

KEIO/KYOTO JOINT
GLOBAL CENTER OF EXCELLENCE PROGRAM
Raising Market Quality-Integrated Design of “Market Infrastructure”

KEIO/KYOTO GLOBAL COE DISCUSSION PAPER SERIES

DP2011-024

Testing for Trend in the Presence of Autoregressive
Error: A Comment

Pierre Perron*
Tomoyoshi Yabu**

Abstract

Roy, Falk and Fuller (2004) presented a procedure aimed at providing a test for the value of the slope of a trend function that has (nearly) controlled size in autoregressive models whether the noise component is stationary or has a unit root. In this note, we document errors in both their theoretical results and the simulations they reported. Once these are corrected for, their procedure delivers a test that has very liberal size in the case with a unit root so that the stated goal is not achieved. Interestingly, the mistakes in the code used to generate the simulated results (which is the basis for the evidence about the reliability of the method) are such that what they report is essentially equivalent to the size and power of the test proposed by Perron and Yabu (2009), which was shown to have the standard Normal distribution whether the noise is stationary or has a unit root.

*Pierre Perron

Department of Economics, Boston University

**Tomoyoshi Yabu

Faculty of Business and Commerce, Keio University

KEIO/KYOTO JOINT GLOBAL COE PROGRAM

Raising Market Quality-Integrated Design of “Market Infrastructure”

Graduate School of Economics and Graduate School of Business and Commerce,
Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan

Institute of Economic Research,
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Testing for Trend in the Presence of Autoregressive Error: A Comment*

Pierre Perron[†] Tomoyoshi Yabu[‡]
Boston University Keio University

October 17, 2011

Abstract

Roy, Falk and Fuller (2004) presented a procedure aimed at providing a test for the value of the slope of a trend function that has (nearly) controlled size in autoregressive models whether the noise component is stationary or has a unit root. In this note, we document errors in both their theoretical results and the simulations they reported. Once these are corrected for, their procedure delivers a test that has very liberal size in the case with a unit root so that the stated goal is not achieved. Interestingly, the mistakes in the code used to generate the simulated results (which is the basis for the evidence about the reliability of the method) are such that what they report is essentially equivalent to the size and power of the test proposed by Perron and Yabu (2009), which was shown to have the standard Normal distribution whether the noise is stationary or has a unit root.

JEL Classification Number: C22.

Keywords: linear trend, unit root, median-unbiased estimates, GLS procedure, super efficient estimates.

*We are especially grateful to Anindya Roy for providing us the code used to generate the simulations reported in Roy, Falk and Fuller (2004). We also thank Chih-nan Chen and Daisuke Nagakura for useful comments.

[†]Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).

[‡]Faculty of Business and Commerce, Keio University, Tokyo 108-8345, Japan (tomoyoshi.yabu@gmail.com)

1 Introduction

Roy, Falk and Fuller (2004) presented a procedure aimed at providing a test for the value of the slope of a trend function that has (nearly) controlled size in autoregressive model, whether the noise component is stationary or has a unit root. In this note, we document errors in both their theoretical results and the simulations they reported. Once these are corrected for, their procedure delivers a test that has very liberal size in the case with a unit root so that the stated goal is not achieved. Interestingly, the mistakes in the code used to generate the simulated results (which is the basis for the evidence about the reliability of the method) are such that what they report is essentially equivalent to the size and power of the test proposed by Perron and Yabu (2009), which was shown to have the standard Normal distribution whether the noise is stationary or has a unit root.

The structure of this note is as follows. Section 2 presents the model and the test statistic proposed by Roy et al. (2004). Section 3 provides the correct limit distribution for the case with a unit root. Section 4 discusses the discrepancies from the suggested procedure in the code they used to generate the finite sample results and how these made the simulated procedure very close to that of Perron and Yabu (2009) for the sample size used.

2 The Model and Test Statistic

Consider the following model for the time series process $\{Y_t\}$:

$$\begin{aligned} Y_t &= \mu + \beta t + y_t \\ y_t &= \alpha y_{t-1} + e_t \end{aligned} \tag{1}$$

where $e_t \sim i.i.d.(0, \sigma^2)$ and $\mu = Y_1 - \beta$ when $\alpha = 1$. The parameter β is of primary interest. The Gauss-Newton estimator of (μ, α, β) is obtained by expanding the regression:

$$Y_1 = \mu + \beta + y_1$$

and

$$Y_t = \mu(1 - \alpha) + \beta[t - \alpha(t - 1)] + \alpha Y_{t-1} + e_t$$

for $t = 2, \dots, T$, in a first-order Taylor expansion about an initial estimator $(\ddot{\mu}, \ddot{\beta}, \ddot{\alpha})$. They recommend for $\ddot{\alpha}$ the Roy-Fuller (2001) estimator with $\tau_{pct} = -2.94$, which corresponds to the value of τ_{pct} that is approximately the 90th percentile of the limiting distribution of the t-statistic of the Weighted Symmetric Least-Squares estimate, $\hat{\tau}_{WSLS}$, when $\alpha = 1$; $\ddot{\beta}$ is the Feasible Generalized Least Squares (FGLS) estimator of β using $\hat{\alpha}$ the Roy-Fuller (2001) estimator with $\tau_{pct} = -1.96$, which corresponds to the value of τ_{pct} that is the median of the limiting distribution of $\hat{\tau}_{WSLS}$ when $\alpha = 1$. Then the Gauss-Newton regression equations, with parameters $\Delta\ddot{\mu}$, $\Delta\ddot{\beta}$, and $\Delta\ddot{\alpha}$, are ¹:

$$\begin{aligned} \tilde{e}_1 &= \begin{cases} \Delta\ddot{\mu}[(1 - \ddot{\alpha}^2)^{1/2}] + \Delta\ddot{\beta}[(1 - \ddot{\alpha}^2)^{1/2}] + v_1, & t = 1 \text{ and } \ddot{\alpha}^2 < 1 \\ \Delta\ddot{\mu} + \Delta\ddot{\beta} + v_1, & t = 1 \text{ and } \ddot{\alpha} = 1 \end{cases} \\ \tilde{e}_t &= \Delta\ddot{\mu}(1 - \ddot{\alpha}) + \Delta\ddot{\beta}[t - \ddot{\alpha}(t - 1)] + \Delta\ddot{\alpha}\ddot{y}_{t-1} + v_t \end{aligned} \quad (2)$$

for $t = 2, \dots, T$, where ²

$$\begin{aligned} \ddot{y}_t &= Y_t - \ddot{\mu} - \ddot{\beta}t, \quad t = 1, \dots, T \\ \tilde{e}_1 &= \begin{cases} (1 - \ddot{\alpha}^2)^{1/2}\ddot{y}_1, & t = 1 \text{ and } \ddot{\alpha}^2 < 1 \\ \ddot{y}_1, & t = 1 \text{ and } \ddot{\alpha} = 1 \end{cases} \end{aligned} \quad (3)$$

and $\tilde{e}_t = \ddot{y}_t - \ddot{\alpha}\ddot{y}_{t-1}$, for $t = 2, \dots, T$. The one step Gauss-Newton estimator is obtained as

$$(\hat{\mu}_{GN}, \hat{\beta}_{GN}, \hat{\alpha}_{GN}) = (\ddot{\mu}, \ddot{\beta}, \ddot{\alpha}) + (\Delta\hat{\mu}, \Delta\hat{\beta}, \Delta\hat{\alpha}),$$

where $(\Delta\hat{\mu}, \Delta\hat{\beta}, \Delta\hat{\alpha})$ are the least squares coefficients for the least squares regression associated with (2). Then the Gauss-Newton test statistic is

$$t_{GN}(\hat{\beta}) = \frac{\hat{\beta}_{GN} - \beta}{se_{GN}(\hat{\beta})} \quad (4)$$

¹Roy et al. (1994) omitted the term $\Delta\ddot{\beta}$ for $t = 1$ and $\ddot{\alpha} = 1$.

²The expression is corrected for a typo in Roy et al. (1994) where $(1 - \ddot{\alpha}^2)^{-1/2}$ appears instead of $(1 - \ddot{\alpha}^2)^{1/2}$.

where they recommend the use of the following standard error

$$se_{GN}(\hat{\beta}) = \min\{s(T-1)^{-1/2}, s\nu_{\beta}^{-1/2}\} \quad (5)$$

where ν_{β} is the (2,2)th element of the inverse of the sums-of-squares-and-products matrix from the Gauss-Newton regressions and s^2 is the residual mean square.

3 The limit distribution

The corrected limit distribution of the Gauss-Newton statistic is presented in the next Theorem for the case $\alpha = 1$ (when $|\alpha| < 1$, all results stated in Roy et al. (2004) are correct).

Theorem 1 *Let $\ddot{\alpha}$ be the initial estimator of α in the Gauss-Newton procedure. Let $\hat{\alpha}$ be the estimator of α used in the FGLS. Assume that $T(\hat{\alpha} - 1) \rightarrow^L \Pi$ if $\alpha = 1$ and $T(\ddot{\alpha} - 1) \rightarrow^L \ddot{\Pi}$ if $\alpha = 1$. If $\alpha = 1$, then $t_{GN}(\hat{\beta}) \rightarrow^L v_{\beta}^{-1/2}\zeta$, where $\zeta = \ddot{\zeta} + \kappa$ with $\ddot{\zeta} = 1(\Pi < 0)\xi + 1(\Pi = 0)W(1)$, ξ defined by (A.3), κ is the second element of the vector $V_{(2)}^{-1}F$ and $v_{\beta} = \min(v_{(2)}^{(22)}, 1)$ where $v_{(2)}^{(22)}$ is the (2,2)th element of $V_{(2)}^{-1}$, as defined in (A.4). As usual, $W(\cdot)$ stands for a Wiener process defined on $[0, 1]$.*

To complete the limit distribution of $t_{GN}(\hat{\beta})$, it remains to specify the limit distributions Π and $\ddot{\Pi}$. These are stated in equation (A.5) of the appendix with $\tau_{pct} = -1.96$ for Π and $\tau_{pct} = -2.94$ for $\ddot{\Pi}$.

To verify the correctness of our results, we evaluated the quantiles and the probability that the stated statistic be greater in absolute value than 1.98 for the following three cases: a) the finite sample distribution of $t_{GN}(\hat{\beta})$; b) the limited distribution of $t_{GN}(\hat{\beta})$ as stated in our Theorem 1 and c) the limit distribution of $t_{GN}(\hat{\beta})$ as stated in Roy et al. (1994). For the latter two we used partial sums of *i.i.d.* $N(0, 1)$ random variables to approximate the Wiener processes and the integrals were evaluated with 2,000 steps. In all cases, 10,000 replications were performed to obtain the stated quantities. The results are presented in

Table 1. Several interesting features emerge from these results. First, our limit distribution is close to the exact finite sample distribution indicating that our results are indeed correct. On the other hand, the limit distribution as stated in Roy et al. (2004) is far from the exact distribution which suggest that it is incorrect. More importantly, the size of the test given by the probability that $|t_{GN}(\hat{\beta})| > 1.98$ is about 30%, very far from the 5% nominal size. This indicates that, contrary to the claim made by Roy et al. (2004), the statistic $t_{GN}(\hat{\beta})$ has limit (and finite sample) distributions that are very different when $\alpha = 1$ and $|\alpha| < 1$. When $\alpha = 1$ the size of the test is very liberal. It remains to understand why the results stated in their Table 2 supported an opposite conclusion to the effect that the size of the test statistic $t_{GN}(\hat{\beta})$ was close to 5% even when $\alpha = 1$. This is addressed in the next section.

4 Sources of the errors in the code used

In the code used by Roy et al. (2004), there are two important departures from the recommended procedure. They all apply to the case $\alpha = 1$ only. The first is that the standard error is set to $s\nu_{\beta}^{-1/2}$ instead of $\min\{s(T-1)^{-1/2}, s\nu_{\beta}^{-1/2}\}$ as stated in (5). We again simulated the limit and finite sample distributions of the test with this change. The results are reported in Table 2. The size distortions are considerably reduced but are still high, with a 16% rejection probability instead of 5%. The second departure is the fact that the test is constructed as $t_{GN}(\hat{\beta}) = (\ddot{\beta} - \beta)/se_{GN}(\hat{\beta})$ so that instead of the Gauss-Newton estimate $\hat{\beta}_{GN}$, they used the FGLS estimate $\ddot{\beta}$ based on the Roy and Fuller (2001) estimate $\hat{\alpha}$ with $\tau_{pct} = -1.96$. This makes the statistic essentially equivalent to the FGLS test since $se_{GN}(\hat{\beta})$ is almost equivalent to $se(\ddot{\beta})$, the standard error of $\ddot{\beta}$. The other mistake was minor: in (3) they defined $\ddot{y}_t = Y_t - \ddot{\beta}t$ instead of $\ddot{y}_t = Y_t - \ddot{\mu} - \ddot{\beta}t$ for $t = 2, \dots, T$. But because of the second discrepancy, this only affects the standard error $se_{GN}(\hat{\beta})$. Table 2 reports the distribution and size of the test with this correction done. As can be seen the results are similar to those they reported (also reproduced). In summary, what they simulate is essentially equivalent to

using the FGLS estimate of α that corresponds to the Roy and Fuller (2001) estimate $\ddot{\alpha}$ with $\tau_{pct} = -2.94$ and the FGLS estimate $\ddot{\beta}$ constructed with the Roy and Fuller (2001) estimate $\hat{\alpha}$ with $\tau_{pct} = -1.96$. Interestingly, it turns out that this procedure is almost identical, at least for $T = 100$ (the sample size used in their simulations), to the procedure of Perron and Yabu (2009).

The approach of Perron and Yabu (2009) is based on a Quasi Feasible Generalized Least Squares approach that uses a superefficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. The estimate of α is the OLS estimate obtained from an autoregression applied to detrended data and is truncated to take a value 1 when the estimate is in a $dT^{-\delta}$ neighborhood of 1 for some $0 < \delta < 1$. This makes the estimate “super-efficient” when $\alpha = 1$ and implies that inference on the slope parameter can be performed using the standard Normal distribution whether $\alpha = 1$ or $|\alpha| < 1$. Theoretical arguments showed that $\delta = 1/2$ is the appropriate choice and based on simulations they recommended using $d = 1$. To improve the finite sample properties of the test, they used the Roy and Fuller (2001) estimate $\hat{\alpha}$ with a value $\tau_{pct} = -1.96$. Hence, the main difference between the procedure of Perron and Yabu (2009) and the one that results from the output of their code is the decision rule to set $\alpha = 1$. It turns out that for $T = 100$, they are essentially the same. Table 3 presents simulation results about the probabilities that $|\hat{\alpha}(\tau_{pct} = -1.96) - 1| < T^{-1/2}$ and $\hat{\tau}_{WSLS} > -2.94$. We again use 10,000 replications. The results show that these probabilities are almost identical when $T = 100$ (they are, however, different for other sample sizes). Hence, their simulations (apart from minor discrepancies) provide evidence about the adequacy of the approach suggested by Perron and Yabu (2009). More details about the validity of this procedure and its superiority over other methods in more general contexts can be found in their paper.

References

- Perron, P., and Yabu, T. (2009), “Estimating Deterministic Trends with an Integrated or Stationary Noise Component,” *Journal of Econometrics*, 151, 56-69.
- Roy, A., Falk, B., and Fuller, W.A. (2004), “Testing for Trend in the Presence of Autoregressive Error,” *Journal of the American Statistical Association*, 99, 1082-1091.
- Roy, A., and Fuller W.A. (2001), “Estimation for Autoregressive Processes with a Root Near One,” *Journal of Business & Economic Statistics*, 19, 482-493.

Appendix

Proof of Theorem 1: Without loss of generality, we assume that $\sigma^2 = 1$. First, we derive the limiting distribution of the FGLS estimator of $(\mu, \beta)'$ constructed using an estimator, $\hat{\alpha}$, of α , such that $T(\hat{\alpha} - 1) \xrightarrow{L} \Pi$ where Π is some nondegenerate limiting distribution. The FGLS estimator is:

$$\begin{bmatrix} \ddot{\mu} - \mu \\ \ddot{\beta} - \beta \end{bmatrix} = \frac{1}{q_{11}q_{22} - q_{12}^2} \begin{bmatrix} q_{22}r_1 - q_{12}r_2 \\ q_{11}r_2 - q_{12}r_1 \end{bmatrix}$$

where the definition of each element is given in (A.6) of Roy et al. (2004). It can be shown that for a sequence of $\hat{\alpha}$ such that $T(\hat{\alpha} - 1) < 0$,

$$T^{-1/2}(\ddot{\mu} - \mu) = \frac{T^{-1}q_{22}T^{1/2}r_1 - q_{12}T^{-1/2}r_2}{Tq_{11}T^{-1}q_{22} - q_{12}^2} \xrightarrow{L} \Xi \quad (\text{A.1})$$

$$T^{1/2}(\ddot{\beta} - \beta) = \frac{Tq_{11}T^{-1/2}r_2 - q_{12}T^{1/2}r_1}{Tq_{11}T^{-1}q_{22} - q_{12}^2} \xrightarrow{L} \xi \quad (\text{A.2})$$

where

$$\begin{aligned} \Xi = & 1(\Pi < 0) [(\Pi^2 - 2\Pi)(1 - 0.5\Pi + 12^{-1}\Pi^2)]^{-1} \left[(1 - \Pi + 3^{-1}\Pi^2)(\Pi^2 \int_0^1 W(t)dt - \Pi W(1)) \right. \\ & \left. - (0.5\Pi^2 - \Pi)(\Pi^2 \int_0^1 tW(t)dt - \Pi \int_0^1 W(t)dt + W(1) - \Pi \int_0^1 t dW(t)) \right], \end{aligned}$$

$$\xi = 1(\Pi < 0) [1 - 0.5\Pi + 12^{-1}\Pi^2]^{-1} \int_0^1 (1 + 0.5\Pi - \Pi s) d\hat{W}(s). \quad (\text{A.3})$$

Note that $\hat{W}(s) = W(s) - \Pi \int_0^s W(\tau)d\tau$ and $W(\cdot)$ is a standard Wiener process. To derive the limiting distributions, we use the following results:

$$Tq_{11} = T^2(\hat{\alpha} - 1)^2 - 2T(\hat{\alpha} - 1) + o_p(1) \xrightarrow{L} \Pi^2 - 2\Pi,$$

$$q_{12} = (1/2) [T^2(\hat{\alpha} - 1)^2 - 2T(\hat{\alpha} - 1)] + o_p(1) \xrightarrow{L} (1/2)(\Pi^2 - 2\Pi),$$

$$T^{-1}q_{22} = 1 - T(\hat{\alpha} - 1) + T^2(\hat{\alpha} - 1)^2/3 + o_p(1) \xrightarrow{L} 1 - \Pi + \Pi^2/3,$$

$$T^{1/2}r_1 = -T(\hat{\alpha} - 1)T^{-1/2} \sum_{t=2}^T (y_t - \hat{\alpha}y_{t-1}) + o_p(1)$$

$$= -T(\hat{\alpha} - 1)T^{-1/2} \sum_{t=2}^T (e_t - (\hat{\alpha} - 1)y_{t-1}) + o_p(1) \xrightarrow{L} \Pi^2 \int_0^1 W(t)dt - \Pi W(1),$$

$$T^{-1/2}r_2 = -T^{-1/2} \sum_{t=2}^T [T(\hat{\alpha} - 1)(t/T) - 1](y_t - \hat{\alpha}y_{t-1}) + o_p(1)$$

$$= -T(\hat{\alpha} - 1)T^{-1/2} \sum_{t=2}^T (t/T)(y_t - \hat{\alpha}y_{t-1}) + T^{-1/2} \sum_{t=2}^T (y_t - \hat{\alpha}y_{t-1}) + o_p(1)$$

$$\xrightarrow{L} \Pi^2 \int_0^1 tW(t)dt - \Pi \int_0^1 t dW(t) + W(1) - \Pi \int_0^1 W(t)dt.$$

Therefore, the limiting distribution of $\ddot{\beta}$ is $T^{1/2}(\ddot{\beta} - \beta) \xrightarrow{L} \ddot{\zeta} = 1(\Pi < 0)\xi + 1(\Pi = 0)W(1)$.

The Gauss-Newton Estimator is:

$$\begin{bmatrix} \Delta \hat{\mu} \\ \Delta \hat{\beta} \\ \Delta \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \ddot{v}_{11} & \ddot{v}_{12} & \ddot{v}_{13} \\ \ddot{v}_{12} & \ddot{v}_{22} & \ddot{v}_{23} \\ \ddot{v}_{13} & \ddot{v}_{23} & \ddot{v}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \ddot{f}_1 \\ \ddot{f}_2 \\ \ddot{f}_3 \end{bmatrix} = \ddot{V}^{-1} \ddot{F}$$

where \ddot{v}_{ij} and \ddot{f}_i are defined in Roy et al. (2004, p. 1088). These elements are functions of $(\hat{\alpha} - 1)$, $\sum_{t=2}^T \ddot{y}_{t-1}$, $\sum_{t=2}^T (t-1)\ddot{y}_{t-1}$ and $\sum_{t=2}^T \ddot{y}_{t-1}^2$. Now $\ddot{y}_{t-1} = y_{t-1} - (\ddot{\mu} - \mu) - (\ddot{\beta} - \beta)(t-1)$. Therefore,

$$\sum_{t=2}^T \ddot{y}_{t-1} = \sum_{t=2}^T y_{t-1} - (\ddot{\beta} - \beta)(0.5T^2) - (\ddot{\mu} - \mu)T + O_p(T),$$

$$\sum_{t=2}^T (t-1)\ddot{y}_{t-1} = \sum_{t=2}^T (t-1)y_{t-1} - (\ddot{\beta} - \beta)(3^{-1}T^3) - 0.5(\ddot{\mu} - \mu)T^2 + O_p(T^2),$$

$$\sum_{t=2}^T \ddot{y}_{t-1}^2 = \sum_{t=2}^T y_{t-1}^2 + (\ddot{\beta} - \beta)^2(3^{-1}T^3) - 2(\ddot{\beta} - \beta) \sum_{t=2}^T (t-1)y_{t-1}$$

$$+ (\ddot{\mu} - \mu)(\ddot{\beta} - \beta)T^2 - 2(\ddot{\mu} - \mu) \sum_{t=2}^T y_{t-1} + (\ddot{\mu} - \mu)^2T + O_p(T^{3/2}).$$

Then,

$$T^{-3/2} \sum_{t=2}^T \ddot{y}_{t-1} = T^{-3/2} \sum_{t=2}^T y_{t-1} - 0.5T^{1/2}(\ddot{\beta} - \beta) - T^{-1/2}(\ddot{\mu} - \mu) + o_p(1)$$

$$\xrightarrow{L} \int_0^1 W(t)dt - 0.5\ddot{\zeta} - \Xi = a_1,$$

$$\begin{aligned}
T^{-2} \sum_{t=2}^T \dot{y}_{t-1}^2 &= T^{-2} \sum_{t=2}^T y_{t-1}^2 + T(\ddot{\beta} - \beta)^2(3^{-1}) - 2T^{1/2}(\ddot{\beta} - \beta)T^{-5/2} \sum_{t=2}^T (t-1)y_{t-1} \\
&+ T^{-1/2}(\ddot{\mu} - \mu)T^{1/2}(\ddot{\beta} - \beta) - 2T^{-1/2}(\ddot{\mu} - \mu)T^{-3/2} \sum_{t=2}^T y_{t-1} + T^{-1}(\ddot{\mu} - \mu)^2 + o_p(1) \\
&\xrightarrow{L} \int_0^1 W^2(t)dt + 3^{-1}\zeta^2 - 2\zeta \int_0^1 tW(t)dt + \Xi\ddot{\zeta} - 2\Xi \int_0^1 W(t)dt + \Xi^2 = \eta_{33},
\end{aligned}$$

$$\begin{aligned}
T^{-5/2} \sum_{t=2}^T (t-1)\dot{y}_{t-1} &= T^{-5/2} \sum_{t=2}^T (t-1)y_{t-1} - 3^{-1}T^{1/2}(\ddot{\beta} - \beta) - 0.5T^{-1/2}(\ddot{\mu} - \mu) + o_p(1) \\
&\xrightarrow{L} \int_0^1 tW(t)dt - 3^{-1}\zeta - 0.5\Xi.
\end{aligned}$$

Let $D_{(2)T} = \text{diag}(\ddot{v}_{11}^{1/2}, T^{1/2}, T)$. Then after some algebra, it can be shown that

$$\begin{aligned}
\begin{bmatrix} \ddot{v}_{11}^{1/2} \Delta \hat{\mu} \\ T^{1/2} \Delta \hat{\beta} \\ T \Delta \hat{\alpha} \end{bmatrix} &= (D_{(2)T}^{-1} \ddot{V} D_{(2)T}^{-1})^{-1} D_{(2)T}^{-1} \ddot{F} \\
&= \begin{bmatrix} 1 & T^{-1/2} \ddot{v}_{11}^{-1/2} \ddot{v}_{12} & T^{-1} \ddot{v}_{11}^{-1/2} \ddot{v}_{13} \\ T_{11}^{-1/2} \ddot{v}_{11}^{-1/2} \ddot{v}_{12} & T^{-1} \ddot{v}_{22} & T^{-3/2} \ddot{v}_{23} \\ T^{-1} \ddot{v}_{11}^{-1/2} \ddot{v}_{13} & T^{-3/2} \ddot{v}_{23} & T^{-2} \ddot{v}_{33} \end{bmatrix}^{-1} \begin{bmatrix} \ddot{v}_{11}^{-1/2} \ddot{f}_1 \\ T^{-1/2} \ddot{f}_2 \\ T^{-1} \ddot{f}_3 \end{bmatrix} \\
&\xrightarrow{L} \begin{bmatrix} 1 & \eta_{12} & \eta_{13} \\ \eta_{12} & \eta_{22} & \eta_{23} \\ \eta_{13} & \eta_{23} & \eta_{33} \end{bmatrix}^{-1} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = V_{(2)}^{-1} F \tag{A.4}
\end{aligned}$$

where the elements of $V_{(2)}$ and F are as follows:

$$\begin{aligned}
(T\ddot{v}_{11})^{-1/2} \ddot{v}_{12} &= (T\{(T-1)(1-\ddot{\alpha})^2 + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)]^2\})^{-1/2} \\
&\times (\sum_{t=2}^T (1-\ddot{\alpha}) [t(1-\ddot{\alpha}) + \ddot{\alpha}] + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)]^2) \\
&\xrightarrow{L} [\ddot{\Pi}^2 - 2\ddot{\Pi}]^{-1/2} (1/2) [\ddot{\Pi}^2 - 2\ddot{\Pi}] 1(\ddot{\Pi} < 0) = (1/2) [\ddot{\Pi}^2 - 2\ddot{\Pi}]^{1/2} 1(\ddot{\Pi} < 0) = \eta_{12}
\end{aligned}$$

$$\begin{aligned}
(T\ddot{v}_{11})^{-1/2} (T^{-1/2} \ddot{v}_{13}) &= (T\{(T-1)(1-\ddot{\alpha})^2 + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)]^2\})^{-1/2} \\
&\times (T^{-1/2} \sum_{t=2}^T (1-\ddot{\alpha}) \dot{y}_{t-1}) \\
&\xrightarrow{L} [\ddot{\Pi}^2 - 2\ddot{\Pi}]^{-1/2} [-\ddot{\Pi} \{ \int_0^1 W(t)dt - 0.5\zeta - \Xi \}] 1(\ddot{\Pi} < 0) = \eta_{13}
\end{aligned}$$

$$\begin{aligned}
T^{-1}\ddot{v}_{22} &= T^{-1}\left\{\sum_{t=2}^T [t(1-\ddot{\alpha}) + \ddot{\alpha}]^2 + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)]^2\right\} \\
&= T^2(\ddot{\alpha} - 1)^2 T^{-3} \sum_{t=2}^T t^2 + 1 - 2T(\ddot{\alpha} - 1)T^{-2} \sum_{t=2}^T t + o_p(1) \xrightarrow{L} 3^{-1}\ddot{\Pi}^2 + 1 - \ddot{\Pi} = \eta_{22}
\end{aligned}$$

$$\begin{aligned}
T^{-3/2}\ddot{v}_{23} &= T^{-3/2}\left\{\sum_{t=2}^T [t(1-\ddot{\alpha}) + \ddot{\alpha}] \ddot{y}_{t-1}\right\} \\
&= -T(\ddot{\alpha} - 1)T^{-5/2} \sum_{t=2}^T t\ddot{y}_{t-1} + T^{-3/2} \sum_{t=2}^T \ddot{y}_{t-1} + o_p(1) \\
&\xrightarrow{L} -\ddot{\Pi}[\int_0^1 tW(t)dt - 3^{-1}\ddot{\zeta} - 0.5\Xi] + [\int_0^1 W(t)dt - 0.5\ddot{\zeta} - \Xi] = \eta_{23}
\end{aligned}$$

and $T^{-2}\ddot{v}_{33} = T^{-2} \sum_{t=2}^T \ddot{y}_{t-1}^2 \xrightarrow{L} \eta_{33}$. Now $\tilde{e}_t = \Delta\ddot{y}_t - (\ddot{\alpha} - 1)\ddot{y}_{t-1} = e_t - (\ddot{\beta} - \beta) - (\ddot{\alpha} - 1)\ddot{y}_{t-1}$. Therefore,

$$\begin{aligned}
(T\ddot{v}_{11})^{-1/2}(T^{1/2}\ddot{f}_1) &= (T\{(T-1)(1-\ddot{\alpha})^2 + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)]^2\})^{-1/2} \\
&\quad \times (T^{1/2}\{(1-\ddot{\alpha}) \sum_{t=2}^T \tilde{e}_t + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)] \tilde{e}_1\}) \\
&\quad \xrightarrow{L} [\ddot{\Pi}^2 - 2\ddot{\Pi}]^{-1/2}[\ddot{\Pi}(\ddot{\zeta} - W(1)) + \ddot{\Pi}^2 a_1 + 2\ddot{\Pi}\Xi]1(\ddot{\Pi} < 0)
\end{aligned}$$

$$\begin{aligned}
T^{-1/2}\ddot{f}_2 &= T^{-1/2}\left\{\sum_{t=2}^T \tilde{e}_t [t(1-\ddot{\alpha}) + \ddot{\alpha}] + [(1-\ddot{\alpha}^2)^{1/2}1(\ddot{\alpha} < 1) + 1(\ddot{\alpha} = 1)] \tilde{e}_1\right\} \\
&= T^{-1/2} \sum_{t=2}^T [e_t - (\ddot{\beta} - \beta) - (\ddot{\alpha} - 1)\ddot{y}_{t-1}] [t(1-\ddot{\alpha}) + \ddot{\alpha}] + o_p(1) \\
&= -T(\ddot{\alpha} - 1)T^{-3/2} \sum_{t=2}^T t e_t + T^{-1/2} \sum_{t=2}^T e_t + T^{1/2}(\ddot{\beta} - \beta)\{T(\ddot{\alpha} - 1)T^{-2} \sum_{t=2}^T t - 1\} \\
&\quad + T(\ddot{\alpha} - 1)\{T(\ddot{\alpha} - 1)T^{-5/2} \sum_{t=2}^T t\ddot{y}_{t-1} - T^{-3/2} \sum_{t=2}^T \ddot{y}_{t-1}\} + o_p(1) \\
&\quad \xrightarrow{L} -\ddot{\Pi} \int_0^1 t dW(t) + W(1) + \ddot{\zeta}(0.5\ddot{\Pi} - 1) + \ddot{\Pi}\{\ddot{\Pi}[\int_0^1 tW(t)dt - 3^{-1}\ddot{\zeta} - 0.5\Xi] - a_1\} \\
&= -\ddot{\Pi} \int_0^1 t dW(t) + W(1) + \ddot{\zeta}(0.5\ddot{\Pi} - 1) - \ddot{\Pi}\eta_{23}
\end{aligned}$$

$$\begin{aligned}
T^{-1}\ddot{f}_3 &= T^{-1} \sum_{t=2}^T \tilde{e}_t \ddot{y}_{t-1} \\
&= T^{-1} \sum_{t=2}^T [e_t - (\ddot{\beta} - \beta) - (\ddot{\alpha} - 1)\ddot{y}_{t-1}] \ddot{y}_{t-1} \\
&= T^{-1} \sum_{t=2}^T \ddot{y}_{t-1} e_t - T^{1/2}(\ddot{\beta} - \beta)T^{-3/2} \sum_{t=2}^T \ddot{y}_{t-1} - T(\ddot{\alpha} - 1)T^{-2} \sum_{t=2}^T \ddot{y}_{t-1}^2
\end{aligned}$$

$$\begin{aligned}
&= T^{-1} \sum_{t=2}^T y_{t-1} e_t - T^{-1/2} (\ddot{\mu} - \mu) T^{-1/2} \sum_{t=2}^T e_t - T^{1/2} (\ddot{\beta} - \beta) T^{-3/2} \sum_{t=2}^T (t-1) e_t \\
&\quad - T^{1/2} (\ddot{\beta} - \beta) T^{-3/2} \sum_{t=2}^T \ddot{y}_{t-1} - T(\ddot{\alpha} - 1) T^{-2} \sum_{t=2}^T \ddot{y}_{t-1}^2 \\
&\xrightarrow{L} \int_0^1 W(t) dW(t) - \Xi W(1) - \zeta \int_0^1 t dW(t) - \zeta a_1 - \ddot{\Pi} \eta_{33} = f_3
\end{aligned}$$

The Gauss-Newton estimator of the slope satisfies

$$T^{1/2}(\hat{\beta}_{GN} - \beta) = T^{1/2}(\ddot{\beta} - \beta) + T^{1/2} \Delta \hat{\beta} \rightarrow \zeta + \kappa := \zeta$$

where κ is the second element of the vector $V_{(2)}^{-1} F$. In addition, the standard error is given by $TV(\hat{\beta}_{GN} - \beta) \rightarrow \min(v_{(2)}^{(22)}, 1) := v_\beta$, where $v_{(2)}^{(22)}$ is the (2,2)th element of $V_{(2)}^{-1}$. Therefore,

$$t_{GN}(\hat{\beta}) = \frac{T^{1/2}(\hat{\beta}_{GN} - \beta)}{\sqrt{TV(\hat{\beta}_{GN} - \beta)}} \rightarrow v_\beta^{-1/2} \zeta.$$

Derivation of Π and $\ddot{\Pi}$: An estimator of α in the Roy and Fuller (2001) class is a function of a test for a unit root of the form $\tilde{\alpha} = \hat{\alpha}_{WSLS} + \tilde{C}(\hat{\tau}_{WSLS}) \hat{\sigma}_{WSLS}$, where

$$\tilde{C}(\hat{\tau}_{WSLS}) = \begin{cases} -\hat{\tau}_{WSLS} & \text{if } \hat{\tau}_{WSLS} > \tau_{pct} \\ T^{-1} \hat{\tau}_{WSLS} - 3[\hat{\tau}_{WSLS} + \kappa(\hat{\tau}_{WSLS} + 5)]^{-1} & \text{if } -5 < \hat{\tau}_{WSLS} \leq \tau_{pct} \\ T^{-1} \hat{\tau}_{WSLS} - 3(\hat{\tau}_{WSLS})^{-1} & \text{if } -(3T)^{0.5} < \hat{\tau}_{WSLS} \leq -5 \\ 0 & \text{if } \hat{\tau}_{WSLS} \leq -(3T)^{0.5} \end{cases}$$

τ_{pct} is a percentile of the limiting distribution of $\hat{\tau}_{WSLS}$ when $\alpha = 1$, $\kappa = [3T - \tau_{pct}^2(1 + T)][\tau_{pct}(5 + \tau_{pct})(1 + T)]^{-1}$, and $\hat{\tau}_{WSLS} = (\hat{\alpha}_{WSLS} - 1)/\hat{\sigma}_{WSLS}$, $\hat{\alpha}_{WSLS} = [\sum_{t=2}^{T-1} \hat{y}_t^2 + T^{-1} \sum_{t=1}^T \hat{y}_t^2]^{-1} [\sum_{t=2}^T \hat{y}_{t-1} \hat{y}_t]$, $\hat{\sigma}_{WSLS}^2 = [\sum_{t=2}^{T-1} \hat{y}_t^2 + T^{-1} \sum_{t=1}^T \hat{y}_t^2]^{-1} \hat{\sigma}^2$, $\hat{\sigma}^2 = (T-3)^{-1} \sum_{t=2}^T [\hat{y}_t - \hat{\alpha}_{WSLS} \hat{y}_{t-1}]^2$ and \hat{y}_t is the OLS detrended value of y_t . Using $T(\tilde{\alpha} - 1) = T(\hat{\alpha}_{WSLS} - 1) + \tilde{C}(\hat{\tau}_{WSLS}) T \hat{\sigma}_{WSLS}$, we obtain from results in Fuller (1996),

$$T(\hat{\alpha}_{WSLS} - 1) \xrightarrow{L} \frac{\{0.5[(W(1) - 2 \int_0^1 W(t) dt) (W(1) - 6K) - 1] + [\int_0^1 W(t) dt - 3K]^2 - [\int_0^1 W^2(t) dt - (\int_0^1 W(t) dt)^2 - 3K^2]\}}{\int_0^1 W^2(t) dt - (\int_0^1 W(t) dt)^2 - 3K^2} = \Psi$$

$$\hat{\tau}_{WSLS} \xrightarrow{L} \frac{\{0.5[(W(1) - 2 \int_0^1 W(t) dt) (W(1) - 6K) - 1] + [\int_0^1 W(t) dt - 3K]^2 - [\int_0^1 W^2(t) dt - (\int_0^1 W(t) dt)^2 - 3K^2]\}}{(\int_0^1 W^2(t) dt - (\int_0^1 W(t) dt)^2 - 3K^2)^{1/2}} = \tau$$

where $K = 2 \int_0^1 tW(t)dt - \int_0^1 W(t)dt$. Therefore, when $\alpha = 1$,

$$T(\tilde{\alpha} - 1) \xrightarrow{L} \Psi + \frac{C^*(\tau)}{\left(\int_0^1 W^2(t)dt - \left(\int_0^1 W(t)dt\right)^2 - 3K^2\right)^{1/2}} \quad (\text{A.5})$$

where

$$C^*(\tau) = \begin{cases} -\tau & \text{if } \tau > \tau_{pct} \\ -3[\tau + \kappa^*(\tau + 5)]^{-1} & \text{if } -5 < \tau \leq \tau_{pct} \\ -3\tau^{-1} & \text{if } -\infty < \tau \leq -5 \end{cases}$$

with $\kappa^* = [3 - \tau_{pct}^2][\tau_{pct}(5 + \tau_{pct})]^{-1}$.

Table 1: Distribution of the Gauss-Newton test

	%				Probability
	0.80	0.90	0.95	0.975	$ t_{GN} > 1.98$
T=100	1.595	2.561	3.538	4.654	0.300
T=250	1.621	2.553	3.481	4.351	0.299
T=500	1.569	2.515	3.428	4.359	0.291
T=1000	1.572	2.495	3.455	4.302	0.294
T=5000	1.573	2.560	3.480	4.370	0.296
$T = \infty$; Corrected	1.571	2.518	3.455	4.428	0.298
$T = \infty$; Roy et al. (2004)	2.291	5.412	9.408	13.31	0.506

Note: “ $T = \infty$; Corrected” stands for the limit distribution in Theorem 1 of this note, “ $T = \infty$; Roy et al. (2004)” stands for the limit distribution stated in Theorem 1 of Roy et al. (2004). The statistic correspond to what is suggested in their Section 4.1 with $se_{GN}(\hat{\beta}) = \min\{s(T-1)^{-1/2}, s\nu_{\beta}^{-1/2}\}$.

Table 2: Distribution of the Gauss-Newton test with $se_{GN}(\hat{\beta}) = s\nu_{\beta}^{-1/2}$.

	%				Probability		
	0.80	0.90	0.95	0.975	$ t_{GN} > 1.98$	$ t_{GN}^1 > 1.98$	$ t_{GN}^2 > 1.98$
T=100	1.270	1.824	2.363	2.979	0.163	0.076	0.080
T=250	1.288	1.833	2.288	2.816	0.152	0.071	0.075
T=500	1.265	1.813	2.297	2.867	0.149	0.065	0.073
T=1000	1.260	1.805	2.239	2.832	0.152	0.067	0.074
T=5000	1.260	1.819	2.269	2.762	0.149	0.063	0.070
$T = \infty$; Corrected	1.260	1.790	2.268	2.843	0.159		
$T = \infty$; Roy et al. (2004)	1.481	3.134	5.210	7.385	0.361		

Note: t_{GN} is computed from the code we constructed, t_{GN}^1 from the code of Roy et al. (2004), and t_{GN}^2 from the code of Roy et al. (2004) with the minor change that we compute $\ddot{y}_t = Y_t - \hat{\mu} - \hat{\beta}t$ (for $t = 2, \dots, T$). “ $T = \infty$; Corrected” stands for the limit distribution in Theorem 1 of this note, “ $T = \infty$; Roy et al. (2004)” stands for the limit distribution stated in Theorem 1 of Roy et al. (2004).

Table 3: Probabilities that the estimate of α takes value 1; $T = 100$.

α	Perron-Yabu (2009)		Gauss-Newton
	$\Pr[\hat{\alpha}(\tau_{pct} = -1.96) - 1 < T^{-0.5}]$		$\Pr[\hat{\tau}_{WLSL} > -2.94]$
1.00	0.890		0.896
0.98	0.870		0.878
0.97	0.854		0.861
0.95	0.799		0.809
0.80	0.550		0.566
0.70	0.079		0.089
0.60	0.001		0.001
0.40	0.000		0.000
0.00	0.000		0.000