

# A Fixed- $b$ Perspective on the Phillips-Perron Unit Root Tests\*

Timothy J. Vogelsang  
Department of Economics  
Michigan State University

Martin Wagner  
Department of Economics and Finance  
Institute for Advanced Studies, Vienna  
&  
Frisch Centre for Economic Research, Oslo

July 2011

## Abstract

In this paper we extend fixed- $b$  asymptotic theory to the nonparametric Phillips-Perron (PP) unit root tests. We show that the fixed- $b$  limits depend on nuisance parameters in a complicated way. These non-pivotal limits provide an alternative theoretical explanation for the well known finite sample problems of PP tests. We also show that the fixed- $b$  limits depend on whether deterministic trends are removed using one-step or two-step approaches. This is in contrast to the asymptotic equivalence of the one- and two-step approaches under a consistency approximation for the long run variance estimator. Based on these results we introduce modified PP tests that allow for fixed- $b$  inference. The theoretical analysis is cast in the framework of near-integrated processes which allows to study the asymptotic behavior both under the unit root null hypothesis as well as for local alternatives. The performance of the original and modified PP tests is compared by means of local asymptotic power and a small finite sample simulation study.

**JEL Classification:** C12, C13, C32

**Keywords:** Nonparametric kernel estimator, long run variance, detrending, one-step, two-step.

---

\*The authors acknowledge financial support from the Jubiläumsfonds of the Oesterreichische Nationalbank (Grant Nr. 13398).

## 1 Introduction

In this paper we extend the fixed- $b$  asymptotic theory of Kiefer and Vogelsang (2005) to the well known unit root tests of Phillips and Perron (1988), i.e. the PP tests. We focus on the case where PP tests are constructed using nonparametric kernel estimators of the long run variance. We find that the fixed- $b$  limits of the PP tests are not pivotal and furthermore also depend on whether deterministic trends are removed using one-step or two-step detrending methods. Our results are in contrast to existing results based on consistency of the long run variance estimator in which case the asymptotic distributions of the PP tests are pivotal and are the same for one- and two-step detrending. Our finding of a non-pivotal fixed- $b$  limit provides an alternative explanation for the often inadequate finite sample performance of PP tests (Perron and Ng, 1996 and Schwert, 1989). We propose a simple adjustment to the PP tests that provides a pivotal fixed- $b$  limit under the unit root null. The theoretical analysis is performed using the framework of near-integrated process (cf. Phillips, 1987) which allows the derivation of limiting distributions both under the unit root null hypothesis as well as under local alternatives (to study local asymptotic power). The remainder of the paper is organized as follows. In the next section we provide the fixed- $b$  limits of the one- and two-step detrended versions of the PP tests. In Section 3 we propose the adjustment that restores an asymptotically pivotal limit. Section 4 provides some limited finite sample results and Section 5 briefly summarizes and concludes. All proofs are relegated to the appendix. Supplementary material available upon request provides fixed- $b$  critical values for five kernel functions (Bartlett, Bohman, Daniell, Parzen and Quadratic Spectral), for the specification without deterministic components, with intercept only and with intercept and linear trend. For the latter two specifications of the deterministic component, critical values are available for both one- and two-step detrending.

## 2 The Fixed- $b$ Limits of the Phillips-Perron Tests

We assume that the data are generated according to

$$y_t = D_t' \theta + y_t^0, \quad t = 1, \dots, T, \quad (1)$$

$$y_t^0 = \left(1 - \frac{c}{T}\right) y_{t-1}^0 + u_t, \quad (2)$$

where  $D_t = [1, t, t^2, \dots, t^q]'$  for some  $0 \leq q < \infty$ . When  $c = 0$ ,  $y_t^0$  is a unit root process and values of  $c > 0$  correspond to near-integrated (in the terminology of Phillips, 1987) stationary (for fixed  $T$ ) alternatives.

The key assumption is that the process  $u_t$  satisfies a functional central limit (FCLT), i.e.

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} u_t \Rightarrow \omega W(r), \quad (3)$$

where  $[rT]$  is the integer part of  $rT$  with  $r \in [0, 1]$ ,  $W(r)$  is a standard Wiener process and  $0 < \omega^2 < \infty$  is the long run variance of  $u_t$ . Assuming for notational simplicity that  $u_t$  is covariance stationary with summable autocovariance function,  $\gamma_j$ , we have

$$\omega^2 = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j. \quad (4)$$

As is common, we use  $\sigma^2$  to denote  $\gamma_0$  and we furthermore define the half long run variance  $\lambda := \frac{1}{2}(\omega^2 - \sigma^2)$ .<sup>1</sup> Sufficient conditions for the FCLT can be found in Phillips and Solo (1992), or more specifically also in Phillips and Perron (1988) and in Sims *et al.* (1990) who consider similar deterministic components as we do. It is well known (see Phillips 1987) that given (2) and (3) it follows that

$$T^{-1/2} y_{[rT]}^0 \Rightarrow \omega V_c(r),$$

where  $V_c(r) = \int_0^r e^{-c(r-s)} dW(s)$ .

Before we can turn to the analysis of the tests' asymptotic behavior we need to define several additional quantities. Define  $D(r) := [1, \dots, r^q]'$  and define the correspondingly detrended process,  $\tilde{V}_c(r)$ , and generalized Brownian bridge,  $\widehat{W}(r)$ , as:

$$\tilde{V}_c(r) := V_c(r) - D(r)' \left( \int_0^1 D(s) D(s)' ds \right)^{-1} \int_0^1 D(s) V_c(s) ds, \quad (5)$$

$$\widehat{W}(r) := W(r) - \int_0^r D(s)' ds \left( \int_0^1 D(s) D(s)' ds \right)^{-1} \int_0^1 D(s) dW(s). \quad (6)$$

A slight variant of  $\tilde{V}_c(r)$  is also needed and is defined as

$$\tilde{\tilde{V}}_c(r) := W(r) - \int_0^r \left( \dot{D}(s) + cD(s) \right)' ds \left( \int_0^1 D(s) D(s)' ds \right)^{-1} \int_0^1 D(s) V_c(s) ds,$$

where  $\dot{D}(r) = \frac{\partial D(r)}{\partial r} = [0, 1, 2r, \dots, qr^{q-1}]'$ . Note that  $\int_0^r \dot{D}(s) ds = [0, r, r^2, \dots, r^q]'$  is simply  $D(r)$  with its first element replaced with 0. For the pure unit root case,  $c = 0$ , because  $V_0(r) = W(r)$ , it follows that  $\tilde{\tilde{V}}_0(r)$  and  $\tilde{V}_0(r)$  are similar but different stochastic processes. As we shall see below this difference is the effect of using either one- or two-step detrending.

In the rest of the paper we will use throughout a subscript  $i = 1, 2$  to refer to quantities related to either one- or two-step detrending, e.g.  $\tilde{y}_{t,1}$  refers to the one-step detrended  $y_t$ .

## 2.1 One-Step Approach

The one-step approach is based on estimating the regression model

$$y_t = D_t' \delta + \alpha y_{t-1} + u_t, \quad t = 2, \dots, T, \quad (7)$$

---

<sup>1</sup>In case  $u_t$  is not assumed to be stationary  $\omega^2$  and  $\sigma^2$  are defined as  $\omega^2 := \lim_{T \rightarrow \infty} \mathbb{E} \left( \frac{1}{T} (\sum_{t=1}^T u_t)^2 \right)$  and  $\sigma^2 := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \mathbb{E}(u_t^2)$ , compare Phillips and Perron (1988).

with the null hypothesis of interest being  $H_0 : \alpha = 1$ . Using the Frisch-Waugh theorem the deterministic components can be eliminated and one can focus on the regression

$$\tilde{y}_{t,1} = \alpha \tilde{y}_{t-1,1} + \tilde{u}_t, \quad t = 2, \dots, T, \quad (8)$$

with  $\tilde{y}_{t,1} := y_t - D'_t(D'_T D_T)^{-1} D'_T Y_T$ ,  $\tilde{y}_{t-1,1} := y_{t-1} - D'_t(D'_T D_T)^{-1} D'_T Y_{T-1}$  and  $\tilde{u}_t := u_t - D'_t(D'_T D_T)^{-1} D'_T U_T$ , using the notation  $Y_T := [y_2, \dots, y_T]'$ ,  $Y_{T-1} := [y_1, \dots, y_{T-1}]'$ ,  $U_T := [u_2, \dots, u_T]'$  and  $D_T := [D_2, \dots, D_T]'$ .

The one-step PP unit root tests are based on the OLS estimator  $\hat{\alpha}_1 := \frac{\sum_{t=2}^T \tilde{y}_{t,1} \tilde{y}_{t-1,1}}{\sum_{t=2}^T \tilde{y}_{t-1,1}^2}$  of  $\alpha$  from (8), respectively the  $t$ -statistic

$$t_{\alpha_1} := \frac{\hat{\alpha}_1 - 1}{\sqrt{\hat{\sigma}^2 (\sum_{t=2}^T \tilde{y}_{t-1,1}^2)}},$$

with  $\hat{\sigma}^2 := \frac{1}{T} \sum_{t=2}^T \hat{u}_{t,1}^2$  and  $\hat{u}_{t,1} := \tilde{y}_{t,1} - \hat{\alpha}_1 \tilde{y}_{t-1,1}$ .<sup>2</sup> Furthermore, denote the estimated long run variance as  $\hat{\omega}^2 := \hat{\gamma}_0 + 2 \sum_{j=1}^{T-2} k(j/M) \hat{\gamma}_j$ , with  $\hat{\gamma}_j := \frac{1}{T} \sum_{t=j+2}^T \hat{u}_{t,1} \hat{u}_{t-j,1}$ . In addition to regularity conditions on  $u_t$ , consistency of  $\hat{\omega}^2$  depends upon the kernel function  $k(\cdot)$  and the rate of divergence of  $M$  such that  $M \rightarrow \infty$  and  $M/T \rightarrow 0$  as  $T \rightarrow \infty$  (for a discussion see Jansson 2002).

The coefficient and  $t$ -statistic based one-step PP unit root tests are given by

$$Z_{\alpha,1} := T(\hat{\alpha}_1 - 1) - \frac{1}{2}(\hat{\omega}^2 - \hat{\sigma}^2) \left( T^{-2} \sum_{t=2}^T \tilde{y}_{t-1,1}^2 \right)^{-1}, \quad (9)$$

$$Z_{t,1} := \frac{\hat{\sigma}}{\hat{\omega}} t_{\alpha_1} - \frac{1}{2}(\hat{\omega}^2 - \hat{\sigma}^2) \left( \hat{\omega}^2 T^{-2} \sum_{t=2}^T \tilde{y}_{t-1,1}^2 \right)^{-1/2}. \quad (10)$$

## 2.2 Two-Step Approach

The two-step detrending approach is very similar, yet there are some subtle differences that will matter. The two-step approach is based on first detrending the series  $y_t$  to then estimate a regression similar to (8) on the first detrended data. Thus, we have  $\tilde{y}_{t,2} := y_t - D'_t \hat{\theta}$ , with  $\hat{\theta} := \left( \sum_{t=1}^T D_t D'_t \right)^{-1} \sum_{t=1}^T D_t y_t$  and  $\tilde{y}_{t-1,2} := y_{t-1} - D'_{t-1} \hat{\theta}$ . Consequently the OLS estimator  $\hat{\alpha}_2$  of  $\alpha$  used in this approach is given by

$$\hat{\alpha}_2 := \frac{\sum_{t=2}^T \tilde{y}_{t,2} \tilde{y}_{t-1,2}}{\sum_{t=2}^T \tilde{y}_{t-1,2}^2},$$

---

<sup>2</sup>We often omit the subscript  $i \in \{1, 2\}$  for the quantities  $\hat{\alpha}$ ,  $t_\alpha$  and  $\hat{\sigma}^2$ , since these have, under the assumptions stated, the same limit for both the one- and two-step detrending methods. Only when the distinction is important a subscript is added. We are confident that this economical use of notation does not lead to any confusion. Note, however, that if one were to consider more general deterministic components the limits of e.g. the one- and two-step OLS estimators of  $\alpha$  can be different, see Remark 1 in Section 2.3.

and the corresponding two-step residuals, used to compute the two-step estimates of  $\sigma^2$  and  $\omega^2$ , are given by  $\hat{u}_{t,2} := \tilde{y}_{t,2} - \hat{\alpha}_2 \tilde{y}_{t-1,2}$ .

The two-step PP tests,  $Z_{\alpha,2}$  and  $Z_{t,2}$  say, are defined as before – with  $\tilde{y}_{t-1,2}$  and the two-step estimates of  $\alpha$ ,  $\sigma^2$  and  $\omega^2$  instead of the corresponding one-step quantities – in (9) and (10).

### 2.3 Asymptotic Results

It is well known that for deterministic polynomial trends the asymptotic distribution of  $T(\hat{\alpha} - 1)$  is the same for both the one-step and two-step approaches. Thus, when a consistent estimator of  $\omega^2$  is used, the asymptotic distributions of the PP tests are identical for the one- and two-step versions of the tests and it holds that

$$Z_{\alpha,i} \Rightarrow -c + \frac{\int_0^1 \tilde{V}_c(r) dW(r)}{\int_0^1 \tilde{V}_c(r)^2 dr}, \quad Z_{t,i} \Rightarrow -c \sqrt{\int_0^1 \tilde{V}_c(r)^2 dr} + \frac{\int_0^1 \tilde{V}_c(r) dW(r)}{\sqrt{\int_0^1 \tilde{V}_c(r)^2 dr}},$$

for  $i = 1, 2$ .

**Remark 1** *Further differences between one- and two-step detrending occur in case of more general deterministic components since then also the one- and two-step limits of  $T(\hat{\alpha} - 1)$  differ. In particular, the numerator of the two-step limit contains an additional term*

$$- \int_0^1 \tilde{V}_c(r) \dot{D}(r) dr \left( \int_0^1 D(r) D(r)' dr \right)^{-1} \int_0^1 D(r) V_c(r) dr,$$

with  $\dot{D}(r)$  denoting the (corresponding quantity in more general cases) first difference of  $D(r)$  as defined before for the polynomial trend case. This term is, clearly, not zero in general, but in case of polynomial trends it holds that  $\int_0^1 \tilde{V}_c(r) \dot{D}(r) dr = 0$ , since in that case the span of  $\dot{D}(r)$  is contained in the span of  $D(r)$ . For an example where this term is not zero, see Perron and Vogelsang (1992) who include a mean shift dummy in the deterministic component.

The fact that these limiting distributions rely upon the consistency of  $\hat{\omega}^2$ , implies that the asymptotic distributions do not capture the influence of the randomness in  $\hat{\omega}^2$  on the resulting test statistics. In particular, the choices with respect to both the kernel function and the bandwidth are not reflected in the asymptotic distribution, yet affect the finite sample performance of  $\hat{\omega}^2$  and thus of the PP test statistics.

This limitation of conventional asymptotic theory is addressed with fixed- $b$  theory by means of deriving an asymptotic approximation for  $\hat{\omega}^2$  under the assumption that  $M = bT$ , where  $b \in (0, 1]$  is held fixed as  $T \rightarrow \infty$ . The fixed- $b$  limit of  $\hat{\omega}^2$  depends on the asymptotic behavior of the scaled partial sums of  $\hat{u}_{t,i}$ , which is shown below in Lemma 1 to differ between the one- and two-step approaches. This in turn implies that also the one- and two-step PP test statistics will have different limits when using the fixed- $b$  approximation.

**Lemma 1** *Assume that the data are generated by (1) with the errors,  $u_t$ , fulfilling the FCLT (3). As  $T \rightarrow \infty$  it holds for  $0 \leq r \leq 1$  that*

$$T^{-1/2} \sum_{t=2}^{[rT]} \hat{u}_{t,1} \Rightarrow \omega H_{1,c}(r), \quad H_{1,c}(r) := \widehat{W}(r) - \left( \frac{\omega^2 \int_0^1 \tilde{V}_c(s) dW(s) + \lambda}{\omega^2 \int_0^1 \tilde{V}_c(s)^2 dr} \right) \int_0^r \tilde{V}_c(s) ds, \quad (11)$$

$$T^{-1/2} \sum_{t=2}^{[rT]} \hat{u}_{t,2} \Rightarrow \omega H_{2,c}(r), \quad H_{2,c}(r) := \tilde{V}_c(r) - \left( \frac{\omega^2 \int_0^1 \tilde{V}_c(s) dW(s) + \lambda}{\omega^2 \int_0^1 \tilde{V}_c(s)^2 dr} \right) \int_0^r \tilde{V}_c(s) ds. \quad (12)$$

Lemma 1 shows that in the limit of the scaled residual partial sums the leading terms differ between the two approaches. The leading terms reflect the impact of the detrending method and the second terms are identical, since they essentially reflect the estimation of  $\alpha$ , which is for the considered deterministic components asymptotically equivalent in both cases.

Based on the above partial sum result the fixed- $b$  limits of  $\hat{\omega}^2$  can be expressed in terms of the processes  $H_{1,c}(r)$  and  $H_{2,c}(r)$ . As is common in fixed- $b$  theory the results depend upon  $b$  and the shape of the kernel function in ways outlined in Definition 1.

**Definition 1** *With  $H(r)$  denoting a generic scalar stochastic process we define the stochastic process  $P(b, k, H)$  as follows:*

1. *If  $k''(x)$  exists and is continuous, then:*

$$P(b, k, H) = -\frac{1}{b^2} \int_0^1 \int_0^1 k'' \left( \frac{r-s}{b} \right) H(r)H(s) dr ds + \frac{2}{b} H(1) \int_0^1 k' \left( \frac{1-r}{b} \right) H(r) dr + H(1)^2$$

2. *If  $k(x)$  is continuous,  $k(x) = 0$  for  $|x| \geq 1$ , and  $k(x)$  is twice continuously differentiable everywhere except for possibly  $|x| = 1$ , then*

$$P(b, k, H) = -\frac{1}{b^2} \int \int_{|r-s| \leq b} k'' \left( \frac{r-s}{b} \right) H(r)H(s) dr ds + \frac{2}{b} k'_-(1) \int_0^{1-b} H(r)H(r+b) dr + \frac{2}{b} H(1) \int_{1-b}^1 k' \left( \frac{1-r}{b} \right) H(r) dr + H(1)^2,$$

$$\text{with } k'_-(1) = \lim_{h \rightarrow 0} \frac{k(1) - k(1-h)}{h}.$$

3. *If  $k(x) = 1 - |x|$  for  $|x| \leq 1$  and  $k(x) = 0$  otherwise, then*

$$P(b, k, H) = \frac{2}{b} \int_0^1 H(r)^2 dr - \frac{2}{b} \int_0^{1-b} H(r)H(r+b) dr - \frac{2}{b} H(1) \int_{1-b}^1 H(r) dr + H(1)^2.$$

Using the quantities just defined the following proposition gives the fixed- $b$  limits of  $\hat{\omega}^2$  and of the PP tests for both the one- and two-step approaches.

**Proposition 1** *Assume that the data are generated by (1) with the errors,  $u_t$ , fulfilling the FCLT (3). Furthermore assume that  $M = bT$ , with  $b \in (0, 1]$  fixed and the subscript  $i \in \{1, 2\}$  refers again to the one- and two-step approaches. Then as  $T \rightarrow \infty$*

$$\widehat{\omega}_i^2 \Rightarrow \omega^2 P(b, k, H_{i,c}),$$

with  $P(b, k, H)$  as defined above in Definition 1 and

$$Z_{\alpha,i} \Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r) dW(r) + \frac{1}{2}(1 - P(b, k, H_{i,c}))}{\int_0^1 \widetilde{V}_c(r)^2 dr},$$

$$Z_{t,i} \Rightarrow -c \sqrt{\frac{\int_0^1 \widetilde{V}_c(r)^2 dr}{P(b, k, H_{i,c})}} + \frac{\int_0^1 \widetilde{V}_c(r) dW(r) + \frac{1}{2}(1 - P(b, k, H_{i,c}))}{\sqrt{P(b, k, H_{i,c}) \int_0^1 \widetilde{V}_c(r)^2 dr}}.$$

The proposition shows that under fixed- $b$  asymptotics the asymptotic null distributions of the PP tests exhibit certain distinct features. First, the limits are non-pivotal given the dependence on  $\sigma^2$  and  $\omega^2$  via the dependence on  $H_{i,c}(r)$ . This asymptotic result indicates that the finite sample performance of the PP tests will be sensitive to the serial correlation structure in  $u_t$  even for moderate to large sample sizes, which matches the well known finite sample problems of the PP statistics documented in the literature. Second, the fixed- $b$  limits are different for the one- and two-step approaches, with these differences occurring via  $\widehat{\omega}_i^2$ . Third, as is common when using fixed- $b$  asymptotics, the choice of bandwidth and kernel exert influence even asymptotically as reflected by  $P(b, k, H_{i,c})$ . Notice that when  $P(b, k, H_{i,c}) = 1$  the standard PP asymptotic distributions are obtained, which is exactly as expected since using a consistent estimator  $\widehat{\omega}^2$  exactly coincides with  $P(b, k, H_{i,c}) = 1$ .

### 3 Modified PP Unit Root Tests

The reason for the non-pivotal limits of the PP tests under fixed- $b$  asymptotics is that the scaled partial sums of the residuals  $\widehat{u}_{t,i}$  are not, as has been shown in Lemma 1, directly proportional to  $\omega^2$ , due to the dependencies in the processes  $H_{i,c}$ . It is, however, straightforward to modify the residuals  $\widehat{u}_{t,i}$  in order to obtain the needed asymptotic proportionality for a pivotal fixed- $b$  limit result. For obtaining this result it is in fact sufficient to construct residuals using modified estimators of  $\alpha$ . Define the modified estimators as

$$\widehat{\alpha}_i^m := \widehat{\alpha}_i + \frac{\frac{1}{2}\widehat{\sigma}_i^2}{T^{-1} \sum_{t=2}^T \widetilde{y}_{t-1,i}^2},$$

for  $i \in \{1, 2\}$ .

Using the above modification, the one-step modified residuals can be written as

$$\begin{aligned}\widehat{u}_{t,1}^m &:= \widetilde{y}_{t,1} - \widehat{\alpha}_1^m \widetilde{y}_{t-1,1} = \widetilde{y}_{t,1} - \left( \widehat{\alpha}_1 + \frac{\frac{1}{2}\widehat{\sigma}_1^2}{T^{-1} \sum_{t=2}^T \widetilde{y}_{t-1,1}^2} \right) \widetilde{y}_{t-1,1} \\ &= \widetilde{u}_t - \left( \frac{T^{-1} \sum_{t=2}^T \widetilde{y}_{t-1,1} u_t + \frac{1}{2}\widehat{\sigma}_1^2}{T^{-2} \sum_{t=2}^T \widetilde{y}_{t-1,1}^2} \right) T^{-1} \widetilde{y}_{t-1,1}.\end{aligned}\quad (13)$$

For the two-step approach the modified residuals can be written as

$$\begin{aligned}\widehat{u}_{t,2}^m &:= \widetilde{y}_{t,2} - \widehat{\alpha}_2^m \widetilde{y}_{t-1,2} = \widetilde{y}_{t,2} - \left( \widehat{\alpha}_2 + \frac{\frac{1}{2}\widehat{\sigma}_2^2}{T^{-1} \sum_{t=2}^T \widetilde{y}_{t-1,2}^2} \right) \widetilde{y}_{t-1,2} \\ &= u_t - (D_t - D_{t-1})'(\widehat{\theta} - \theta) - \left( \frac{T^{-1} \sum_{t=2}^T \widetilde{y}_{t-1,2} u_t + \frac{1}{2}\widehat{\sigma}_2^2}{T^{-2} \sum_{t=2}^T \widetilde{y}_{t-1,2}^2} \right) T^{-1} \widetilde{y}_{t-1,2}.\end{aligned}\quad (14)$$

The following lemma gives the limit of the scaled partial sums of the modified residuals.

**Lemma 2** *Assume that the data are generated by (1) with the errors,  $u_t$ , fulfilling the FCLT (3). Consider the modified residuals,  $\widehat{u}_{t,i}^m$ , as given by (13) and (14) for the one- and two-step approaches. Then as  $T \rightarrow \infty$  it holds that*

$$T^{-1/2} \sum_{t=2}^{[rT]} \widehat{u}_{t,1}^m \Rightarrow \omega H_{1,c}^m(r), \quad H_{1,c}^m(r) := \widehat{W}(r) - \left( \frac{\int_0^1 \widetilde{V}_c(s) dW(s) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(s)^2 dr} \right) \int_0^r \widetilde{V}_c(s) ds, \quad (15)$$

$$T^{-1/2} \sum_{t=2}^{[rT]} \widehat{u}_{t,2}^m \Rightarrow \omega H_{2,c}^m(r), \quad H_{2,c}^m(r) := \widetilde{V}_c(r) - \left( \frac{\int_0^1 \widetilde{V}_c(s) dW(s) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(s)^2 dr} \right) \int_0^r \widetilde{V}_c(s) ds. \quad (16)$$

Thus, we see that the processes  $H_{i,c}^m(r)$  are free of nuisance parameters and the limit of the scaled partial sums of the modified residuals is proportional to  $\omega$ . These results now form the basis for modified PP tests using a fixed- $b$  estimate of  $\omega^2$  using the modified residuals rather than the original ones. We denote the corresponding estimators of the long run variance as  $\widetilde{\omega}_i^2$  in the sequel and using  $\widetilde{\omega}_i^2$  instead of  $\widehat{\omega}_i^2$  defines the modified PP tests.

It is interesting to note that for the coefficient test, the estimation of the long run variance can be avoided altogether using a very simple transformation based on the modified estimator of  $\alpha$  as follows:

$$Z_{\alpha,i}^* := T(\widehat{\alpha}_i^m - 1) = T(\widehat{\alpha}_i - 1) + \frac{\frac{1}{2}\widehat{\sigma}^2}{T^{-2} \sum_{t=2}^T \widetilde{y}_{t-1}^2}.$$

Because no estimate of  $\omega^2$  is required, there is no asymptotic distinction between one- and two-step detrending.

The following proposition provides the asymptotic limits of the modified statistics.

**Proposition 2** *Assume that the data are generated by (1) with the errors  $u_t$  fulfilling the FCLT (3). Furthermore assume that  $M = bT$ , with  $b \in (0, 1]$  fixed and the subscript  $i \in \{1, 2\}$  refers again to the one- and two-step approaches. Then as  $T \rightarrow \infty$*

$$Z_{\alpha,i}^* \Rightarrow -c + \frac{\int_0^1 \tilde{V}_c(r) dW(r) + \frac{1}{2}}{\int_0^1 \tilde{V}_c(r)^2 dr},$$

$$\tilde{\omega}_i^2 \Rightarrow \omega^2 P(b, k, H_{i,c}^m)$$

and

$$Z_{\alpha,i}^m \Rightarrow -c + \frac{\int_0^1 \tilde{V}_c(r) dW(r) + \frac{1}{2} \left(1 - P(b, k, H_{i,c}^m)\right)}{\int_0^1 \tilde{V}_c(r)^2 dr},$$

$$Z_{t,i}^m \Rightarrow -c \sqrt{\frac{\int_0^1 \tilde{V}_c(r)^2 dr}{P(b, k, H_{i,c}^m)}} + \frac{\int_0^1 \tilde{V}_c(r) dW(r) + \frac{1}{2} \left(1 - P(b, k, H_{i,c}^m)\right)}{\sqrt{P(b, k, H_{i,c}^m) \int_0^1 \tilde{V}_c(r)^2 dr}}.$$

Because the processes  $H_{i,c}^m(r)$  do not depend upon nuisance parameters, the processes  $P(b, k, H_{i,c}^m)$  are free of nuisance parameters which leads to pivotal fixed- $b$  limiting distributions of the modified PP statistics. Thus, under the null hypothesis of a unit root ( $c = 0$ ), critical values can be simulated for given deterministic components,  $b$  and kernel function. These are, as already mentioned in the introduction, available in a supplement for five kernels (Bartlett, Bohman, Daniell, Parzen and Quadratic Spectral) for the specifications without deterministic component, with intercept only and with intercept and linear trend. For the latter two specifications of the deterministic component the fixed- $b$  critical values differ between one- and two-step detrending. The values for  $b$  in the tables range from 0.02 to 1 with a mesh of size 0.02.<sup>3</sup>

Because the focus of this paper is to understand the impact of kernel and bandwidth choices and detrending method used on the performance of the PP tests and the modified versions  $Z_{\alpha,i}^m$  and  $Z_{t,i}^m$ , we do not further analyze the test  $Z_{\alpha,i}^*$ . The  $Z_{\alpha,i}^*$  test is unaffected in all samples by choice of  $b$  and kernel and is asymptotically unaffected (for our deterministic) by the detrending method, although in finite samples the detrending method will generate small differences. A detailed analysis of the performance of the  $Z_{\alpha,i}^*$  test will be performed elsewhere.

For non-zero values of  $c$  we can use the results of Proposition 2 to compute local asymptotic power (LAP) of the modified statistics. Because the limits in Proposition 2 depend on the kernel, bandwidth and form of detrending, we can use LAP to make predictions about the impact of kernel, bandwidth and detrending choices on finite sample power. We simulate LAP for the same set of five kernels and a selection of values of  $b$ . The simulations are performed using partial sums of 1,000

---

<sup>3</sup>Also `MATLAB` code to compute the modified test statistics as well as to perform inference using the fixed- $b$  critical values is available from the authors upon request.

i.i.d.  $N(0, 1)$  random errors to approximate the Wiener process that drives the limits in Proposition 2 and 10,000 replications. Local asymptotic power is computed for a grid over  $c$  running from 0 to 80 with steps of size 2. Rejections are computed using the  $c = 0$  asymptotic critical value for a given kernel, bandwidth, detrending combination.

We report results for the Bartlett and QS kernels. Results for the other kernels are qualitatively similar. Figures 1-4 plot LAP for the intercept only model whereas Figures 5-8 plot LAP for the intercept and linear trend model. The first notable pattern in these figures is the sensitivity of power to the choice of  $b$ . In many cases power decreases as  $b$  increases, but in other cases power is non-monotonic in  $b$ . For example, in the intercept and linear trend model,  $Z_t^m$  using the QS kernel has good power when  $b = 0.02$ , but power drops very quickly when  $b$  is increased to 0.1. Then, as  $b$  is increased further, power increases but stays well below power when  $b = 0.02$ . The second notable pattern is that, except for  $b = 0.02$ , there are clear differences in power between the one and two-step detrending approaches with the greatest differences occurring for larger values of  $b$ . The third notable pattern is that power is often very low, often close to zero, when  $b$  is not small and  $c$  takes on small to medium values. The fourth notable pattern is that the kernel matters for power unless  $b = 0.02$  in which case power is similar for both kernels in all cases. Finally, there are substantial differences in power between the  $Z_\alpha^m$  and  $Z_t^m$  statistics except when  $b = 0.02$ . The local asymptotic power analysis suggests that small bandwidths are much preferable to non-small bandwidths.

#### 4 Finite Sample Behavior

For the sake of brevity we only include a small selection of finite sample results obtained by performing extensive simulations. In particular we only report some results for the sample size  $T = 200$  for the QS kernel for the  $t$ -statistic tests. Qualitatively similar results are available also for the coefficient tests, sample size  $T = 100$  and the Bartlett kernel. The number of replications is 5,000 for each experiment. The selected results, for a narrow set of statistics and nuisance parameters, are meant to be illustrative in capturing the main observations in relation to the finite sample predictions of the asymptotic theory. In particular, the fixed- $b$  theory suggests that under the unit root null hypothesis: **i)** the traditional PP unit root tests will be sensitive to nuisance parameters and the choice of kernel and bandwidth, **ii)** the modified PP tests will be more robust to nuisance parameters and will be robust to the choice of kernel and bandwidth when the asymptotic fixed- $b$  critical value is used, and **iii)** there can be a difference between the one- and two-step detrending approaches.

The data generating process is given by (1) and (2) where we set  $\theta = \mathbf{0}$  without loss of generality. We generate data for  $u_t$  according to the  $ARMA(1, 1)$  model:  $u_t = \rho u_{t-1} + \varepsilon_t + \varphi \varepsilon_{t-1}$  where  $\varepsilon_t$  is a sequence of i.i.d.  $N(0, 1)$  random variables for  $t = 0, 1, \dots, T$  and  $u_0 = 0$ . The standard PP test uses

the usual unit root asymptotic distribution critical values and is labeled  $PP$ . The modified PP test, labeled  $PP(fb)$ , uses the fixed- $b$  asymptotic critical values corresponding to the limits given in Proposition 2 for  $c = 0$ .

In Figures 9–16 we display the empirical null rejection probabilities at the 5% nominal level. The results are reported for a grid of bandwidths given by  $M = 2, 4, 6, \dots, 198, 200$ , indexed by the corresponding value of  $b = M/200$ , with this grid corresponding to the grid for which fixed- $b$  critical values are available. In all of these figures except for Figure 9 (only intercept) an intercept and a linear time trend are included. In each figure, the rejections are reported for both the one- and two-step detrending. Figures 9 and 10 give results for the case where  $u_t$  has no serial correlation ( $\rho = 0, \varphi = 0$ ). Figure 9 gives results for the intercept case and Figure 10 gives results for the intercept and linear trend case. Several patterns stand out in the figures. The rejections of the  $PP(fb)$  tests are close to 0.05 regardless of the bandwidth which indicates that the fixed- $b$  critical values are doing an adequate job of capturing the dependence of the finite sample distribution on the bandwidth. In contrast, unless a small bandwidth is used, the  $PP$  tests have rejections that are not close to 0.05 and the rejections show a sensitivity to the bandwidth. This is consistent with the predictions of Proposition 1. Comparing the one-step with two-step approaches, we see that for the  $PP$  tests, there are stark differences in rejection probabilities between the two approaches as predicted by Proposition 1. In contrast, the  $PP(fb)$  tests have similar rejections for both one-step and two-step detrending, which is consistent with the fact that the fixed- $b$  limits of the  $PP(fb)$  statistics depend on whether one-step or two-step detrending is used.

Figures 11 and 12 give results for MA errors with  $\varphi = 0.4, -0.4$ . Rejections of the  $PP$  statistics are systematically different from 0.05 regardless of the bandwidth. This is consistent with the dependence on serial correlation nuisance parameters indicated by the fixed- $b$  limits of the  $PP$  tests. There is a noticeable difference between the one-step and two-step approaches, especially when  $\varphi = 0.4$ . The rejections of the  $PP$  statistics are even more distorted in the  $\varphi = -0.4$  case and substantial over-rejections are possible for some bandwidths. The rejections of the  $PP(fb)$  statistics are very different. When  $\varphi = 0.4$ , rejections are close to 0.05 regardless of bandwidth. In contrast, when  $\varphi = -0.4$ , rejections are inflated above 0.05 for the  $PP(fb)$  statistics. This general tendency to over-reject when there is a negative MA component is well documented in the literature, see Perron and Ng (1996).

Figures 13 and 14 give results for AR errors with  $\rho = 0.4, -0.4$ . Similar patterns hold as in the MA case with patterns in the  $\rho = 0.4$  case being very similar to the  $\varphi = 0.4$  case and likewise for the  $\rho = -0.4$  and  $\varphi = -0.4$  cases. Intuitively, as either  $\rho$  or  $\varphi$  approach -1, the spectral density at frequency zero of  $u_t$  approaches zero and this counteracts the unit root in  $y_t^0$  leading to over-rejections. Figures 15 and 16 give some results for ARMA errors. When both  $\rho$  and  $\varphi$  are positive, rejections of the  $PP(fb)$  statistics are close to 0.05 regardless of bandwidth whereas rejections of

the  $PP$  statistics are highly distorted. When both  $\rho$  and  $\varphi$  are negative<sup>4</sup>, all statistics tend to over-reject substantially.

We now turn to some limited finite sample power results to assess the adequacy of the LAP results for the  $PP(fb)$  statistics. We only report results for  $\rho = 0$ ,  $\varphi = 0$  and power is size adjusted in all cases. Figure 17 depicts power of the  $Z_\alpha^m$  statistics and Figure 18 depicts power for the  $Z_t^m$  statistics for the intercept only model. Figures 19 and 20 give results for the intercept and linear trend model. We use the same values of  $b$  as for the LAP results. The general patterns in Figures 17-20 are similar to the LAP results. Power is highest for  $b = 0.02$  and power can be much lower for other values of  $b$ . There are noticeable differences in power between the one- and two-step approaches and there are also clear differences in power between the  $Z_\alpha^m$  and  $Z_t^m$  statistics. The one notable difference between the LAP power curves and finite sample power curves is that power with  $T = 200$  is generally higher than what is predicted by the LAP analysis. The values of  $\alpha = 0.9, 0.8, 0.7, 0.6$  with  $T = 200$  correspond to values of  $c = 20, 40, 60, 80$ . Comparing Figure 8 with Figure 20, notice that power with  $c = 40$  is very low in Figure 8 (except for  $b = 0.02$ ) whereas power with  $\alpha = 0.8$  is much higher in Figure 20. While the LAP analysis adequately captures the general patterns of power with respect to dependency on bandwidth, model, one- and two-step detrending, etc., the accuracy of the magnitudes of power predicted by the LAP analysis is not so impressive.

## 5 Summary and Conclusions

The fixed- $b$  theory developed in this paper provides an alternative theoretical explanation for the finite sample dependence of the traditional PP unit root tests on serial correlation in the errors driving the unit root. Unlike the traditional consistency approximation for the long run variance estimators used in the PP tests, the fixed- $b$  theory also indicates a finite sample difference between one-step and two-step detrending. Both local asymptotic power simulations as well as finite sample simulations show that there can be large differences between the one-step and two-step approaches in terms of both null rejection probabilities and power. We propose modified PP statistics that have asymptotically pivotal fixed- $b$  limits. The fixed- $b$  limits depend on the kernel and bandwidth used for the long run variance estimators and the fixed- $b$  limits are different for the one-step and two-step approaches. In finite samples, the modified PP statistics, when used with fixed- $b$  critical values, have null rejections close to the nominal level unless the serial correlation in the innovations to the unit root process behave similarly to an over-differenced stationary process. While the modified PP tests are a clear improvement over the traditional PP tests, further modifications would probably be necessary to make them reliable in practice in terms of adequate size control.

---

<sup>4</sup>We do not report results where  $\rho$  and  $\varphi$  are equal with opposite signs. In this case, the AR and MA components cancel making  $u_t$  uncorrelated and patterns are essentially identical to those depicted in Figure 10.

## 6 Appendix: Proofs

### Proof of Lemma 1

With both methods of detrending, the residuals can be written as

$$\hat{u}_{t,i} = \tilde{y}_{t,i} - \alpha \tilde{y}_{t-1,i} - T(\hat{\alpha}_i - \alpha) \frac{\tilde{y}_{t-1}}{T},$$

with differences occurring asymptotically only in the first part  $\tilde{y}_{t,i} - \alpha \tilde{y}_{t-1,i}$ .

We first consider one-step detrending, in which case we obtain

$$\tilde{y}_{t,1} - \alpha \tilde{y}_{t-1,1} = \tilde{u}_t = u_t - D'_t(D'_T D_T)^{-1} D'_T U_T.$$

This implies that  $T^{-1/2} \sum_{t=2}^{[rT]} \tilde{y}_{t,1} - \alpha \tilde{y}_{t-1,1} \Rightarrow \omega \widehat{W}(r) dr$ . Under the assumptions stated it is well known that  $T(\hat{\alpha}_1 - \alpha) \Rightarrow \left( \frac{\omega^2 \int_0^1 \tilde{V}_c(s) dW(s) + \lambda}{\omega^2 \int_0^1 \tilde{V}_c(s)^2 ds} \right)$  and that  $T^{-3/2} \sum_{t=2}^{[rT]} \tilde{y}_{t-1,1} \Rightarrow \omega \int_0^r \tilde{V}_c(s) ds$ . Combining these three results establishes (11).

Let us now look at two-step detrending. In this case we can write

$$\begin{aligned} \tilde{y}_{t,2} - \alpha \tilde{y}_{t-1,2} &= y_t - D'_t \hat{\theta} - \alpha(y_{t-1} - D'_{t-1} \hat{\theta}) \\ &= y_t^0 + D'_t \theta - D'_t \hat{\theta} - \alpha(y_{t-1}^0 + D'_{t-1} \theta - D'_{t-1} \hat{\theta}) \\ &= y_t^0 - \alpha y_{t-1}^0 - D'_t (\hat{\theta} - \theta) + \alpha D'_{t-1} (\hat{\theta} - \theta) \\ &= u_t - D'_t (\hat{\theta} - \theta) + (1 - cT^{-1}) D'_{t-1} (\hat{\theta} - \theta) \\ &= u_t - (D_t - D_{t-1})' (\hat{\theta} - \theta) - cT^{-1} D'_{t-1} (\hat{\theta} - \theta) \\ &= u_t - (D_t - D_{t-1})' (D'_T D_T)^{-1} D'_T Y_T^0 - cT^{-1} D'_{t-1} (D'_T D_T)^{-1} D'_T Y_T^0 \end{aligned}$$

with  $Y_T^0 = [y_2^0, \dots, y_T^0]'$ . Defining  $G_D = \text{diag}(1, T, \dots, T^q)$ , straightforward calculations give the standard results

$$\begin{aligned} T^{-1/2} G_D (D'_T D_T)^{-1} D'_T Y_T^0 &= (T^{-1} G_D^{-1} D'_T D_T G_D^{-1})^{-1} T^{-3/2} G_D^{-1} D'_T Y_T^0 \\ &\Rightarrow \omega \left( \int_0^1 D(s) D(s)' ds \right)^{-1} \int_0^1 D(s) V_c(s) ds, \end{aligned}$$

and

$$T^{-1} \sum_{t=2}^{[rT]} D'_{t-1} G_D^{-1} \Rightarrow \int_0^r D(s)' ds.$$

Less standard, but straightforward, is the result

$$\sum_{t=2}^{[rT]} (D_t - D_{t-1})' G_D^{-1} \Rightarrow \int_0^r \dot{D}(s)' ds.$$

Using these limits and (3), it easily follows that  $T^{-1/2} \sum_{t=2}^{[rT]} (\tilde{y}_{t,2} - \alpha \tilde{y}_{t-1,2}) \Rightarrow \omega \tilde{V}_c(r)$ . Combining this result with the unchanged results (compared to the one-step detrending) for the other two components establishes (12).

### Proof of Proposition 1

The results of the proposition follow from the asymptotic result for the partial sum processes of the residuals established in Lemma 1, using similar arguments as in Hashimzade and Vogelsang (2008). Once the fixed- $b$  limits for  $\widehat{\omega}_i^2$ ,  $i \in \{1, 2\}$  are established, the fixed- $b$  limit distributions of the test statistics follow from using these when calculating the limiting distributions of the test statistics, as given in (9) and (10), with these expressions referring to one-step detrending and the two-step detrending versions of the test statistics similarly defined.

### Proof of Lemma 2

The proof of the lemma builds heavily on the proof of Lemma 1, with in fact only the term comprising  $\widehat{\alpha}_i$  being different in the expressions for the modified residuals compared to the residuals previously considered. For both detrending approaches it holds that

$$\begin{aligned} T(\widehat{\alpha}_i^m - \alpha) &= T(\widehat{\alpha}_i - \alpha) + \frac{\frac{1}{2}\widehat{\sigma}_i^2}{T^{-2}\sum_{t=2}^T \widetilde{y}_{t-1,i}^2} \\ &\Rightarrow \frac{\omega^2 \int_0^1 \widetilde{V}_c(r)dW(r) + \lambda}{\omega^2 \int_0^1 \widetilde{V}_c(r)^2 dr} + \frac{\frac{1}{2}\sigma^2}{\omega^2 \int_0^1 \widetilde{V}_c(r)^2 dr} \\ &= \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(r)^2 dr}, \end{aligned}$$

from which the results of the lemma follow, since all other terms are unchanged compared to Lemma 1.

### Proof of Proposition 2

Using the result from Lemma 2 the limit of  $Z_{\alpha,i}^*$  follows immediately:

$$\begin{aligned} T(\widehat{\alpha}_i^m - 1) &= T(\widehat{\alpha}_i^m - \alpha + \alpha - 1) = T(\widehat{\alpha}_i^m - \alpha) + T(\alpha - 1) \\ &= T(\widehat{\alpha}_i^m - \alpha) - c \\ &\Rightarrow -c + \frac{\int_0^1 \widetilde{V}_c(r)dW(r) + \frac{1}{2}}{\int_0^1 \widetilde{V}_c(r)^2 dr}. \end{aligned}$$

The other results follow, similar to the results of Proposition 1, from the asymptotic result for the partial sum processes of the modified residuals established in Lemma 2 using now the fixed- $b$  limit for the modified long run variance estimators  $\widetilde{\omega}_i^2$  in place of  $\widehat{\omega}_i^2$ .

## References

- Hashimzade, N. and T.J. Vogelsang (2008). Fixed-b Asymptotic Approximation of the Sampling Behavior of Nonparametric Spectral Density Estimators. *Journal of Time Series Analysis* **29**, 142–162.
- Jansson, M. (2002). Consistent Covariance Matrix Estimation for Linear Processes. *Econometric Theory* **18**, 1449–1459.
- Kiefer, N.M. and T.J. Vogelsang (2005). A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests. *Econometric Theory* **21**, 1130–1164.
- Perron, P. and S. Ng (1996). Useful Modifications to some Unit Root Tests with Dependent Errors and their Local Asymptotic Properties. *Review of Economic Studies* **63**, 435–463.
- Perron, P. and T.J. Vogelsang (1992). Testing for a Unit Root in a Time Series with a Changing Mean: Corrections and Extensions. *Journal of Business and Economic Statistics* **10**, 467–470.
- Phillips, P.C.B. (1987). Towards a Unified Asymptotic Theory for Autoregression. *Biometrika* **74**, 535–547.
- Phillips, P.C.B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* **75**, 335–346.
- Phillips, P.C.B. and V. Solo (1992). Asymptotics for Linear Processes. *The Annals of Statistics* **20**, 971–1001.
- Schwert, W. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics* **7**, 147–159.
- Sims, C.A., J.H. Stock and M.W. Watson (1990). Inference in Linear Time Series Models with some Unit Roots. *Econometrica* **58**, 113–144.

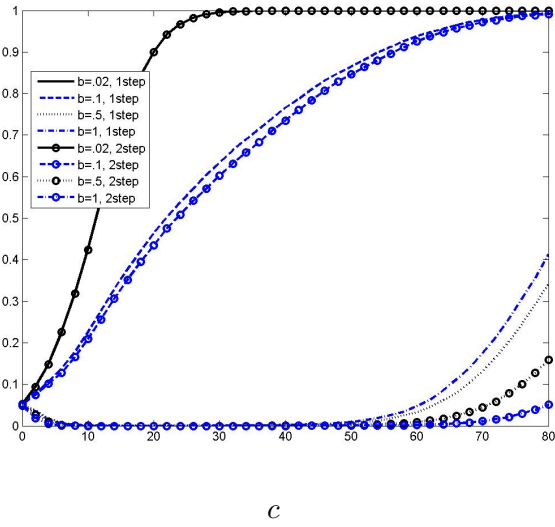


Figure 1: Local Asymptotic Power,  $Z_{\alpha}^m$ , Bartlett Intercept Only

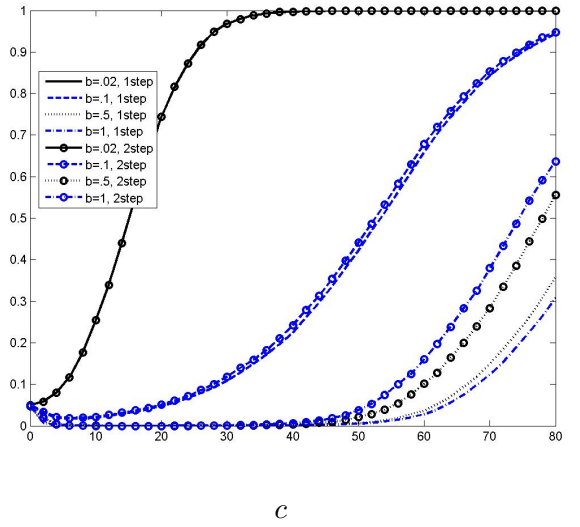


Figure 2: Local Asymptotic Power,  $Z_t^m$ , Bartlett Intercept Only

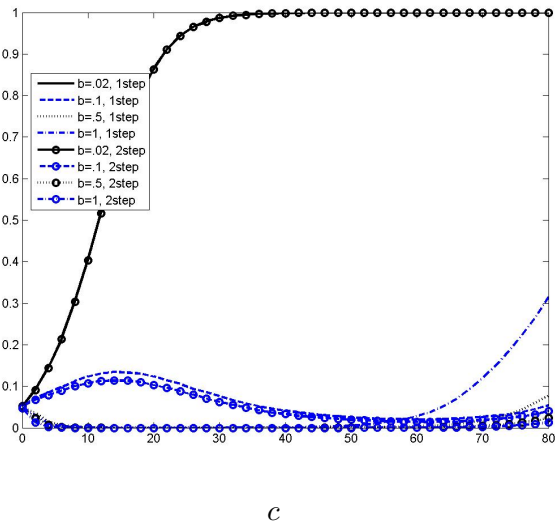


Figure 3: Local Asymptotic Power,  $Z_{\alpha}^m$ , QS Intercept Only

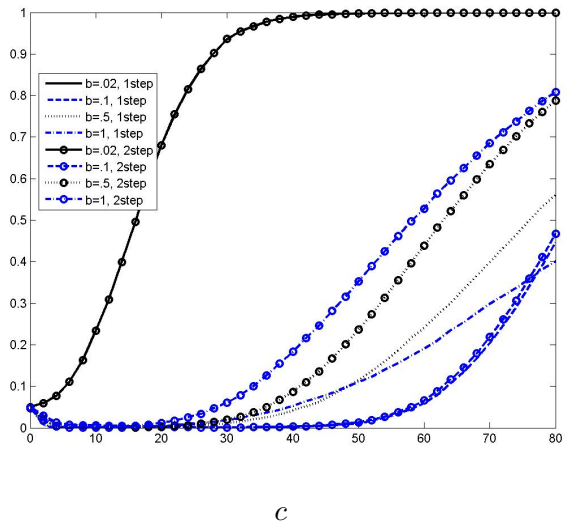


Figure 4: Local Asymptotic Power,  $Z_t^m$ , QS Intercept Only

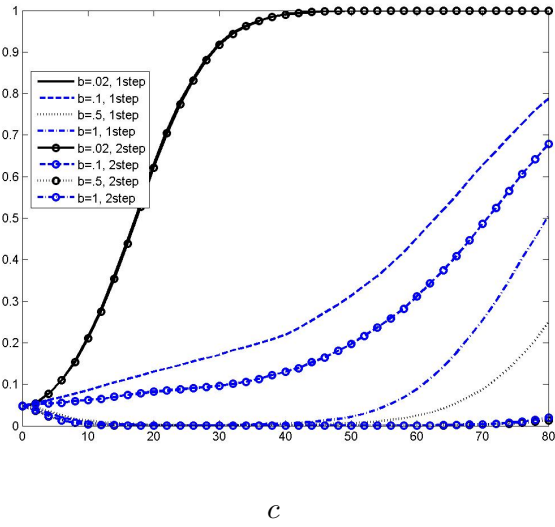


Figure 5: Local Asymptotic Power,  $Z_{\alpha}^m$ , Bartlett Intercept+Linear Trend

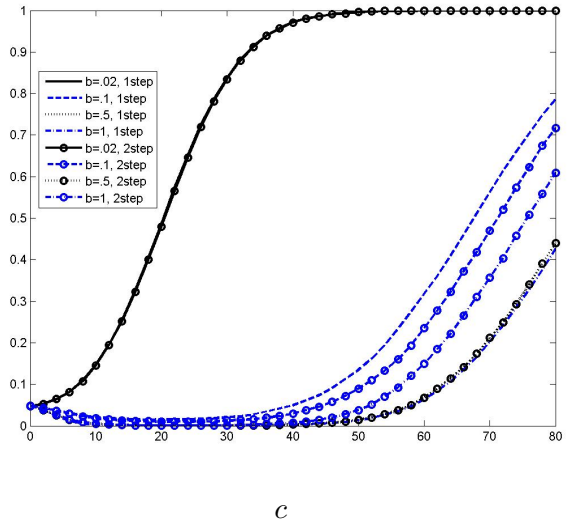


Figure 6: Local Asymptotic Power,  $Z_t^m$ , Bartlett Intercept+Linear Trend

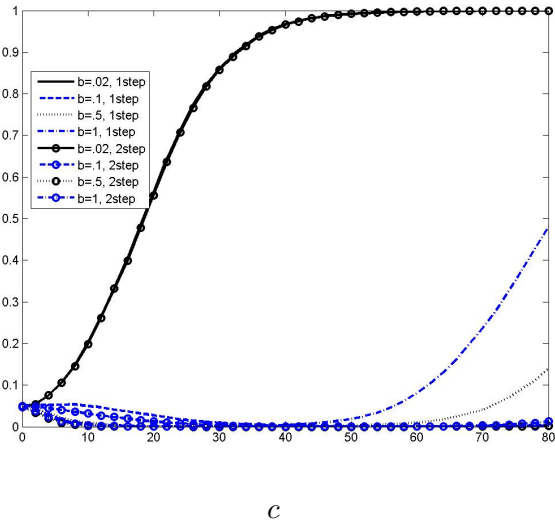


Figure 7: Local Asymptotic Power,  $Z_{\alpha}^m$ , QS Intercept+Linear Trend

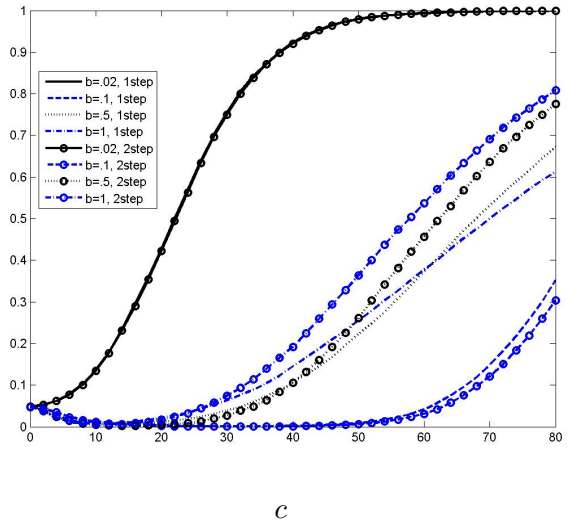


Figure 8: Local Asymptotic Power,  $Z_t^m$ , QS Intercept+Linear Trend

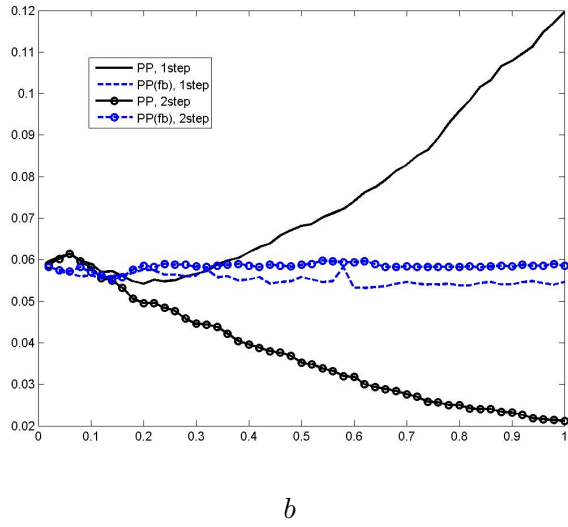


Figure 9: Empirical Null Rejections,  $Z_t^m$ , QS Intercept Only,  $\rho = 0$ ,  $\varphi = 0$

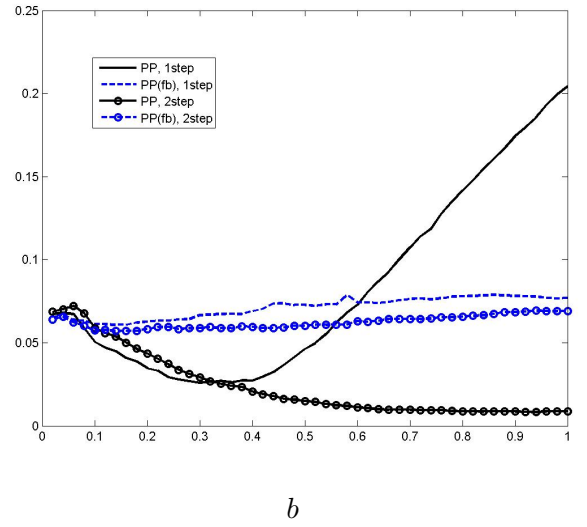


Figure 10: Empirical Null Rejections,  $Z_t^m$ , QS Intercept+Linear Trend,  $\rho = 0$ ,  $\varphi = 0$

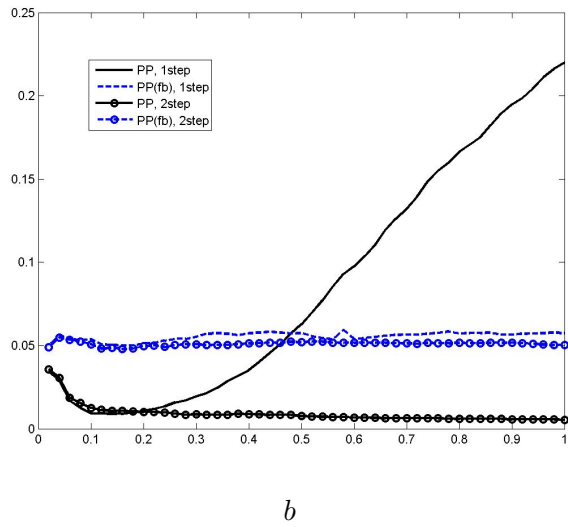


Figure 11: Empirical Null Rejections,  $Z_t^m$ , QS Intercept+Linear Trend,  $\rho = 0$ ,  $\varphi = 0.4$

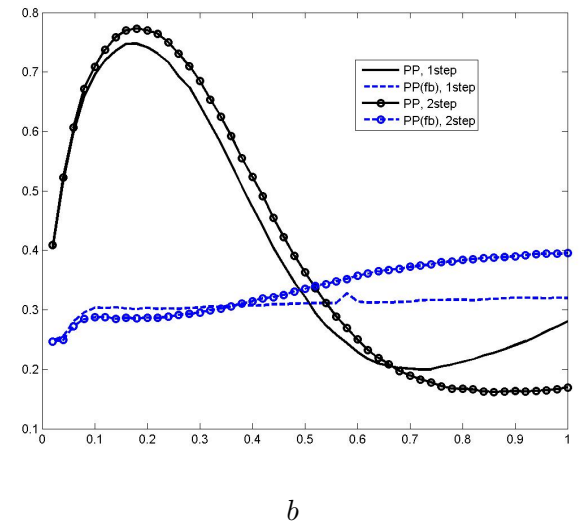


Figure 12: Empirical Null Rejections,  $Z_t^m$ , QS Intercept+Linear Trend,  $\rho = 0$ ,  $\varphi = -0.4$

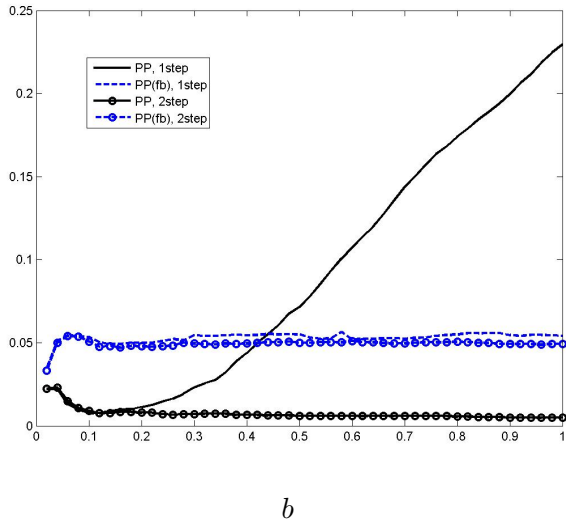


Figure 13: Empirical Null Rejections,  $Z_t^m$ , QS Inter.+Lin. Trend,  $\rho = 0.4$ ,  $\varphi = 0$ ,  $T = 200$

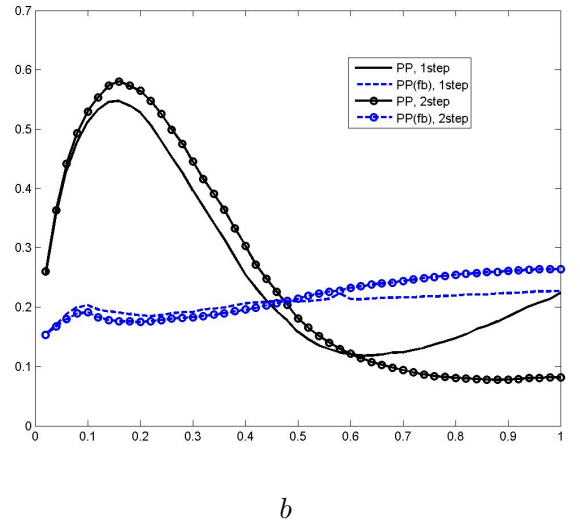


Figure 14: Empirical Null Rejections,  $Z_t^m$ , QS Inter.+ Lin. Trend,  $\rho = -0.4$ ,  $\varphi = 0$ ,  $T = 200$

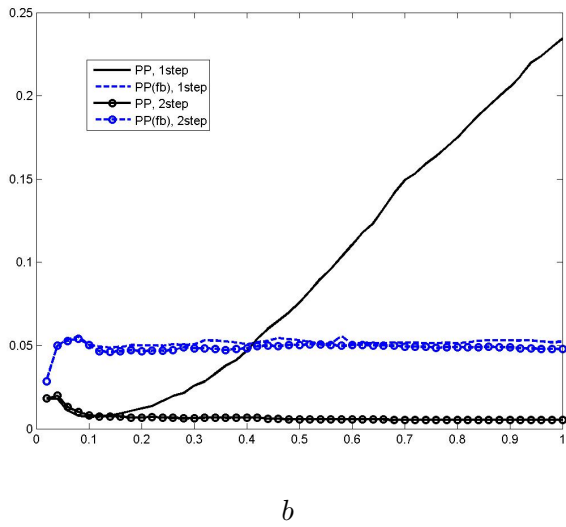


Figure 15: Empirical Null Rejections,  $Z_t^m$ , QS Inter.+Lin. Trend,  $\rho = 0.4$ ,  $\varphi = 0.4$ ,  $T = 200$

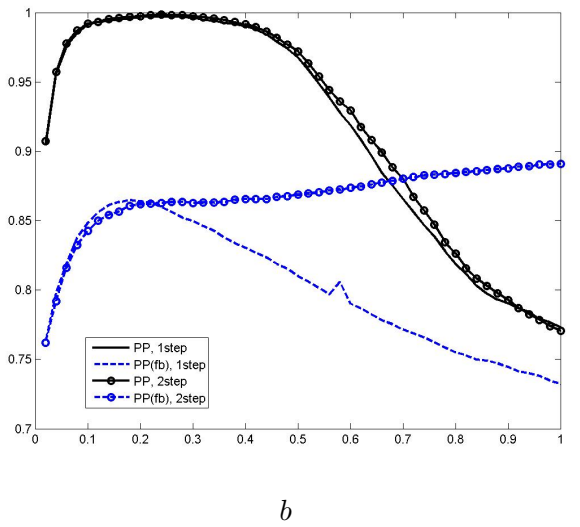


Figure 16: Empirical Null Rejections,  $Z_t^m$ , QS Inter.+Lin. Trend,  $\rho = -0.4$ ,  $\varphi = -0.4$ ,  $T = 200$

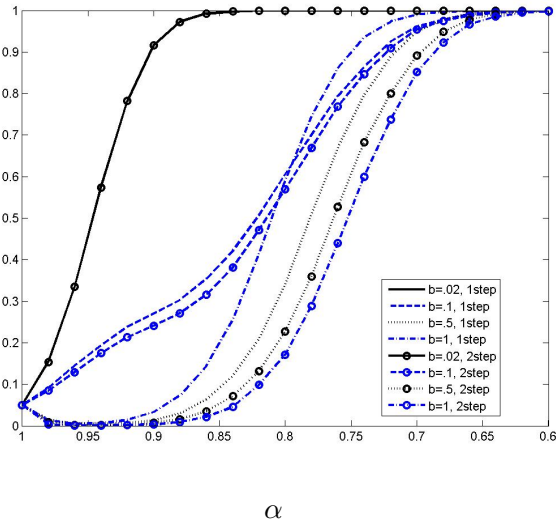


Figure 17: Finite Sample Power (size adjusted)  
 $Z_{\alpha}^m$ , QS, Intercept Only,  $\rho = 0$ ,  $\varphi = 0$ ,  $T = 200$

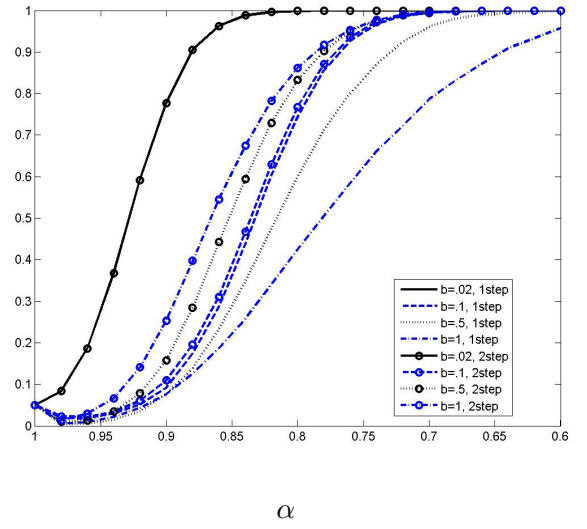


Figure 18: Finite Sample Power (size adjusted)  
 $Z_t^m$ , QS, Intercept Only,  $\rho = 0$ ,  $\varphi = 0$ ,  $T = 200$

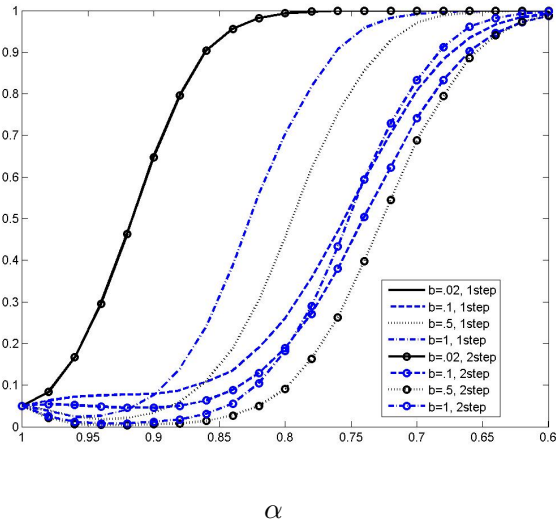


Figure 19: Finite Sample Power (size adjusted)  
 $Z_{\alpha}^m$ , QS, Inter.+Lin. Trend,  $\rho = 0$ ,  $\varphi = 0$ ,  $T = 200$

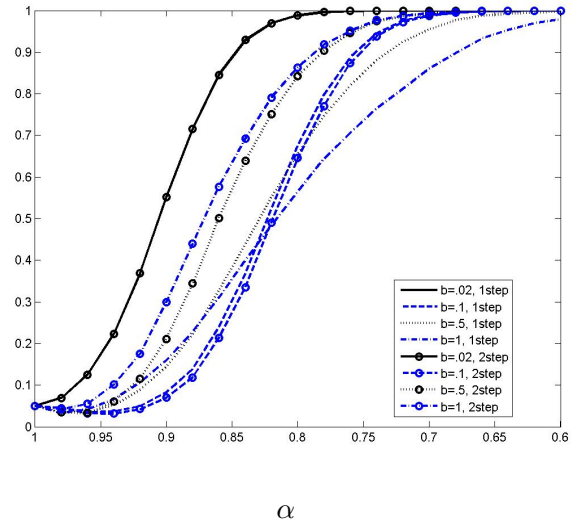


Figure 20: Finite Sample Power (size adjusted)  
 $Z_t^m$ , QS, Inter.+Lin. Trend,  $\rho = 0$ ,  $\varphi = 0$ ,  $T = 200$