease in magnitude. The bias-corrected OLS t-statistic, t_{BC}, is less subject to over-rejection problems especially when T is not small, although for $T = 50$, t_{BC} shows nontrivial over-rejection problems. In contrast both t_{IV} and t_{θ_0} have null rejections close to 0.05 for medium sized trend slopes. When the trend slopes are small or very small, $\beta_1 \leq 0.04$, $\beta_2 \leq 0.02$, the t_{OLS} and t_{BC} statistics have severe over-rejection problems and can reject 100% of the time. While t_{IV} has less over-rejection problems in this case, the over-rejections are nontrivial and are problematic. In contrast, t_{θ_0} has null rejections that are close to 0.05 regardless of the magnitudes of β_1, β_2 including the case of $\beta_1 = \beta_2 = 0$. In fact, the rejections are identical for t_{θ_0} across values of

β_1, β_2. This is because t_{θ_0} is exactly invariant to the values of β_1, β_2. It is clear in terms of null rejection probabilities that t_{θ_0} is the preferred test statistic.

Given that t_{θ_0} is the preferred statistic in terms of size, we computed, for each of the parameter configurations in Table 2, the proportions of replications that lead to the three possible shapes of confidence intervals one obtains by inverting t_{θ_0}. Recall that the cases are given by Case 1: $\theta_0 \in [r_1, r_2]$, Case 2: $\theta_0 \in (-\infty, r_1] \cup [r_2, \infty)$ and Case 3: $\theta_0 \in (-\infty, \infty)$. Table 3 gives these results. For large trend slopes Case 1 occurs 100% of the time. As the trend slopes decrease in magnitude, Case 2 occurs some of the time and as the trend slopes decrease further, Case 3 can occur frequently if the trend slopes are very small. As T increases, the likelihood of Case 1 increases for all trend slope magnitudes. The relative frequencies of the three cases also depends on the bandwidth but the relationship appears complicated. This is not surprising given the complex manner in which the bandwidth affects the null distribution of t_{θ_0}. Overall, unless trends slopes are small or very small, Case 1 is the most likely confidence interval shape.

While t_{θ_0} is the preferred test in terms of size, how do the t-statistics compare in terms of power? Table 4 reports power results for a subset of the grid of β_2 as used in Tables 2,3. For a given value of β_2, we specify a grid of six equally spaced values for θ in the range $\theta \in [2, \theta_{\max}]$ where $\theta = \theta_0 = 2$ is the null value and $\theta_{\max} = 0.01/\beta_2$. By construction $\beta_1 = \theta\beta_2$ in all cases. Given the way we define the grid for θ, we ensure that $\overline{\beta}_2\overline{\theta}_\Delta$ is the same for all values of β_2. Results are reported for $T = 100$. Results for other values of T are qualitatively similar and are omitted.

The patterns in power given in Table 4 are what we would expect given the local asymptotic limiting distributions.

For large trend slopes ($\beta_2 = 10, 2$), power of the four tests is essentially the same as predicted by Theorem 4. As the bandwidth increases, power of all the tests decreases. This inverse relationship between power and bandwidth is well known in the fixed-b literature (see Kiefer and Vogelsang 2005).

For medium sized trend slopes ($\beta_2 = 0.2, 0.1$) differences in power begin to emerge with t_{OLS} having substantially lower power than the other tests. This lower power occurs even though t_{OLS} over-rejects under the null when a small bandwidth is used ($b = 0.1$). With larger bandwidths, t_{OLS} has no power at all. Both t_{BC} and t_{IV} have good power that is somewhat lower than t_{θ_0}, which, according to Theorem 5 would result from differences between $\lambda^2_{\theta_0}$ and $\lambda^2_{\theta_1}$.

For small and very small trends slopes ($\beta_2 = 0.01, 0.001$), both t_{OLS} and t_{BC} are distorted under the null with severe over-rejections although t_{OLS} severely under-rejects with a large bandwidth ($b = 1.0$) when $\beta_2 = 0.01$. While t_{IV} is less size distorted than t_{OLS} and t_{BC}, it has no power regardless of bandwidth. In contrast to other three statistics, t_{θ_0} continues to have excellent size and power. The superior power of t_{θ_0} relative to the other statistics is completely in line with the predictions of Theorems 6 and 7.

In summary, the patterns in the finite sample simulations are consistent with the predictions of Theorems 4-7. Clearly t_{θ_0} is the recommended statistic given its superior behavior under the null and its higher power under the alternative.

1.6 Practical Recommendations

For point estimation, we recommend the IV estimator given its relative robustness to the magnitude of the trend slopes. OLS and bias-corrected OLS are not recommended given that they can become severely biased for small to very small trend slopes. For inference, we strongly recommend the t_{θ_0} statistic given its superior behavior under the null and the alternative both theoretically and in our limited finite sample simulations. Good empirical practice would be to report the IV estimator, $\widehat{\theta}$, along with the confidence interval constructed by inverting t_{θ_0}. Because this confidence interval must contain $\widehat{\theta}$, we avoid situations where the recommended point estimator lies outside the recommended confidence interval.

For confidence interval construction, there is also the practical need to choose a kernel and bandwidth. We do not explore this choice here but encourage empirical researchers to use the fixed-b critical values provided by Bunzel and Vogelsang (2005) once a kernel and bandwidth have been chosen.

1.7 Conclusion and Directions for Future Research

In this paper we analyze estimation and inference of the ratio of trend slopes of two time series with linear deterministic trend functions. We consider three estimators of the trend slope ratio: OLS, bias-corrected OLS, and IV. Asymptotic theory indicates that when the magnitude of the trend slopes are large relative to the noise in the series, the three estimators are approximately unbiased and have essentially equivalent sampling distributions. For small trend slopes, the IV estimator tends to remain unbiased whereas OLS and bias-corrected OLS can have substantial bias. For very small trend slopes all three estimators become poor estimators of the trend slopes ratio.

We analyze four t-statistics for testing hypotheses about the trend slopes ratio. We consider t-statistics based on each of the three estimators of the trend slopes ratio and we propose a fourth t-statistic based on an alternative testing approach. Asymptotic theory indicates that the alternative test dominates the other three tests in terms of size and power regardless of the magnitude of the trend slopes.

Finite sample simulations show that the predictions of the asymptotic theory tend to hold in practice. Based on the asymptotic theory and finite sample evidence we recommend that the IV estimator be used to estimate the trend slopes ratio and that confidence intervals be computed using our alternative test statistic. A nice property of our recommendation is that the IV estimator is always contained in the confidence interval even though the confidence interval is not constructed using the IV estimator itself.

A future research direction related to extension of the results in this paper is as follows: There may be empirical settings with more than two trending time series in which case more than one trend slope ratio can be estimated. It would be interesting to investigate whether panel methods can deliver better estimators of the trend slope ratios than applying the methods in this paper on a pairwise basis.

APPENDIX

Proofs of Theorems

Before giving proofs of the theorems, we prove a series of lemmas for each of the trend slope magnitude cases: large, medium, small and very small. The lemmas establish the limits of the scaled sums that appear in the estimators of θ and the HAC estimators. Using the results of the lemmas, the theorems are easy to establish using straightforward algebra and the continuous mapping theorem (CMT). We begin with a lemma that has limits of scaled sums that are exactly invariant to the magnitudes of the trend slopes followed by four lemmas for each of the trend slope cases. Throughout the appendix, we use ϵ_t to denote $\epsilon_t(\theta)$.

Lemma 1 *Suppose that (3) and (6) hold. The following hold as $T \to \infty$ for any values of β_1, β_2 :*

$$T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2 \to \int_0^1 (s-\frac{1}{2})^2 ds = \frac{1}{12},$$

$$T^{-2}\sum_{t=1}^{[rT]}(t-\bar{t}) \to \int_0^r (s-\frac{1}{2}) ds = L(r),$$

$$T^{-3/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon}) \Rightarrow \lambda_\theta \int_0^1 (s-\frac{1}{2}) dw(s),$$

$$T^{-1}\sum_{t=1}^{T}(u_{2t}-\bar{u}_2)(\epsilon_t-\bar{\epsilon}) \xrightarrow{p} E\left(u_{2t}\epsilon_t(\theta)\right),$$

$$T^{-1}\sum_{t=1}^{T}(u_{2t}-\bar{u}_2)^2 \xrightarrow{p} E\left(u_{2t}^2\right),$$

$$T^{-3/2}\sum_{t=1}^{T}(t-\bar{t})(u_{2t}-\bar{u}_2) \Rightarrow \int_0^1 (s-\frac{1}{2}) dB_2(s),$$

$$T^{3/2}\left(\widehat{\beta}_2-\beta_2\right) \Rightarrow \left(\int_0^1 (r-\frac{1}{2})^2 dr\right)^{-1} \int_0^1 (s-\frac{1}{2}) dB_2(s).$$

Proof: The results in this lemma are standard given the FCLTs (3) and (6) and ergodicity of u_{1t} and u_{2t}. See **?**.

Lemma 2 *(Large trend slopes) Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with* *respect to T. The following hold as $T \to \infty$ for $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) \Rightarrow \overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s),$$

$$T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) \Rightarrow \overline{\beta}_2\int_0^1(s-\frac{1}{2})dB_2(s),$$

$$T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 \xrightarrow{p} \overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds = \frac{\overline{\beta}_2^2}{12},$$

$$T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \xrightarrow{p} \overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds = \frac{\overline{\beta}_2}{12},$$

$$T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \Rightarrow \overline{\beta}_2L(r).$$

Proof: The results of the lemma are easy to establish once we substitute $y_{2t}-\overline{y}_2 = \overline{\beta}_2\left(t-\overline{t}\right)+(u_{2t}-\overline{u}_2)$ into each expression:

$$\begin{aligned}
T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) &= \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})+T^{-3/2}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})\\
&= \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})+o_p(1),\\
T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) &= \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2)+T^{-3/2}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2\\
&= \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2)+o_p(1),\\
T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 &= \overline{\beta}_2^2T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2+o_p(1),\\
T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) &= \overline{\beta}_2T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2+o_p(1),\\
T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) &= \overline{\beta}_2T^{-2}\sum_{t=1}^{[rT]}(t-\overline{t})+o_p(1).
\end{aligned}$$

The limits follow from Lemma 1.

Lemma 3 *(Medium trend slopes) Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed* *with respect to T. The following hold as $T \to \infty$ for $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,*

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) \Rightarrow \overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)+E\left(u_{2t}\epsilon_t\left(\theta\right)\right),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) \Rightarrow \overline{\beta}_2\int_0^1(s-\frac{1}{2})dB_2(s)+E\left(u_{2t}^2\right),$$

$$T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 \xrightarrow{p} \overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds = \frac{\overline{\beta}_2^2}{12},$$

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \xrightarrow{p} \overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds = \frac{\overline{\beta}_2}{12},$$

$$T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \Rightarrow \overline{\beta}_2L(r).$$

Proof: The results of the lemma are easy to establish once we substitute $y_{2t}-\overline{y}_2 = \overline{\beta}_2\left(t-\overline{t}\right)+(u_{2t}-\overline{u}_2)$ into each expression:

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) = \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})+T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) = \overline{\beta}_2T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2)+T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2,$$

$$T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 = \overline{\beta}_2^2T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2+o_p(1),$$

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) = \overline{\beta}_2T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2+o_p(1),$$

$$T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) = \overline{\beta}_2T^{-2}\sum_{t=1}^{[rT]}(t-\overline{t})+o_p(1).$$

The limits follow from Lemma 1.

Lemma 4 *(Small trend slopes) Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with* *respect to T. The following hold as $T \to \infty$ for $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,*

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)(\epsilon_t - \overline{\epsilon}) \xrightarrow{p} E\left(u_{2t}\epsilon_t\left(\theta\right)\right),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)(u_{2t} - \overline{u}_2) \xrightarrow{p} E\left(u_{2t}^2\right),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2 \xrightarrow{p} \overline{\beta}_2^2 \int_0^1 (s - \frac{1}{2})^2 ds + E\left(u_{2t}^2\right) = \frac{\overline{\beta}_2^2}{12} + E\left(u_{2t}^2\right),$$

$$T^{-2}\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2) \xrightarrow{p} \overline{\beta}_2 \int_0^1 (s - \frac{1}{2})^2 ds = \frac{\overline{\beta}_2}{12},$$

$$T^{-1}\sum_{t=1}^{[rT]}(y_{2t} - \overline{y}_2) \Rightarrow \overline{\beta}_2 L(r).$$

Proof: The results of the lemma are easy to establish once we substitute $y_{2t} - \overline{y}_2 = \overline{\beta}_2\left(t - \overline{t}\right) + (u_{2t} - \overline{u}_2)$ into each expression:

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)(\epsilon_t - \overline{\epsilon}) = \overline{\beta}_2 T^{-2}\sum_{t=1}^{T}(t - \overline{t})(\epsilon_t - \overline{\epsilon}) + T^{-1}\sum_{t=1}^{T}(u_{2t} - \overline{u}_2)(\epsilon_t - \overline{\epsilon})$$

$$= T^{-1}\sum_{t=1}^{T}(u_{2t} - \overline{u}_2)(\epsilon_t - \overline{\epsilon}) + o_p(1),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)(u_{2t} - \overline{u}_2) = \overline{\beta}_2 T^{-2}\sum_{t=1}^{T}(t - \overline{t})(u_{2t} - \overline{u}_2) + T^{-1}\sum_{t=1}^{T}(u_{2t} - \overline{u}_2)^2$$

$$= T^{-1}\sum_{t=1}^{T}(u_{2t} - \overline{u}_2)^2 + o_p(1),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2 = \overline{\beta}_2^2 T^{-3}\sum_{t=1}^{T}(t - \overline{t})^2 + T^{-1}\sum_{t=1}^{T}(u_{2t} - \overline{u}_2)^2 + o_p(1),$$

$$T^{-2}\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2) = \overline{\beta}_2 T^{-3}\sum_{t=1}^{T}(t - \overline{t})^2 + o_p(1),$$

$$T^{-1}\sum_{t=1}^{[rT]}(y_{2t} - \overline{y}_2) = \overline{\beta}_2 T^{-2}\sum_{t=1}^{[rT]}(t - \overline{t}) + o_p(1).$$

The limits follow from Lemma 1.

Lemma 5 *(Very small trend slopes) Suppose that (3) and (6) hold and $\overline{\beta}_1, \overline{\beta}_2$ are fixed* *with respect to T. The following hold as $T \to \infty$ for $\beta_1 = T^{-3/2}\overline{\beta}_1$, $\beta_2 = T^{-3/2}\overline{\beta}_2$,*

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) \xrightarrow{p} E\left(u_{2t}\epsilon_t\left(\theta\right)\right),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) \xrightarrow{p} E\left(u_{2t}^2\right),$$

$$T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 \xrightarrow{p} E\left(u_{2t}^2\right),$$

$$T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \xrightarrow{p} \overline{\beta}_2\int_0^1(s-\frac{1}{2})^2 ds + \int_0^1(s-\frac{1}{2})dB_2(s) = \frac{\overline{\beta}_2}{12} + \int_0^1(s-\frac{1}{2})dB_2(s),$$

$$T^{-1/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \Rightarrow \overline{\beta}_2 L(r) + \widetilde{B}_2(r).$$

Proof: The results of the lemma are easy to establish once we substitute $y_{2t}-\overline{y}_2 = \overline{\beta}_2\left(t-\overline{t}\right)+(u_{2t}-\overline{u}_2)$ into each expression:

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) &= \overline{\beta}_2 T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon}) + T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})\\
&= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}) + o_p(1),\\
T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) &= \overline{\beta}_2 T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2) + T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2\\
&= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2 + o_p(1),\\
T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2 &= \overline{\beta}_2^2 T^{-4}\sum_{t=1}^{T}(t-\overline{t})^2 + T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2 + o_p(1)\\
&= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)^2 + o_p(1),\\
T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) &= \overline{\beta}_2 T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2 + T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2),\\
T^{-1/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) &= \overline{\beta}_2 T^{-2}\sum_{t=1}^{[rT]}(t-\overline{t}) + T^{-1/2}\sum_{t=1}^{[rT]}(u_{2t}-\overline{u}_2).
\end{aligned}$$

The limits follow from Lemma 1.

Proof of Theorem 1. The proof follows directly from Lemmas 1, 2 and the CMT.

For the case of large trend slopes it follows that

$$\begin{aligned}
T^{3/2}\left(\widetilde{\theta}-\theta\right) &= \left(T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1} T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) \\
&\Rightarrow \left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1} \overline{\beta}_2\lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s), \\
T^{3/2}\left(\widehat{\theta}-\theta\right) &= \left(T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon}) \\
&\Rightarrow \left(\overline{\beta}_2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1} \lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s).
\end{aligned}$$

For the case of medium trend slopes it follows that

$$\begin{aligned}
T\left(\widetilde{\theta}-\theta\right) &= \left(T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1} T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) \\
&\Rightarrow \left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1} \left(\overline{\beta}_2\lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s) + E\left(u_{2t}\epsilon_t\left(\theta\right)\right)\right), \\
T\left(\widehat{\theta}-\theta\right) &= \left(T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon}) \\
&\Rightarrow \left(\overline{\beta}_2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1} \lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s).
\end{aligned}$$

Proof of Theorem 2. We first need to derive the limit of $T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t$ for the large and medium trend slope cases. For large trend slopes we have, using Lemmas 1, 2 and Theorem 1:

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t &= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}) - T^{-1}T^{3/2}\left(\widetilde{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2) \\
&- T^{-1}T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon}) \\
&+ T^{-1}T^{3/2}\left(\widetilde{\theta}-\theta\right)T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \\
&= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}) + o_p(1) \xrightarrow{p} E\left(u_{2t}\epsilon_t\left(\theta\right)\right). \qquad (26)
\end{aligned}$$

For medium trend slopes we have, using Lemmas 1,3 and Theorem 2:

$$T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t = T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}) - T^{-1}T\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2)$$

$$- T^{-1}T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})$$

$$+ T^{-1}T\left(\widetilde{\theta}-\theta\right)T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)$$

$$= T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon}) + o_p(1) \xrightarrow{p} E\left(u_{2t}\epsilon_t\left(\theta\right)\right). \tag{27}$$

The results for $\widetilde{\theta}^c$ are as follows. For large trend slopes

$$T^{3/2}\left(\widetilde{\theta}^c-\theta\right) = \left(T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}\left(T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) - T^{-3/2}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t\right)$$

$$= \left(T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}\left(T^{-3/2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) + o_p(1)\right)$$

$$\Rightarrow \left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s),$$

using Lemma 1, (26) and the CMT. For medium trend slopes

$$T\left(\widetilde{\theta}^c-\theta\right) = \left(T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon}) - T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t(\theta)\right)$$

$$\Rightarrow \left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\left(\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s) + E\left(u_{2t}\epsilon_t\left(\theta\right)\right) - E\left(u_{2t}\epsilon_t\left(\theta\right)\right)\right)$$

$$= \left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s),$$

using Lemma 2, (27) and the CMT.

Proof of Theorem 3. We first give the proof for the case of small trend slopes. Using

Lemmas 1, 4 and the CMT it follows that

$$\widetilde{\theta}-\theta=\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon})$$

$$\xrightarrow{p}\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds+E(u_{2t}^2)\right)^{-1}E\left(u_{2t}\epsilon_t\left(\theta\right)\right)=\Re,$$

$$T^{1/2}\left(\widehat{\theta}-\theta\right)=\left(T^{-2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right)^{-1}T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})$$

$$\Rightarrow\left(\overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s).$$

To establish the result for $\widetilde{\theta}^c$ we first need to derive the limit of $T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t$. Using Lemmas 1, 4 and the results for OLS

$$\begin{aligned}T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t(\theta)&=T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})-\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2)\\&-T^{-1}T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})\\&+T^{-1/2}\left(\widetilde{\theta}-\theta\right)T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\\&=T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})-\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2)+o_p(1)\\&\xrightarrow{p}E\left(u_{2t}\epsilon_t\left(\theta\right)\right)-\Re E(u_{2t}^2)=\overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds\Re.\end{aligned}\tag{28}$$

Using Lemma 4, (28) and the CMT, it follows that

$$\widetilde{\theta}^c-\theta=\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon})-T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t(\theta)\right)$$

$$\xrightarrow{p}\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds+E(u_{2t}^2)\right)^{-1}\left(E\left(u_{2t}\epsilon_t\left(\theta\right)\right)-\overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds\Re\right)$$

$$=\frac{\Re^2E(u_{2t}^2)}{E\left(u_{2t}\epsilon_t\left(\theta\right)\right)}=\Re_c.$$

The proof for very small slopes is similar. Using Lemmas 1, 5 and the CMT it follows that

$$\widetilde{\theta}-\theta=\frac{T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon})}{T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2}\xrightarrow{p}\frac{E\left(u_{2t}\epsilon_t\left(\theta\right)\right)}{E(u_{2t}^2)},$$

$$\widehat{\theta}-\theta=\left(T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right)^{-1}T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})$$

$$\Rightarrow\left(\overline{\beta}_2\int_0^1(s-\frac{1}{2})^2ds+\int_0^1(s-\frac{1}{2})dB_2(s)\right)^{-1}\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s).$$

To establish the result for $\widetilde{\theta}^c$ we first need to derive the limit of $T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t$. Using Lemmas 1, 5 and the results for OLS

$$\begin{aligned}T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t(\theta)&=T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})-\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2)\\&-T^{-1}T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon})\\&+T^{-1}\left(\widetilde{\theta}-\theta\right)T^{3/2}\left(\widehat{\beta}_2-\beta_2\right)T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\\&=T^{-1}\sum_{t=1}^{T}(u_{2t}-\overline{u}_2)(\epsilon_t-\overline{\epsilon})-\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(u_{2t}-\overline{u}_2)+o_p(1)\\&\xrightarrow{p}E\left(u_{2t}\epsilon_t\left(\theta\right)\right)-\frac{E\left(u_{2t}\epsilon_t\left(\theta\right)\right)}{E(u_{2t}^2)}E(u_{2t}^2)=0.\end{aligned}\tag{29}$$

Using Lemma 4, (29) and the CMT, it follows that

$$\begin{aligned}\widetilde{\theta}^c-\theta&=\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right)^{-1}\left(T^{-1}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)(\epsilon_t-\overline{\epsilon})-T^{-1}\sum_{t=1}^{T}\widehat{u}_{2t}\widetilde{\epsilon}_t(\theta)\right)\\&=\widetilde{\theta}-\theta+o_p(1)\xrightarrow{p}\frac{E\left(u_{2t}\epsilon_t\left(\theta\right)\right)}{E(u_{2t}^2)}.\end{aligned}$$

We now provide proofs of Theorems 4-7. The proofs involve two steps. For a given t-statistic, we first establish the fixed-b limit of the relevant HAC estimator and then we establish the limit of the t-statistic using Theorems 1-3 and the CMT. To establish the fixed-b limit of a HAC estimator we rewrite the HAC estimator in terms of scaled partial sum processes using the fixed-b algebra of Kiefer and Vogelsang (2005). We do not provide details of the fixed-b algebra here and we simply establish the scaling and limits of the relevant partial sum processes which are given by $\widetilde{S}_{[rT]}=\sum_{t=1}^{[rT]}\widetilde{\epsilon}_t$, $\widetilde{S}^c_{[rT]}=\sum_{t=1}^{[rT]}\left(\widetilde{\epsilon}^c_t-\overline{\widetilde{\epsilon}^c}\right)$ and $\widehat{S}_{[rT]}=\sum_{t=1}^{[rT]}\widehat{\epsilon}_t$.

Note that no proofs are needed for the limit of t_{θ_0} because t_{θ_0} is computed using regression model (18) under the local alternative for $\pi_1(\theta_0)$ given by (25). This is a special case of Theorem 2 in Bunzel and Vogelsang (2005).

Proof of Theorem 4. With large trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from the Lemmas 1, 2 and Theorems 1, 2. For the OLS estimator we have

$$\begin{aligned}
T^{-1/2}\widetilde{S}_{[rT]} &= T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - T^{3/2}\left(\widetilde{\theta}-\theta\right)T^{-2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \\
&\Rightarrow \lambda_\theta \widetilde{w}(r) - \left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1}\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)\overline{\beta}_2 L(r) \\
&= \lambda_\theta\left[\widetilde{w}(r) - 12L(r)\int_0^1(s-\frac{1}{2})dw(s)\right] = \lambda_\theta Q(r).
\end{aligned}$$

The forms of $\widetilde{S}^c_{[rT]}$ and $\widehat{S}_{[rT]}$ are identical with $\widetilde{\theta}$ replaced by $\widetilde{\theta}^c$ or $\widehat{\theta}$. Because the three estimators have the same asymptotic distributions, it follows that

$$T^{-1/2}\widetilde{S}^c_{[rT]}, T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \lambda_\theta Q(r).$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$\widetilde{\lambda}^2_\theta, \widetilde{\lambda}^2_{\theta c}, \widehat{\lambda}^2_\theta \Rightarrow \lambda^2_\theta P_b\left(Q(r)\right).$$

The limit of t_{OLS} is as follows. The arguments for t_{BC} and t_{IV} are similar and are omitted. Simple algebra gives

$$\begin{aligned}
t_{OLS} &= \frac{T^{3/2}\left(\widetilde{\theta}-\theta_0\right)}{\sqrt{\widetilde{\lambda}^2_\theta\left[T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right]^{-1}}} = \frac{T^{3/2}\left(\widetilde{\theta}-\theta_1\right)+\overline{\theta}_\Delta}{\sqrt{\widetilde{\lambda}^2_\theta\left[T^{-3}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right]^{-1}}} \\
&\Rightarrow \frac{\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\overline{\beta}_2\lambda_{\theta_1}\int_0^1(s-\frac{1}{2})dw(s)+\overline{\theta}_\Delta}{\sqrt{\lambda^2_{\theta_1}P_b\left(Q(r)\right)\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}}} \\
&= \frac{\sqrt{12}\int_0^1(s-\frac{1}{2})dw(s)}{\sqrt{P_b\left(Q(r)\right)}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_1}P_b\left(Q(r)\right)}} = \frac{Z}{\sqrt{P_b\left(Q(r)\right)}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_1}P_b\left(Q(r)\right)}},
\end{aligned}$$

using the fact that $\int_0^1(s-\frac{1}{2})dw(s) \sim N\left(0, \frac{1}{12}\right)$.

Proof of Theorem 5. With medium trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from the Lemmas 1, 3 and

Theorems 1, 2. For the OLS estimator we have

$$
\begin{aligned}
T^{-1/2}\widetilde{S}_{[rT]} &= T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t-\overline{\epsilon}) - T\left(\widetilde{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \\
&\Rightarrow \lambda_\theta \widetilde{w}(r) - \left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1}\left(\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)+E\left(u_{2t}\epsilon_t\left(\theta\right)\right)\right)\overline{\beta}_2 L(r) \\
&= \lambda_\theta\left(Q(r)-12\left(\lambda_\theta\overline{\beta}_2\right)^{-1}L(r)E\left(u_{2t}\epsilon_t\left(\theta\right)\right)\right) = \lambda_\theta H_1(r).
\end{aligned}
$$

Because the bias-corrected OLS estimator and IV have the same limit, the limits of the scaled partial sums are the same and we only give details for IV:

$$
\begin{aligned}
T^{-1/2}\widehat{S}_{[rT]} &= T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t-\overline{\epsilon}) - T\left(\widehat{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2) \\
&\Rightarrow \lambda_\theta \widetilde{w}(r) - \left(\overline{\beta}_2^2\int_0^1 (s-\frac{1}{2})^2 ds\right)^{-1}\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)\overline{\beta}_2 L(r) = \lambda_\theta Q(r).
\end{aligned}
$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$
\widetilde{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(H_1(r)\right), \quad \widetilde{\lambda}_{\theta c}^2, \widehat{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(Q(r)\right).
$$

The limit of t_{OLS} is as follows:

$$
\begin{aligned}
t_{OLS} &= \frac{T\left(\widetilde{\theta}-\theta_0\right)}{\sqrt{\widetilde{\lambda}_\theta^2\left[T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right]^{-1}}} = \frac{T\left(\widetilde{\theta}-\theta_1\right)+\overline{\theta}_\Delta}{\sqrt{\widetilde{\lambda}_\theta^2\left[T^{-2}\sum_{t=1}^{T}(y_{2t}-\overline{y}_2)^2\right]^{-1}}} \\
&\Rightarrow \frac{\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\left(\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)+E\left(u_{2t}\epsilon_t\left(\theta\right)\right)\right)+\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(H_1(r)\right)\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}}}.
\end{aligned}
$$

Straightforward algebraic manipulation gives the form given by Theorem 5. Because the limits of t_{BC} and t_{IV} are identical, details are only given for t_{IV}:

$$
\begin{aligned}
t_{IV} &= \frac{T\left(\widehat{\theta}-\theta_0\right)}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}} \\
&= \frac{T\left(\widehat{\theta}-\theta_1\right)+\overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}} \\
&\Rightarrow \frac{\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\overline{\beta}_2\lambda_{\theta_1}\int_0^1(s-\frac{1}{2})dw(s)+\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2P_b\left(Q(r)\right)\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}}},
\end{aligned}
$$

which is the same limit obtained in the proof of Theorem 4.

Proof of Theorem 6. With small trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from the Lemmas 1, 4 and Theorem 3. For the OLS and bias-corrected OLS estimators we have

$$
\begin{aligned}
T^{-1}\widetilde{S}_{[rT]} &= T^{-1}\sum_{t=1}^{[rT]}\left(\epsilon_t-\overline{\epsilon}\right)-\left(\widetilde{\theta}-\theta\right)T^{-1}\sum_{t=1}^{[rT]}\left(y_{2t}-\overline{y}_2\right)\xrightarrow{p}-\Re\overline{\beta}_2L(r), \\
T^{-1}\widetilde{S}_{[rT]}^c &= T^{-1}\sum_{t=1}^{[rT]}\left(\epsilon_t-\overline{\epsilon}\right)-\left(\widetilde{\theta}^c-\theta\right)T^{-1}\sum_{t=1}^{[rT]}\left(y_{2t}-\overline{y}_2\right)\xrightarrow{p}-\Re_c\overline{\beta}_2L(r),
\end{aligned}
$$

whereas for the IV estimator we have

$$
\begin{aligned}
T^{-1/2}\widehat{S}_{[rT]} &= T^{-1/2}\sum_{t=1}^{[rT]}\left(\epsilon_t-\overline{\epsilon}\right)-T^{1/2}\left(\widehat{\theta}-\theta\right)T^{-1}\sum_{t=1}^{[rT]}\left(y_{2t}-\overline{y}_2\right) \\
&\Rightarrow \lambda_\theta\widetilde{w}(r)-\left(\overline{\beta}_2^2\int_0^1(s-\frac{1}{2})^2ds\right)^{-1}\overline{\beta}_2\lambda_\theta\int_0^1(s-\frac{1}{2})dw(s)\overline{\beta}_2L(r)=\lambda_\theta Q(r).
\end{aligned}
$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$
T^{-1}\widetilde{\lambda}_\theta^2\xrightarrow{p}\Re^2\overline{\beta}_2^2P_b\left(L(r)\right),\quad T^{-1}\widetilde{\lambda}_{\theta c}^2\xrightarrow{p}\Re_c^2\overline{\beta}_2^2P_b\left(L(r)\right),\quad \widehat{\lambda}_\theta^2\Rightarrow\lambda_\theta^2P_b\left(Q(r)\right).
$$

The limits of t_{OLS} and t_{BC} are identical and details are only given for t_{OLS}:

$$t_{OLS} = \frac{\left(\widetilde{\theta} - \theta_0\right)}{\sqrt{T^{-1}\widetilde{\lambda}_\theta^2 \left[T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}} = \frac{\left(\widetilde{\theta} - \theta_1\right) + T^{-1/2}\overline{\theta}_\Delta}{\sqrt{T^{-1}\widetilde{\lambda}_\theta^2 \left[T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}}$$

$$\Rightarrow \frac{\Re}{\sqrt{\Re^2 \overline{\beta}_2^2 P_b\left(L(r)\right)\left(\overline{\beta}_2^2 \int_0^1 (s - \frac{1}{2})^2 ds + E\left(u_{2t}^2\right)\right)^{-1}}}$$

$$= \frac{1}{\sqrt{\overline{\beta}_2^2 P_b\left(L(r)\right)\left(\overline{\beta}_2^2 \int_0^1 (s - \frac{1}{2})^2 ds + E\left(u_{2t}^2\right)\right)^{-1}}}.$$

The limit of t_{IV} is given by

$$t_{IV} = \frac{T^{1/2}\left(\widehat{\theta} - \theta_0\right)}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-2}\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t - \overline{t})^2}}$$

$$= \frac{T^{1/2}\left(\widehat{\theta} - \theta_1\right) + \overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-2}\sum_{t=1}^{T}(t - \overline{t})(y_{2t} - \overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t - \overline{t})^2}}$$

$$\Rightarrow \frac{\left(\overline{\beta}_2^2 \int_0^1 (s - \frac{1}{2})^2 ds\right)^{-1} \overline{\beta}_2 \lambda_{\theta_1} \int_0^1 (s - \frac{1}{2}) dw(s) + \overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(Q(r)\right)\left(\overline{\beta}_2^2 \int_0^1 (s - \frac{1}{2})^2 ds\right)^{-1}}},$$

which is the same limit obtained as in the proof of Theorem 4.

Proof of Theorem 7. With very small trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from the Lemmas 1, 5 and Theorem 3. Results for OLS and bias-corrected OLS are the same and we only give details for OLS:

$$T^{-1/2}\widetilde{S}_{[rT]} = T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - \left(\widetilde{\theta} - \theta\right) T^{-1/2}\sum_{t=1}^{[rT]}(y_{2t} - \overline{y}_2)$$

$$\Rightarrow \lambda_\theta \widetilde{w}(r) - \frac{E(u_{2t}\epsilon_t(\theta))}{E(u_{2t}^2)}\left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right) = H_2(r),$$

For the IV estimator we have

$$T^{-1/2}\widehat{S}_{[rT]} = T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta} - \theta\right) T^{-1/2}\sum_{t=1}^{[rT]}(y_{2t} - \overline{y}_2)$$

$$\Rightarrow \lambda_\theta \widetilde{w}(r) - \left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)^{-1} \lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s) \left(\overline{\beta}_2 L(r) + \widetilde{B}_2(r)\right)$$

$$= \lambda_\theta H_3(r).$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$\widetilde{\lambda}_\theta^2, \widetilde{\lambda}_{\theta c}^2 \Rightarrow P_b\left(H_2(r)\right), \quad \widehat{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(H_3(r)\right).$$

The limits of t_{OLS} and t_{BC} are identical and details are only given for t_{OLS}:

$$T^{-1/2}t_{OLS} = \frac{\widetilde{\theta} - \theta_0}{\sqrt{\widetilde{\lambda}_\theta^2 \left[T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}} = \frac{\left(\widetilde{\theta} - \theta_1\right) + \overline{\theta}_\Delta}{\sqrt{\widetilde{\lambda}_\theta^2 \left[T^{-1}\sum_{t=1}^{T}(y_{2t} - \overline{y}_2)^2\right]^{-1}}}$$

$$\Rightarrow \frac{\left(E(u_{2t}^2)\right)^{-1} E(u_{2t}\epsilon_t(\theta)) + \overline{\theta}_\Delta}{\sqrt{P_b\left(H_2(r)\right)\left[E\left(u_{2t}^2\right)\right]^{-1}}}.$$

The limit of t_{IV} is as follows:

$$t_{IV} = \frac{\widehat{\theta} - \theta_0}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t} - \overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}}$$

$$= \frac{\left(\widehat{\theta} - \theta_1\right) + \overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-3/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t} - \overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}}$$

$$\Rightarrow \frac{\left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)^{-1} \lambda_\theta \int_0^1 (s-\frac{1}{2})dw(s) + \overline{\theta}_\Delta}{\sqrt{\lambda_\theta^2 P_b\left(H_3(r)\right)\left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)^{-2} \int_0^1 (s-\frac{1}{2})^2 ds}}$$

$$= \frac{\int_0^1 (s-\frac{1}{2})dw(s) + \left(\overline{\beta}_2 \int_0^1 (s-\frac{1}{2})^2 ds + \int_0^1 (s-\frac{1}{2})dB_2(s)\right)\lambda_\theta^{-1}\overline{\theta}_\Delta}{\sqrt{P_b\left(H_3(r)\right)\int_0^1 (s-\frac{1}{2})^2 ds}}.$$

Table 1.1: Finite Sample Means and Standard Deviations.
10,000 Replications, $\theta = 2$.

			Mean			Standard Deviation		
T	β_1	β_2	OLS	OLS_{BC}	IV	OLS	OLS_{BC}	IV
50	20	10	2.000	2.000	2.000	0.003	0.003	0.003
	10	5	2.000	2.000	2.000	0.006	0.006	0.006
	4	2	1.998	2.000	2.000	0.015	0.015	0.015
	.4	.2	1.816	2.014	2.014	0.117	0.156	0.156
	.3	.15	1.697	2.024	2.024	0.132	0.214	0.214
	.2	.1	1.441	2.058	2.058	0.145	0.353	0.353
	.1	.05	0.911	2.242	2.288	0.185	1.051	14.117
	.04	.02	0.553	1.409	1.897	0.175	1.454	75.360
	.004	.002	0.463	0.544	1.931	0.155	0.860	191.451
	0	0	0.462	0.534	2.114	0.155	0.840	160.777
100	20	10	2.000	2.000	2.000	0.001	0.001	0.001
	10	5	2.000	2.000	2.000	0.002	0.002	0.002
	4	2	1.999	2.000	2.000	0.005	0.005	0.005
	.4	.2	1.946	2.002	2.002	0.050	0.054	0.054
	.3	.15	1.906	2.003	2.003	0.063	0.073	0.073
	.2	.1	1.802	2.007	2.007	0.082	0.110	0.110
	.1	.05	1.422	2.028	2.028	0.102	0.232	0.232
	.04	.02	0.778	2.192	2.205	0.130	0.799	3.965
	.004	.002	0.469	0.602	0.263	0.110	0.719	73.150
	0	0	0.465	0.528	0.631	0.109	0.619	89.149
200	20	10	2.000	2.000	2.000	0.000	0.000	0.000
	10	5	2.000	2.000	2.000	0.001	0.001	0.001
	4	2	2.000	2.000	2.000	0.002	0.002	0.002
	.4	.2	1.985	2.000	2.000	0.019	0.019	0.019
	.3	.15	1.974	2.000	2.000	0.025	0.025	0.025
	.2	.1	1.944	2.001	2.001	0.035	0.038	0.038
	.1	.05	1.796	2.003	2.003	0.058	0.077	0.077
	.04	.02	1.244	2.021	2.021	0.074	0.201	0.201
	.004	.002	0.484	0.968	1.644	0.078	0.731	88.359
	0	0	0.469	0.517	0.862	0.076	0.437	50.018

Note: OLS, OLS_{BC} and IV denote the estimators given by (7), (11), and (13) respectively.

Table 1.2: Empirical Null Rejection Probabilities, 5% Nominal Level, 10,000 Rep. $H_0: \theta = \theta_0 = 2$, $H_1: \theta \neq 2$.

			$b = 0.1$		$b = 0.5$		$b = 1.0$	
T	β_1	β_2	t_{OLS}	t_{BC}	t_{OLS}	t_{BC}	t_{OLS}	t_{BC}
50	20	10	.065	.065	.050	.051	.049	.053
	10	5	.065	.065	.049	.051	.038	.053
	4	2	.066	.065	.043	.051	.010	.053
	.4	.2	.236	.075	.001	.056	.000	.055
	.3	.15	.356	.087	.001	.062	.000	.059
	.2	.1	.610	.119	.001	.075	.000	.071
	.1	.05	.981	.261	.004	.131	.000	.122
	.04	.02	1.00	.538	.203	.340	.031	.250
	.004	.002	1.00	.874	.708	.719	.258	.546
	0	0	1.00	.882	.721	.730	.262	.552
100	20	10	.054	.054	.054	.053	.049	.052
	10	5	.054	.054	.054	.053	.042	.052
	4	2	.054	.054	.047	.053	.018	.052
	.4	.2	.140	.057	.001	.055	.000	.054
	.3	.15	.198	.059	.000	.056	.000	.055
	.2	.1	.341	.067	.000	.060	.000	.059
	.1	.05	.811	.106	.000	.080	.000	.076
	.04	.02	1.00	.334	.001	.183	.000	.161
	.004	.002	1.00	.908	.855	.807	.356	.627
	0	0	1.00	.947	.908	.856	.412	.679
200	20	10	.047	.047	.046	.045	.050	.049
	10	5	.047	.047	.045	.045	.046	.049
	4	2	.046	.047	.044	.045	.027	.049
	.4	.2	.090	.048	.001	.046	.000	.049
	.3	.15	.121	.048	.000	.046	.000	.049
	.2	.1	.197	.050	.000	.047	.000	.050
	.1	.05	.504	.061	.000	.053	.000	.056
	.04	.02	.995	.135	.000	.087	.000	.087
	.004	.002	1.00	.846	.818	.680	.183	.453
	0	0	1.00	.991	.989	.953	.605	.789

Note: The formulas t_{OLS}, t_{BC} are given by (15), (16) respectively.

Table 1.2 (cont'd)

T	β_1	β_2	$b = 0.1$ t_{IV}	$b = 0.1$ t_{θ_0}	$b = 0.5$ t_{IV}	$b = 0.5$ t_{θ_0}	$b = 1.0$ t_{IV}	$b = 1.0$ t_{θ_0}
50	20	10	.065	.065	.051	.051	.053	.053
	10	5	.065	.065	.051	.051	.053	.053
	4	2	.065	.065	.051	.051	.053	.053
	.4	.2	.059	.065	.050	.051	.050	.053
	.3	.15	.059	.065	.050	.051	.049	.053
	.2	.1	.061	.065	.049	.051	.049	.053
	.1	.05	.082	.065	.052	.051	.051	.053
	.04	.02	.138	.065	.079	.051	.074	.053
	.004	.002	.232	.065	.129	.051	.116	.053
	0	0	.231	.065	.132	.051	.118	.053
100	20	10	.054	.054	.053	.053	.052	.052
	10	5	.054	.054	.053	.053	.052	.052
	4	2	.054	.054	.053	.053	.052	.052
	.4	.2	.054	.054	.054	.053	.052	.052
	.3	.15	.054	.054	.054	.053	.053	.052
	.2	.1	.053	.054	.054	.053	.052	.052
	.1	.05	.055	.054	.053	.053	.054	.052
	.04	.02	.072	.054	.054	.053	.056	.052
	.004	.002	.211	.054	.121	.053	.109	.052
	0	0	.226	.054	.124	.053	.114	.052
200	20	10	.047	.047	.045	.045	.049	.049
	10	5	.047	.047	.045	.045	.049	.049
	4	2	.047	.047	.045	.045	.049	.049
	.4	.2	.046	.047	.045	.045	.049	.049
	.3	.15	.047	.047	.046	.045	.049	.049
	.2	.1	.046	.047	.046	.045	.049	.049
	.1	.05	.047	.047	.046	.045	.049	.049
	.04	.02	.049	.047	.047	.045	.050	.049
	.004	.002	.149	.047	.092	.045	.080	.049
	0	0	.216	.047	.122	.045	.113	.049

Note: The formulas t_{IV}, and t_{θ_0} are given by (17) and (19) respectively.

Table 1.3: Finite Sample Proportions of Confidence Interval Shapes Based on t_{θ_0}. 5% Nominal Level, 10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta \neq 2$.

			$b = 0.1$		$b = 0.5$		$b = 1.0$	
T	β_1	β_2	Case1	Case2	Case1	Case2	Case1	Case2
50	20	10	1.000	.000	1.000	.000	1.000	.000
	10	5	1.000	.000	1.000	.000	1.000	.000
	4	2	1.000	.000	1.000	.000	1.000	.000
	.4	.2	1.000	.000	.987	.013	.922	.078
	.3	.15	1.000	.000	.928	.072	.812	.187
	.2	.1	.999	.001	.744	.255	.612	.384
	.1	.05	.723	.277	.339	.605	.293	.650
	.02	.01	.202	.514	.109	.517	.104	.610
	.002	.001	.071	.131	.053	.270	.053	.378
	0	0	.066	.127	.051	.271	.052	.376
100	20	10	1.000	.000	1.000	.000	1.000	.000
	10	5	1.000	.000	1.000	.000	1.000	.000
	4	2	1.000	.000	1.000	.000	1.000	.000
	.4	.2	1.000	.000	1.000	.000	1.000	.000
	.3	.15	1.000	.000	1.000	.000	1.000	.000
	.2	.1	1.000	.000	.999	.001	.986	.014
	.1	.05	1.000	.000	.897	.103	.772	.227
	.02	.01	.761	.239	.386	.574	.332	.623
	.002	.001	.063	.165	.053	.304	.055	.404
	0	0	.053	.122	.050	.279	.051	.382
200	20	10	1.000	.000	1.000	.000	1.000	.000
	10	5	1.000	.000	1.000	.000	1.000	.000
	4	2	1.000	.000	1.000	.000	1.000	.000
	.4	.2	1.000	.000	1.000	.000	1.000	.000
	.3	.15	1.000	.000	1.000	.000	1.000	.000
	.2	.1	1.000	.000	1.000	.000	1.000	.000
	.1	.05	1.000	.000	1.000	.000	.999	.001
	.02	.01	1.000	.000	.931	.069	.819	.181
	.002	.001	.113	.412	.085	.449	.081	.547
	0	0	.043	.120	.046	.274	.051	.376

Notes: Case 1 is $\theta_0 \in [r_1, r_2]$, Case 2 is $\theta_0 \in (-\infty, r_1] \cup [r_2, \infty)$.

Table 1.3 (cont'd)

			$b = 0.1$	$b = 0.5$	$b = 1.0$
T	β_1	β_2	Case3	Case3	Case3
50	20	10	.000	.000	.000
	10	5	.000	.000	.000
	4	2	.000	.000	.000
	.4	.2	.000	.000	.000
	.3	.15	.000	.000	.000
	.2	.1	.000	.001	.004
	.1	.05	.000	.056	.057
	.02	.01	.284	.374	.287
	.002	.001	.799	.677	.569
	0	0	.806	.678	.572
100	20	10	.000	.000	.000
	10	5	.000	.000	.000
	4	2	.000	.000	.000
	.4	.2	.000	.000	.000
	.3	.15	.000	.000	.000
	.2	.1	.000	.000	.000
	.1	.05	.000	.000	.001
	.02	.01	.000	.040	.045
	.002	.001	.772	.643	.541
	0	0	.825	.671	.567
200	20	10	.000	.000	.000
	10	5	.000	.000	.000
	4	2	.000	.000	.000
	.4	.2	.000	.000	.000
	.3	.15	.000	.000	.000
	.2	.1	.000	.000	.000
	.1	.05	.000	.000	.000
	.02	.01	.000	.000	.000
	.002	.001	.475	.467	.372
	0	0	.837	.680	.573

Notes: Case 3 is $\theta_0 \in (-\infty, \infty)$.

Table 1.4: Finite Sample Power, 5% Nominal Level, $T = 100$, Two-sided Tests. 10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta = \theta_1$, $\beta_1 = \theta_1\beta_2$.

		$b = 0.1$		$b = 0.5$		$b = 1.0$	
β_2	θ_1	t_{OLS}	t_{BC}	t_{OLS}	t_{BC}	t_{OLS}	t_{BC}
10	2.000	.054	.054	.054	.053	.049	0.052
	2.001	.133	.138	.090	.093	.082	.085
	2.002	.383	.389	.200	.205	.173	.180
	2.003	.691	.699	.346	.350	.295	.300
	2.004	.906	.910	.500	.503	.413	.419
	2.005	.984	.985	.642	.644	.522	.529
2	2.000	.054	.054	.047	.053	.018	.052
	2.005	.119	.136	.072	.092	.034	.085
	2.010	.350	.387	.176	.203	.092	.179
	2.015	.662	.694	.321	.345	.195	.297
	2.020	.889	.907	.470	.499	.320	.416
	2.025	.979	.984	.615	.638	.442	.524
.2	2.00	.140	.057	.001	.055	.000	.054
	2.05	.036	.123	.000	.086	.000	.081
	2.10	.073	.356	.000	.188	.000	.171
	2.15	.246	.662	.000	.317	.000	.278
	2.20	.529	.886	.003	.451	.000	.381
	2.25	.789	.977	.009	.574	.000	.474
.1	2.0	.341	.067	.000	.060	.000	.059
	2.1	.089	.121	.000	.087	.000	.083
	2.2	.120	.349	.000	.183	.000	.165
	2.3	.014	.649	.000	.302	.000	.265
	2.4	.074	.878	.000	.426	.000	.353
	2.5	.226	.973	.000	.529	.000	.441
.01	2	1.00	.517	.149	.307	.007	.225
	3	1.00	.536	.001	.250	.000	.170
	4	.999	.741	.000	.348	.000	.242
	5	.979	.891	.000	.445	.000	.320
	6	.792	.959	.000	.512	.000	.370
	7	.405	.979	.000	.560	.000	.406
.001	2	1.00	.936	.894	.844	.396	.663
	12	1.00	.718	.337	.529	.039	.334
	22	1.00	.777	.014	.460	.000	.268
	32	1.00	.899	.000	.497	.000	.286
	42	.998	.943	.000	.525	.000	.308
	52	.957	.954	.000	.540	.000	.318

Note: The formulas t_{OLS}, t_{BC} are given by (15), (16) respectively.

Table 1.4 (cont'd)

β_2	θ_1	$b = 0.1$ t_{IV}	$b = 0.1$ t_{θ_0}	$b = 0.5$ t_{IV}	$b = 0.5$ t_{θ_0}	$b = 1.0$ t_{IV}	$b = 1.0$ t_{θ_0}
10	2.000	.054	.054	.053	.053	0.052	0.052
	2.001	.138	.138	.093	.093	.085	.085
	2.002	.389	.390	.205	.205	.180	.180
	2.003	.699	.699	.350	.351	.300	.300
	2.004	.910	.910	.503	.504	.419	.419
	2.005	.985	.985	.644	.646	.529	.530
2	2.000	.054	.054	.053	.053	.052	.052
	2.005	.136	.138	.092	.093	.085	.085
	2.010	.386	.390	.203	.205	.179	.180
	2.015	.694	.699	.345	.351	.297	.300
	2.020	.907	.910	.499	.504	.416	.419
	2.025	.984	.985	.638	.646	.524	.530
.2	2.00	.054	.054	.054	.053	.052	.052
	2.05	.116	.138	.084	.093	.079	.085
	2.10	.342	.390	.181	.205	.166	.180
	2.15	.650	.699	.310	.351	.271	.300
	2.20	.878	.910	.442	.504	.372	.419
	2.25	.974	.985	.562	.646	.466	.530
.1	2.0	.053	.054	.054	.053	.052	.052
	2.1	.094	.138	.077	.093	.072	.085
	2.2	.290	.390	.162	.205	.149	.180
	2.3	.585	.699	.274	.351	.241	.300
	2.4	.834	.910	.386	.504	.330	.419
	2.5	.956	.985	.487	.646	.408	.530
.01	2	.110	.054	.068	.053	.064	.052
	3	.014	.138	.021	.093	.023	.085
	4	.003	.390	.028	.205	.029	.180
	5	.007	.699	.043	.351	.046	.300
	6	.021	.910	.057	.504	.060	.419
	7	.040	.985	.068	.646	.072	.530
.001	2	.221	.054	.127	.053	.111	.052
	12	.114	.138	.067	.093	.067	.085
	22	.070	.390	.047	.205	.046	.180
	32	.055	.699	.043	.351	.043	.300
	42	.048	.910	.043	.504	.044	.419
	52	.043	.985	.043	.646	.045	.530

Note: The formulas t_{IV}, and t_{θ_0} are given by (17) and (19) respectively.

2 ESTIMATION AND INFERENCE OF LINEAR TREND SLOPE RATIOS WITH I(1) ERRORS

2.1 Introduction

In this chapter, the analysis of chapter one is extended to the case where the series can have unit root errors. We carry out an extensive theoretical analysis of the IV estimator, the residuals from the IV estimator and inference procedures when the stationarity assumption is relaxed, and the stochastic parts of the trending series follow an I(1) process. We consider the case when the individual series have a unit root and the noise term in the IV regression equation also has a unit root. We also consider the case of cointegration in the two series leading to an I(0) error in the IV regression equation. Our theoretical framework explicitly captures the impact of the magnitude of the trend slopes on the estimation and inference about the trend slope ratio. If the trend slopes are relatively large in magnitude, the IV estimator is consistent for both I(1) and I(0) regression errors. For medium and small trend slopes, the IV estimator is inconsistent for the I(1) case, but consistent for an I(0) regression error. For inference, the test based on the IV estimator is compared with the alternative testing approach. Finite sample simulations suggest that the alternative testing approach is superior both under the null and under the alternative. Whether the noise term in the IV regression equation is I(0) or I(1) has an impact on the power performance of the test for the trend slopes ratio.

The remainder of this chapter is organized as follows: Section 2.2 describes the model and analyzes the asymptotic properties of the IV estimator of the trend slope ratio. Section 2.3 provides some finite sample evidence on the performance of the IV estimator. Section 2.4 investigates inference regarding the trend slope ratio. We show how to construct heteroskedasticity autocorrelation (HAC) robust tests using the IV estimator and analyze the alternative testing approach. We derive asymptotic results of the tests under the null and under local alternatives. The finite sample simulations in section 2.5 clearly show that our alternative testing approach is superior under both the null and local alternatives. Section 2.6 presents the test of no cointegration in the trend slope regression. Section 2.7 gives the finite sample null rejection probabilities and power of unit root tests applied to regression residuals. Section 2.8 concludes. All proofs are given in the Appendix.

2.2 The Model and Estimation

2.2.1 Model and Assumptions

Suppose the univariate time series y_{1t} and y_{2t} are given by

$$y_{1t} = \mu_1 + \beta_1 t + u_{1t}, \tag{30}$$

$$y_{2t} = \mu_2 + \beta_2 t + u_{2t}. \tag{31}$$

Suppose that $\beta_2 \neq 0$, then by substituting t from eq. (31) into (30), and defining $\delta = \mu_1 - \theta\mu_2$ and $\epsilon_t(\theta) = u_{1t} - \theta u_{2t}$ gives the regression model

$$y_{1t} = \delta + \theta y_{2t} + \epsilon_t(\theta). \tag{32}$$

In eqs. (30) and (31), u_{1t} and u_{2t} are assumed to be I(1) processes as follows:

$$\begin{aligned} u_{1t} &= u_{1t-1} + w_{1t}, \\ u_{2t} &= u_{2t-1} + w_{2t}, \\ w_{1t} &\sim (0, \sigma_1^2), \\ w_{2t} &\sim (0, \sigma_2^2), \\ T^{-1/2} u_{2[sT]} &\Rightarrow \sigma_2 w_2(s), \end{aligned} \tag{33}$$

where w_{1t} and w_{2t} both have zero mean with variances σ_1^2 and σ_2^2 respectively. Whether or not there is cointegration in the two trending series, leads to the following two cases for $\epsilon_t(\theta)$:

Case 1: If there is no cointegration between the two trending series, then

$$\epsilon_t(\theta) = \epsilon_{t-1}(\theta) + \eta_t,$$

where

$$\eta_t = (w_{1t} - \theta w_{2t})$$

Assume that

$$T^{-1/2}\epsilon_{[sT]} = T^{-1/2}\sum_{t=1}^{[rT]}\eta_t \Rightarrow \sigma_{\eta,\theta}w(s), \tag{34}$$

where $w(s)$ is a univariate standard Wiener process and $\sigma^2_{\eta,\theta}$ is the long run variance of η_t.

Case 2: If there is cointegration between the two series, and the cointegrating vector for w_{1t}, w_{2t} is $\begin{pmatrix} 1 & -\theta \end{pmatrix}$, then ϵ_t is an I(0) process. In this case, we assume that

$$T^{-1/2}\sum_{t=1}^{[rT]}\epsilon_t \Rightarrow \lambda_\theta w(r), \tag{35}$$

where $w(r)$ is a univariate standard Wiener process and λ^2_θ is the long run variance of ϵ_t.

2.2.2 Estimation of the Trend Slope Ratio

Using regression (32), the IV estimator of theta, with time as an instrument for y_{2t} has been defined in chapter 1 by eq. (11). Standard algebra gives the following relationship:

$$\left(\widehat{\theta} - \theta\right) = \left(\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right)^{-1}\sum_{t=1}^{T}(t-\overline{t})(\epsilon_t-\overline{\epsilon}). \tag{36}$$

2.2.3 Asymptotic Properties of IV when $\epsilon_t(\theta)$ is an I(1) Process

We first explore the asymptotic properties of $\widehat{\theta}$, the IV estimator of θ when $\epsilon_t(\theta)$ is an I(1) process. The asymptotic behavior of $\widehat{\theta}$ depends on the magnitude of the trend slope parameters relative to the variation in the random components, u_{1t} and u_{2t}, i.e. the noise. The following theorem summarizes the asymptotic behavior of $\widehat{\theta}$ for fixed βs and for βs that are modeled as local to zero at rates $T^{-1/2}$ and T^{-1}.

Theorem 8 *Suppose that u_{1t}, u_{2t}, $\epsilon_t(\theta)$ are I(1) processes, (33) and (34) hold, and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{1/2}\left(\widehat{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr\right)^{-1} \sigma_{\eta,\theta} \int_0^1 (r - \frac{1}{2}) w(r) dr.$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$\left(\widehat{\theta} - \theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_{\eta,\theta} \int_0^1 (r - \frac{1}{2}) w(r) dr.$$

Case 3 (small trend slopes): For $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,

$$\left(\widehat{\theta} - \theta\right) \Rightarrow \left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_{\eta,\theta} \int_0^1 (r - \frac{1}{2}) w(r) dr.$$

Theorem 1 makes some interesting predictions about the sampling properties of the IV estimator. When the trend slopes are fixed, i.e. when the trend slopes are large relative to the noise, $\widehat{\theta}$ converges to the true value of θ at the rate $T^{1/2}$ and is asymptotically normal. The precision of the IV estimator improves when there is less noise ($\sigma^2_{\eta,\theta}$ is smaller) or when the magnitude of the trend slope parameter for y_{2t} increases (β_2 is larger). As β_2 approaches zero or $\sigma^2_{\eta,\theta}$ approaches infinity, $\widehat{\theta}$ centered around θ is indeterminate. When the trend slopes are modeled as local to zero at rate $T^{-1/2}$ and T^{-1}, i.e. when trend slopes are medium and small sized respectively relative to the noise, the IV estimator is biased and inconsistent. Case 3 is a special case of

Phillips (1986) framework.

2.2.4 Asymptotic Properties of IV when $\epsilon_t(\theta)$ is an I(0) Process

We next discuss the asymptotic properties of the IV estimator of θ when $\epsilon_t(\theta)$ is an I(0) process but u_{1t} and u_{2t} are I(1) processes. As before, the asymptotic behavior of $\widehat{\theta}$ depends on the magnitude of the trend slope parameters relative to the variation in the random components, u_{1t} and u_{2t}, i.e. the noise. The following theorem summarizes the asymptotic behavior of $\widehat{\theta}$ for fixed βs and for βs that are modeled as local to zero at rate $T^{-1/2}$ and T^{-1}.

Theorem 9 *Suppose that u_{1t} and u_{2t} are I(1) processes, $\epsilon_t(\theta)$ is I(0), (33) and (35) hold, and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{3/2}\left(\widehat{\theta}-\theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr\right)^{-1} \lambda_\theta \int_0^1 (r-\frac{1}{2}) dw(r).$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$T\left(\widehat{\theta}-\theta\right) \Rightarrow \left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-1} \lambda_\theta \int_0^1 (r-\frac{1}{2}) dw(r).$$

Case 3 (small trend slopes): For $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,

$$T\left(\widehat{\theta}-\theta\right) \Rightarrow \left(\sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-1} \lambda_\theta \int_0^1 (r-\frac{1}{2}) dw(r).$$

When the trend slopes are fixed, i.e. when the trend slopes are large relative to the noise, $\widehat{\theta}$ converges to the true value of θ at the rate $T^{3/2}$ whereas for the cases when trend slopes are local to zero at the rate $T^{-1/2}$ and T^{-1}, i.e. when trend slopes are medium and small sized respectively relative to the noise, the IV estimator converges to θ at the rate T. In the asymptotic limits for cases 2 and 3 for I(0) errors, $w_2(r)$ is correlated with $w(r)$ as long as u_{2t} is correlated with $\epsilon_t(\theta)$. Because $\epsilon_t(\theta)$ is an I(0) process, $\widehat{\theta}$ remains consistent for medium and small trend slopes in contrast to the case where $\epsilon_t(\theta)$ is an I(1) process.

2.2.5 Implications (Predictions) of Asymptotics for Finite Samples

Theorems 8 and 9 make clear predictions about the finite sample behavior of the IV estimator of θ when u_{1t} and u_{2t} are I(1) and $\epsilon_t(\theta)$ is either an I(0) or I(1) process. When $\epsilon_t(\theta)$ is an I(0) process, $\widehat{\theta}$ is consistent for large, medium and small trend slopes, whereas $\widehat{\theta}$ is consistent only for large trend slopes when $\epsilon_t(\theta)$ is an I(1) process. Therefore $\widehat{\theta}$ should exhibit some finite sample bias for medium and small trend slopes in the later case. For medium trend slopes, $\widehat{\theta}$ can also show some finite sample bias for the case when $\epsilon_t(\theta)$ is an I(0) process, because $w_2(r)$ is correlated with $w(r)$ as long as u_{2t} is correlated with $\epsilon_t(\theta)$. With large trend slopes and $\epsilon_t(\theta)$ an I(0) process (with u_{1t} and u_{2t} still I(1)), $\widehat{\theta}$ is asymptotically equivalent to the OLS, IV and bias corrected OLS estimators in large trend slopes case when $\epsilon_t(\theta)$, u_{1t} and u_{2t} are all I(0) processes (see Chapter 1).

2.3 Finite Sample Means and Standard Deviations of IV

In this section we illustrate the finite sample performance of the IV estimator via a Monte Carlo simulation study. For the data generating processes (DGP's) that we consider, the finite sample behavior of the IV estimator closely follows the predictions suggested by Theorem 8 and 9.

The following DGP was used for Theorem 8. The y_{1t} and y_{2t} variables were generated by models (30) and (31) where the noise is given by

$$u_{1t} = u_{1t-1} + \varepsilon_{1t},$$

$$u_{2t} = u_{2t-1} + \varepsilon_{2t},$$

$$[\varepsilon_{1t}, \varepsilon_{2t}]' \sim i.i.d.\ N(0, I_2),$$

$$u_{10} = u_{20} = 0.$$

For Theorem 9, the DGP is as follows:

$$u_{2t} = u_{2t-1} + \varepsilon_{2t},$$

$$u_{1t} = \theta u_{2t} + \varepsilon_{1t}$$

$$\varepsilon_{1t} = 0.2\varepsilon_{1t-1} + \zeta_t,$$

$$\varepsilon_{2t}, \zeta_t \sim i.i.d.\ N(0, I_2),$$

$$u_{10} = u_{20} = 0.$$

Given that $\widehat{\theta}$ is exactly invariant to the values of μ_1 and μ_2, we set $\mu_1 = 0, \mu_2 = 0$ without loss of generality. We report results for various magnitudes of β_1 and β_2 where it is almost always the case that $\theta = \beta_1/\beta_2 = 2$. The exception is when $\beta_1 = 0$, $\beta_2 = 0$ in which case θ is not defined. We report results for $T = 50, 100, 200$ with $10,000$ replications used in all cases.

Table 2.1a reports estimated means and standard deviations of the IV estimator across $10,000$ replications for I(1) case. Focusing on the $T = 50$ case we see that when the trend slopes are large ($\beta_1 = 6, \beta_2 = 3$), the IV estimator does not show any bias. For medium to small sized trend slopes ($\beta_2 = 2$ to $\beta_2 = 0.1$), the IV estimator shows some bias, which becomes more and more severe as the trend slopes get smaller in magnitude. The standard deviations increase as β_2 decreases as expected and the IV becomes very imprecise as the trend slopes approach zero.

Results for the cases of $T = 100, 200$ are similar to the $T = 50$ case. The only difference is that the bias kicks in more slowly as β_2 decreases. For the case $\beta_1 = 0$, $\beta_2 = 0$, θ is not well defined, and the IV estimator is based on an instrument that has zero correlation with y_{2t}. It is well known in the literature that weak instruments have important implications for $\widehat{\theta}$ and estimation of θ is no exception. Furthermore, when u_{1t}, u_{2t} are I(1),

$$\widehat{\theta} = \frac{T^{-5/2}\Sigma_{t=2}^{T}(t-\bar{t})u_{1t}}{T^{-5/2}\Sigma_{t=2}^{T}(t-\bar{t})u_{2t}} \Rightarrow \frac{\sigma_1 \int_0^1 (r-\frac{1}{2})w_1(r)dr}{\sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr},$$

which is a Cauchy random variable when $w_1(r)$ and $w_2(r)$ are uncorrelated. This explains why the standard deviation for the $\beta_1 = 0$, $\beta_2 = 0$ case in table 2.1a does not steadily decrease as the sample size increases.

Table 2.1b reports estimated means and standard deviations of the IV estimator across $10,000$ replications for the case of I(0) regression errors. As predicted by Theorem 9, $\widehat{\theta}$ remains unbiased even for small trend slopes.

Overall, the finite sample means and variances exhibit patterns as predicted by the asymptotic theory.

2.4 Inference

In this section we analyze test statistics for testing simple hypotheses about θ. Suppose we are interested in testing the null hypothesis

$$H_0 : \theta = \theta_0, \tag{37}$$

against the alternative hypothesis

$$H_1 : \theta = \theta_1 \neq \theta_0.$$

It is straightforward to construct a HAC robust statistic using the IV estimator of θ as

$$t_{IV} = \frac{(\widehat{\theta} - \theta_0)}{\sqrt{\widehat{\lambda}_\theta^2 \left[\sum_{t=1}^{T}(t-\bar{t})(y_{2t} - \bar{y}_2)\right]^{-2} \sum_{t=1}^{T}(t-\bar{t})^2}}, \tag{38}$$

where the estimated long run variance estimator is given by

$$\widehat{\lambda}_\theta^2 = \widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widehat{\gamma}_j, \quad \widehat{\gamma}_j = T^{-1}\sum_{t=j+1}^{T}\widehat{\epsilon}_t\widehat{\epsilon}_{t-j},$$

with $\widehat{\epsilon}_t = y_{1t} - \widehat{\delta} - \widehat{\theta} y_{2t}$ being the IV residuals from (32). The long run variance estimator is constructed using the kernel weighting function $k(x)$ and M is the bandwidth tuning parameter.

2.4.1 Linear in Slopes Approach

We investigate the linear in slopes approach when there are unit roots in data. The test statistic has already been defined by eq. (19) in Chapter 1. The critical values for this approach have been derived in Bunzel and Vogelsang (2005) for the test robust to I(0) or I(1) errors, as well as that specific to I(1) errors.

2.4.2 Asymptotic Results for t-statistics

In this section we provide asymptotic limits of the t-statistic based on the IV estimator. The asymptotic limit of the t-statistic along with the critical values, based on linear in slopes approach can be found in Bunzel and Vogelsang (2005) specifically for I(1) errors as well as the one robust to the nature of serial correlation in errors. We derive asymptotic limits under alternatives that are local to the null given by (37). Suppose that $\beta_2 = T^{-\kappa}\overline{\beta}_2$. Then the alternative value of θ_1 is modeled local to θ_0 as

$$\theta_1 = \theta_0 + T^{-l+\kappa}\overline{\theta}_\Delta. \tag{39}$$

The parameter $\overline{\theta}_\Delta$ measures the magnitude of the departure from the null under the local alternative, and the value of l differs for I(0) and I(1) errors. In the results presented below, the asymptotic null distributions of the t-statistics are obtained by setting $\overline{\theta}_\Delta = 0$.

We derive the limits of the various HAC estimators using fixed-b theory following Bunzel and Vogelsang (2005). The form of these limits depends on the type of kernel function used to compute the HAC estimator. We follow Bunzel and Vogelsang (2005) and use the following definitions.

Definition 3 *A kernel is labelled Type 1 if $k(x)$ is twice continuously differentiable everywhere and as a Type 2 kernel if $k(x)$ is continuous, $k(x) = 0$ for $|x| \geq 1$ and $k(x)$ is twice continuously differentiable everywhere except at $|x| = 1$.*

We also consider the Bartlett kernel (which is neither Type 1 or 2) separately. The fixed-b limiting distributions are expressed in terms of the following stochastic functions.

Definition 4 *Let $Q(r)$ be a generic stochastic process. Define the random variable $P_b(Q(r))$ as*

$$P_b(Q(r)) = \begin{cases} \int_0^1 \int_0^1 -k^{*\prime\prime}(r-s)\,Q(r)Q(s)drds & \text{if } k(x) \text{ is Type 1} \\ \int\int_{|r-s|<b} -k^{*\prime\prime}(r-s)\,Q(r)\,Q(s)\,drds & \\ \quad +2k_{-}^{*\prime}(b)\int_0^{1-b} Q(r+b)\,Q(r)\,dr & \text{if } k(x) \text{ is Type 2} \\ \frac{2}{b}\int_0^1 Q(r)^2\,dr - \frac{2}{b}\int_0^{1-b} Q(r+b)\,Q(r)\,dr & \text{if } k(x) \text{ is Bartlett} \end{cases}$$

where $k^*(x) = k\left(\frac{x}{b}\right)$ and $k_{-}^{*\prime}$ is the first derivative of k^* from below.

The following theorem summarizes the asymptotic limits of the t-statistics for testing (37) when the alternative is given by (39).

Theorem 10 *(Large Trend Slopes) Let $M = bT$ where $b \in (0,1]$ is fixed. Let $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$ where $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T, and let $\theta_1 = \theta_0 + T^{-1/2}\overline{\theta}_\Delta$ for case 1 and $\theta_1 = \theta_0 + T^{-3/2}\overline{\theta}_\Delta$ for case 2. Then as $T \to \infty$, (Case 1: ϵ_t, u_{1t} and u_{2t} are all I(1) processes.)*

$$t_{IV} \Rightarrow \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1(s))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\sigma^2_{\eta,\theta_1}P_b(Q_1(s))}},$$

$$t_{\theta_0} \Rightarrow \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1(s))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\sigma^2_{\eta,\theta_0}P_b(Q_1(s))}}.$$

Case 2: u_{1t} and u_{2t} are I(1) processes, whereas ϵ_t is an I(0) process.

$$t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_1}P_b(Q(r))}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_0} P_b(Q(r))}},$$

where $Z \sim N(0,1)$, $Q_1(s) = \sigma_\eta \int_0^s w(r)dr - s\sigma_\eta \int_0^1 w(r)dr - L(s)\left(\int_0^1 (r-\frac{1}{2})^2 dr\right)^{-1} \sigma_\eta \int_0^1 (r-\frac{1}{2})w(r)dr$, $Q(r) = \widetilde{w}(r) - 12L(r)\int_0^1 \left(s-\frac{1}{2}\right)dw(s)$, $\widetilde{w}(r) = w(r) - rw(1)$, $L(r) = \int_0^r \left(s-\frac{1}{2}\right)ds$ *and* Z *and* $Q(r)$ *are independent.*

Theorem 11 *(Medium Trend Slopes): Let* $M = bT$ *where* $b \in (0,1]$ *is fixed. Let* $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$ *where* $\overline{\beta}_1, \overline{\beta}_2$ *are fixed with respect to* T*, and let* $\theta_1 = \theta_0 + \overline{\theta}_\Delta$ *for case 1 and* $\theta_1 = \theta_0 + T^{-1}\overline{\theta}_\Delta$ *for case 2. Then as* $T \to \infty$*, (Case 1:* ϵ_t, u_{1t} *and* u_{2t} *are all I(1) processes.)*

$$t_{IV} \Rightarrow \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1^*(s))}}$$

$$+\frac{\overline{\theta}_\Delta}{\sqrt{\sigma^2_{\eta,\theta_1} Q(P^*(s))\left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1 (r-\frac{1}{2})^2 dr\right)}},$$

$$t_{\theta_0} \Rightarrow \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1(s))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\sigma^2_{\eta,\theta_0} P_b(Q_1(s))}}.$$

Case 2: u_{1t} *and* u_{2t} *are I(1) processes, whereas* ϵ_t *is an I(0) process.*

$$t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b\left(Q'(s)\right)}}$$

$$+\frac{\overline{\theta}_\Delta}{\sqrt{\lambda^2_{\theta_1} P_b\left(Q'(s)\right)\left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1 (r-\frac{1}{2})^2 dr\right)}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2 \overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_0} P_b(Q(r))}},$$

where $Q_1^*(s) = \int_0^s w(r)dr - s\int_0^1 w(r)dr - \overline{\beta}_2 L(s)\left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1} * \int_0^1 (r-\frac{1}{2})w(r)dr$, *and* $Q'(s) = \widetilde{w}(s) - \overline{\beta}_2 L(s)\left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1} \int_0^1 (r-\frac{1}{2})dw(r)$.

Theorem 12 *(Small Trend Slopes): Let* $M = bT$ *where* $b \in (0,1]$ *is fixed. Let* $\beta_1 =$

$T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$ *where* $\overline{\beta}_1, \overline{\beta}_2$ *are fixed with respect to* T*, and let* $\theta_1 = \theta_0 + \overline{\theta}_\Delta$ *for case 1 and* $\theta_1 = \theta_0 + T^{-1}\overline{\theta}_\Delta$ *for case 2. Then as* $T \to \infty$*, (Case 1:* ϵ_t, u_{1t} *and* u_{2t} *are all I(1) processes.)*

$$t_{IV} \Rightarrow \frac{\sqrt{12}\int_0^1 (r - \frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1^{**}(s))}} + \frac{\overline{\theta}_\Delta}{\sqrt{\sigma^2_{\eta,\theta_1} P_b(Q_1^{**}(s)) \left(\sigma_2 \int_0^1 (r - \frac{1}{2})w_2(r)dr\right)^{-2} \left(\int_0^1 (r - \frac{1}{2})^2 dr\right)}},$$

$$t_{\theta_0} \Rightarrow \frac{\sqrt{12}\int_0^1 (r - \frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1(s))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\sigma^2_{\eta,\theta_0} P_b(Q_1(s))}}.$$

Case 2: u_{1t} *and* u_{2t} *are I(1) processes, whereas* ϵ_t *is an I(0) process.*

$$t_{IV} \Rightarrow \frac{Z}{\sqrt{P_b\left(\widetilde{K}(s)\right)}} + \frac{\overline{\theta}_\Delta}{\sqrt{\lambda^2_{\theta_1} P_b\left(\widetilde{K}(s)\right) \left(\sigma_2 \int_0^1 (r - \frac{1}{2})w_2(r)dr\right)^{-2} \left(\int_0^1 (r - \frac{1}{2})^2 dr\right)}},$$

$$t_{\theta_0} \Rightarrow \frac{Z}{\sqrt{P_b(Q(r))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda^2_{\theta_0} P_b(Q(r))}},$$

where $Q_1^{**}(s) = \int_0^s w(r)dr - s\int_0^1 w(r)dr - \overline{\beta}_2 L(s) \left(\sigma_2 \int_0^1 (r - \frac{1}{2})w_2(r)dr\right)^{-1} \int_0^1 (r - \frac{1}{2})w(r)dr$*, and* $\widetilde{K}(s) = \widetilde{w}(r) - \sigma_2 J(s) \left(\sigma_2 \int_0^1 (r - \frac{1}{2})w_2(r)dr\right)^{-1} \int_0^1 (r - \frac{1}{2})dw(r)$*.*

Some interesting results and predictions about the finite sample behavior of the t-statistics are given by Theorems 10-12. First examine the limiting null distributions that are obtained when $\overline{\theta}_\Delta = 0$. The bias in the IV estimator for medium and small trend slopes affect the asymptotic distribution of the t-statistic when $\epsilon_t(\theta)$ is an I(1) process. We get differernt limits for large, medium and small trend slopes both for the cases when $\epsilon_t(\theta)$ is I(1) as well as I(0). The unit roots in u_{1t} and u_{2t} do not change the critical values if ϵ_t is an I(0) process, and β_1, β_2 are large as is evident from the limit of the t-statistics for large trend slopes which is exactly the same as in Chapter one.

When $\overline{\theta}_\Delta \neq 0$, in which case we are under the alternative, the t-statistics have additional terms in their limits which push the distributions away from the null distributions giving the tests power.

As the limiting distributions for t_{θ_0} are different for I(0) and I(1) regression errors, it may be useful to have a test robust to both I(0) and I(1) errors. A robust test has been proposed in Bunzel and Vogelsang (2005), and the t-statistic is defined as follows:

$$t_{\theta_0} = \frac{\widehat{\pi}_1(\theta_0)}{\sqrt{\widehat{\lambda}^2_{\theta_0}\left(\sum_{t=1}^{T}(t-\bar{t})^2\right)^{-1}}}\exp(-cUR),$$

$\exp(-cUR)$ is a scaling factor, UR denotes either the J or BG unit root statistics and c is a constant. J is defined as follows:

Consider the regression

$$z_t\left(\theta_0\right) = \pi_0(\theta_0) + \pi_1(\theta_0)t + \sum_{i=2}^{9}\alpha_i t^i + v_t\left(\theta_0\right),$$

then the J statistic is defined as

$$J = \frac{SSR_{(1)} - SSR_{(2)}}{SSR_{(2)}},$$

where $SSR_{(2)}$ is the sum of squared residuals obtained from the estimation of above equation by OLS, and $SSR_{(1)}$ be the sum of squared residuals from the OLS estimation of (18). The value of scalar c can be found as follows:

$$c(b) = \lambda_0 + \lambda_1 b + \lambda_2 b^2 + \lambda_3 b^3 + \lambda_4 b^4 + \lambda_5 b^5 + \lambda_6 b^6 + \lambda_7 b^7,$$

where the values of λ_i's and the methodology for the data dependent bandwidth rule can be found in Bunzel and Vogelsang (2005).

For confidence interval construction, however, a lot more work needs to be done as the scaling factor depends on the true value of θ which is unknown in practice. This is left as a future research task.

2.5 Finite Sample Null Rejection Probabilities and Power

Using the same DGP's as used in Section 2.3 we simulated finite sample null rejection probabilities and power of the IV t-statistics both for $\epsilon_t(\theta)$ as I(0) and I(1) cases and the t-statistics for linear hypothsis (t_{θ_0}) based on I(1) errors as well as robust to the nature of serial correlation in the error term. Tables 2.2a and 2.2b report null rejection probabilities for 5% nominal level tests for testing $H_0 : \theta = \theta_0 = 2$ against the two-sided alternative $H_1 : \theta \neq 2$ for I(1) and I(0) regression errors respectively. Results are reported for the same values of β_1, β_2 as used in Tables 2.1a and b, for $T = 50, 100, 200$ and $10,000$ replications are used in all cases. The HAC estimators are implemented using the Daniell kernel. Results for three bandwidth sample size ratios are provided: $b = 0.1, 0.5, 1.0$. For a given sample size, T, we use the bandwidth $M = bT$ for each of the three values of b. We compute empirical rejections using fixed-b asymptotic critical values as given by Bunzel and Vogelsang (2005) for the Daniell kernel. For IV, and the linear hypothesis for I(0) errors, the critical values specific to I(0) errors have been used, whereas for IV, and the linear hypothesis based on I(1) errors, the critical values specific to the I(1) errors have been used. For the robust linear test, I(0) critical values have been used.

The patterns in the empirical null rejections closely match the predictions of the asymptotic results as in Theorems 10-12. When the trend slopes are large, $\beta_1 \geq 4$, $\beta_2 \geq 2$, null rejections are essentially the same for all t-statistics and are close to 0.05 even when $T = 50$. This is true for all three bandwidth choices which illustrates the effectiveness of the fixed-b critical values. For medium and small trend slopes, $0.2 \leq \beta_1 \leq 2$, $0.1 \leq \beta_2 \leq 1$, IV begins to show over-rejection problem for the I(1) case, that becomes very severe as the trend slopes decrease in magnitude. In contrast, t_{θ_0} has null rejections that are close to 0.05 regardless of the magnitudes of β_1, β_2 including the case of $\beta_1 = \beta_2 = 0$. It is clear in terms of null rejection probabilities that t_{θ_0} is the preferred test statistic. For I(0) case, the IV shows under-rejections for small trend slopes, whereas t_{θ_0} has null rejections that are close to 0.05 regardless of the magnitudes of trend slopes.

Tables 2.3a and b report power results for a subset of the grid of β_2 as used in Tables 2.1a and b. For a given value of β_2, we specify a grid of six equally spaced values for θ in the range $\theta \in [2, \theta_{\max}]$ where $\theta = \theta_0 = 2$ is the null value. By construction

$\beta_1 = \theta\beta_2$ in all cases. Results are reported for $T = 100$. Results for other values of T are qualitatively similar and are omitted.

The power of t_{θ_0} for I(1) errors is the highest among all three test statistics. For large trend slopes ($\beta_2 = 10, 2$), power of the IV test is lower than that of t_{θ_0} for I(1) errors, however, it is higher than the robust t_{θ_0}. As the trend slopes decrease in magnitude, the power of IV test decreases substantially. The power of the robust t_{θ_0} test is the lowest among all three. The power of the test for I(0) errors decreases as the trend slopes decrease in magnitude. This is evident from the results reported in table 2.3b. These results suggest that if empirical researchers have a reason to believe that the noise term is an I(1) process, it is better to use the test based on I(1) errors as compared to the one which is robust to the nature of serial correlation in the noise term. As the bandwidth increases, power of all the tests decreases. This inverse relationship between power and bandwidth is well known in the fixed-b literature (see Kiefer and Vogelsang 2005).

In summary, the patterns in the finite sample simulations are consistent with the predictions of Theorems 10-12. Clearly t_{θ_0} is the recommended statistic given its superior behavior under the null and its higher power under the alternative.

2.6 Unit Root Tests for $\epsilon_t(\theta)$

The inference section suggests that it is important for the empirical researchers to conduct a unit root test on $\epsilon_t(\theta)$ to determine which critical values they should use for inference on θ. In this section, we show how to carry out a unit root test on $\epsilon_t(\theta)$, and how to improve the power of the unit root test through the ADF-GLS transformation. We focus on the case where $\epsilon_t(\theta)$ is an AR(1) process. The additional serial correlation can be handled in the usual way by including lagged first differences. In order to test the null hypothesis of a unit root in ϵ_t, we need to regress $\widehat{\epsilon}_t$ on $\widehat{\epsilon}_{t-1}$ in the following regression equation:

$$\widehat{\epsilon}_t = \rho\widehat{\epsilon}_{t-1} + \xi_t.$$

OLS gives

$$\widehat{\rho} = \frac{\Sigma_{t=2}^{T}\widehat{\epsilon}_t\widehat{\epsilon}_{t-1}}{\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2}.$$

Centering the estimator $\widehat{\rho}$ around one and scaling it by T, we obtain

$$T(\widehat{\rho}-1) = \frac{T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}(\widehat{\epsilon}_t - \widehat{\epsilon}_{t-1})}{T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2} = \frac{T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t}{T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2}.$$

The t-statistic for testing the null hypothesis of a unit root in $\epsilon_t(\theta)$ is given as follows:

$$t_{\widehat{\rho}=1} = \frac{T(\widehat{\rho}-1)}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2\right)^{-1}}},$$

where

$$s^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1})^2.$$

The next theorem provides the limit of $t_{\widehat{\rho}=1}$ under the null that $\epsilon_t(\theta)$ is I(1).

Theorem 13 *Suppose that $\epsilon_t(\theta)$ is an I(1) process and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following hold as $T \to \infty$. Case 1 (large trend slopes): For $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$t_{\widehat{\rho}=1} \Rightarrow \frac{\int_0^1 \widehat{w}(r)dw(r)}{\sqrt{\int_0^1 \widehat{w}(r)^2 dr}}.$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$t_{\widehat{\rho}=1} \Rightarrow \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^*(1)^2 - \sigma_{\eta,\theta}^2\widehat{w}^*(0)^2 - L\right]}{\sigma_\eta\sqrt{L\int_0^1 \widehat{w}^*(r)^2 dr}}.$$

Case 3 (small trend slopes): For $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,

$$t_{\widehat{\rho}=1} \Rightarrow \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^{**}(1)^2 - \sigma_{\eta,\theta}^2\widehat{w}^{**}(0)^2 - L'\right]}{\sigma_\eta\sqrt{L'\int_0^1 \widehat{w}^{**}(r)^2 dr}}.$$

Case 4 (zero trend slopes): For $\beta_1 = 0$, $\beta_2 = 0$,

$$t_{\widehat{\rho}=1} \Rightarrow \frac{\frac{1}{2}\left[\sigma_1^2 \widehat{w}^{***}(1)^2 - \sigma_1^2 \widehat{w}^{***}(0)^2 - L''\right]}{\sigma_1 \sqrt{L'' \int_0^1 \widehat{w}^{***}(r)^2 dr}},$$

where

$$\widehat{w}(r) = w(s) - \int_0^1 w(r)dr - (s - \frac{1}{2}) \left(\int_0^1 (r - \frac{1}{2})^2 dr \right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr,$$

$$\widehat{w}^*(r) = \frac{w(s) - \int_0^1 w(r)dr - \left(\overline{\beta}_2 (s - \frac{1}{2}) + \sigma_2 \left[w(s) - \int_0^1 w(r)dr\right]\right) *}{\left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr},$$

$$\widehat{w}^{**}(r) = \frac{w(s) - \int_0^1 w(r)dr - \sigma_2 \left[w(s) - \int_0^1 w(r)dr\right] *}{\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr}$$

$$\widehat{w}^{***}(s) = \frac{\sigma_1 \left[w_1(s) - \int_0^1 w_1(r)dr\right] - \left[w_2(s) - \int_0^1 w_2(r)dr\right] *}{\left[\left(\int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr\right]},$$

$$L = \sigma_{\eta,\theta}^2 \left[1 + \sigma_2^2 \left[\begin{array}{c} \left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \\ * \int_0^1 (r - \frac{1}{2}) w(r) dr \end{array}\right]^2\right] +$$
$$2\theta \sigma_{\eta,\theta} \sigma_2^2 \left[\left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr\right],$$

$$L' = \sigma_{\eta,\theta}^2 \left[1 + \sigma_2^2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr\right]^2\right] +$$
$$2\theta \sigma_{\eta,\theta} \sigma_2^2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \int_0^1 (r - \frac{1}{2}) w(r) dr\right].$$

$$L'' = \sigma_1^2 + \left[\left(\int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr\right]^2$$

When the trend slopes are fixed, i.e. when the trend slopes are large relative to the noise, the limit of the unit root test statistic is the same as ADF limit when an intercept and time trend are included in the ADF regression equation. When the trend slopes are not large in magnitude, the asymptotic limits are no longer the same.

We can boost the power of the unit root test through the ADF-GLS transformation. Straight forward algebra allows us to write:

$$\begin{aligned}
\widehat{\epsilon}_t &= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta} - \theta\right)(y_{2t} - \overline{y}_2), \\
&= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta} - \theta\right)\left[\overline{\beta}_2 (t - \overline{t}) + (u_{2t} - \overline{u}_2)\right], \\
&= (\epsilon_t - \overline{\epsilon}) - \overline{\beta}_2 \left(\widehat{\theta} - \theta\right)(t - \overline{t}) - \left(\widehat{\theta} - \theta\right)(u_{2t} - \overline{u}_2).
\end{aligned} \tag{40}$$

Now let

$$\widehat{\epsilon}_t^{GLS} = \widehat{\epsilon}_t - \widehat{\varrho}_1 - \widehat{\varrho}_2 t.$$

The estimators $\widehat{\varrho}_1, \widehat{\varrho}_2$ are GLS estimators obtained from regression of $\Delta_{\overline{\alpha}}\widehat{\epsilon}_t$ on $\Delta_{\overline{\alpha}} d_t$, where

$$\begin{aligned}
\Delta_{\overline{\alpha}}\widehat{\epsilon}_t &= \widehat{\epsilon}_t - \overline{\alpha}\widehat{\epsilon}_{t-1},\ t = 2, ..., T \\
\Delta_{\overline{\alpha}}\widehat{\epsilon}_1 &= \widehat{\epsilon}_1, \\
d_t &= (1, t)', \\
\varrho' &= (\varrho_1, \varrho_2), \\
\Delta_{\overline{\alpha}} d_t &= d_t - \overline{\alpha} d_{t-1},\ t = 2, ..., T \\
\Delta_{\overline{\alpha}} d_1 &= d_1.
\end{aligned}$$

Define GLS detrended residuals as

$$\epsilon_t^{GLS} = \epsilon_t - \widehat{\gamma}_1 - \widehat{\gamma}_2 t.$$

The estimators $\widehat{\gamma}_1, \widehat{\gamma}_2$ are GLS estimators obtained from regression of $\Delta_{\overline{\alpha}}\epsilon_t$ on $\Delta_{\overline{\alpha}} d_t$.

Similarly, $\widehat{\phi}_1, \widehat{\phi}_2$ are GLS estimators obtained from regression of $\Delta_{\overline{\alpha}} u_{2t}$ on $\Delta_{\overline{\alpha}} d_t$ leading to the GLS residuals

$$u_{2t}^{GLS} = u_{2t} - \widehat{\phi}_1 - \widehat{\phi}_2 t.$$

Applying GLS detrending to both sides of eq. (40) gives

$$\widehat{\epsilon}_t^{GLS} = \epsilon_t^{GLS} - \left(\widehat{\theta} - \theta\right) u_{2t}^{GLS},$$

noting that GLS detrending eliminates anything that is constant (like a sample average) or proportional to t and replaces OLS demeaned (or detrended) quantities with GLS detrended quantities. To test the null hypothesis of unit root in ϵ_t, we regress $\Delta\widehat{\epsilon}_t^{GLS}$ on $\widehat{\epsilon}_{t-1}^{GLS}$ as follows:

OLS gives

$$\widehat{\pi} = \frac{\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}}{\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS}},$$
$$T\widehat{\pi} = \frac{T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}}{T^{-2}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS}}.$$

The t-statistic for testing the null hypothesis $H_0 : \pi = 0$ is as follows:

$$t_{\widehat{\pi}=0}^{GLS} = \frac{T\widehat{\pi}}{\sqrt{s^2 \left(T^{-2}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS}\right)^{-1}}},$$

where

$$s^2 = \frac{1}{T-2}\Sigma_{t=2}^T \left(\Delta\widehat{\epsilon}_t^{GLS} - \widehat{\pi}\widehat{\epsilon}_{t-1}^{GLS}\right)^2$$

The asymptotic limits of ADF-GLS unit root tests for $\epsilon_t(\theta)$ for various magnitudes of beta's are presented in the following theorem.

Theorem 14 *Suppose that $\epsilon_t(\theta)$ is an I(1) process and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to*

T. *The following hold as* $T \to \infty$. *Case 1 (large trend slopes): For* $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,

$$t^{GLS}_{\widehat{\pi}=0} \Rightarrow \frac{\frac{1}{2}\left(K_c(1,\overline{c})^2 - K_c(0,\overline{c})^2 - 1\right)}{\sqrt{\int_0^1 K_c(r,\overline{c})^2 dr}}.$$

Case 2 (medium trend slopes): For $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,

$$t^{GLS}_{\widehat{\pi}=0} \Rightarrow \frac{\frac{1}{2}\left[\sigma_\eta^2 L_c(1,\overline{c},\sigma_2)^2 - \sigma_\eta^2 L_c(0,\overline{c},\sigma_2)^2 - F\right]}{\sigma_\eta\sqrt{F\int_0^1 L_c(r,\overline{c},\sigma_2)^2 dr}}.$$

Case 3 (zero trend slopes): For $\beta_1 = 0$, $\beta_2 = 0$,

$$t^{GLS}_{\widehat{\pi}=0} \Rightarrow \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c^*(1,\overline{c},\sigma_1,\sigma_2,\sigma_\eta)^2 - \sigma_{\eta,\theta}^2 L_c^*(0,\overline{c},\sigma_1,\sigma_2,\sigma_\eta)^2 - F^*\right]}{\sigma_{\eta,\theta}\sqrt{F^*\int_0^1 L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_\eta)^2 dr}}.$$

where

$$K_c(\overline{c}) = 3\varpi w_c(1) + 3(1-\varpi)\int_0^1 r w_c(r)dr,$$

$$\varpi = (1-\overline{c})/(1-\overline{c}+\overline{c}^2),$$

$$1-\varpi = \overline{c}^2/(1-\overline{c}+\overline{c}^2).$$

$$K_c(s,\overline{c}) = [w_c(s) - sK_c(\overline{c})]$$

$$L_c(s,\overline{c},\sigma_2) = \left[\begin{array}{c} w_c(s) - sK_c(\overline{c}) - \\ \sigma_2 w_c(s) * \left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1 (r-\frac{1}{2})w(r)dr + \\ s\sigma_2 K_c(\overline{c})\left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1 (r-\frac{1}{2})w(r)dr \end{array}\right]$$

$$F = \sigma_{\eta,\theta}^2 + \sigma_2^2\left[\left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_\eta\int_0^1 (r-\frac{1}{2})w(r)dr\right]^2$$
$$+ 2\theta\sigma_2^2\left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_\eta\int_0^1 (r-\frac{1}{2})w(r)dr$$

$$F^* = \sigma^2_{\eta,\theta} + \sigma^2_2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr \right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr - \theta \right]^2$$
$$+ 2\theta\sigma^2_2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr \right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr - \theta \right].$$

When trend slopes are large in magnitude, the asymptotic limit of $t^{GLS}_{\widehat{\pi}=0}$ is the same as the usual ADF-GLS limit obtained by Elliott, Rothenberg and Stock (1996). When the trend slopes are not large in magnitude, the asymptotic limits are no longer the same.

2.7 Finite Sample Null Rejection Probabilities and Power of Unit Root Tests

Using the same DGP as used in section 2.3 we simulated finite sample null rejection probabilities of the unit root tests. For power of the ADF and ADF-GLS tests for $\epsilon_t(\theta)$ for the cointegration case, the DGP is the same as used in section 2.3 for theorem 9, i.e.

$$u_{2t} = u_{2t-1} + \varepsilon_{2t},$$
$$u_{1t} = \theta u_{2t} + \varepsilon_{1t}$$
$$\varepsilon_{1t} = \rho\varepsilon_{1t-1} + \zeta_t,$$
$$\varepsilon_{2t}, \zeta_t \sim i.i.d.\ N(0, I_2),$$
$$u_{10} = u_{20} = 0,$$

whereas for the case where both u_{1t} and u_{2t} are I(0), the DGP is as follows:

$$u_{1t} = \alpha_1 u_{1t-1} + \varepsilon_{1t},$$

$$u_{2t} = \alpha_2 u_{2t-1} + \varepsilon_{2t},$$

$$[\varepsilon_{1t}, \varepsilon_{2t}]' \sim i.i.d.\ N(0, I_2),$$

$$u_{10} = u_{20} = 0.$$

Table 2.4 reports null rejection probabilities for 5% nominal level for testing the null hypothesis of a unit root in $\epsilon_t(\theta)$ (ADF critical value $= -3.451$, ADF-GLS critical value $= -3.03$). Results are reported for the same values of β_1, β_2 as used in Tables 2.1a and b, for $T = 50, 100, 200$ and $10,000$ replications are used in all cases. The null rejection probabilities are very similar for both ADF and ADF-GLS and are close to 0.05 in all cases. Table 2.5 reports power results for six equally spaced values of $\alpha_1 = \alpha_2 = \rho$, which clearly shows that ADF-GLS test performs better than ADF in terms of power.

2.8 Conclusion

In this chapter we analyze estimation and inference of the ratio of trend slopes of two time series with linear deterministic trend functions under the assumption that the stochastic parts of both series are I(1) processes. We consider the IV estimator of the trend slopes ratio both for the cases where the noise term in the IV regression equation follows an I(0) and I(1) process. Asymptotic theory indicates that when the magnitude of the trend slopes are large relative to the noise, the IV estimator tends to remain unbiased, whereas for medium to smaller trend slopes, the IV estimator becomes a poor estimator of the trend slopes ratio when the noise term is an I(1) process. In contrast (see Chapter 1), the IV estimator is consistent for all three cases, i.e. large, medium and small trend slopes when the noise term is an I(0) process.

We analyze t-statistics based on the IV estimator for testing hypotheses about the trend slopes ratio both for the I(0) and I(1) regression errors. We also consider the

t-statistic based on the alternative linear in slopes testing approach. Simulations show that the alternative test dominates the test based on the IV estimator in terms of size and power regardless of the magnitude of the trend slopes.

We propose an ADF-GLS test for a unit root in the noise term of the IV regression equation. The power of the ADF-GLS test is higher than that of ADF test as shown by the finite sample simulations. The test may help empirical researchers choose a test for the trend slopes ratio with higher power as compared to a test which is robust to the nature of serial correlation in the noise term. Finite sample simulations show that the predictions of the asymptotic theory tend to hold in practice.

APPENDIX

Proofs of Theorems

Before giving proofs of the theorems, we prove a series of lemmas for each of the trend slope magnitude cases: large, medium, small and very small. The lemmas establish the limits of the scaled sums that appear in the estimators of θ and the HAC estimators. Using the results of the lemmas, the theorems are easy to establish using straightforward algebra and the continuous mapping theorem (CMT). We begin with a lemma that has limits of scaled sums that are exactly invariant to the magnitudes of the trend slopes followed by four lemmas for each of the trend slope cases. Throughout the appendix, we use ϵ_t to denote $\epsilon_t(\theta)$.

Lemma 6 *Suppose that (34) holds. The following holds as $T \to \infty$ for any values of β_1, β_2 :*

$$T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon}) \Rightarrow \sigma_{\eta,\theta}\int_0^1 (r-\frac{1}{2})w(r)dr.$$

Proof: The result in this lemma is standard given the FCLT (34). See Hamilton (1994).

Lemma 7 *(Large trend slopes when $\epsilon_t(\theta)$ is an I(1) process) Suppose that (34) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following holds as $T \to \infty$ for $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$,*

$$T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2) \to \overline{\beta}_2\int_0^1 (s-\frac{1}{2})^2 ds.$$

Proof: The result of the lemma is easy to establish once we substitute $y_{2t}-\overline{y}_2 = \overline{\beta}_2\left(t-\bar{t}\right)+(u_{2t}-\overline{u}_2)$ into the above expression:

$$T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2 + o_p(1).$$

The limit follows from Lemma 1 in chapter one.

Lemma 8 *(Medium trend slopes when $\epsilon_t(\theta)$ is an I(1) process) Suppose that (34) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following holds as $T \to \infty$ for $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,*

$$T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2) \Rightarrow \overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr.$$

Proof: The result of the lemma is easy to establish once we substitute $y_{2t}-\overline{y}_2 = T^{-1/2}\overline{\beta}_2\left(t-\bar{t}\right)+(u_{2t}-\overline{u}_2)$ into the above expression:

$$T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2 + T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(u_{2t}-\overline{u}_2).$$

The limits follow from Lemma 1 in chapter one.

Lemma 9 *(Small trend slopes when $\epsilon_t(\theta)$ is an I(1) process) Suppose that (34) holds and $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following holds as $T \to \infty$ for $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,*

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \Rightarrow \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr.$$

Proof: The result of the lemma is easy to establish once we substitute $y_{2t} - \overline{y}_2 = T^{-1}\overline{\beta}_2(t-\overline{t}) + (u_{2t} - \overline{u}_2)$ into the above expression:

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-7/2}\sum_{t=1}^{T}(t-\overline{t})^2 + T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2).$$

The limits follow from Lemma 1 in chapter one and two.

Lemma 10 *(Medium trend slopes when $\epsilon_t(\theta)$ is an I(0) process) Suppose that $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following holds as $T \to \infty$ for $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$,*

$$T^{-3/2}\sum_{t=1}^{[sT]}(y_{2t}-\overline{y}_2) \xrightarrow{p} L(s).$$

Proof: The result of the lemma is easy to establish once we substitute $y_{2t} - \overline{y}_2 = T^{-1/2}\overline{\beta}_2\left(t-\overline{t}\right) + (u_{2t} - \overline{u}_2)$ into the above expression:

$$T^{-3/2}\sum_{t=1}^{[sT]}(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-2}\sum_{t=1}^{[sT]}(t-\overline{t}) + T^{-3/2}\sum_{t=1}^{[sT]}(u_{2t}-\overline{u}_2).$$

The limits follow from Lemma 1 in chapter one.

Lemma 11 *(Small trend slopes when $\epsilon_t(\theta)$ is an I(0) process) Suppose that $\overline{\beta}_1, \overline{\beta}_2$ are fixed with respect to T. The following holds as $T \to \infty$ for $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$,*

$$T^{-3/2}\sum_{t=1}^{[sT]}(y_{2t}-\overline{y}_2) \Rightarrow \sigma_2\left[\int_0^s w_2(r)dr - s\int_0^1 w_2(r)dr\right] \equiv \sigma_2 J(s),$$

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) \Rightarrow \sigma_2 \int_0^1 (r-\frac{1}{2})w_2(r)dr.$$

Proof: The result of the lemma is easy to establish once we substitute $y_{2t} - \overline{y}_2 = T^{-1}\overline{\beta}_2\left(t-\overline{t}\right) + (u_{2t} - \overline{u}_2)$ into the above expression:

$$T^{-3/2}\sum_{t=1}^{[sT]}(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-5/2}\sum_{t=1}^{[sT]}(t-\overline{t}) + T^{-3/2}\sum_{t=1}^{[sT]}(u_{2t}-\overline{u}_2)$$

$$T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2) = \overline{\beta}_2 T^{-7/2}\sum_{t=1}^{T}(t-\overline{t})^2 + T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(u_{2t}-\overline{u}_2)$$

The limits follow from Lemma 1 in chapter one and two.

Proof of Theorem 8. The proof follows directly from Lemmas 1, 2, 3, 4 and the CMT. For the case of large trend slopes it follows that

$$T^{1/2}\left(\widehat{\theta}-\theta\right) = \left(T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon}),$$
$$\Rightarrow \left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr\right)^{-1} \sigma_{\eta,\theta}\int_0^1 (r-\frac{1}{2})w(r)dr.$$

For the case of medium trend slopes it follows that

$$\left(\widehat{\theta}-\theta\right) = \left(T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon}),$$
$$\Rightarrow \left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr+\sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1} \sigma_{\eta,\theta}\int_0^1 (r-\frac{1}{2})w(r)dr.$$

For the case of small trend slopes it follows that

$$\left(\widehat{\theta}-\theta\right) = \left(T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon}),$$
$$\Rightarrow \left(\sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1} \sigma_{\eta,\theta}\int_0^1 (r-\frac{1}{2})w(r)dr.$$

Proof of Theorem 9. The proof follows directly from Lemmas 1, 2, 3, 4, 5, 6 and the CMT. For the case of large trend slopes it follows that

$$T^{3/2}\left(\widehat{\theta}-\theta\right) = \left(T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-3/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon})$$
$$\Rightarrow \left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr\right)^{-1} \lambda_\theta\int_0^1 (r-\frac{1}{2})dw(r).$$

For the case of medium trend slopes it follows that

$$T\left(\widehat{\theta}-\theta\right) = \left(T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-3/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon})$$
$$\Rightarrow \left(\overline{\beta}_2\int_0^1 (r-\frac{1}{2})^2 dr+\sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1} \lambda_\theta\int_0^1 (r-\frac{1}{2})dw(r).$$

For the case of small trend slopes it follows that

$$T\left(\widehat{\theta}-\theta\right) = \left(T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right)^{-1} T^{-3/2}\sum_{t=1}^{T}(t-\bar{t})(\epsilon_t-\bar{\epsilon})$$
$$\Rightarrow \left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\lambda_\theta\int_0^1(r-\frac{1}{2})dw(r).$$

Proof of Theorem 10. With large trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from Lemmas 1, 2 and Theorems 1, 2.

Case 1: $T^{-3/2}\widehat{S}_{[sT]} = T^{-3/2}\sum_{t=1}^{[rT]}(\epsilon_t-\bar{\epsilon}) - T^{1/2}\left(\widehat{\theta}-\theta\right)T^{-2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2)$

$$\Rightarrow \sigma_{\eta,\theta}\int_0^s w(r)dr - s\sigma_{\eta,\theta}\int_0^1 w(r)dr - L(s)\left(\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}\sigma_{\eta,\theta}\int_0^1(r-\frac{1}{2})w(r)dr$$
$$= \sigma_{\eta,\theta}Q_1(s).$$

$$\text{Case 2: } T^{-1/2}\widehat{S}_{[rT]} \Rightarrow \lambda_\theta Q(r).$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$\text{Case 1: } T^{-2}\widehat{\lambda}_\theta^2 \Rightarrow \sigma_{\eta,\theta}^2 P_b(Q_1(s)),$$
$$\text{Case 2: } \widehat{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(Q(r)\right).$$

The limits of t_{IV} for case 1 and 2 are as follows:

$$\text{Case 1: } t_{IV} = \frac{T^{1/2}(\widehat{\theta}-\theta_0)}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2}},$$
$$= \frac{T^{1/2}(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-3}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2}},$$
$$\Rightarrow \frac{\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}\sigma_{\eta,\theta_1}\int_0^1(r-\frac{1}{2})w(r)dr+\overline{\theta}_\Delta}{\sqrt{\sigma_{\eta,\theta_1}^2P_b(Q_1(s))\left(\overline{\beta}_2^2\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}}},$$
$$= \frac{\sqrt{12}\int_0^1(r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1(s))}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\sigma_{\eta,\theta_1}^2P_b(Q_1(s))}}.$$

$$\text{Case 2: } t_{IV} = \frac{T^{3/2}(\widehat{\theta}-\theta_0)}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$= \frac{T^{3/2}(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-3}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$\Rightarrow \frac{\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}\lambda_{\theta_1}\int_0^1(r-\frac{1}{2})dw(r)+\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(Q(s)\right)\left(\overline{\beta}_2^2\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}}},$$

$$= \frac{\sqrt{12}\int_0^1(r-\frac{1}{2})dw(r)}{\sqrt{P_b\left(Q(s)\right)}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b\left(Q(s)\right)}}$$

$$= \frac{Z}{\sqrt{P_b\left(Q(s)\right)}} + \frac{\overline{\beta}_2\overline{\theta}_\Delta}{\sqrt{12\lambda_{\theta_1}^2 P_b\left(Q(s)\right)}},$$

using the fact that $\int_0^1(r-\frac{1}{2})dw(r) \sim N\left(0, \frac{1}{12}\right)$.

Proof of Theorem 11. With medium trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from Lemmas 1, 2, 3, 4 and Theorems 1, 2.

$$\text{Case 1: } T^{-3/2}\widehat{S}_{[sT]} = T^{-3/2}\sum_{t=1}^{[rT]}(\epsilon_t-\overline{\epsilon}) - \left(\widehat{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2)$$

$$\Rightarrow \sigma_{\eta,\theta}\int_0^s w(r)dr - s\sigma_{\eta,\theta}\int_0^1 w(r)dr$$

$$-\overline{\beta}_2 L(s)\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_{\eta,\theta}\int_0^1(r-\frac{1}{2})w(r)dr$$

$$= \sigma_{\eta,\theta}Q_1^*(s).$$

$$\text{Case 2: } T^{-1/2}\widehat{S}_{[rT]} = T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t-\overline{\epsilon}) - T\left(\widehat{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2)$$

$$\Rightarrow \lambda_\theta\widetilde{w}(s) - \overline{\beta}_2 L(s)\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\lambda_\theta\int_0^1(r-\frac{1}{2})dw(r)$$

$$= \lambda_\theta Q'(s).$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$\text{Case 1: } T^{-2}\widehat{\lambda}_\theta^2 \Rightarrow \sigma_{\eta,\theta}^2 P_b(Q_1^*(s)),$$
$$\text{Case 2: } \widehat{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(Q'(s)\right).$$

The limits of t_{IV} are as follows:

$$\begin{aligned}
\text{Case 1: } t_{IV} &= \frac{(\widehat{\theta}-\theta_0)}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}}, \\
&= \frac{(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}}, \\
&\Rightarrow \frac{\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_{\eta,\theta}\int_0^1(r-\frac{1}{2})w(r)dr+\overline{\theta}_\Delta}{\sqrt{\sigma_{\eta,\theta}^2 P_b(Q_1^*(s))\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}}, \\
&= \frac{\sqrt{12}\int_0^1(r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1^*(s))}} \\
&+ \frac{\overline{\theta}_\Delta}{\sqrt{\sigma_{\eta,\theta}^2 P_b(Q_1^*(s))\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}}.
\end{aligned}$$

$$\text{Case 2: } t_{IV} = \frac{T(\widehat{\theta}-\theta_0)}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$= \frac{T(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2 \left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2} T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$\Rightarrow \frac{\left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-1} \lambda_{\theta_1} \int_0^1 (r-\frac{1}{2}) dw(r) + \overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(Q'(s)\right) \left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-2} \left(\int_0^1 (r-\frac{1}{2})^2 dr\right)}},$$

$$= \frac{\sqrt{12}\int_0^1 (r-\frac{1}{2}) dw(r)}{\sqrt{P_b\left(Q'(s)\right)}}$$

$$+ \frac{\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(Q'(s)\right) \left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-2} \left(\int_0^1 (r-\frac{1}{2})^2 dr\right)}},$$

$$= \frac{Z}{\sqrt{P_b\left(Q'(s)\right)}}$$

$$+ \frac{\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(Q'(s)\right) \left(\overline{\beta}_2 \int_0^1 (r-\frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-2} \left(\int_0^1 (r-\frac{1}{2})^2 dr\right)}}.$$

Proof of Theorem 12. With small trend slopes, the scaling and limits of the partial sum processes are given as follows where the limits follow from Lemmas 1, 2, 3, 4, 5, 6 and Theorems 1, 2.

$$\text{Case 1: } T^{-3/2}\widehat{S}_{[sT]} = T^{-3/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta}-\theta\right) T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2)$$

$$\Rightarrow \sigma_{\eta,\theta}\int_0^s w(r)dr - s\sigma_\eta \int_0^1 w(r)dr$$

$$- \overline{\beta}_2 L(s)\left(\sigma_2 \int_0^1 (r-\frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_{\eta,\theta}\int_0^1 (r-\frac{1}{2}) w(r) dr$$

$$= \sigma_{\eta,\theta} Q_1^{**}(s).$$

$$\text{Case 2: } T^{-1/2}\widehat{S}_{[rT]} = T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - T\left(\widehat{\theta}-\theta\right)T^{-3/2}\sum_{t=1}^{[rT]}(y_{2t}-\overline{y}_2)$$

$$= T^{-1/2}\sum_{t=1}^{[rT]}(\epsilon_t - \overline{\epsilon}) - \sigma_2 J(s)\left(\sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\lambda_\theta\int_0^1(r-\frac{1}{2})dw(r)$$

$$\Rightarrow \lambda_\theta \widetilde{K}(s)$$

Using fixed-b algebra and arguments from Kiefer and Vogelsang (2005), it follows that

$$\text{Case 1: } T^{-2}\widehat{\lambda}_\theta^2 \Rightarrow \sigma_{\eta,\theta}^2 P_b(Q_1^{**}(s)),$$

$$\text{Case 2: } \widehat{\lambda}_\theta^2 \Rightarrow \lambda_\theta^2 P_b\left(\widetilde{K}(s)\right).$$

The limits of t_{IV} are as follows:

$$\text{Case 1: } t_{IV} = \frac{(\widehat{\theta}-\theta_0)}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$= \frac{(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{T^{-2}\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\overline{t})(y_{2t}-\overline{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\overline{t})^2}},$$

$$\Rightarrow \frac{\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_{\eta,\theta}\int_0^1(r-\frac{1}{2})w(r)dr+\overline{\theta}_\Delta}{\sqrt{\sigma_{\eta,\theta}^2 P_b(Q_1^{**}(s))\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}},$$

$$= \frac{\sqrt{12}\int_0^1(r-\frac{1}{2})w(r)dr}{\sqrt{P_b(Q_1^{**}(s))}} + \frac{\overline{\theta}_\Delta}{\sqrt{\sigma_{\eta,\theta}^2 P_b(Q_1^{**}(s))\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}}.$$

$$\text{Case 2: } t_{IV} = \frac{T(\widehat{\theta}-\theta_0)}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\bar{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2}},$$

$$= \frac{T(\widehat{\theta}-\theta_1)+\overline{\theta}_\Delta}{\sqrt{\widehat{\lambda}_\theta^2\left[T^{-5/2}\sum_{t=1}^{T}(t-\bar{t})(y_{2t}-\bar{y}_2)\right]^{-2}T^{-3}\sum_{t=1}^{T}(t-\bar{t})^2}},$$

$$\Rightarrow \frac{\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\lambda_{\theta_1}\int_0^1(r-\frac{1}{2})dw(r)+\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(\widetilde{K}(s)\right)\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}},$$

$$= \frac{\sqrt{12}\int_0^1(r-\frac{1}{2})dw(r)}{\sqrt{P_b\left(\widetilde{K}(s)\right)}} + \frac{\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(\widetilde{K}(s)\right)\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}},$$

$$= \frac{Z}{\sqrt{P_b\left(\widetilde{K}(s)\right)}} + \frac{\overline{\theta}_\Delta}{\sqrt{\lambda_{\theta_1}^2 P_b\left(\widetilde{K}(s)\right)\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-2}\left(\int_0^1(r-\frac{1}{2})^2dr\right)}}.$$

Proof of Theorem 13. The asymptotic limit of the unit root test statistic for $\epsilon_t(\theta)$ for all the three cases is derived as follows:

Case 1, $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$:

$$\begin{aligned} T^{-5/2}\Sigma_{t=1}^T(\epsilon_t-\bar{\epsilon})(y_{2t}-\bar{y}_2) &= T^{-5/2}\Sigma_{t=1}^T(\epsilon_t-\bar{\epsilon})\left[\overline{\beta}_2(t-\bar{t})+(u_{2t}-\bar{u}_2)\right] \\ &= \overline{\beta}_2 T^{-5/2}\Sigma_{t=1}^T(t-\bar{t})(\epsilon_t-\bar{\epsilon})+o_p(1) \\ &\Rightarrow \overline{\beta}_2\sigma_{\eta,\theta}\int_0^1(r-\frac{1}{2})w(r)dr, \end{aligned}$$

as $\Sigma_{t=1}^T(\epsilon_t-\bar{\epsilon})(u_{2t}-\bar{u}_2) = O_p(T^2)$.

Also, the following holds:

$$\begin{aligned} T^{-1/2}\left(\epsilon_{[sT]}-\bar{\epsilon}\right) &= T^{-1/2}\epsilon_{[sT]} - T^{-3/2}\Sigma_{t=2}^T\epsilon_{[sT]} \\ \Rightarrow \sigma_{\eta,\theta}\left[w(s)-\int_0^1 w(r)dr\right] &\equiv \sigma_{\eta,\theta}N(s), \\ T^{-2}\Sigma_{t=2}^T(\epsilon_{t-1}-\bar{\epsilon})^2 &= T^{-1}\Sigma_{t=1}^T\left[T^{-1/2}(\epsilon_{t-1}-\bar{\epsilon})\right]^2 \Rightarrow \sigma_{\eta,\theta}^2\int_0^1 N(r)^2dr \\ T^{-1}\left(y_{2[sT]}-\bar{y}_2\right) &= T^{-1}\left[\overline{\beta}_2([sT]-\bar{t})+(u_{2[sT]}-\bar{u}_2)\right] \\ &\to \overline{\beta}_2(s-\frac{1}{2}). \end{aligned}$$

For the limit of the unit root test statistic, we need to compute the limit of the following expression:

$$T(\widehat{\rho}-1)=\frac{T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_{t}}{T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2}}.$$

Let us first derive the asymptotic limit of the denominator as follows:

$$\widehat{\epsilon}_t=(\epsilon_t-\overline{\epsilon})-\left(\widehat{\theta}-\theta\right)(y_{2t}-\overline{y}_2).$$

After scaling the partial sums of $\widehat{\epsilon}_t$ by $T^{-1/2}$ we get

$$\begin{aligned}
T^{-1/2}\widehat{\epsilon}_{[sT]} &= T^{-1/2}\left(\epsilon_{[sT]}-\overline{\epsilon}\right)-T^{1/2}\left(\widehat{\theta}-\theta\right)T^{-1}\left(y_{2[sT]}-\overline{y}_2\right)\\
&\Rightarrow \sigma_{\eta,\theta}\left[w(s)-\int_0^1 w(r)dr-(s-\frac{1}{2})\left(\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr\right]\\
&\equiv \sigma_{\eta,\theta}\widehat{w}(s),
\end{aligned}$$

where

$$\widehat{w}(s)=w(s)-\int_0^1 w(r)dr-(s-\frac{1}{2})\left(\int_0^1(r-\frac{1}{2})^2dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr.$$

It follows that

$$T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2}=T^{-1}\Sigma_{t=2}^{T}\left(T^{-1/2}\widehat{\epsilon}_{t-1}\right)^2\Rightarrow\sigma_{\eta,\theta}^2\int_0^1\widehat{w}(r)^2dr.$$

The asymptotic limit of the numerator is obtained as follows. Straightforward calculations give

$$\begin{aligned}
y_{2t} &= \mu_2+\overline{\beta}_2 t+u_{2t},\\
y_{2t-1} &= \mu_2+\overline{\beta}_2(t-1)+u_{2t-1},\\
y_{2t} &= \overline{\beta}_2+y_{2t-1}+w_{2t}.
\end{aligned}$$

In order to compute the limit of $T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2$, we proceed as follows:

$$\widehat{\epsilon}_t=(\epsilon_t-\overline{\epsilon})-\left(\widehat{\theta}-\theta\right)(y_{2t}-\overline{y}_2).$$

First difference both sides to obtain

$$\widehat{\epsilon}_t-\widehat{\epsilon}_{t-1}=\Delta\widehat{\epsilon}_t=\Delta\epsilon_t-\left(\widehat{\theta}-\theta\right)\Delta y_{2t}=\eta_t-\left(\widehat{\theta}-\theta\right)(w_{2t}+\overline{\beta}_2).$$

Using the formula for $\Delta\widehat{\epsilon}_t$ gives

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 &= T^{-1}\Sigma_{t=2}^{T}\eta_t^2 + T^{-1}\Sigma_{t=2}^{T}\left(\widehat{\theta}-\theta\right)^2(w_{2t}+\overline{\beta}_2)^2 - 2T^{-1}\Sigma_{t=2}^{T}\eta_t\left(\widehat{\theta}-\theta\right)(w_{2t}+\overline{\beta}_2) \\
&= T^{-1}\Sigma_{t=2}^{T}\eta_t^2 + T\left(\widehat{\theta}-\theta\right)^2 T^{-2}\Sigma_{t=2}^{T}(w_{2t}+\overline{\beta}_2)^2 \\
&- 2T^{1/2}\left(\widehat{\theta}-\theta\right)T^{-3/2}\Sigma_{t=2}^{T}\eta_t(w_{2t}+\overline{\beta}_2) \\
&= T^{-1}\Sigma_{t=2}^{T}\eta_t^2 + o_p(1) \xrightarrow{p} \sigma_\eta^2.
\end{aligned}
$$

Using the limits of $T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2$, and $T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2$, we obtain the limit of $T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t$ as follows:

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_t^2 &= T^{-1}\Sigma_{t=2}^{T}(\widehat{\epsilon}_{t-1}+\Delta\widehat{\epsilon}_t)^2 = T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2 + T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 + 2T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t \\
T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t &= \frac{1}{2}\left[T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_t^2 - T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2 - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2\right] \\
&= \frac{1}{2}\left[T^{-1}(\widehat{\epsilon}_T^2-\widehat{\epsilon}_1^2) - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2\right] \\
&\Rightarrow \frac{1}{2}\left[\sigma_\eta^2\widehat{w}(1)^2 - \sigma_\eta^2\widehat{w}(0)^2 - \sigma_\eta^2\right] = \frac{\sigma_\eta^2}{2}\left[\widehat{w}(1)^2-\widehat{w}(0)^2-1\right].
\end{aligned}
$$

Now the asymptotic distribution of t-statistic is as follows:

$$
\begin{aligned}
t_{\widehat{\rho}=1} &= \frac{T(\widehat{\rho}-1)}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2\right)^{-1}}} \\
&\Rightarrow \frac{\frac{\frac{\sigma_{\eta,\theta}^2}{2}\left[\widehat{w}(1)^2-\widehat{w}(0)^2-1\right]}{\sigma_{\eta,\theta}^2\int_0^1\widehat{w}(r)^2dr}}{\sqrt{\sigma_{\eta,\theta}^2\left(\sigma_{\eta,\theta}^2\int_0^1\widehat{w}(r)^2dr\right)^{-1}}} = \frac{\frac{1}{2}\left(\widehat{w}(1)^2-\widehat{w}(0)^2-1\right)}{\sqrt{\int_0^1\widehat{w}(r)^2dr}} = \frac{\int_0^1\widehat{w}(r)dw(r)}{\sqrt{\int_0^1\widehat{w}(r)^2dr}},
\end{aligned}
$$

which is the usual DF limit for $t_{\widehat{\rho}=1}$ when an intercept and time trend are in DF regression. It is easy to show that $\text{plim}s^2 = \sigma_{\eta,\theta}^2$ as follows:

$$
\begin{aligned}
s^2 &= \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1})^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\epsilon}_{t-1} - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}(\Delta\widehat{\epsilon}_t - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t T(\widehat{\rho}-1) + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2 T^2(\widehat{\rho}-1)^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 + o_p(1) \xrightarrow{p} \sigma_{\eta,\theta}^2.
\end{aligned}
$$

Case 2, $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$

$$T^{-1/2}\left(y_{2[sT]} - \overline{y}_2\right) = T^{-1}\overline{\beta}_2(t-\overline{t}) + T^{-1/2}(u_{2[sT]} - \overline{u}_2)$$
$$\Rightarrow \overline{\beta}_2(s-\frac{1}{2}) + \sigma_2\left[w(s) - \int_0^1 w(r)dr\right].$$

In the expression for $T(\widehat{\rho}-1)$, the asymptotic limit of the denominator is as follows:

$$\widehat{\epsilon}_t = (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta}-\theta\right)(y_{2t} - \overline{y}_2)$$

Scaling the partial sums of $\widehat{\epsilon}_t$ by $T^{-1/2}$, we get

$$T^{-1/2}\widehat{\epsilon}_{[sT]} = T^{-1/2}\left(\epsilon_{[sT]} - \overline{\epsilon}\right) - \left(\widehat{\theta}-\theta\right)T^{-1/2}\left(y_{2[sT]} - \overline{y}_2\right)$$
$$\Rightarrow \sigma_{\eta,\theta}\left[\begin{array}{c} w(s) - \int_0^1 w(r)dr - \left(\overline{\beta}_2(s-\frac{1}{2}) + \sigma_2\left[w(s) - \int_0^1 w(r)dr\right]\right) * \\ \left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr + \sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr \end{array}\right]$$
$$\equiv \sigma_{\eta,\theta}\widehat{w}^*(s).$$

This implies that

$$T^{-2}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}^2 = T^{-1}\Sigma_{t=2}^T\left(T^{-1/2}\widehat{\epsilon}_{t-1}\right)^2 \Rightarrow \sigma_{\eta,\theta}^2\int_0^1\widehat{w}^*(r)^2dr.$$

The asymptotic limit of the numerator is obtained as follows:

$$T^{-1}\Sigma_{t=2}^T\eta_t w_{2t} = T^{-1}\Sigma_{t=2}^T(w_{1t} - \theta w_{2t})w_{2t}$$
$$= T^{-1}\Sigma_{t=2}^T w_{1t}w_{2t} - \theta T^{-1}\Sigma_{t=2}^T w_{2t}^2 \to -\theta\sigma_2^2.$$

In order to compute the limit of $T^{-1}\Sigma_{t=2}^T\Delta\widehat{\epsilon}_t^2$, we proceed as follows:

$$\widehat{\epsilon}_t - \widehat{\epsilon}_{t-1} = \Delta\widehat{\epsilon}_t = \Delta\epsilon_t - \left(\widehat{\theta}-\theta\right)\Delta y_{2t} = \eta_t - \left(\widehat{\theta}-\theta\right)(w_{2t} + T^{-1/2}\overline{\beta}_2).$$

Using the formula for $\Delta\widehat{\epsilon}_t$, we obtain

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 &= T^{-1}\Sigma_{t=2}^{T}\eta_t^2 + T^{-1}\Sigma_{t=2}^{T}\left(\widehat{\theta}-\theta\right)^2(w_{2t}+T^{-1/2}\overline{\beta}_2)^2 \\
&- 2T^{-1}\Sigma_{t=2}^{T}\eta_t\left(\widehat{\theta}-\theta\right)(w_{2t}+T^{-1/2}\overline{\beta}_2) \\
&= T^{-1}\Sigma_{t=2}^{T}\eta_t^2 + \left(\widehat{\theta}-\theta\right)^2 T^{-1}\Sigma_{t=2}^{T}(w_{2t}+T^{-1/2}\overline{\beta}_2)^2 \\
&- 2\left(\widehat{\theta}-\theta\right)T^{-1}\Sigma_{t=2}^{T}\eta_t(w_{2t}+T^{-1/2}\overline{\beta}_2) \\
&\Rightarrow \sigma_{\eta,\theta}^2\left[1+\sigma_2^2\left[\begin{array}{c}\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1} \\ *\int_0^1(r-\frac{1}{2})w(r)dr\end{array}\right]^2\right]+ \\
&2\theta\sigma_{\eta,\theta}\sigma_2^2\left[\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr+\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr\right] \\
&\equiv L.
\end{aligned}
$$

Using the limits of $T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2$, and $T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2$, we obtain the limit of $T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t$ as follows:

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t &= \frac{1}{2}\left[T^{-1}(\widehat{\epsilon}_T^2-\widehat{\epsilon}_1^2) - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2\right] \\
&\Rightarrow \frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^*(1)^2 - \sigma_{\eta,\theta}^2\widehat{w}^*(0)^2 - L\right]
\end{aligned}
$$

Now the asymptotic distribution of t-statistic is as follows:

$$
\begin{aligned}
t_{\widehat{\rho}=1} &= \frac{T(\widehat{\rho}-1)}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2\right)^{-1}}} \\
&\Rightarrow \frac{\frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^*(1)^2-L\right]}{\sigma_{\eta,\theta}^2\int_0^1\widehat{w}^*(r)^2dr}}{\sqrt{L\left(\sigma_{\eta,\theta}^2\int_0^1\widehat{w}^*(r)^2dr\right)^{-1}}} = \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^*(1)^2-\sigma_{\eta,\theta}^2\widehat{w}^*(0)^2-L\right]}{\sigma_{\eta,\theta}\sqrt{L\int_0^1\widehat{w}^*(r)^2dr}},
\end{aligned}
$$

where

$$
\begin{aligned}
s^2 &= \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t-\widehat{\rho\epsilon}_{t-1})^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t-\widehat{\epsilon}_{t-1}-(\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}(\Delta\widehat{\epsilon}_t-(\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_tT(\widehat{\rho}-1) + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2T^2(\widehat{\rho}-1)^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 + o_p(1) \xrightarrow{p} L.
\end{aligned}
$$

Case 3, $\beta_1 = T^{-1}\overline{\beta}_1$, $\beta_2 = T^{-1}\overline{\beta}_2$

$$T^{-1/2}\left(y_{2[sT]}-\overline{y}_2\right)=T^{-3/2}\overline{\beta}_2(t-\overline{t})+T^{-1/2}(u_{2[sT]}-\overline{u}_2)$$
$$\Rightarrow \sigma_2\left[w(s)-\int_0^1 w(r)dr\right].$$

In the expression for $T(\widehat{\rho}-1)$, the asymptotic limit of the denominator is obtained as follows:

We know that

$$\widehat{\epsilon}_t=(\epsilon_t-\overline{\epsilon})-\left(\widehat{\theta}-\theta\right)(y_{2t}-\overline{y}_2).$$

Scaling the partial sums of $\widehat{\epsilon}_t$ by $T^{-1/2}$, we get

$$\begin{aligned}T^{-1/2}\widehat{\epsilon}_{[sT]}&=T^{-1/2}\left(\epsilon_{[sT]}-\overline{\epsilon}\right)-\left(\widehat{\theta}-\theta\right)T^{-1/2}\left(y_{2[sT]}-\overline{y}_2\right)\\&\Rightarrow \sigma_{\eta,\theta}\begin{bmatrix} w(s)-\int_0^1 w(r)dr-\sigma_2\left[w(s)-\int_0^1 w(r)dr\right]* \\ \left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr\end{bmatrix}\\&\equiv \sigma_{\eta,\theta}\widehat{w}^{**}(s).\end{aligned}$$

It follows that

$$T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2=T^{-1}\Sigma_{t=2}^{T}\left(T^{-1/2}\widehat{\epsilon}_{t-1}\right)^2\Rightarrow\sigma_{\eta,\theta}^2\int_0^1\widehat{w}^{**}(r)^2dr.$$

The asymptotic limit of the numerator is obtained as follows:

$$\widehat{\epsilon}_t-\widehat{\epsilon}_{t-1}=\Delta\widehat{\epsilon}_t=\Delta\epsilon_t-\left(\widehat{\theta}-\theta\right)\Delta y_{2t}=\eta_t-\left(\widehat{\theta}-\theta\right)(w_{2t}+T^{-1}\overline{\beta}_2).$$

Using the formula for $\Delta\widehat{\epsilon}_t$, we obtain

$$\begin{aligned}T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2&=T^{-1}\Sigma_{t=2}^{T}\eta_t^2+T^{-1}\Sigma_{t=2}^{T}\left(\widehat{\theta}-\theta\right)^2(w_{2t}+T^{-1}\overline{\beta}_2)^2\\&-2T^{-1}\Sigma_{t=2}^{T}\eta_t\left(\widehat{\theta}-\theta\right)(w_{2t}+T^{-1}\overline{\beta}_2)\\&=T^{-1}\Sigma_{t=2}^{T}\eta_t^2+\left(\widehat{\theta}-\theta\right)^2T^{-1}\Sigma_{t=2}^{T}(w_{2t}+T^{-1}\overline{\beta}_2)^2\\&-2\left(\widehat{\theta}-\theta\right)T^{-1}\Sigma_{t=2}^{T}\eta_t(w_{2t}+T^{-1}\overline{\beta}_2)\\&\Rightarrow\sigma_{\eta,\theta}^2\left[1+\sigma_2^2\left[\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr\right]^2\right]+\\&2\theta\sigma_{\eta,\theta}\sigma_2^2\left[\left(\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr\right]\equiv L'.\end{aligned}$$

Using the limits of $T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2$, and $T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2$, we obtain the limit of $T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t$ as follows:

$$T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t = \frac{1}{2}\left[T^{-1}(\widehat{\epsilon}_T^2 - \widehat{\epsilon}_1^2) - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2\right]$$
$$\Rightarrow \frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^{**}(1)^2 - \sigma_{\eta,\theta}^2\widehat{w}^{**}(0)^2 - L'\right].$$

Now the asymptotic distribution of t-statistic is as follows:

$$t_{\widehat{\rho}=1}^{GLS} = \frac{T(\widehat{\rho}-1)}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2\right)^{-1}}}$$
$$\Rightarrow \frac{\frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^{**}(1)^2-\sigma_{\eta,\theta}^2\widehat{w}^{**}(0)^2-L'\right]}{\sigma_{\eta,\theta}^2\int_0^1\widehat{w}^{**}(r)^2dr}}{\sqrt{L'\left(\sigma_{\eta,\theta}^2\int_0^1\widehat{w}^{**}(r)^2dr\right)^{-1}}} = \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2\widehat{w}^{**}(1)^2 - \sigma_{\eta,\theta}^2\widehat{w}^{**}(0)^2 - L'\right]}{\sigma_{\eta,\theta}\sqrt{L'\int_0^1\widehat{w}^{**}(r)^2dr}},$$

where

$$s^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1})^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\epsilon}_{t-1} - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}(\Delta\widehat{\epsilon}_t - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t T(\widehat{\rho}-1) + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2T^2(\widehat{\rho}-1)^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 + o_p(1) \xrightarrow{p} L'$$

Case 4, $\beta_1 = 0$, $\beta_2 = 0$. The asymptotic limit of the IV estimator of θ is derived as follows:

$$\widehat{\theta} = \frac{T^{-5/2}\Sigma_{t=2}^{T}(t-\bar{t})u_{1t}}{T^{-5/2}\Sigma_{t=2}^{T}(t-\bar{t})u_{2t}} \Rightarrow \frac{\sigma_1\int_0^1(r-\frac{1}{2})w_1(r)dr}{\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr}.$$

Now

$$T^{-1/2}(u_{1[sT]} - \overline{u}_1) \Rightarrow \sigma_1\left[w_1(s) - \int_0^1 w_1(r)dr\right].$$

and

$$T^{-1/2}(u_{2[sT]} - \overline{u}_2) \Rightarrow \sigma_2\left[w_2(s) - \int_0^1 w_2(r)dr\right].$$

In the expression for $T(\widehat{\rho}-1)$, the asymptotic limit of the denominator is obtained as follows:

$$\widehat{\epsilon}_t = (y_{1t} - \overline{y}_1) - \widehat{\theta}\,(y_{2t} - \overline{y}_2)\,,$$
$$= (u_{1t} - \overline{u}_1) - \widehat{\theta}\,(u_{2t} - \overline{u}_2)$$

Scaling the partial sums of $\widehat{\epsilon}_t$ by $T^{-1/2}$, we get

$$\begin{aligned}
T^{-1/2}\widehat{\epsilon}_{[sT]} &= T^{-1/2}\left(u_{1[sT]}-\overline{u}_1\right)-\widehat{\theta}T^{-1/2}\left(u_{2[sT]}-\overline{u}_2\right)\\
&\Rightarrow \begin{array}{c}\sigma_1\left[w_1(s)-\int_0^1 w_1(r)dr\right]-\left[w_2(s)-\int_0^1 w_2(r)dr\right]*\\ \left[\left(\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_1\int_0^1(r-\frac{1}{2})w_1(r)dr\right]\end{array}\\
&\equiv \sigma_1\widehat{w}^{***}(s).
\end{aligned}$$

It follows that

$$T^{-2}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}^2 = T^{-1}\Sigma_{t=2}^T\left(T^{-1/2}\widehat{\epsilon}_{t-1}\right)^2 \Rightarrow \sigma_1^2\int_0^1\widehat{w}^{***}(r)^2dr.$$

The asymptotic limit of the numerator is obtained as follows:

$$\begin{aligned}
\widehat{\epsilon}_t-\widehat{\epsilon}_{t-1} &= \Delta\widehat{\epsilon}_t = \Delta u_{1t}-\widehat{\theta}\Delta u_{2t} = w_{1t}-\widehat{\theta}w_{2t}\\
T^{-1}\Sigma_{t=2}^T\Delta\widehat{\epsilon}_t^2 &= T^{-1}\Sigma_{t=2}^T w_{1t}^2+\widehat{\theta}^2T^{-1}\Sigma_{t=2}^T w_{2t}^2-2\widehat{\theta}T^{-1}\Sigma_{t=2}^T w_{1t}w_{2t}\\
&\Rightarrow \sigma_1^2+\left[\left(\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_1\int_0^1(r-\frac{1}{2})w_1(r)dr\right]^2\\
&-2\varsigma\sigma_2\left(\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_1\int_0^1(r-\frac{1}{2})w_1(r)dr\\
&\equiv L'',
\end{aligned}$$

assuming $w_{1t}=\varsigma w_{2t}+v_{1t}$, where $v_{1t}\sim(0,\sigma_v^2)$. Using the limits of $T^{-2}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}^2$, and $T^{-1}\Sigma_{t=2}^T\Delta\widehat{\epsilon}_t^2$, we obtain the limit of $T^{-1}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t$ as follows:

$$\begin{aligned}
T^{-1}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t &= \frac{1}{2}\left[T^{-1}(\widehat{\epsilon}_T^2-\widehat{\epsilon}_1^2)-T^{-1}\Sigma_{t=2}^T\Delta\widehat{\epsilon}_t^2\right]\\
&\Rightarrow \frac{1}{2}\left[\sigma_1^2\widehat{w}^{***}(1)^2-\sigma_1^2\widehat{w}^{***}(0)^2-L''\right].
\end{aligned}$$

Now the asymptotic distribution of t-statistic is as follows:

$$\begin{aligned}
t_{\widehat{\rho}=1} &= \frac{T(\widehat{\rho}-1)}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^T\widehat{\epsilon}_{t-1}^2\right)^{-1}}}\\
&\Rightarrow \frac{\frac{\frac{1}{2}\left[\sigma_1^2\widehat{w}^{***}(1)^2-\sigma_1^2\widehat{w}^{***}(0)^2-L''\right]}{\sigma_1^2\int_0^1\widehat{w}^{***}(r)^2dr}}{\sqrt{L''\left(\sigma_1^2\int_0^1\widehat{w}^{***}(r)^2dr\right)^{-1}}} = \frac{\frac{1}{2}\left[\sigma_1^2\widehat{w}^{***}(1)^2-\sigma_1^2\widehat{w}^{***}(0)^2-L''\right]}{\sigma_1\sqrt{L''\int_0^1\widehat{w}^{***}(r)^2dr}},
\end{aligned}$$

where

$$s^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\rho}\widehat{\epsilon}_{t-1})^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}(\widehat{\epsilon}_t - \widehat{\epsilon}_{t-1} - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}(\Delta\widehat{\epsilon}_t - (\widehat{\rho}-1)\widehat{\epsilon}_{t-1})^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}\Delta\widehat{\epsilon}_t T(\widehat{\rho}-1) + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^2 T^2(\widehat{\rho}-1)^2$$
$$= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^2 + o_p(1) \xrightarrow{p} L''.$$

Proof of Theorem 14. The asymptotic limit of the ADF-GLS test statistic for $\epsilon_t(\theta)$ for large and medium trend slopes is derived as follows:

Case 1, $\beta_1 = \overline{\beta}_1$, $\beta_2 = \overline{\beta}_2$:

We already know that

$$\begin{aligned}
\widehat{\epsilon}_t &= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta} - \theta\right)(y_{2t} - \overline{y}_2), \\
&= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta} - \theta\right)\left[\overline{\beta}_2(t - \overline{t}) + (u_{2t} - \overline{u}_2)\right], \\
&= (\epsilon_t - \overline{\epsilon}) - \overline{\beta}_2\left(\widehat{\theta} - \theta\right)(t - \overline{t}) - \left(\widehat{\theta} - \theta\right)(u_{2t} - \overline{u}_2).
\end{aligned}$$

Let the regression equation of ϵ_t on d_t be as follows:

$$\begin{aligned}
\epsilon_t &= \gamma' d_t + \upsilon_t = \gamma_1 + \gamma_2 t + \upsilon_t, \\
\gamma' &= (\gamma_1, \gamma_2), \\
d_t &= (1, t)', \\
\Delta\upsilon_t &= \varphi\upsilon_{t-1} + \eta_t, \\
\varphi &= \frac{c}{T},
\end{aligned}$$

then we can write

$$\Delta_{\overline{\alpha}}\epsilon_t = \gamma'\Delta_{\overline{\alpha}}d_t + \Delta_{\overline{\alpha}}\upsilon_t. \tag{41}$$

In the true data generating process, $(\gamma_1, \gamma_2) = (0, 0)$, and $\upsilon_t = \epsilon_t$. Through OLS applied to (41), we obtain

$$\widehat{\gamma} = \gamma + \left(\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)'\right)^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\Delta_{\overline{\alpha}}\upsilon_t.$$

Now

$$\begin{aligned}
\Delta_{\overline{\alpha}}d_t &= \Delta d_t + (1 - \overline{\alpha})\, d_{t-1} = \Delta d_t - \overline{c}d_{t-1}/T, \\
\Delta_{\overline{\alpha}}d_1 &= (1, 1)', \\
\Delta_{\overline{\alpha}}d_t &= [-\overline{c}/T, 1 - \overline{c}(t-1)/T] \ \text{ for } t \geq 2.
\end{aligned}$$

This implies that

$$\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' = (1,1)'(1,1)+\sum_{t=2}^{T}\left[\begin{array}{cc}(\overline{c}/T)^2 & -\overline{c}\left(1-\overline{c}(t-1)/T\right)/T \\ -\overline{c}\left(1-\overline{c}(t-1)/T\right)/T & \left(1-\overline{c}(t-1)/T\right)^2\end{array}\right].$$

Now let $D_T = \left[\begin{array}{cc}1 & 0 \\ 0 & T^{1/2}\end{array}\right]$, then

$$\begin{aligned}D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' D_T^{-1} &= \left[\begin{array}{cc}1 & 1/\sqrt{T} \\ 1/\sqrt{T} & 1/T\end{array}\right] \\ &+\sum_{t=2}^{T}\left[\begin{array}{cc}(\overline{c}/T)^2 & -\overline{c}\left(1-\overline{c}(t-1)/T\right)/T^{3/2} \\ -\overline{c}\left(1-\overline{c}(t-1)/T\right)/T^{3/2} & \left(1-\overline{c}(t-1)/T\right)^2/T\end{array}\right] \\ &= \left[\begin{array}{cc}1 & 0 \\ 0 & 0\end{array}\right]+\sum_{t=2}^{T}\left[\begin{array}{cc}0 & 0 \\ 0 & \left(1-\overline{c}(t-1)/T\right)^2/T\end{array}\right]+\frac{1}{\sqrt{T}}O_p(1).\end{aligned}$$

Therefore

$$\begin{aligned}&D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' D_T^{-1} \xrightarrow{p} \left[\begin{array}{cc}1 & 0 \\ 0 & 0\end{array}\right]+\int_0^1\left[\begin{array}{cc}0 & 0 \\ 0 & (1-\overline{c}r)^2\end{array}\right]dr \\ &= \left[\begin{array}{cc}1 & 0 \\ 0 & \left(1-\overline{c}+\overline{c}^2/3\right)\end{array}\right], \text{ and}\end{aligned}$$

$$\left(D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' D_T^{-1}\right)^{-1} \xrightarrow{p} \left[\begin{array}{cc}1 & 0 \\ 0 & \left(3/(1-\overline{c}+\overline{c}^2)\right)\end{array}\right].$$

Also

$$\begin{aligned}D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\Delta_{\overline{\alpha}}\upsilon_t &= D_T^{-1}(1,1)'\upsilon_1 + D_T^{-1}\sum_{t=2}^{T}\left[\begin{array}{c}-(\overline{c}/T)\Delta_{\overline{\alpha}}\upsilon_t \\ (1-\overline{c}(t-1)/T)\Delta_{\overline{\alpha}}\upsilon_t\end{array}\right] \\ &= D_T^{-1}(1,1)'\upsilon_1 + \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left[\begin{array}{c}-(\overline{c}\Delta_{\overline{\alpha}}\upsilon_t)/\sqrt{T} \\ (1-\overline{c}(t-1)/T)\Delta_{\overline{\alpha}}\upsilon_t\end{array}\right].\end{aligned}$$

As $\Delta_{\overline{\alpha}}\upsilon_t = \Delta\upsilon_t - \overline{c}\upsilon_{t-1}/T$ for $t \geq 2$, $\Delta_{\overline{\alpha}}\upsilon_1 = \upsilon_1$, and $\Sigma_{t=2}^{T}\Delta\upsilon_t = \upsilon_T - \upsilon_1$, therefore

$$\begin{aligned}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}\upsilon_t &= \frac{1}{\sqrt{T}}\upsilon_1 + \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\Delta\upsilon_t - \frac{\overline{c}}{T^{3/2}}\sum_{t=2}^{T}\upsilon_{t-1} \\ &= \frac{1}{\sqrt{T}}\upsilon_T - \frac{\overline{c}}{T^{3/2}}\sum_{t=2}^{T}\upsilon_{t-1}.\end{aligned}$$

Let $\alpha = 1 + c/T$, then

$$\begin{aligned}
\frac{1}{\sqrt{T}}\upsilon_t &= \frac{1}{\sqrt{T}}\sum_{j=1}^{T}\alpha^{t-j}\eta_j, \\
\frac{1}{\sqrt{T}}\upsilon_{[sT]} &= \frac{1}{\sqrt{T}}\sum_{j=1}^{[sT]}\alpha^{[sT]-j}\eta_j \Rightarrow \sigma_\eta w_c(s), \\
w_c(s) &= \int_0^s \exp(c\,(s-u))dw_0(u).
\end{aligned}$$

This implies that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}\upsilon_t = \frac{1}{\sqrt{T}}\upsilon_T - \frac{\overline{c}}{T^{3/2}}\sum_{t=2}^{T}\upsilon_{t-1} \Rightarrow \sigma_{\eta,\theta}\left[w_c(1) - \overline{c}\int_0^1 w_c(r)dr\right].$$

As $\Sigma_{t=2}^{T}(t-1)\Delta\upsilon_t = -\upsilon_1 + (T-1)\upsilon_T - \Sigma_{t=2}^{T-1}\upsilon_t$, therefore

$$\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left(1-\frac{\overline{c}(t-1)}{T}\right)\Delta_{\overline{\alpha}}\upsilon_t &= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\left(1-\frac{\overline{c}(t-1)}{T}\right)\left(\Delta\upsilon_t - \frac{\overline{c}}{T}\upsilon_{t-1}\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\Delta\upsilon_t - \frac{\overline{c}}{T^{3/2}}\sum_{t=2}^{T}(t-1)\Delta\upsilon_t - \frac{\overline{c}}{T^{3/2}}\sum_{t=2}^{T}\upsilon_{t-1} \\
&+ \frac{\overline{c}^2}{T^{5/2}}\sum_{t=2}^{T}(t-1)\upsilon_{t-1} \\
&= \frac{(1-\overline{c})}{\sqrt{T}}\upsilon_T + \frac{\overline{c}}{T^{3/2}}\left(\sum_{t=2}^{T-1}\upsilon_t - \sum_{t=2}^{T}\upsilon_{t-1}\right) + \frac{\overline{c}^2}{T^{5/2}}\sum_{t=2}^{T}(t-1)\upsilon_{t-1} \\
&+ o_p(1) \\
&= \frac{(1-\overline{c})}{\sqrt{T}}\upsilon_T + \frac{\overline{c}^2}{T^{5/2}}\sum_{t=2}^{T}t\upsilon_{t-1} + o_p(1) \\
&\Rightarrow \sigma_{\eta,\theta}\left[(1-\overline{c})\,w_c(1) + \overline{c}^2\int_0^1 rw_c(r)dr\right].
\end{aligned}$$

It follows that

$$\begin{aligned}
D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\Delta_{\overline{\alpha}}\upsilon_t &= D_T^{-1}\,(1,1)'\,\upsilon_1 + \frac{1}{\sqrt{T}}\sum_{t=2}^{T}\begin{bmatrix} -(\overline{c}\Delta_{\overline{\alpha}}\upsilon_t)/\sqrt{T} \\ (1-\overline{c}(t-1)/T)\Delta_{\overline{\alpha}}\upsilon_t \end{bmatrix} \\
&= \begin{bmatrix} \upsilon_1 + \frac{1}{\sqrt{T}}O_p(1) \\ o_p(1) + (1-\overline{c}(t-1)/T)\Delta_{\overline{\alpha}}\upsilon_t/\sqrt{T} \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} \upsilon_1 \\ \sigma_{\eta,\theta}\left[(1-\overline{c})\,w_c(1) + \overline{c}^2\int_0^1 rw_c(r)dr\right] \end{bmatrix}
\end{aligned}$$

This implies that

$$
\begin{aligned}
D_T\left(\widehat{\gamma}-\gamma\right) &= \left(D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' D_T^{-1}\right)^{-1} D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\Delta_{\overline{\alpha}}\upsilon_t, \\
&\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & \left(3/(1-\overline{c}+\overline{c}^2)\right) \end{bmatrix} \begin{bmatrix} \upsilon_1 \\ \sigma_{\eta,\theta}\left[\left(1-\overline{c}\right)w_c(1)+\overline{c}^2\int_0^1 rw_c(r)dr\right] \end{bmatrix}, \\
&= \begin{bmatrix} \upsilon_1 \\ \sigma_{\eta,\theta}K_c(\overline{c}) \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\text{where } K_c(\overline{c}) &= 3\varpi w_c(1)+3(1-\varpi)\int_0^1 rw_c(r)dr, \\
\varpi &= \left(1-\overline{c}\right)/(1-\overline{c}+\overline{c}^2), \\
1-\varpi &= \overline{c}^2/(1-\overline{c}+\overline{c}^2).
\end{aligned}
$$

Similarly, let the regression equation of u_{2t} on d_t be as follows:

$$
\begin{aligned}
u_{2t} &= \phi' d_t+\kappa_t = \phi_1+\phi_2 t+\kappa_t, \\
\phi' &= \left(\phi_1,\phi_2\right), \\
d_t &= (1,t)', \\
\Delta\kappa_t &= \varkappa\kappa_{t-1}+w_{2t}, \\
\varkappa &= \frac{c}{T}.
\end{aligned}
$$

This implies that

$$
\begin{aligned}
D_T\left(\widehat{\phi}-\phi\right) &= \left(D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\left(\Delta_{\overline{\alpha}}d_t\right)' D_T^{-1}\right)^{-1} D_T^{-1}\sum_{t=1}^{T}\Delta_{\overline{\alpha}}d_t\Delta_{\overline{\alpha}}\kappa_t \\
&\Rightarrow \begin{bmatrix} \kappa_1 \\ \sigma_2 K_c(\overline{c}) \end{bmatrix}.
\end{aligned}
$$

Let us first derive the asymptotic limit of the denominator in the expression of $T\widehat{\pi}$ as follows:

$$
\widehat{\epsilon}_t^{GLS} = \epsilon_t^{GLS}-\left(\widehat{\theta}-\theta\right)u_{2t}^{GLS},
$$

Substituting the expressions for ϵ_t^{GLS}, and u_{2t}^{GLS}, we get

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \epsilon_t-\widehat{\gamma}_1-\widehat{\gamma}_2 t-\left(\widehat{\theta}-\theta\right)\left[u_{2t}-\widehat{\phi}_1-\widehat{\phi}_2 t\right] \\
&= \upsilon_t-(\widehat{\gamma}_1-\gamma_1)-(\widehat{\gamma}_2-\gamma_2)t-\left(\widehat{\theta}-\theta\right)\left[\kappa_t-(\widehat{\phi}_1-\phi_1)-(\widehat{\phi}_2-\phi_2)t\right].
\end{aligned}
$$

Scaling the partial sum of $\widehat{\epsilon}_t^{GLS}$ by $T^{-1/2}$, we get

$$
\begin{aligned}
T^{-1/2}\widehat{\epsilon}_t^{GLS} &= T^{-1/2}\upsilon_t - T^{-1/2}(\widehat{\gamma}_1-\gamma_1) - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{t}{T} - T^{-1}\kappa_t T^{1/2}\left(\widehat{\theta}-\theta\right)\\
&\quad + T^{-1}T^{1/2}\left(\widehat{\theta}-\theta\right)(\widehat{\phi}_1-\phi_1) + T^{-1/2}T^{1/2}\left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\frac{t}{T}\\
&= T^{-1/2}\upsilon_t - T^{-1/2}O_p(1) - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{t}{T} + o_p(1),\\
T^{-1/2}\widehat{\epsilon}_{[sT]}^{GLS} &= T^{-1/2}\upsilon_{[sT]} - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{[sT]}{T} + o_p(1)\\
&\Rightarrow \sigma_{\eta,\theta}\left[w_c(s) - sK_c(\overline{c})\right] \equiv \sigma_{\eta,\theta}K_c(s,\overline{c}).
\end{aligned}
$$

This implies that

$$
T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS} = T^{-1}\Sigma_{t=2}^{T}\left(T^{-1/2}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \Rightarrow \sigma_{\eta,\theta}^2\int_0^1 K_c(r,\overline{c})^2dr.
$$

The asymptotic limit of the numerator in the expression of $T\widehat{\pi}$ is obtained as follows:

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \upsilon_t - (\widehat{\gamma}_1-\gamma_1) - (\widehat{\gamma}_2-\gamma_2)t - \left(\widehat{\theta}-\theta\right)\left[\kappa_t - (\widehat{\phi}_1-\phi_1) - (\widehat{\phi}_2-\phi_2)t\right],\\
\widehat{\epsilon}_{t-1}^{GLS} &= \upsilon_{t-1} - (\widehat{\gamma}_1-\gamma_1) - (\widehat{\gamma}_2-\gamma_2)(t-1)\\
&\quad - \left(\widehat{\theta}-\theta\right)\left[\kappa_{t-1} - (\widehat{\phi}_1-\phi_1) - (\widehat{\phi}_2-\phi_2)(t-1)\right],\\
\widehat{\epsilon}_t^{GLS} - \widehat{\epsilon}_{t-1}^{GLS} &= \Delta\widehat{\epsilon}_t^{GLS} = \Delta\upsilon_t - (\widehat{\gamma}_2-\gamma_2) - \left(\widehat{\theta}-\theta\right)\Delta\kappa_t + \left(\widehat{\theta}-\theta\right)(\widehat{\phi}_2-\phi_2)\\
T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^{T}\Delta\upsilon_t^2 + T^{-2}\Sigma_{t=2}^{T}T(\widehat{\gamma}_2-\gamma_2)^2 + T^{-1}T\left(\widehat{\theta}-\theta\right)^2 T^{-1}\Sigma_{t=2}^{T}\Delta\kappa_t^2\\
&\quad + T^{-3}\Sigma_{t=2}^{T}T\left(\widehat{\theta}-\theta\right)^2 T(\widehat{\phi}_2-\phi_2)^2 - 2T^{-3/2}T^{1/2}(\widehat{\gamma}_2-\gamma_2)\Sigma_{t=2}^{T}\Delta\upsilon_t\\
&\quad - 2T^{-3/2}T^{1/2}\left(\widehat{\theta}-\theta\right)\Sigma_{t=2}^{T}\Delta\upsilon_t\Delta\kappa_t\\
&\quad + 2T^{-2}T^{1/2}\left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\Sigma_{t=2}^{T}\Delta\upsilon_t\\
&\quad + 2T^{-2}T^{1/2}(\widehat{\gamma}_2-\gamma_2)T^{1/2}\left(\widehat{\theta}-\theta\right)\Sigma_{t=2}^{T}\Delta\kappa_t\\
&\quad - 2T^{-5/2}\Sigma_{t=2}^{T}T^{1/2}(\widehat{\gamma}_2-\gamma_2)T^{1/2}\left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\\
&\quad - 2T^{-5/2}T\left(\widehat{\theta}-\theta\right)^2 T^{1/2}(\widehat{\phi}_2-\phi_2)\Sigma_{t=2}^{T}\Delta\kappa_t\\
&\xrightarrow{p} \sigma_{\eta,\theta}^2.
\end{aligned}
$$

Using the limits of $T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS}$, and $T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS}$, the limit of $T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS}$ is obtained as follows:

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^{T}\left(\widehat{\epsilon}_{t-1}^{GLS}+\Delta\widehat{\epsilon}_t^{GLS}\right)^2 = T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS}+T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} \\
&\quad + 2T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS}, \\
T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS} &= \frac{1}{2}\left[T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_t^{2GLS} - T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS} - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&= \frac{1}{2}\left[T^{-1}\left(\widehat{\epsilon}_T^{2GLS}-\widehat{\epsilon}_1^{2GLS}\right) - T^{-1}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&\Rightarrow \frac{1}{2}\left[\sigma_{\eta,\theta}^2 K_c(1,\overline{c})^2 - \sigma_{\eta,\theta}^2 K_c(0,\overline{c})^2 - \sigma_{\eta,\theta}^2\right] \\
&= \frac{\sigma_{\eta,\theta}^2}{2}\left[K_c(1,\overline{c})^2 - K_c(0,\overline{c})^2 - 1\right].
\end{aligned}
$$

Now the asymptotic distribution of t-statistic is as follows:

$$
\begin{aligned}
t_{\widehat{\rho}=1} &= \frac{T\widehat{\pi}}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widetilde{\epsilon}_{t-1}^2\right)^{-1}}} \\
&\Rightarrow \frac{\frac{\frac{\sigma_{\eta,\theta}^2}{2}\left[K_c(1,\overline{c})^2-K_c(0,\overline{c})^2-1\right]}{\sigma_{\eta,\theta}^2\int_0^1 K_c(r,\overline{c})^2dr}}{\sqrt{\sigma_{\eta,\theta}^2\left(\sigma_{\eta,\theta}^2\int_0^1 K_c(r,\overline{c})^2dr\right)^{-1}}} = \frac{\frac{1}{2}\left(K_c(1,\overline{c})^2-K_c(0,\overline{c})^2-1\right)}{\sqrt{\int_0^1 K_c(r,\overline{c})^2dr}}.
\end{aligned}
$$

It is easy to show that plim$s^2=\sigma_{\eta,\theta}^2$ as follows:

$$
s^2 = \frac{1}{T-2}\Sigma_{t=2}^{T}\left(\Delta\widehat{\epsilon}_t^{GLS} - \widehat{\pi}\widehat{\epsilon}_{t-1}^{GLS}\right)^2
$$

$$
\begin{aligned}
s^2 = s^2 &= \frac{1}{T-2}\Sigma_{t=2}^{T}\left(\Delta\widehat{\epsilon}_t^{GLS} - \widehat{\pi}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS}T\widehat{\pi} + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS}T^2\widehat{\pi}^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} + o_p(1) \xrightarrow{p} \sigma_{\eta,\theta}^2.
\end{aligned}
$$

Case 2, $\beta_1 = T^{-1/2}\overline{\beta}_1$, $\beta_2 = T^{-1/2}\overline{\beta}_2$:

We already know that

$$
\begin{aligned}
\widehat{\epsilon}_t &= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta}-\theta\right)(y_{2t}-\overline{y}_2), \\
&= (\epsilon_t - \overline{\epsilon}) - \left(\widehat{\theta}-\theta\right)\left[T^{-1/2}\overline{\beta}_2(t-\overline{t}) + (u_{2t}-\overline{u}_2)\right], \\
&= (\epsilon_t - \overline{\epsilon}) - T^{-1/2}\overline{\beta}_2\left(\widehat{\theta}-\theta\right)(t-\overline{t}) - \left(\widehat{\theta}-\theta\right)(u_{2t}-\overline{u}_2).
\end{aligned}
$$

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^{T}\eta_t w_{2t} &= T^{-1}\Sigma_{t=2}^{T}(w_{1t}-\theta w_{2t})w_{2t}\\
&= T^{-1}\Sigma_{t=2}^{T}w_{1t}w_{2t}-\theta T^{-1}\Sigma_{t=2}^{T}w_{2t}^2 \to -\theta\sigma_2^2.
\end{aligned}
$$

$$
T^{-1}\Sigma_{t=2}^{T}\Delta\upsilon_t\Delta\kappa_t = T^{-1}\Sigma_{t=2}^{T}(\frac{c}{T}\epsilon_{t-1}+\eta_t)(\frac{c}{T}u_{2t-1}+w_{2t}) \xrightarrow{p} -\theta\sigma_2^2.
$$

The asymptotic limit of the denominator in the expression of $T\widehat{\pi}$ can be derived as follows:

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \epsilon_t - \widehat{\gamma}_1 - \widehat{\gamma}_2 t - \left(\widehat{\theta}-\theta\right)\left[u_{2t}-\widehat{\phi}_1-\widehat{\phi}_2 t\right]\\
&= \upsilon_t - (\widehat{\gamma}_1-\gamma_1) - (\widehat{\gamma}_2-\gamma_2)t - \left(\widehat{\theta}-\theta\right)\left[\kappa_t-(\widehat{\phi}_1-\phi_1)-(\widehat{\phi}_2-\phi_2)t\right].
\end{aligned}
$$

Scaling the partial sum of $\widehat{\epsilon}_t^{GLS}$ by $T^{-1/2}$, we get

$$
\begin{aligned}
T^{-1/2}\widehat{\epsilon}_t^{GLS} &= T^{-1/2}\upsilon_t - T^{-1/2}(\widehat{\gamma}_1-\gamma_1) - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{t}{T} - T^{-1/2}\kappa_t\left(\widehat{\theta}-\theta\right)\\
&+ T^{-1/2}\left(\widehat{\theta}-\theta\right)(\widehat{\phi}_1-\phi_1) + \left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\frac{t}{T}\\
&= T^{-1/2}\upsilon_t - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{t}{T} - T^{-1/2}\kappa_t\left(\widehat{\theta}-\theta\right) + \left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\frac{t}{T}\\
&+ o_p(1),\\
T^{-1/2}\widehat{\epsilon}_{[sT]}^{GLS} &= T^{-1/2}\upsilon_{[sT]} - T^{1/2}(\widehat{\gamma}_2-\gamma_2)\frac{[sT]}{T} - T^{-1/2}\kappa_{[sT]}\left(\widehat{\theta}-\theta\right)\\
&+ \left(\widehat{\theta}-\theta\right)T^{1/2}(\widehat{\phi}_2-\phi_2)\frac{[sT]}{T} + o_p(1)\\
&\Rightarrow \sigma_{\eta,\theta}\left[\begin{array}{c}
w_c(s) - sK_c(\overline{c}) - \\
\sigma_2 w_c(s) * \\
\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr + \sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1}\int_0^1(r-\frac{1}{2})w(r)dr + \\
s\sigma_2 K_c(\overline{c})\left(\overline{\beta}_2\int_0^1(r-\frac{1}{2})^2dr + \sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr\right)^{-1} * \\
\int_0^1(r-\frac{1}{2})w(r)dr
\end{array}\right]\\
&\equiv \sigma_{\eta,\theta}L_c(s,\overline{c},\sigma_2).
\end{aligned}
$$

It follows that

$$
T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS} = T^{-1}\Sigma_{t=2}^{T}\left(T^{-1/2}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \Rightarrow \sigma_{\eta,\theta}^2\int_0^1 L_c(r,\overline{c},\sigma_2)^2dr.
$$

The asymptotic limit of the numerator is obtained as follows:

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \upsilon_t - (\widehat{\gamma}_1 - \gamma_1) - (\widehat{\gamma}_2 - \gamma_2)t - \left(\widehat{\theta} - \theta\right)\left[\kappa_t - (\widehat{\phi}_1 - \phi_1) - (\widehat{\phi}_2 - \phi_2)t\right], \\
\widehat{\epsilon}_{t-1}^{GLS} &= \upsilon_{t-1} - (\widehat{\gamma}_1 - \gamma_1) - (\widehat{\gamma}_2 - \gamma_2)(t-1) \\
&\quad - \left(\widehat{\theta} - \theta\right)\left[\kappa_{t-1} - (\widehat{\phi}_1 - \phi_1) - (\widehat{\phi}_2 - \phi_2)(t-1)\right], \\
\widehat{\epsilon}_t^{GLS} - \widehat{\epsilon}_{t-1}^{GLS} &= \Delta\widehat{\epsilon}_t^{GLS} = \Delta\upsilon_t - (\widehat{\gamma}_2 - \gamma_2) - \left(\widehat{\theta} - \theta\right)\Delta\kappa_t + \left(\widehat{\theta} - \theta\right)(\widehat{\phi}_2 - \phi_2), \\
T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^T \Delta\upsilon_t^2 + T^{-2}\Sigma_{t=2}^T T(\widehat{\gamma}_2 - \gamma_2)^2 + \left(\widehat{\theta} - \theta\right)^2 T^{-1}\Sigma_{t=2}^T \Delta\kappa_t^2 \\
&\quad + T^{-2}\Sigma_{t=2}^T \left(\widehat{\theta} - \theta\right)^2 T(\widehat{\phi}_2 - \phi_2)^2 - 2T^{-3/2}T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\Sigma_{t=2}^T \Delta\upsilon_t \\
&\quad - 2T^{-1}\left(\widehat{\theta} - \theta\right)\Sigma_{t=2}^T \Delta\upsilon_t \Delta\kappa_t + 2T^{-3/2}\left(\widehat{\theta} - \theta\right)T^{1/2}(\widehat{\phi}_2 - \phi_2)\Sigma_{t=2}^T \Delta\upsilon_t \\
&\quad + 2T^{-3/2}T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\left(\widehat{\theta} - \theta\right)\Sigma_{t=2}^T \Delta\kappa_t \\
&\quad - 2T^{-2}\Sigma_{t=2}^T T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\left(\widehat{\theta} - \theta\right)T^{1/2}(\widehat{\phi}_2 - \phi_2) \\
&\quad - 2T^{-3/2}\left(\widehat{\theta} - \theta\right)^2 T^{1/2}(\widehat{\phi}_2 - \phi_2)\Sigma_{t=2}^T \Delta\kappa_t \\
&\quad \xrightarrow{p} \sigma_{\eta,\theta}^2 + \sigma_2^2 \left[\begin{array}{c} \left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} * \\ \sigma_{\eta,\theta} \int_0^1 (r - \frac{1}{2}) w(r) dr \end{array}\right]^2 \\
&\quad + 2\theta\sigma_2^2 \left(\overline{\beta}_2 \int_0^1 (r - \frac{1}{2})^2 dr + \sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr\right)^{-1} \sigma_{\eta,\theta} \int_0^1 (r - \frac{1}{2}) w(r) dr \\
&\quad \equiv F.
\end{aligned}
$$

Using the limits of $T^{-2}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS}$, and $T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}$, the limit of $T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}$ can be found as follows:

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^T \left(\widehat{\epsilon}_{t-1}^{GLS} + \Delta\widehat{\epsilon}_t^{GLS}\right)^2 = T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS} + T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS} \\
&\quad + 2T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}, \\
T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS} &= \frac{1}{2}\left[T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_t^{2GLS} - T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS} - T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&= \frac{1}{2}\left[T^{-1}\left(\widehat{\epsilon}_T^{2GLS} - \widehat{\epsilon}_1^{2GLS}\right) - T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&\Rightarrow \frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c(1, \overline{c}, \sigma_2)^2 - \sigma_{\eta,\theta}^2 L_c(0, \overline{c}, \sigma_2)^2 - F\right].
\end{aligned}
$$

Now the asymptotic distribution of t-statistic is as follows:

$$t_{\widehat{\rho}=1} = \frac{T\widehat{\pi}}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widetilde{\epsilon}_{t-1}^{2}\right)^{-1}}}$$

$$\Rightarrow \frac{\frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c(1,\overline{c},\sigma_2)^2 - \sigma_{\eta,\theta}^2 L_c(0,\overline{c},\sigma_2)^2 - \digamma\right]}{\sigma_{\eta,\theta}^2\int_0^1 L_c(r,\overline{c},\sigma_2)^2 dr}}{\sqrt{\digamma\left(\sigma_{\eta,\theta}^2\int_0^1 L_c(r,\overline{c},\sigma_2)^2 dr\right)^{-1}}} = \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c(1,\overline{c},\sigma_2)^2 - \sigma_{\eta,\theta}^2 L_c(0,\overline{c},\sigma_2)^2 - \digamma\right]}{\sigma_{\eta,\theta}\sqrt{\digamma\int_0^1 L_c(r,\overline{c},\sigma_2)^2 dr}}.$$

It is easy to show that plim$s^2 = \digamma$ as follows:

$$\begin{aligned} s^2 = s^2 &= \frac{1}{T-2}\Sigma_{t=2}^{T}\left(\Delta\widehat{\epsilon}_t^{GLS} - \widehat{\pi}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \\ &= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} - \frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS}T\widehat{\pi} + \frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS}T^2\widehat{\pi}^2 \\ &= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS} + o_p(1) \xrightarrow{p} \digamma. \end{aligned}$$

Case 3, $\beta_1 = 0$, $\beta_2 = 0$:

The asymptotic limit of the IV estimator of θ is as follows:

$$\widehat{\theta} \Rightarrow \frac{\sigma_1\int_0^1(r-\frac{1}{2})w_1(r)dr}{\sigma_2\int_0^1(r-\frac{1}{2})w_2(r)dr}.$$

The asymptotic limit of the denominator in the expression of $T\widehat{\pi}$ can be obtained as follows:

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \epsilon_t - \widehat{\gamma}_1 - \widehat{\gamma}_2 t - \left(\widehat{\theta} - \theta\right)\left[u_{2t} - \widehat{\phi}_1 - \widehat{\phi}_2 t\right] \\
&= \upsilon_t - (\widehat{\gamma}_1 - \gamma_1) - (\widehat{\gamma}_2 - \gamma_2)t - \left(\widehat{\theta} - \theta\right)\left[\kappa_t - (\widehat{\phi}_1 - \phi_1) - (\widehat{\phi}_2 - \phi_2)t\right] \\
T^{-1/2}\widehat{\epsilon}_t^{GLS} &= T^{-1/2}\upsilon_t - T^{-1/2}(\widehat{\gamma}_1 - \gamma_1) - T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\frac{t}{T} - T^{-1/2}\kappa_t\left(\widehat{\theta} - \theta\right) \\
&+ T^{-1/2}\left(\widehat{\theta} - \theta\right)(\widehat{\phi}_1 - \phi_1) + \left(\widehat{\theta} - \theta\right)T^{1/2}(\widehat{\phi}_2 - \phi_2)\frac{t}{T} \\
&= T^{-1/2}\upsilon_t - T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\frac{t}{T} - T^{-1/2}\kappa_t\left(\widehat{\theta} - \theta\right) + \left(\widehat{\theta} - \theta\right)T^{1/2}(\widehat{\phi}_2 - \phi_2)\frac{t}{T} \\
&+ o_p(1) \\
T^{-1/2}\widehat{\epsilon}_{[sT]}^{GLS} &= T^{-1/2}\upsilon_{[sT]} - T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\frac{[sT]}{T} - T^{-1/2}\kappa_{[sT]}\left(\widehat{\theta} - \theta\right) \\
&+ \left(\widehat{\theta} - \theta\right)T^{1/2}(\widehat{\phi}_2 - \phi_2)\frac{[sT]}{T} + o_p(1) \\
&\Rightarrow \sigma_{\eta,\theta}\left[\begin{array}{c} w_c(s) - sK_c(\overline{c}) - \\ \sigma_{\eta,\theta}^{-1}w_c(s) * \left[\left(\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_1\int_0^1 (r-\frac{1}{2})w_1(r)dr - \theta\right] + \\ s\sigma_{\eta,\theta}^{-1}\sigma_2 K_c(\overline{c})\left[\left(\sigma_2\int_0^1 (r-\frac{1}{2})w_2(r)dr\right)^{-1}\sigma_1\int_0^1 (r-\frac{1}{2})w_1(r)dr - \theta\right] \end{array}\right] \\
&\equiv \sigma_{\eta,\theta}L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta}).
\end{aligned}
$$

It follows that

$$
T^{-2}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS} = T^{-1}\Sigma_{t=2}^T\left(T^{-1/2}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \Rightarrow \sigma_{\eta,\theta}^2\int_0^1 L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2 dr.
$$

The asymptotic limit of the numerator is obtained as follows:

$$
\begin{aligned}
\widehat{\epsilon}_t^{GLS} &= \upsilon_t - (\widehat{\gamma}_1 - \gamma_1) - (\widehat{\gamma}_2 - \gamma_2)t - \left(\widehat{\theta} - \theta\right) \left[\kappa_t - (\widehat{\phi}_1 - \phi_1) - (\widehat{\phi}_2 - \phi_2)t\right], \\
\widehat{\epsilon}_{t-1}^{GLS} &= \upsilon_{t-1} - (\widehat{\gamma}_1 - \gamma_1) - (\widehat{\gamma}_2 - \gamma_2)(t-1) \\
&\quad - \left(\widehat{\theta} - \theta\right) \left[\kappa_{t-1} - (\widehat{\phi}_1 - \phi_1) - (\widehat{\phi}_2 - \phi_2)(t-1)\right], \\
\widehat{\epsilon}_t^{GLS} - \widehat{\epsilon}_{t-1}^{GLS} &= \Delta\widehat{\epsilon}_t^{GLS} = \Delta\upsilon_t - (\widehat{\gamma}_2 - \gamma_2) - \left(\widehat{\theta} - \theta\right) \Delta\kappa_t + \left(\widehat{\theta} - \theta\right) (\widehat{\phi}_2 - \phi_2) \\
T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^T \Delta\upsilon_t^2 + T^{-2}\Sigma_{t=2}^T T(\widehat{\gamma}_2 - \gamma_2)^2 + \left(\widehat{\theta} - \theta\right)^2 T^{-1}\Sigma_{t=2}^T \Delta\kappa_t^2 \\
&\quad + T^{-2}\Sigma_{t=2}^T \left(\widehat{\theta} - \theta\right)^2 T(\widehat{\phi}_2 - \phi_2)^2 - 2T^{-3/2}T^{1/2}(\widehat{\gamma}_2 - \gamma_2)\Sigma_{t=2}^T \Delta\upsilon_t \\
&\quad - 2T^{-1} \left(\widehat{\theta} - \theta\right) \Sigma_{t=2}^T \Delta\upsilon_t \Delta\kappa_t + 2T^{-3/2} \left(\widehat{\theta} - \theta\right) T^{1/2}(\widehat{\phi}_2 - \phi_2)\Sigma_{t=2}^T \Delta\upsilon_t \\
&\quad + 2T^{-3/2}T^{1/2}(\widehat{\gamma}_2 - \gamma_2) \left(\widehat{\theta} - \theta\right) \Sigma_{t=2}^T \Delta\kappa_t \\
&\quad - 2T^{-2}\Sigma_{t=2}^T T^{1/2}(\widehat{\gamma}_2 - \gamma_2) \left(\widehat{\theta} - \theta\right) T^{1/2}(\widehat{\phi}_2 - \phi_2) \\
&\quad - 2T^{-3/2} \left(\widehat{\theta} - \theta\right)^2 T^{1/2}(\widehat{\phi}_2 - \phi_2)\Sigma_{t=2}^T \Delta\kappa_t \\
&\quad \xrightarrow{p} \sigma_{\eta,\theta}^2 + \sigma_2^2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr \right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr - \theta \right]^2 \\
&\quad + 2\theta\sigma_2^2 \left[\left(\sigma_2 \int_0^1 (r - \frac{1}{2}) w_2(r) dr \right)^{-1} \sigma_1 \int_0^1 (r - \frac{1}{2}) w_1(r) dr - \theta \right] \\
&\quad \equiv F^*.
\end{aligned}
$$

Using the limit of $T^{-2}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS}$, and $T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}$, the limit of $T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}$ can be obtained as follows:

$$
\begin{aligned}
T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_t^{2GLS} &= T^{-1}\Sigma_{t=2}^T \left(\widehat{\epsilon}_{t-1}^{GLS} + \Delta\widehat{\epsilon}_t^{GLS}\right)^2 = T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS} + T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS} \\
&\quad + 2T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS}, \\
T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{GLS} \Delta\widehat{\epsilon}_t^{GLS} &= \frac{1}{2} \left[T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_t^{2GLS} - T^{-1}\Sigma_{t=2}^T \widehat{\epsilon}_{t-1}^{2GLS} - T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&= \frac{1}{2} \left[T^{-1} \left(\widehat{\epsilon}_T^{2GLS} - \widehat{\epsilon}_1^{2GLS}\right) - T^{-1}\Sigma_{t=2}^T \Delta\widehat{\epsilon}_t^{2GLS}\right] \\
&\Rightarrow \frac{1}{2} \left[\sigma_{\eta,\theta}^2 L_c^*(1, \bar{c}, \sigma_1, \sigma_2, \sigma_{\eta,\theta})^2 - \sigma_{\eta,\theta}^2 L_c^*(0, \bar{c}, \sigma_1, \sigma_2, \sigma_{\eta,\theta})^2 - F^*\right].
\end{aligned}
$$

Now the asymptotic distribution of t-statistic is as follows:

$$
\begin{aligned}
t_{\widehat{\rho}=1} &= \frac{T\widehat{\pi}}{\sqrt{s^2\left(T^{-2}\Sigma_{t=2}^{T}\widetilde{\epsilon}_{t-1}^2\right)^{-1}}} \\
&\Rightarrow \frac{\frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c^*(1,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2-\sigma_{\eta,\theta}^2 L_c^*(0,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2-F^*\right]}{\sigma_{\eta,\theta}^2\int_0^1 L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2dr}}{\sqrt{F^*\left(\sigma_{\eta,\theta}^2\int_0^1 L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2dr\right)^{-1}}} \\
&= \frac{\frac{1}{2}\left[\sigma_{\eta,\theta}^2 L_c^*(1,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2-\sigma_{\eta,\theta}^2 L_c^*(0,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2-F^*\right]}{\sigma_{\eta,\theta}\sqrt{F^*\int_0^1 L_c^*(s,\overline{c},\sigma_1,\sigma_2,\sigma_{\eta,\theta})^2dr}}.
\end{aligned}
$$

It is easy to show that $\text{plim}s^2 = F^*$ as follows:

$$
\begin{aligned}
s^2 = s^2 &= \frac{1}{T-2}\Sigma_{t=2}^{T}\left(\Delta\widehat{\epsilon}_t^{GLS}-\widehat{\pi}\widehat{\epsilon}_{t-1}^{GLS}\right)^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS}-\frac{2}{T-2}T^{-1}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{GLS}\Delta\widehat{\epsilon}_t^{GLS}T\widehat{\pi}+\frac{1}{T-2}T^{-2}\Sigma_{t=2}^{T}\widehat{\epsilon}_{t-1}^{2GLS}T^2\widehat{\pi}^2 \\
&= \frac{1}{T-2}\Sigma_{t=2}^{T}\Delta\widehat{\epsilon}_t^{2GLS}+o_p(1)\xrightarrow{p}F^*.
\end{aligned}
$$

Table 2.1a: Finite Sample Mean and Standard Deviation, $\epsilon_t \sim I(1)$
$u_{1t}, u_{2t} \sim I(1)$, 10,000 Replications, $\theta = 2$.

			Mean	Standard Deviation
T	β_1	β_2	IV	IV
50	20	10	2.000886	.0346252
	14	7	1.999809	.0484356
	10	5	1.998806	.0668215
	8	4	2.004179	.0856235
	6	3	2.004114	.1190519
	4	2	2.010776	.1805444
	2	1	2.038802	.355008
	.4	.2	1.854142	27.48011
	.2	.1	.3588796	53.2373
	0	0	3.625984	84.86976
100	20	10	1.999302	.0236247
	14	7	1.999138	.0337489
	10	5	2.001098	.0503667
	8	4	2.001684	.0630423
	6	3	2.000639	.0810775
	4	2	2.002888	.1222339
	2	1	2.029194	.2628126
	.4	.2	2.003915	18.49181
	.2	.1	1.833327	42.03659
	0	0	.088921	12.9427
200	20	10	2.000063	.0171978
	14	7	1.999372	.0245572
	10	5	2.000412	.0354813
	8	4	2.004122	.0421344
	6	3	2.001642	.0594053
	4	2	2.001649	.0845398
	2	1	2.006317	.1758927
	.4	.2	2.540382	2.56661
	.2	.1	2.845896	22.6069
	0	0	2.227734	102.2849

Note: IV denotes the estimator given by (11)

Table 2.1b: Finite Sample Mean and Standard Deviation, $\epsilon_t \sim I(0)$
$u_{1t}, u_{2t} \sim I(1)$, 10,000 Replications, $\theta = 2$.

			Mean	Standard Deviation
T	β_1	β_2	IV	IV
50	20	10	2.000047	.0011734
	14	7	2.000029	.0016992
	10	5	1.999854	.0024306
	8	4	2.000055	.003043
	6	3	1.999914	.0039232
	4	2	1.999913	.0060052
	2	1	1.999915	.0122615
	.4	.2	2.063545	2.008758
	.2	.1	2.043028	.8976636
	0	0	2.089652	3.950261
100	20	10	1.999994	.0004378
	14	7	1.999968	.0006101
	10	5	1.999987	.0008715
	8	4	1.999995	.0010907
	6	3	1.999947	.0014223
	4	2	1.999985	.0022623
	2	1	1.999895	.0042304
	.4	.2	2.001467	.0905835
	.2	.1	2.083594	2.250663
	0	0	2.000129	2.027314
200	20	10	2.000008	.0001443
	14	7	2.000004	.0002254
	10	5	2.000009	.0003098
	8	4	1.999989	.0003664
	6	3	2.000035	.0005234
	4	2	1.999999	.000771
	2	1	2.000043	.0015436
	.4	.2	2.001051	.0344767
	.2	.1	2.004925	.4323694
	0	0	1.955252	.9218907

Note: IV denotes the estimator given by (11)

Table 2.2a: Empirical Null Rejection Probabilities, 5% Nominal Level, $\epsilon_t \sim I(1)$.
10,000 Replications, b for t_{θ_0}(R≡Robust)is data dependent, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta \neq 2$.

			$b = 0.1$			$b = 0.5$			$b = 1.0$		
T	β_1	β_2	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)
50	20	10	.053	.054	.053	.051	.052	.049	.047	.047	.048
	14	7	.048	.049	.039	.053	.049	.048	.055	.056	.053
	10	5	.040	.037	.045	.050	.050	.049	.046	.045	.045
	8	4	.050	.050	.049	.055	.053	.054	.050	.051	.047
	6	3	.059	.055	.051	.049	.052	.051	.045	.046	.040
	4	2	.054	.051	.055	.055	.049	.054	.050	.053	.052
	2	1	.044	.039	.044	.059	.052	.049	.056	.062	.053
	.4	.2	.093	.048	.049	.065	.035	.051	.044	.043	.033
	.2	.1	.144	.042	.037	.111	.058	.044	.085	.045	.047
	0	0	.218	.049	.048	.141	.058	.056	.124	.039	.040
100	20	10	.045	.051	.052	.051	.051	.052	.049	.050	.051
	14	7	.046	.052	.052	.053	.051	.047	.047	.052	.050
	10	5	.046	.052	.052	.046	.046	.046	.052	.052	.053
	8	4	.047	.046	.052	.043	.043	.050	.048	.049	.050
	6	3	.043	.047	.046	.064	.068	.066	.050	.049	.048
	4	2	.037	.049	.054	.046	.049	.053	.046	.047	.046
	2	1	.035	.052	.050	.052	.049	.050	.051	.053	.052
	.4	.2	.218	.051	.049	.053	.051	.049	.051	.052	.049
	.2	.1	.190	.049	.048	.096	.039	.038	.068	.049	.048
	0	0	.207	.052	.052	.135	.050	.054	.122	.046	.049
200	20	10	.044	.041	.042	.053	.050	.050	.052	.049	.048
	14	7	.046	.045	.046	.050	.049	.048	.059	.056	.049
	10	5	.054	.053	.057	.043	.052	.051	.047	.046	.043
	8	4	.054	.053	.060	.052	.051	.049	.053	.050	.051
	6	3	.049	.050	.045	.046	.050	.050	.051	.052	.049
	4	2	.051	.053	.043	.043	.045	.040	.049	.048	.049
	2	1	.040	.045	.046	.053	.049	.048	.053	.051	.049
	.4	.2	.077	.053	.057	.061	.054	.053	.046	.047	.046
	.2	.1	.099	.053	.060	.074	.049	.045	.075	.061	.057
	0	0	.225	.050	.045	.147	.043	.043	.123	.049	.060

Table 2.2b: Empirical Null Rejection Probabilities, 5% Nominal Level, $\epsilon_t \sim I(0)$. 10,000 Replications, b for t_{θ_0}(Robust)is data dependent, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta \neq 2$.

			$b = 0.1$			$b = 0.5$			$b = 1.0$		
T	β_1	β_2	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)
50	20	10	.047	.053	.033	.047	.052	.036	.052	.058	.036
	14	7	.047	.051	.039	.051	.060	.035	.052	.059	.038
	10	5	.046	.049	.038	.044	.048	.033	.051	.054	.039
	8	4	.055	.058	.032	.047	.048	.036	.051	.060	.031
	6	3	.051	.049	.035	.047	.048	.031	.046	.054	.030
	4	2	.051	.056	.039	.050	.053	.034	.047	.051	.039
	2	1	.047	.059	.032	.049	.065	.039	.048	.055	.031
	.4	.2	.034	.067	.030	.032	.055	.038	.031	.060	.031
	.2	.1	.033	.059	.032	.026	.058	.030	.023	.055	.039
	0	0	.028	.053	.035	.027	.065	.035	.025	.061	.038
100	20	10	.047	.048	.035	.048	.051	.038	.047	.052	.035
	14	7	.051	.054	.033	.051	.054	.032	.046	.048	.039
	10	5	.050	.056	.039	.055	.057	.032	.051	.049	.038
	8	4	.051	.055	.030	.049	.051	.038	.053	.057	.039
	6	3	.046	.053	.037	.051	.054	.030	.056	.058	.032
	4	2	.046	.056	.038	.048	.055	.036	.052	.058	.035
	2	1	.045	.055	.028	.040	.047	.039	.044	.054	.039
	.4	.2	.038	.051	.039	.041	.045	.038	.032	.054	.035
	.2	.1	.032	.040	.039	.030	.063	.035	.023	.061	.036
	0	0	.021	.048	.023	.024	.060	.028	.025	.065	.030
200	20	10	.048	.054	.031	.051	.053	.032	.045	.045	.035
	14	7	.046	.056	.038	.049	.049	.039	.048	.050	.030
	10	5	.056	.057	.035	.051	.053	.034	.052	.055	.033
	8	4	.049	.052	.031	.049	.049	.031	.056	.053	.030
	6	3	.050	.055	.033	.049	.051	.032	.055	.056	.031
	4	2	.047	.055	.031	.052	.056	.039	.046	.052	.033
	2	1	.045	.056	.037	.048	.055	.031	.047	.057	.039
	.4	.2	.038	.055	.035	.046	.053	.039	.035	.060	.037
	.2	.1	.019	.054	.039	.045	.061	.033	.025	.048	.039
	0	0	.026	.050	.034	.035	.060	.030	.023	.049	.029

Table 2.3a: Finite Sample Power, 5% Nominal Level, $\epsilon_t \sim I(1)$. $T = 100$.
Two-sided Tests, 10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta = \theta_1$, $\beta_1 = \theta_1 \beta_2$.

		$b(IV, t_{\theta_0}) = 0.1$			$b(IV, t_{\theta_0}) = 0.5$			$b(IV, t_{\theta_0}) = 1.0$		
β_2	θ_1	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)
10	2.000	.045	.052	.052	.051	.051	.052	.051	.052	.050
	2.030	.200	.445	.106	.115	.123	.101	.110	.116	.106
	2.060	.528	.750	.192	.295	.308	.192	.233	.243	.193
	2.090	.876	.971	.293	.484	.506	.262	.377	.388	.263
	2.120	.962	.999	.308	.588	.619	.308	.474	.498	.325
	2.150	.997	1.00	.347	.730	.770	.378	.578	.613	.348
2	2.000	.046	.047	.046	.052	.049	.050	.046	.047	.046
	2.150	.137	.393	.099	.119	.140	.112	.076	.092	.090
	2.300	.432	.784	.192	.229	.275	.187	.193	.249	.183
	2.450	.751	.947	.267	.358	.441	.245	.327	.383	.273
	2.600	.902	.996	.323	.524	.635	.310	.401	.496	.315
	2.750	.984	1.00	.362	.639	.777	.403	.468	.605	.341
1	2.000	.059	.055	.056	.052	.049	.050	.051	.053	.052
	2.300	.092	.375	.098	.086	.118	.091	.104	.146	.115
	2.600	.331	.766	.185	.193	.281	.167	.172	.239	.186
	2.900	.653	.960	.285	.313	.483	.273	.263	.375	.248
	3.200	.804	.994	.318	.403	.637	.308	.328	.494	.324
	3.500	.920	1.00	.347	.494	.747	.381	.386	.604	.356
.4	2.0	.053	.052	.048	.052	.051	.053	.051	.049	.047
	2.75	.027	.406	.088	.053	.135	.106	.057	.123	.098
	3.5	.102	.750	.192	.113	.288	.193	.106	.272	.202
	4.25	.264	.834	.263	.176	.481	.263	.147	.371	.255
	5.00	.373	.963	.325	.207	.622	.325	.208	.518	.330
	5.75	.486	.995	.348	.221	.747	.348	.231	.605	.356
.2	2	.067	.043	.042	.053	.051	.049	.051	.052	.049
	3.5	.050	.190	.088	.016	.104	.090	.023	.097	.106
	5.0	.080	.524	.180	.055	.299	.183	.051	.254	.202
	6.5	.048	.851	.241	.078	.465	.264	.083	.378	.263
	8.0	.108	.956	.323	.097	.650	.323	.089	.479	.300
	9.5	.134	.998	.362	.129	.763	.362	.117	.622	.346
.1	2	.130	.045	.046	.096	.039	.038	.068	.049	.048
	5	.012	.188	.099	.023	.118	.098	.021	.146	.115
	8	.009	.551	.210	.035	.290	.185	.020	.239	.186
	11	.026	.822	.261	.054	.505	.285	.040	.375	.248
	14	.022	.964	.301	.043	.640	.318	.065	.494	.324
	17	.043	.991	.361	.044	.767	.347	.047	.604	.356

Table 2.3b: Finite Sample Power, 5% Nominal Level, $\epsilon_t \sim I(0)$. $T = 100$. Two-sided Tests, 10,000 Replications, $H_0 : \theta = \theta_0 = 2$, $H_1 : \theta = \theta_1$, $\beta_1 = \theta_1\beta_2$.

		$b(IV, t_{\theta_0}) = 0.1$			$b(IV, t_{\theta_0}) = 0.5$			$b(IV, t_{\theta_0}) = 1.0$		
β_2	θ_1	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)	t_{IV}	t_{θ_0}	t_{θ_0}(R)
10	2.0000	.046	.056	.038	.045	.040	.030	.050	.054	.036
	2.0005	.153	.156	.143	.131	.133	.120	.171	.179	.154
	2.0010	.583	.579	.491	.283	.297	.470	.263	.259	.504
	2.0015	.856	.859	.860	.532	.547	.850	.453	.456	.843
	2.0020	.994	.995	.991	.702	.709	.991	.553	.539	.998
	2.0025	1.00	1.00	1.00	,732	.749	1.00	.624	.629	1.00
2	2.0000	.045	.064	.036	.058	.059	.038	.064	.068	.029
	2.0025	.172	.175	.153	.092	.126	.172	.069	.070	.152
	2.0050	.495	.509	.375	.283	.299	.496	.182	.205	.446
	2.0075	.895	.899	.834	.412	.418	.812	.494	.448	.886
	2.0100	.971	.969	.980	.662	.678	1.00	.492	.472	1.00
	2.0125	1.00	1.00	1.00	.802	.822	1.00	.682	.694	1.00
1	2.0000	.046	.053	.030	.046	.050	.032	.056	.050	.029
	2.0050	.187	.189	.123	.124	.137	.179	.108	.089	.120
	2.0100	.511	.534	.452	.223	.241	.513	.192	.215	.459
	2.0150	.901	.916	.882	.342	.363	.774	.405	.457	.887
	2.0200	.980	.985	.971	.621	.689	1.00	.546	.579	.981
	2.0250	.991	.995	.981	.824	.828	1.00	.634	.628	1.00
.4	2.0000	.053	.056	.029	.048	.049	.031	.050	.046	.029
	2.0125	.112	.097	.091	.102	.099	.140	.102	.116	.125
	2.0250	.472	.485	.430	.261	.287	.489	.201	.217	.459
	2.0375	.851	.863	.850	.456	.524	.903	.327	.419	.889
	2.0500	.981	.986	.990	.532	.694	.983	.456	.558	.978
	2.0625	1.00	1.00	1.00	.569	.742	.995	.487	.564	.982
.2	2.0000	.029	.040	.021	.040	.041	.027	.040	.064	.028
	2.0250	.165	.262	.161	.059	.143	.197	.054	.102	.098
	2.0500	.351	.423	.401	.151	.274	.498	.136	.194	.439
	2.0750	.760	.839	.831	.293	.408	.873	.256	.392	.857
	2.1000	.890	1.00	.981	.309	.681	.978	.309	.517	.992
	2.1250	.885	1.00	1.00	.445	.768	1.00	.318	.551	1.00
.1	2.0000	.040	.042	.030	.035	.065	.030	.040	.050	.025
	2.0500	.142	.167	.170	.065	.154	.146	.039	.132	.162
	2.1000	.362	.558	.534	.112	.183	.485	.115	.205	.408
	2.1500	.611	.902	.885	.191	.423	.801	.246	.461	.839
	2.2000	.683	.994	.981	.264	.618	1.00	.258	.497	.990
	2.2500	.725	1.00	1.00	.291	.867	1.00	.249	.671	1.00

Table 2.4: Finite Sample Performance of ADF and ADF-GLS Unit Root Test for $\widehat{\epsilon}_t$.

			Null Rejections	
T	β_1	β_2	DF	DF-GLS
50	20	10	.054	.042
	14	7	.058	.041
	10	5	.053	.061
	8	4	.068	.059
	6	3	.054	.057
	4	2	.044	.045
	2	1	.052	.058
	.4	.2	.051	.052
	.2	.1	.033	.051
	0	0	.055	.060
100	20	10	.049	.053
	14	7	.053	.050
	10	5	.052	.049
	8	4	.055	.068
	6	3	.059	.053
	4	2	.055	.053
	2	1	.052	.056
	.4	.2	.043	.053
	.2	.1	.050	.067
	0	0	.056	.059
500	20	10	.057	.059
	14	7	.060	.059
	10	5	.053	.058
	8	4	.056	.062
	6	3	.039	.041
	4	2	.046	.064
	2	1	.040	.047
	.4	.2	.050	.054
	.2	.1	.047	.051
	0	0	.049	.057

Table 2.5: Finite Sample Power of ADF & ADF-GLS Unit Root Test for $\widehat{\epsilon}_t$. 5% Nominal Level, $T = 100$.

			No Cointegration		Cointegration	
β_1	β_2	ρ	DF	DF-GLS	DF	DF-GLS
20	10	1.00	.049	.053		
		0.98	.059	.058	.071	.062
		0.96	.074	.098	.059	.105
		0.94	.107	.139	.085	.126
		0.92	.142	.201	.089	.139
		0.90	.194	.260	.179	.286
14	7	1.00	.053	.050		
		0.98	.054	.047	.063	.073
		0.96	.079	.101	.081	.065
		0.94	.098	.131	.083	.154
		0.92	.134	.189	.130	.219
		0.90	.202	.280	.187	.285
10	5	1.00	.052	.049		
		0.98	.042	.060	.057	.063
		0.96	.065	.089	.143	.139
		0.94	.105	.143	.139	.190
		0.92	.132	.205	.187	.244
		0.90	.201	.283	.210	.218
8	4	1.00	.055	.068		
		0.98	.064	.078	.049	.058
		0.96	.074	.102	.107	.143
		0.94	.117	.147	.089	.171
		0.92	.143	.194	.152	.159
		0.90	.224	.296	.234	.269
6	3	1.00	.059	.053		
		0.98	.061	.066	.038	.045
		0.96	.086	.109	.054	.063
		0.94	.099	.151	.131	.185
		0.92	.143	.203	.140	.216
		0.90	.178	.257	.192	.286
4	2	1.00	.055	.053		
		0.98	.061	.066	.049	.055
		0.96	.074	.096	.059	.086
		0.94	.099	.153	.103	.155
		0.92	.150	.210	.152	.220
		0.90	.203	.278	.189	.268

3 SPEED OF ECONOMIC CONVERGENCE OF U.S. REGIONS

3.1 Introduction

There is a large literature in economics regarding the economic convergence of countries, regions, etc. in terms of per-capita income. β-convergence occurs when poor economies grow faster than rich ones. The empirical findings on convergence are mixed, e.g. Baumol 1986 finds some evidence in the developed economies. Barro and Sala-i Martin (1990) find evidence of β-convergence in USA, whereas Brown, Coulson and Engle (1990) find no convergence in US states. Quah (1993) does not find any evidence of convergence. Carvalho and Harvey (2005) show that all but the two richest US regions are converging to the average. DeJuan and Tomljanovich (2005) find support for both stochastic convergence (convergence in growth rates) and β-convergence (convergence in levels) on the basis of personal income data for the majority of Canadian provinces, after allowing for a structural break in the data. Koμcenda, Kutan and Yigit (2006) show slow but steady per-capita real income convergence of 10 European Union (EU) members toward EU standards. Cuñado and de Gracia (2006) examine the real convergence hypothesis in 43 African countries (both toward an African average and the U.S. economy). They find convergence both toward the African average and the US economy. Rodríguez (2006) provides some evidence of β-convergence in Canada. Cuñado and de Gracia (2006) find evidence of convergence during the nineties-2003 period for Poland, Czech Republic and Hungary toward Germany and only for Poland toward the US economy. Kutan and Yigit (2007) also provide some evidence in the support of convergence in Eurpoean Union. Galvao Jr and Reis Gomes (2007) investigate the occurrence of per capita income convergence in 19 Latin American countries. Their results indicate that there is substantial evidence in favor of conditional convergence in Latin America. Dawson and Sen (2007) provide evidence that the relative income series of 21 countries are consistent with stochastic convergence, and that β-convergence has occurred in at least 16 countries at some point during the twentieth century. Heckelman (2013) performs convergence tests on the U.S. states for per capita income from 1930 to 2009, and find that about half of the states exhibit stochastic and β-convergence. Ayala, Cu-

nado and Gil-Alana (2013) investigate the real convergence of 17 Latin American countries to the US economy for the period 1950 to 2011. They find real convergence (productivity catch-up) to the US for three Latin American countries: Chile, Costa Rica and Trinidad and Tobago, with these countries also presenting evidence of stochastic and β-convergence.

Where β-convergence is found, it is interesting to obtain a measure of the speed of convergence and estimate it. This chapter develops a simple measure of the speed of convergence that can be expressed as a ratio of two trend slopes. Using the methodology developed in the first two chapters, we estimate the speed of convergence in practice. We apply our approach to U.S. regions and document the speed of convergence for regions that exhibit β-convergence.

The remainder of this chapter is organized as follows: Section 3.2 describes the model. Section 3.3 presents the estimation results and section 3.4 concludes.

3.2 Model

Suppose there are two countries; country 1 is poor and country 2 is rich. The initial incomes of the two countries are Y_{10} and Y_{20} respectively, and $Y_{10} < Y_{20}$. Assume income grows as follows:

$$Y_{1t} = (1+\beta_1)^t Y_{10},$$

$$Y_{2t} = (1+\beta_2)^t Y_{20},$$

$$\beta_2 < \beta_1.$$

The above inequalities suggest that the richer country with higher initial income must have a growth rate lower than that of the poorer country for β-convergence to occur. We now develop a measure of the speed of β-convergence. Suppose incomes of the two countries equalize at date $t = t^*$, then at $t = t^*$,

$$Y^*_{1t} = Y^*_{2t},$$

giving

$$(1 + \beta_1)^{t^*} Y_{10} = (1 + \beta_2)^{t^*} Y_{20}.$$

Solving for t^* gives

$$t^* = \frac{\log (Y_{20}/Y_{10})}{\log ((1 + \beta_1)/(1 + \beta_2))}.$$

Now the speed of convergence, SC, of income of country 1 to income of country 2 is given by

$$SC = \frac{(Y_{20} - Y_{10})}{t^*} = \log \left(\frac{1 + \beta_1}{1 + \beta_2} \right) \frac{(Y_{20} - Y_{10})}{\log (Y_{20}/Y_{10})}.$$

For given initial income levels, the speed of convergence obviously depends on the magnitude of

$$\theta = \frac{1 + \beta_1}{1 + \beta_2},$$

which can be expressed as a ratio of linear trend slopes as follows. Taking the natural log of the expressions for Y_{1t}, Y_{2t}, and denoting $y_{1t} = \log Y_{1t}$ and $y_{2t} = \log Y_{2t}$, we obtain

$$y_{1t} = t \log(1 + \beta_1) + y_{10} \approx y_{10} + \beta_1 t,$$
$$y_{2t} = t \log(1 + \beta_2) + y_{20} \approx y_{20} + \beta_2 t.$$

The estimation equations for y_{1t} and y_{2t} can be written as

$$y_{1t} = y_{10} + \beta_1 t + u_{1t} \equiv \mu_1 + \beta_1 t + u_{1t},$$

$$y_{2t} = y_{20} + \beta_2 t + u_{2t} \equiv \mu_2 + \beta_2 t + u_{2t}.$$

Adding t to both sides of each equation gives

$$y^*_{1t} = y_{1t} + t = \mu_1 + (1 + \beta_1)t + u_{1t},$$

$$y^*_{2t} = y_{2t} + t = \mu_2 + (1 + \beta_2)t + u_{2t}.$$

Given the analysis in Chapter 1 and 2, the regression of y^*_{1t} on y^*_{2t} estimates θ. The regression

$$y^*_{1t} = \delta + \theta y^*_{2t} + \epsilon_t,$$

estimates the parameter θ given by

$$\theta = \frac{1 + \beta_1}{1 + \beta_2}.$$

The speed of convergence is directly proportional to the parameter θ. For given initial income levels, i.e. Y_{10} and Y_{20}, the speed of convergence is greater for higher values of θ and vice versa.

3.3 Estimation Results

We collected annual per capita income series from 1929 – 2013 for eight US regions: Southeast, Southwest, Rocky Mountains, Plains, Great Lakes, Far West, Mideast and New England. According to Carvalho and Harvey (2005), all regions but the Mideast and New England are converging (β-convergence). Figure 1 shows plots of natural log of per capita income of all regions against time. Table 3.1 reports OLS estimated parameters from the regression of each log series (in order from poorest to richest) on an intercept and time trend. Each series was tested for a unit root (around the linear time trend) using the ADF-GLS test. Those results are also given in table 3.1. For the trend slope estimators in Table 3.1, the standard errors have been calculated using the following formula:

$$se(\widehat{\beta}) = \sqrt{\frac{\widehat{\lambda^2}}{\Sigma_{t=1}^{T}(t-\bar{t})^2}}$$

where

$$\bar{t} = T^{-1}\sum_{t=1}^{T} t,$$
$$\widehat{\lambda^2} = \widehat{\gamma}_0 + 2\sum_{j=1}^{T-1} k\left(\frac{j}{M}\right)\widehat{\gamma}_j,$$

and

$$\widehat{\gamma}_j = T^{-1}\Sigma_{t=j+1}^{T}\widehat{\epsilon}_t\widehat{\epsilon}_{t-j}.$$

$\widehat{\epsilon}_t$ are the OLS residuals from the regression of the regions' series on time variable. The fixed-b critical values from Bunzel and Vogelsang (2005) have been used to assess statistical significance. Because of great depression of 1946 and 1973-75 recession, subsamples of data, i.e. post 1946 and post 1973 have been considered separately for analysis. The trend slopes are significant at 95% confidence level. The poorer regions tend to have $\widehat{\beta}'s$ that are bigger than the more wealthy regions which is consistent with β-convergence. The confidence intervals in table 3.1 are reported for the trend slopes using I(1) critical values except for Plains (post 1946) for which I(0) critical value has been used, as we can reject the null of a unit root for the series of Plains (post 1946).

In Table 3.2a, the IV point estimates of θ and the 95% confidence intervals are shown for the following regression equation:

$$y^*_{1t} = \delta + \theta \overline{y}^*_{2t} + \epsilon_t,$$

where

$$y^*_{1t} = y_{1t} + t,$$

for a given region and

$$\overline{y}^*_{2t} = \overline{y}_{2t} + t,$$

where $\overline{y}_{2t}$ is the average across all regions. ADF-GLS statistic is also reported for the residuals. The purpose of carrying out a unit root test for the residuals is to use the right critical value which is different for I(0) and I(1) errors. Although the ADF-GLS statistic values suggest that the residuals (for all but Great Lakes) are integrated of order one, however, the confidence intervals based on both I(0) and I(1) errors have been reported. The null hypothesis of $\theta = 1$, is rejected for Southeast, Great Lakes and Far West, which is a statistical evidence of convergence of these regions to the average income level.

In Table 3.2b, the IV point estimates of θ and the 95% confidence intervals similar to those in Table 3.2a are reported for the subsample of post 1946 period. As in Table 3.2a, for IV residuals, the null of unit root is rejected only for Great Lakes. The null of $\theta = 1$, is rejected for Great Lakes and Far West.

In Table 3.2c, the IV point estimates of θ and the 95% confidence intervals similar to those in Tables 3.2a and b are reported for the subsample of post 1973 period. As in Tables 3.2a and b, for IV residuals, the null of unit root is rejected only for Great Lakes. The null of $\theta = 1$, is rejected for Great Lakes and Far West.

In Table 3.3a, pairwise IV point estimates and the 95% confidence intervals for all the series are reported. The regions in the descending order of initial income are as follows: Mideast, Far West, New England, Great Lakes, Rocky Mountain, Plains, Southwest, Southeast. For Southeast the value of $\widehat{\theta}$ increases from Southwest to Great Lakes, then it falls slightly for New England; it is higher for Far West and decreases for Mideast again. For Southwest again, the estimates increase across the row up to Great Lakes, then the estimate decreases for New England, increases for Far West and decreases for Mideast. This pattern remains the same for each and every row in Table 3.3a, i.e. the estimates increase up to Great Lakes, they all decrease for New England, then increase for Far West and decrease for Mideast. Therefore except for New England and Mideast, as we move across a row, we see increasing $\widehat{\theta}'s$, which is an evidence that convergence is occurring faster, the larger the initial income gap. The estimates are statistically different from one for majority of the estimates in Table 3.3a for I(0) critical values, which is a statistical evidence of β-convergence of regions. For I(1) critical values, the estimated ratio of trend slopes is statistically different from one only for Southeast when regressed against Great Lakes and Far West. There is no evidence of convergence of other regions based on I(1) critical values.

In Table 3.3b, i.e. for post 1946 subsample, the pattern of convergence for I(0) critical values is pretty similar to that in Table 3.3a, however, for I(1) critical values, there is evidence of convergence of only Plains and Far West.

In Table 3.3c, where post 1973 subsample has been used for estimation of θ, the pattern of increasing $\widehat{\theta}'s$ as we move across a row remains the same as that in Table 3.3a except for Southeast when regressed against Southwest and Plains; the estimates decrease from Southwest to Plains instead of increasing. There is no evidence of convergence for any of the regions based on I(1) critical values, whereas based on I(0) critical values, there is evidence of convergence of Southeast to Great Lakes, Plains to Far West, New England to Far West and Mideast, and Far West to Mideast.

3.4 Conclusion

The speed of convergence of two different regions' per capita income has been shown to be proportional to the ratio of trend slopes. This ratio has been estimated for all US regions using the methodology developed in the first two chapters. For all regions, the

IV point estimates and the 95% confidence intervals have been computed. Unit root tests have been applied to the IV residuals, and the results suggest that the residuals (for all but Great Lakes) are integrated of order one. The confidence intervals of θ based on both I(0) and I(1) errors have been reported to compare their performance based on the type of noise. The model suggests that the speed of convergence is higher for the cases where the estimated ratio of trend slopes is larger in magnitude, all else the same. The estimates suggest that for U.S. regions, the convergence is occurring faster, the larger the initial income gap.

APPENDIX

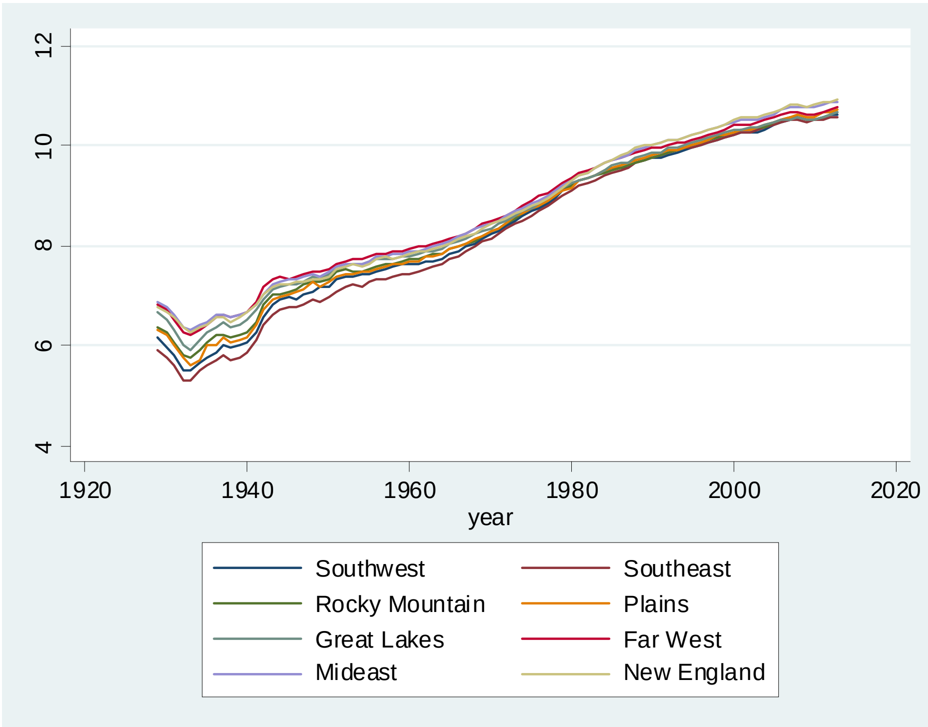

Figure 1: Natural Log of Per Capita Income of US Regions
Data Series: 1929-2013

Table 3.1: Regression Results for Regions against Time, and ADF-GLS for Regions, $b = 0.1$, Critical value (10.97500) for I(1) errors, except for Plains (Post 1946) with $b = 0.25$, Critical value (4.2027536). The regression is $y_t = \mu + \beta t + u_t$.

	$\widehat{\beta}$	DF-GLS	$\widehat{\beta}$ (P.1946)	DF-GLS	$\widehat{\beta}$ (P.1973)	DF-GLS
$S.east$	0.0664*	-1.61	0.0636*	-1.64	0.0529*	-0.64
	(0.0023)		(0.0026)		(0.0065)	
	[.041, .091]		[.035, .092]		[−.019, .124]	
$S.west$	0.0631*	-2.42	0.0601*	-1.66	0.0507*	-0.93
	(0.0023)		(0.0026)		(0.0061)	
	[.038, .088]		[.031, .089]		[−.016, .118]	
$Plains$	0.0621*	-2.13	0.0636*	-3.91#	0.0516*	-0.74
	(0.0021)		(0.0027)		(0.0053)	
	[.039, .085]		[.055, .072]		[−.006, .109]	
$R.Mnts$	0.0603*	-2.05	0.0584*	-2.09	0.0511*	-0.57
	(0.0019)		(0.0021)		(0.0047)	
	[.039, .082]		[.035, .082]		[−.001, .103]	
$G.Lakes$	0.0579*	-1.52	0.0568*	-2.57	0.0494*	-1.18
	(0.0019)		(0.0019)		(0.0045)	
	[.037, .079]		[.035, .079]		[−.000, .099]	
$N.Eng$	0.0596*	-1.26	0.0614*	-2.09	0.0559*	-0.77
	(0.0018)		(0.0019)		(0.0028)	
	[.039, .079]		[.040, .082]		[.025, .087]	
$F.West$	0.0567*	-1.48	0.0561*	-2.36	0.0486*	-0.77
	(0.0018)		(0.0018)		(0.0043)	
	[.037, .076]		[.036, .076]		[.002, .096]	
$Mideast$	0.0581*	-1.13	0.0594*	-2.35	0.0535*	-0.37
	(0.0018)		(0.0018)		(0.0031)	
	[.039, .078]		[.039, .079]		[.019, .087]	

* significant at 5% level, # Unit root rejected at 5% level.

Table 3.2a: IV Point Estimates and the 95% Confidence Intervals. $t = Time$
$y^*_{it} = y_{it} + t$, $y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	I(0) Errors ($b = 0.25$)	I(1) Errors ($b = 0.1$)	ADF-GLS for Residuals
	Average	*Average*	
y^*_{1t}	1.006011*	1.006011	-0.793
	$[1.0016623, 1.0103534]$	$[.99753952, 1.0143694]$	
y^*_{2t}	1.002954	1.002954	-1.563
	$[0.99855986, 1.0073596]$	$[.99361615, 1.0121836]$	
y^*_{3t}	1.001926	1.001926	-1.775
	$[0.99925281, 1.0046108]$	$[.99625521, 1.0075406]$	
y^*_{4t}	1.000278	1.000278	-1.412
	$[0.99819446, 1.0023757]$	$[.99539561, 1.0051402]$	
y^*_{5t}	0.9980236*	0.9980236	-4.012
	$[0.99728663, .99876237]$	$[.99591106, 1.0001218]$	
y^*_{6t}	0.9995676	0.9995676	-1.179
	$[0.99481678, 1.0043051]$	$[.98979901, 1.0094321]$	
y^*_{7t}	0.996802*	0.996802*	-2.834
	$[0.99657843, .99702737]$	$[.99510981, .9985193]$	
y^*_{8t}	0.9981557	0.9981557	-1.093
	$[0.99453362, 1.0017706]$	$[.99055351, 1.0058457]$	

* $H_0 : \theta = 1$ is rejected at 5% level.

Table 3.2b: IV Point Estimates and the 95% Confidence Intervals (Post 1946).
$y^*_{it} = y_{it} + t$, $y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	I(0) Errors ($b = 0.25$)	I(1) Errors ($b = 0.1$)	ADF-GLS for Residuals
	Average	*Average*	
y^*_{1t}	1.00403	1.00403	-0.804
	$[0.99793392, 1.0101014]$	$[.99236877, 1.0155019]$	
y^*_{2t}	1.000696	1.000696	-1.576
	$[0.99434166, 1.0070414]$	$[.987664, 1.0135328]$	
y^*_{3t}	1.000447	1.000447	-1.789
	$[0.99632049, 1.0045694]$	$[.99208869, 1.0086889]$	
y^*_{4t}	0.9991468	0.9991468	-1.420
	$[0.99614358, 1.0021536]$	$[.99254768, 1.0056822]$	
y^*_{5t}	0.9976508*	0.9976508	-4.009
	$[0.99658115, .99871865]$	$[.99504994, 1.0002227]$	
y^*_{6t}	1.001998	1.001998	-1.179
	$[0.99508796, 1.0089175]$	$[.98821934, 1.0159674]$	
y^*_{7t}	0.9969806*	0.9969806*	-2.833
	$[0.99643026, 0.99753492]$	$[.99509792, .99889623]$	
y^*_{8t}	1.000052	1.000052	-1.094
	$[0.99477325, 1.0053404]$	$[.98940668, 1.0108576]$	

* $H_0 : \theta = 1$ is rejected at 5% level.

Table 3.2c: IV Point Estimates and the 95% Confidence Intervals (Post 1973).
$y^*_{it} = y_{it} + t$, $y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	I(0) Errors ($b = 0.25$)	I(1) Errors ($b = 0.1$)	ADF-GLS for Rersiduals
	Average	*Average*	
y^*_{1t}	1.001059	1.001059	-0.821
	$[0.9873747, 1.014141]$	$[.97619774, 1.0239544]$	
y^*_{2t}	0.9989959	0.9989959	-1.586
	$[0.98827432, 1.0092778]$	$[.97843847, 1.0179989]$	
y^*_{3t}	0.9998583	0.9998583	-1.794
	$[0.99418384, 1.0053132]$	$[.98877025, 1.0101407]$	
y^*_{4t}	0.9993285	0.9993285	-1.419
	$[0.99657007, 1.0020165]$	$[.99302216, 1.0053122]$	
y^*_{5t}	0.9977157*	0.9977157	-4.009
	$[0.99674732, .99865932]$	$[.99521983, 1.0001008]$	
y^*_{6t}	1.003979	1.003979	-1.180
	$[0.99245282, 1.0160021]$	$[.98289965, 1.0267798]$	
y^*_{7t}	0.9969747*	0.9969747*	-2.833
	$[0.99642172, .99755378]$	$[.99509723, .99894963]$	
y^*_{8t}	1.001634	1.001634	-1.095
	$[0.99264083, 1.0110192]$	$[.98514176, 1.0194875]$	

* $H_0 : \theta = 1$ is rejected at 5% level.

Table 3.3a: Pairwise IV Point Estimates and the 95% Confidence Intervals.
Top Row: I(1) Errors ($b = 0.1$), Bottom Row: I(0) Errors ($b = 0.25$).
$y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$,
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	y^*_{2t}	y^*_{3t}	y^*_{4t}	y^*_{5t}	y^*_{6t}	y^*_{7t}	y^*_{8t}
y^*_{1t}	1.003$^{\#}$ [.9987, 1.014] [1.002, 1.004]	1.004$^{\#}$ [.9997, 1.008] [1.002, 1.006]	1.006$^{\#}$ [.9992, 1.01] [1.002, 1.01]	1.008$^{*\#}$ [1.001, 1.015] [1.004, 1.012]	1.006 [.989, 1.024] [.997, 1.015]	1.009$^{*\#}$ [1.000, 1.018] [1.005, 1.014]	1.008 [.992, 1.0237] [.999, 1.016]
y^*_{2t}		1.001 [.9967, 1.005] [.9993, 1.003]	1.0027 [.9969, 1.01] [.9999, 1.01]	1.005$^{\#}$ [.9965, 1.013] [1.001, 1.009]	1.003 [.984, 1.022] [.994, 1.013]	1.006$^{\#}$ [.9965, 1.0157] [1.002, 1.011]	1.005 [.988, 1.022] [.997, 1.013]
y^*_{3t}			1.002$^{\#}$ [.9985, 1.005] [1.0004, 1.003]	1.004$^{\#}$ [.9992, 1.008] [1.002, 1.006]	1.002 [.987, 1.018] [.995, 1.009]	1.005$^{\#}$ [.9988, 1.011] [1.002, 1.008]	1.004 [.991, 1.017] [.997, 1.0101]
y^*_{4t}				1.002$^{\#}$ [.9984, 1.006] [1.001, 1.004]	1.001 [.986, 1.015] [.994, 1.007]	1.003$^{\#}$ [.9983, 1.009] [1.001, 1.006]	1.002 [.989, 1.014] [.997, 1.008]
y^*_{5t}					.9984 [.987, 1.009] [.993, 1.004]	1.001$^{\#}$ [.9979, 1.004] [1.000, 1.002]	.9999 [.991, 1.0091] [.996, 1.0041]
y^*_{6t}						1.003 [.9929, 1.013] [.9981, 1.007]	1.001$^{\#}$ [.9988, 1.004] [1.0002, 1.003]
y^*_{7t}							.9986 [.9909, 1.006] [.9951, 1.002]

* $H_0 : \theta = 1$ is rejected at 5% level for I(1), $^{\#}$ $H_0 : \theta = 1$ is rejected at 5% level for I(0).

Table 3.3b: Pairwise IV Point Estimates and the 95% Confidence Intervals (Post 1946).
Top Row: I(1) Errors ($b = 0.1$), Bottom Row: I(0) Errors ($b = 0.25$).
$y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$,
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	y^*_{2t}	y^*_{3t}	y^*_{4t}	y^*_{5t}	y^*_{6t}	y^*_{7t}	y^*_{8t}
y^*_{1t}	1.003# [.9970, 1.01] [1.001, 1.01]	1.003# [.9977, 1.009] [1.001, 1.006]	1.005# [.9966, 1.013] [1.001, 1.009]	1.006# [.9999, 1.013] [1.003, 1.010]	1.002 [.992, 1.012] [.997, 1.007]	1.007*# [1.000, 1.014] [1.004, 1.010]	1.004 [.995, 1.0127] [.999, 1.009]
y^*_{2t}		1.0002 [.9958, 1.005] [.9989, 1.001]	1.001 [.9963, 1.007] [.9989, 1.004]	1.003# [.9969, 1.009] [1.000, 1.006]	.9987 [.989, 1.008] [.994, 1.003]	1.004# [.9992, 1.008] [1.002, 1.006]	1.001 [.992, 1.009] [.997, 1.004]
y^*_{3t}			1.001# [.9978, 1.005] [1.000, 1.002]	1.003# [.9995, 1.006] [1.001, 1.004]	.9984 [.992, 1.005] [.995, 1.002]	1.003*# [1.001, 1.006] [1.003, 1.004]	1.000 [.995, 1.005] [.998, 1.003]
y^*_{4t}				1.001# [.9976, 1.005] [1.000, 1.003]	.9971 [.989, 1.004] [.995, .9994]	1.002# [.9986, 1.006] [1.001, 1.003]	.9991 [.993, 1.005] [.998, 1.000]
y^*_{5t}					.9957 [.991, 1.001] [.994, .9975]	1.001 [.9971, 1.004] [.9995, 1.002]	.9976 [.993, 1.002] [.996, .9988]
y^*_{6t}						1.005# [.9993, 1.011] [1.002, 1.008]	1.002# [.9993, 1.004] [1.0009, 1.003]
y^*_{7t}							.9969# [.9925, 1.001] [.9951, .9987]

* $H_0 : \theta = 1$ is rejected at 5% level for I(1), # $H_0 : \theta = 1$ is rejected at 5% level for I(0).

Table 3.3c: Pairwise IV Point Estimates and the 95% Confidence Intervals (Post 1973).
Top Row: I(1) Errors ($b = 0.1$), Bottom Row: I(0) Errors ($b = 0.25$).
$y_{1t} = Southeast$, $y_{2t} = Southwest$, $y_{3t} = Plains$, $y_{4t} = RockyMountains$,
$y_{5t} = GreatLakes$, $y_{6t} = NewEngland$, $y_{7t} = FarWest$, $y_{8t} = Mideast$.

	y^*_{2t}	y^*_{3t}	y^*_{4t}	y^*_{5t}	y^*_{6t}	y^*_{7t}	y^*_{8t}
y^*_{1t}	1.002 [.9905, 1.01] [.9959, 1.01]	1.001 [.994, 1.008] [.997, 1.005]	1.002 [.9918, 1.011] [.9962, 1.007]	1.003# [.9989, 1.008] [1.002, 1.005]	.9971 [.992, 1.003] [.995, .9990]	1.004 [.9968, 1.011] [.9999, 1.008]	.9994 [.995, 1.004] [.998, 1.001]
y^*_{2t}		.9991 [.992, 1.006] [.996, 1.002]	.9997 [.9923, 1.007] [.9957, 1.003]	1.001 [.9899, 1.013] [.9953, 1.007]	.9950 [.982, 1.009] [.988, 1.002]	1.002 [.9956, 1.008] [.9999, 1.004]	.9974 [.986, 1.009] [.992, 1.003]
y^*_{3t}			1.0005 [.9945, 1.006] [.9976, 1.003]	1.002 [.9962, 1.008] [.9989, 1.005]	.9959 [.986, 1.006] [.991, 1.001]	1.003# [.9981, 1.008] [1.001, 1.004]	.9982 [.992, 1.005] [.995, 1.001]
y^*_{4t}				1.002 [.9935, 1.010] [.9972, 1.006]	.9954 [.982, 1.009] [.988, 1.002]	1.002 [.9958, 1.009] [.9992, 1.005]	.9977 [.987, 1.009] [.992, 1.004]
y^*_{5t}					.9938 [.986, 1.002] [.990, .9974]	1.001 [.9933, 1.008] [.9967, 1.005]	.9961 [.989, 1.003] [.993, .9991]
y^*_{6t}						1.007# [.9976, 1.016] [1.002, 1.012]	1.002# [.9972, 1.007] [1.0005, 1.004]
y^*_{7t}							.9953# [.9882, 1.003] [.9916, .9991]

* $H_0 : \theta = 1$ is rejected at 5% level for I(1), # $H_0 : \theta = 1$ is rejected at 5% level for I(0).

REFERENCES

REFERENCES

Ayala, A., Cunado, J. and Gil-Alana, L. A.: (2013), Real convergence: empirical evidence for latin america, Applied Economics 45(22), 3220–3229.

Barro, R. J. and Sala-i Martin, X.: (1990), Economic growth and convergence across the united states, Technical report, National Bureau of Economic Research.

Baumol, W. J.: 1986, Productivity growth, convergence, and welfare: what the long-run data show, The American Economic Review pp. 1072–1085.

Brown, S. J., Coulson, N. E. and Engle, R. F.: (1990), Non-cointegration and econometric evaluation of models of regional shift and share, Technical report, National Bureau of Economic Research.

Bunzel, H. and Vogelsang, T. J.: (2005), Powerful trend function tests that are robust to strong serial correlation with an application to the prebisch-singer hypothesis, Journal of Business and Economic Statistics 23, 381–394.

Canjels, E. and Watson, M. W.: (1997), Estimating deterministic trends in the presence of serially correlated errors, Review of Economics and Statistics May, 184–200.

Carvalho, V. M. and Harvey, A. C.: (2005), Growth, cycles and convergence in us regional time series, International Journal of Forecasting 21(4), 667–686.

Cuñado, J. and de Gracia, F. P.: (2006), Real convergence in africa in the second-half of the 20th century, Journal of Economics and Business 58(2), 153–167.

Dawson, J. W. and Sen, A.: (2007), New evidence on the convergence of international income from a group of 29 countries, Empirical Economics 33(2), 199–230.

DeJuan, J. and Tomljanovich, M.: (2005), Income convergence across canadian provinces in the 20th century: Almost but not quite there, The Annals of Regional Science 39(3), 567–592.

Elliott, G., Rothenberg, T. J. and Stock, J. H.: (1996), Efficient tests for an autoregressive unit root.

Evans, P.: (1997), How fast do economies converge?, Review of Economics and Statistics 79(2), 219–225.

Fagerberg, J.: (1994), Technology and international differences in growth rates, Journal of Economic Literature 32, 1147–75.

Galvao Jr, A. and Reis Gomes, F.: (2007), Convergence or divergence in latin america? a time series analysis, Applied Economics 39(11), 1353–1360.

Hamilton, J. D.: (1994), Time Series Analysis, Princeton University Press, Princeton, N.J.

Harvey, D., Leybourne, S. J. and Taylor, A. M. R.: (2009), Simple, robust and powerful tests of the breaking trend hypothesis, Econometric Theory 25, 995–1029.

Heckelman, J. C.: (2013), Income convergence among us states: cross-sectional and time series evidence, Canadian Journal of Economics/Revue canadienne d'économique 46(3), 1085–1109.

Kiefer, N. M. and Vogelsang, T. J.: (2005), A new asymptotic theory for heteroskedasticity autocorrelation robust tests, Econometric Theory 21, 1130–1164.

Klotzbach, P. J., Pielke, R. A., Christy, J. R. and McNider, R. T.: (2009), An alternative explanation for differential temperature trends at the surface and in the lower troposphere, Journal of Geophysical Research: Atmospheres (1984–2012) 114(D21).

Koμcenda, E., Kutan, A. M. and Yigit, T. M.: (2006), Pilgrims to the eurozone: How far, how fast?, Economic Systems 30(4), 311–327.

Kutan, A. M. and Yigit, T. M.: (2007), European integration, productivity growth and real convergence, European Economic Review 51(6), 1370–1395.

Perron, P. and Yabu, T.: (2009), Estimating deterministic trends with an integrated or stationary noise component, Journal of Econometrics 151(1), 56–69.

Phillips, P. C.: (1986), Understanding spurious regressions in econometrics, Journal of econometrics 33(3), 311–340.

Quah, D.: (1993), Galton's fallacy and tests of the convergence hypothesis, The Scandinavian Journal of Economics pp. 427–443.

Rodríguez, G.: (2006), The role of the interprovincial transfers in the β-convergence process: further empirical evidence for canada, Journal of Economic Studies 33(1), 12–29.

Santer, B., Wigley, T., Mears, C., Wentz, F., Klein, S., Seidel, D., Taylor, K., Thorne, P., Wehner, M., Gleckler, P. et al.: (2005), Amplification of surface temperature trends and variability in the tropical atmosphere, Science 309(5740), 1551–1556.

Staiger, D. and Stock, J. H.: (1997), Instrumental variables regression with weak instruments, Econometrica 65, 557–586.

Tomljanovich, M. and Vogelsang, T. J.: (2002), Are u.s. regions converging? using new econometric methods to examine old issues, Empirical Economics 27, 49–62.

Vogelsang, T. J.: (1998), Trend function hypothesis testing in the presence of serial correlation correlation parameters, Econometrica 66, 123–148.

Vogelsang, T. J. and Franses, P. H.: (2005), Testing for common deterministic trend slopes, Journal of Econometrics 126, 1–24.