Article:

Hualde, Javier and Iacone, Fabrizio orcid.org/0000-0002-2681-9036 (2015) Small-b and fixed-b asymptotics for weighted covariance estimation in fractional cointegration. Journal of Time Series Analysis. pp. 528-540. ISSN 1467-9892

https://doi.org/10.1111/jtsa.12113

Small-b and fixed-b asymptotics for weighted covariance estimation in fractional cointegration*

Javier Hualde
Universidad Pública de Navarra

Fabrizio Iacone†
University of York

November 14, 2014

Abstract

In a standard cointegrating framework, Phillips (1991) introduced the weighted covariance (WC) estimator of cointegrating parameters. Later, Marinucci (2000) applied this estimator to various fractional circumstances and, like Phillips (1991), analyzed the so-called small-b asymptotic approximation to its sampling distribution. Recently, an alternative limiting theory has been successfully employed to approximate the sampling distribution of nonparametric estimators of spectral densities and long run covariance matrices more accurately than by traditional asymptotics. This has been named fixed-b asymptotics, and the particular form of the WC estimator makes it an ideal candidate for the application of this type of theory. Thus, in this paper we derive the fixed-b limit of WC estimators in a fractional setting, filling also some gaps in the traditional (small-b) theory. Additionally, we compare the small-b and fixed-b limiting approximations to the sampling distribution of a WC estimator by means of a Monte Carlo experiment, finding that the fixed-b limit is more accurate.

Keywords: Fractional cointegration, first stage methods, small-b and fixed-b asymptotic theory.

JEL classification: C32

*We thank Peter Phillips and Robert Taylor for helpful suggestions and discussions. We are also grateful for the comments of two referees which have led to crucial improvements in the paper. Javier Hualde's research is supported by the Spanish Ministerio de Ciencia e Innovación through ECO2011-24304.

†Corresponding author: Fabrizio Iacone, Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, UK.

1. INTRODUCTION

Fixed-b asymptotics have been employed to approximate the sampling distribution of nonparametric estimators of spectral densities and long-run covariance matrices more accurately than by traditional asymptotics. These estimators depend typically on the sample size T and on a bandwidth M. The idea, introduced by Neave (1970), is to analyze the limit of the estimators when $M/T \to b \in (0, 1]$ as $T \to \infty$. This is because, although setting $M/T \to 0$ makes it possible to obtain consistent estimators of the spectral density (or of the long run variance) of weakly dependent processes, in any practical situation a non-zero fraction M/T is used: fixing the proportion M/T could therefore yield a better approximation to the limit distribution of the estimator. Kiefer and Vogelsang (2005) derived the fixed-b limit of the estimator of the long run variance and showed that in most scenarios it represents an improvement over the approximation obtained by setting $M/T \to 0$ (the so-called small-b asymptotics). Bunzel and Vogelsang (2005) showed that the fixed-b approximation to the long run variance is also convenient when making inference on a trend in presence of a potential (but not certain) unit root, because the Wald statistic has non-degenerate limit distribution regardless of the effective existence of said unit root. Iacone, Leybourne and Taylor (2013) exploited this self-normalization property of the Wald statistic in the context of testing for a break in the slope in the presence of residuals that may also be fractionally integrated.

Focusing on the single-equation standard cointegration setting (with unit root observables and weak dependent cointegrating errors), few works have employed fixed-b asymptotics in the cointegration literature. First, Bunzel (2006) analyzed the fixed-b limit of a Wald test statistic based on the dynamic ordinary least squares (DOLS) estimator of the cointegration parameter and on a weighted covariance estimator of the corresponding long-run variance, ignoring, however, the impact of lead and lag choices on the implementation of the DOLS. Jin, Phillips and Sun (2006) derived a fixed-b theory for tests based on the fully modified-OLS (FM-OLS) estimator, although requiring standard consistency results which, in particular, ignore the impact of choices of tuning parameters. Thus the type of fixed-b theory developed by these two papers has been denoted by Vogelsang and Wagner (2014) as "partial". In fact, Vogelsang and Wagner (2014) developed the "complete" fixed-b limit of FM-OLS, which depends on a complicated manner upon nuisance parameters. Additionally, they proposed a new estimator of the cointegration parameter (integrated modified OLS), which does not depend on choices of tuning parameters, and discussed fixed-b inference for Wald statistics based on this estimator.

In the spirit of Vogelsang and Wagner (2014), we examine in the present paper the fixed-b approximation to the distribution of an alternative estimator of the cointegration parameter: the weighted covariance (WC) estimator proposed by Phillips (1991) for the standard cointegration case, and later analyzed by Marinucci (2000) in various fractional circumstances. This estimator is motivated by the consideration that a cointegration parameter is a ratio between appropriate long-run covariance and variance, and therefore its estimation can be naturally based on weighted covariances. Thus, while Phillips (1991) and Marinucci (2000) analyzed the traditional (small-b) limiting approximation to the distribution of the estimator imposing (at least) $M = o(T)$, the form of the estimator opens the door to considering also its fixed-b approximation. In view of the evidence provided by the fixed-b literature, comparing the fixed-b approximation to the traditional, small-b, one appears to be an interesting exercise, which might shed additional light on the advantages of the fixed-b limits in terms of the accuracy.

The purpose of this paper is therefore twofold. First, we derive the fixed-b limits of two WC estimators in a fractional setting, filling also some gaps in the traditional (small-b) limit theory which were not covered by Marinucci (2000). This is of practical importance on its own because the WC estimator may be included in the class of first stage estimators (see Hualde and Iacone, 2012). Second, we compare the accuracy of the small-b and fixed-b approximations relative to the sampling distribution of one of the WC estimators by means of a Monte Carlo experiment. Nicely, we find that, at least in the different scenarios covered by our experiment, the conjecture supported by the literature that the fixed-b limit is more accurate is verified. These results might appear to be of limited empirical relevance, because the limit distribution of the WC estimator is not free of nuisance parameters and therefore it is not suitable for statistical inference. However, evidence showing that the fixed-b approximation is more accurate than the traditional one might support the empirical relevance of appropriate modifications of the WC estimator, which would lead to Wald test statistics with pivotal fixed-b limits. This appears to be especially relevant in view of the size problems displayed by the Wald test statistics based on second stage estimators (like those of Robinson and Hualde, 2003 or Hualde and Robinson, 2010), whose standard limiting distribution is pivotal. We address briefly this issue in Remark 6 below.

Incidentally, our results are connected to a related problem: the determination of the limiting behaviour of the narrow band least squares (NBLS) estimator with fixed bandwidth. The traditional limit theory (with bandwidth tending to ∞) is provided by Robinson and Marinucci (2001, 2003), who conjectured that a faster convergence rate (like that in Theorem 1 below) is attainable in certain circumstances by holding the bandwidth fixed in NBLS estimation. Chen and Hurvich (2003) verified this conjecture for their tapered NBLS estimator using differenced data, and we provide further (heuristic) evidence.

We introduce the WC estimators in Section 2, where we also derive their fixed-b limiting distributions, and show that they always achieve the fastest convergence rate in the class of first stage semiparametric estimators. In Section 3, we compare the small-b and fixed-b limiting approximations to the sampling distribution of a WC estimator by means of a Monte Carlo experiment. In Section 4, we conclude. The proofs of the theorems are given in the Appendix.

2. SMALL-b AND FIXED-b LIMITS OF WC ESTIMATORS

We consider a single-equation fractional cointegration framework. For $t = 0, \pm 1, ...,$, let $\eta_t = \left(\eta_{1,t}, \eta_{2,t}'\right)'$, prime denoting transposition, be a $p \times 1$ zero mean covariance stationary process such that $p \geq 2$, where $\eta_{1,t}$ is scalar and η_t has spectral density finite and nonsingular at all frequencies; define e_t and x_t as

$$e_t = \Delta^{-\delta_1} \left\{\eta_{1,t} 1(t > 0)\right\}, \; x_t = \mu + \Delta^{-\delta_2} \left\{\eta_{2,t} 1(t > 0)\right\}, \tag{1}$$

where μ is a generic constant vector, $1(S)$ denotes the indicator function, which takes value 1 if the statement S is true, 0 otherwise, and L is the lag operator, so that $\Delta = 1 - L$. For $d \geq 0$, Δ^{-d} can be expanded as $(1 - L)^{-d} = \sum_{t=0}^{\infty} \Delta_t^{(d)} L^t$, where $\Delta_t^{(d)} = \Gamma(t + d) / (\Gamma(d) \Gamma(t + 1))$, $\Gamma(.)$ being the Gamma function (with the conventions $\Gamma(0) = \infty$, $\Gamma(0) / \Gamma(0) = 1$). We assume that the random vector $(y_t, x_t')'$ is observable at $t = 1, ..., T$, and

$$y_t = \alpha + \nu' x_t + e_t, \text{ with } 0 \leq \delta_1 < \delta_2. \tag{2}$$

Note that both μ and α play an important role in the model. In particular μ allows x_t to potentially have a non-zero mean (if $\mu \neq 0$). Similarly, the presence of α gives flexibility to the model. For example, if α were not present, $\mu = 0$ would immediately imply that not only x_t has zero mean but also y_t. This might be restrictive, and we allow for more generality. Also, y_t and x_t are fractionally integrated (see, e.g., Hualde and Iacone, 2012). In particular, the individual components of x_t are $I(\delta_2)$ and, if $\nu \neq 0$, y_t is also $I(\delta_2)$, whereas e_t is $I(\delta_1)$. Furthermore, y_t and x_t are fractionally cointegrated, because the linear combination $y_t - \nu' x_t$ reduces the integration order of the observables. Notice that if $\nu = 0$, y_t is $I(\delta_1)$ and the cointegration is trivial.

Our assumptions imply that if $p > 2$, given the nonsingularity of the spectral density of $\eta_{2,t}$, the individual components of x_t cannot cointegrate (see, e.g., Nielsen and Frederiksen, 2011, p.83). Thus, (2) implies that the cointegrating rank is 1. Extensions to more complicated settings, allowing for higher cointegrating ranks and the possibility of multicointegration, can be accounted for as in Hualde and Robinson (2010) or Hualde and Iacone (2012). However, for simplicity, we just consider a single-equation model, which, in any case, is very standard in the literature (see, e.g., Marinucci, 2000, Robinson and Marinucci, 2001, Robinson and Hualde, 2003, Nielsen and Frederiksen, 2011).

In view of the truncations on the right-hand sides of (1), x_t and e_t are nonstationary processes. When $\delta_2 < 1/2$, it is possible to avoid the truncations in (1) and define $e_t^{\circ} = \Delta^{-\delta_1}\eta_{1,t}$, $x_t^{\circ} = \mu + \Delta^{-\delta_2}\eta_{2,t}$, so that processes x_t° and e_t° are stationary: x_t and e_t are usually referred to as Type II fractionally integrated processes, while x_t° and e_t° are Type I. Similarities and differences for these two types have been analyzed by, e.g., Marinucci and Robinson (1999). Notice that for $\delta_2 > 1/2$ a different truncation is still necessary to define the Type I x_t°, and this also holds true to define e_t° for $\delta_1 > 1/2$. We then prefer the notation for Type II fractionally integrated process because it allows a more uniform treatment. Admittedly, our model, which sets all initial conditions to zero (or constants), lacks empirical plausibility, but setting appropriately bounded initial values as in Johansen and Nielsen (2012a) leads equally to our Type II limiting results.

We introduce the WC estimators of ν. Let $k(x)$ be a kernel function satisfying $k(x) = k(-x)$, $k(0) = 1$, $|k(x)| \leq 1$, $k(x)$ continuous at $x = 0$ and $\int_0^{\infty} k(x)^2 dx < \infty$. For two generic sequences ξ_t, ζ_t, with sample means $\overline{\xi} = T^{-1}\sum_{t=1}^{T}\xi_t$, $\overline{\zeta} = T^{-1}\sum_{t=1}^{T}\zeta_t$, consider $\xi_t^* = \xi_t$, $\zeta_t^* = \zeta_t$, or $\xi_t^* = \xi_t - \overline{\xi}$, $\zeta_t^* = \zeta_t - \overline{\zeta}$, and define sample covariances $c_{\xi\zeta}^*(l) = T^{-1}\sum_{t=1}^{T-l}\xi_t^*\zeta_{t+l}^{*\prime}$ for $l \geq 0$; $= T^{-1}\sum_{t=1-l}^{T}\xi_t^*\zeta_{t+l}^{*\prime}$ for $l < 0$. Then if $\xi_t^* = \xi_t$, $\zeta_t^* = \zeta_t$, let $c_{\xi\zeta}(l) = c_{\xi\zeta}^*(l)$, whereas if $\xi_t^* = \xi_t - \overline{\xi}$, $\zeta_t^* = \zeta_t - \overline{\zeta}$, let $\widetilde{c}_{\xi\zeta}(l) = c_{\xi\zeta}^*(l)$. Define

$$\nu^* = \left(\sum_{l=-T+1}^{T-1} k(l/M)\, c_{xx}^*(l)\right)^{-1} \sum_{l=-T+1}^{T-1} k(l/M)\, c_{xy}^*(l), \tag{3}$$

where $1 \leq M \leq T$ and ν^* defines $\widehat{\nu}$ or $\widetilde{\nu}$, depending on whether sample covariances c or $\widetilde{c}$ are used, respectively. Note that $\widehat{\nu}$ is a simple multivariate extension of Marinucci's (2000) estimator, $\widetilde{\nu}$ accounting for the possibility that α in (2) might be nonzero. The parameter M is called bandwidth, and it may be a truncation lag in those kernels that are truncated.

We introduce some notation and regularity conditions to derive the fixed-b limiting approximation to the sampling distributions of $\widehat{\nu}$, $\widetilde{\nu}$.

Assumption 1. Let ε_t be independent and identically distributed (iid) $p \times 1$ vectors, with $E(\varepsilon_t) = 0$, $E(\varepsilon_t\varepsilon_t') = \Sigma$, where Σ is positive definite, and $E\|\varepsilon_t\|^q < \infty$ for $q > 2$. Let $\eta_t = A(L)\varepsilon_t$, where $A(s) = I_p + \sum_{j=1}^{\infty} A_j s^j$ (I_s is the s-rowed identity matrix), and the A_j are $p \times p$ matrices such that $\det(A(s)) \neq 0$, $|s| = 1$, and $A\left(e^{i\lambda}\right)$ is differentiable in λ with derivative in $Lip(\varrho)$, $\varrho > 1/2$.

Assumption 1 implies that the derivative of $A(e^{i\lambda})$ has Fourier coefficients $jA_j = O(j^{-\varrho})$ as $j \to \infty$, so in particular $\sum_{l=1}^{\infty} l^{1/2} \|A_l\| < \infty$. Let $\Omega = A(1) \Sigma A(1)'$, $G(r,s) = diag\left\{\Gamma^{-1}(\delta_1)(r-s)^{\delta_1 - 1}, \widetilde{1}'\Gamma^{-1}(\delta_2)(r-s)^{\delta_2-1}\right\}$, $\widetilde{1}$ being a $(p-1) \times 1$ vector of ones, $\delta = (\delta_1, \delta_2)'$. Let $B(r)$ be the p-dimensional Brownian motion with covariance matrix Ω and $B(r;\delta) = \int_0^r G(r,s)\, dB(s)$ be a Type II fractional Brownian motion, noting that, through $B(r)$, $B(r;\delta)$ depends on Ω. Partition $B(r;\delta) = \left(B_1(r;\delta_1), B_2'(r;\delta_2)\right)'$, $B_1(r;\delta_1)$, $B_2(r;\delta_2)$ collecting the first and last $p-1$ components of $B(r;\delta)$, respectively. For $N(\delta) = diag\left\{T^{-1/2-\delta_1}, \widetilde{1}'T^{-1/2-\delta_2}\right\}$, $u_t = (e_t, x_t' - \mu')'$, under Assumption 1 and $\delta_1 \geq 0$, $\delta_2 \geq 0$, Marinucci and Robinson (2000) derived the functional central limit theorem (FCLT)

$$N(\delta) \sum_{t=1}^{\lfloor rT \rfloor} u_t \Rightarrow \left(B_1\left(r;\delta_1^+\right), B_2'\left(r;\delta_2^+\right)\right)' \text{ for } r \in [0,1], \tag{4}$$

where $\Rightarrow$ denotes weak convergence and $\delta_i^+ = \delta_i + 1$, $i = 1,2$; note that (4) also holds for $-1/2 < \delta_1 < 0$, $-1/2 < \delta_2 < 0$, strengthening appropriately the moment conditions in Assumption 1 (see Johansen and Nielsen, 2012b).

As in Bunzel and Vogelsang (2005), we consider the following kernels: Type 1: $k(.)$ is twice continuously differentiable everywhere, with second derivative $k''(.)$; Type 2: $k(.)$ is twice continuously differentiable everywhere, except for $|x| = 1$; moreover, $k(x) = 0$ if $|x| > 1$. The second derivative is $k''(.)$; for $x \to 1$ define the derivative from the left at $x = 1$, $k'_-(1) = \lim_{h\to 0}((k(1) - k(1-h))/h)$; Type 3: $k(.)$ is the Bartlett kernel. Examples of Type 1 kernel are the Daniell and the Quadratic Spectral; examples of Type 2 are the Parzen and the Bohmann. Formulae for $k(.)$ and $k''(.)$ are given on pp. 393, 394 of Bunzel and Vogelsang (2005). Finally, define $\widetilde{B}_1(r;\delta_1) = B_1(r;\delta_1) - rB_1(1;\delta_1)$, $\widetilde{B}_2(r;\delta_2) = B_2(r;\delta_2) - rB_2(1;\delta_2)$.

Theorem 1. Under Assumption 1 and $0 \leq \delta_1 < \delta_2$, $M = bT$, $b \in (0,1]$,

$$T^{\delta_2 - \delta_1}(\widetilde{\nu} - \nu) \Rightarrow \left\{\widetilde{Q}_{xx}(b;\delta)\right\}^{-1} \widetilde{Q}_{xe}(b;\delta), \tag{5}$$

where

$$\widetilde{Q}_{xe}(b;\delta) = -\frac{1}{b^2}\int_0^1 \int_0^1 k''\left(\frac{r-s}{b}\right) \widetilde{B}_2\left(s;\delta_2^+\right) \widetilde{B}_1\left(r;\delta_1^+\right) drds$$

for Type 1 kernels,

$$\begin{aligned}\widetilde{Q}_{xe}(b;\delta) &= -\frac{1}{b^2}\int\int_{|r-s|<b} k''\left(\frac{r-s}{b}\right) \widetilde{B}_2\left(s;\delta_2^+\right) \widetilde{B}_1\left(r;\delta_1^+\right) drds \\ &\quad + \frac{1}{b}k'_-(1)\int_0^{1-b} \left(\widetilde{B}_2\left(r;\delta_2^+\right)\widetilde{B}_1\left(r+b;\delta_1^+\right) + \widetilde{B}_2\left(r+b;\delta_2^+\right)\widetilde{B}_1\left(r;\delta_1^+\right)\right) dr\end{aligned}$$

for Type 2 kernels, and

$$\begin{aligned}\widetilde{Q}_{xe}(b;\delta) &= \frac{2}{b}\int_0^1 \widetilde{B}_2\left(r;\delta_2^+\right)\widetilde{B}_1\left(r;\delta_1^+\right)dr \\ &-\frac{1}{b}\int_0^{1-b}\left(\widetilde{B}_2\left(r;\delta_2^+\right)\widetilde{B}_1\left(r+b;\delta_1^+\right)+\widetilde{B}_2\left(r+b;\delta_2^+\right)\widetilde{B}_1\left(r;\delta_1^+\right)\right)dr\end{aligned}$$

for the Barlett kernel; $\widetilde{Q}_{xx}(b;\delta)$ is defined by replacing $\widetilde{B}_1\left(r;\delta_1^+\right)$ by $\widetilde{B}_2'\left(r;\delta_2^+\right)$ in all the formulae for $\widetilde{Q}_{xe}(b;\delta)$ above; also if in (1), (2) $\delta_1 > 1/2$, or $0 \le \delta_1 \le 1/2$, $\delta_2 > 1/2$ and $\alpha = 0$, or $0 \le \delta_1, \delta_2 \le 1/2$ and $\mu = 0$, $\alpha = 0$, then

$$T^{\delta_2-\delta_1}(\widehat{\nu}-\nu) \Rightarrow \{Q_{xx}(b;\delta)\}^{-1} Q_{xe}(b;\delta), \tag{6}$$

where

$$\begin{aligned}Q_{xe}(b;\delta) &= B_2\left(1;\delta_2^+\right)B_1\left(1;\delta_1^+\right)-\frac{1}{b^2}\int_0^1\int_0^1 k''\left(\frac{r-s}{b}\right)B_2\left(s;\delta_2^+\right)B_1\left(r;\delta_1^+\right)drds \\ &+\frac{1}{b}\int_0^1 k'\left(\frac{1-r}{b}\right)\left(B_2\left(1;\delta_2^+\right)B_1\left(r;\delta_1^+\right)+B_2\left(r;\delta_2^+\right)B_1\left(1;\delta_1^+\right)\right)dr\end{aligned}$$

for Type 1 kernels,

$$\begin{aligned}Q_{xe}(b;\delta) &= B_2\left(1;\delta_2^+\right)B_1\left(1;\delta_1^+\right)-\frac{1}{b^2}\int\int_{|r-s|<b} k''\left(\frac{r-s}{b}\right)B_2\left(s;\delta_2^+\right)B_1\left(r;\delta_1^+\right)drds \\ &+\frac{1}{b}k_-'(1)\int_0^{1-b}\left(B_2\left(r;\delta_2^+\right)B_1\left(r+b;\delta_1^+\right)+B_2\left(r+b;\delta_2^+\right)B_1\left(r;\delta_1^+\right)\right)dr \\ &+\frac{1}{b}\int_{1-b}^1 k'\left(\frac{1-r}{b}\right)\left(B_2\left(r;\delta_2^+\right)B_1\left(1;\delta_1^+\right)+B_2\left(1;\delta_2^+\right)B_1\left(r;\delta_1^+\right)\right)dr\end{aligned}$$

for Type 2 kernels, and

$$\begin{aligned}Q_{xe}(b;\delta) &= B_2\left(1;\delta_2^+\right)B_1\left(1;\delta_1^+\right)+\frac{2}{b}\int_0^1 B_2\left(r;\delta_2^+\right)B_1\left(r;\delta_1^+\right)dr \\ &-\frac{1}{b}\int_0^{1-b}\left(B_2\left(r+b;\delta_2^+\right)B_1\left(r;\delta_1^+\right)+B_2\left(r;\delta_2^+\right)B_1\left(r+b;\delta_1^+\right)\right)dr \\ &-\frac{1}{b}\int_{1-b}^1\left(B_2\left(r;\delta_2^+\right)B_1\left(1;\delta_1^+\right)+B_2\left(1;\delta_2^+\right)B_1\left(r;\delta_1^+\right)\right)dr\end{aligned}$$

for the Barlett kernel; $Q_{xx}(b;\delta)$ is defined by replacing $B_1\left(r;\delta_1^+\right)$ by $B_2'\left(r;\delta_2^+\right)$ in all the formulae for $Q_{xe}(b;\delta)$ above.

Remark 1. Theorem 1 implies that $B(r;\delta)$, together with $k(x)$ and b (as it is standard in the fixed-b literature), fully characterize the fixed-b limiting distributions of $\widehat{\nu}$, $\widetilde{\nu}$.

Remark 2. Result (6) implies that when x_t and/or e_t are truly nonstationary, the limiting distribution of $\widehat{\nu}$ is invariant to nonzero μ and/or α, because their contribution is of smaller order.

Remark 3. Letting M grow proportional to T, the WC estimators attain the rate $T^{\delta_2-\delta_1}$, the fastest among the first stage estimators (see Hualde and Iacone, 2012).

Remark 4. Theorem 1 provides results for Type II processes; the results for Type I processes are similar (see Johansen and Nielsen, 2012b, for a summary of the regularity conditions, and for the characterization of the limit of the partial sums).

Remark 5. Marinucci (2000) discussed the traditional limiting behaviour of $\widehat{\nu}$ for cases $0 \leq \delta_1, \delta_2 < 1/2$; $\delta_1 = 0$, $\delta_2 = 1$; $1 < \delta_2 < 3/2$, $0 < \delta_1 < 1/2$, when observables are Type I processes. Without the aim of covering all possible cases, we present in Theorem 2 some results from which the small-b limiting approximations to the sampling distributions of $\widehat{\nu}$, $\widetilde{\nu}$, can be straightforwardly derived by the results for the OLS estimator given in Robinson and Marinucci (2001). Note that Assumption 1 strengthened to finite fourth moment is sufficient to derive the different results given in Robinson and Marinucci (2001), which will be used throughout the proof of Theorem 2. In particular, it is sufficient for the conditions related to the cumulant spectral density, it implies square integrability of the univariate spectra of the components of η_t and also fourth-order stationarity of η_t. Note also that we relax conditions in Marinucci (2000), like Gaussianity.

Theorem 2. Let Assumption 1 with $q \geq 4$, $M^{-1} + M/T \to 0$ as $T \to \infty$, and $0 \leq \delta_1 < \delta_2$ hold. Also, let $k(x)$ be nonnegative and bounded, $k(x) = 0$ for $|x| > 1$, $\int_{-1}^{1} k(x)\,dx = 1$. Then, if $\delta_2 > 1/2$ and $M/T^{2\delta_2-1} \to 0$, for $c^*_{xx} = c_{xx}, \widetilde{c}_{xx}$,

$$\frac{1}{M}\sum_{l=-M}^{M} k(l/M)\, c^*_{xx}(l) = c^*_{xx}(0) + o_p\left(T^{2\delta_2-1}\right). \tag{7}$$

Also, for $c^*_{xe} = c_{xe}, \widetilde{c}_{xe}$, if $(\delta_1, \delta_2) = (0, 1)$

$$\frac{1}{M}\sum_{l=-M}^{M} k(l/M)\, c^*_{xe}(l) = c^*_{xe}(0) + \frac{1}{2}\left(\sum_{l=1}^{\infty}\gamma_{-l} - \sum_{l=0}^{\infty}\gamma_l\right) + o_p(1), \tag{8}$$

where $\gamma_l = E\left(\eta_{1,t}\eta_{2,t-l}\right)$, whereas if for $\delta_1 \geq 0$, $\delta_1+\delta_2 > 1$, $T^{-1}M\log T + T^{1-\delta_1-\delta_2}M \to 0$, then

$$\frac{1}{M}\sum_{l=-M}^{M} k(l/M)\, c^*_{xe}(l) = c^*_{xe}(0) + o_p\left(T^{\delta_1+\delta_2-1}\right). \tag{9}$$

Remark 6. Both the small-b and fixed-b limiting distributions of the WC estimators depend on nuisance parameters δ, Ω. This dependence makes the WC estimators unsuitable for statistical inference. However, a modified version of these estimators could be the basis of asymptotically pivotal Wald statistics. As usual in the cointegration literature, the main challenge is to transform the estimators to remove the endogeneity caused by the correlation between $\eta_{1,t}$ and $\eta_{2,t}$, and, without the aim of providing a complete discussion, we explore here this issue. For the sake of an easy presentation, let $\mu = 0$, $\alpha = 0$ in (1), (2). Also, for a scalar or vector process ξ_t and real number a, let $\xi_t(a) = \Delta^a \{\xi_t 1(t > 0)\}$ and, partitioning Ω according to η_t, so

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

let $\theta = \Omega_{22}^{-1}\Omega_{21}$. Clearly, (2) can be written as

$$y_t(\delta_1) = \nu' x_t(\delta_1) + \theta' x_t(\delta_2) + \eta_{1.2,t}, \tag{10}$$

where $\eta_{1.2,t} = \eta_{1,t} - \Omega_{12}\Omega_{22}^{-1}\eta_{2,t}$, so, denoting $z_t(\delta,\theta) = y_t(\delta_1) - \theta' x_t(\delta_2)$ and assuming for the moment that δ and θ are known, we could estimate ν by

$$\overline{\nu}(\delta,\theta) = \left(\sum_{l=-T+1}^{T-1} k(l/M)\, c_{x(\delta_1)x(\delta_1)}(l)\right)^{-1} \sum_{l=-T+1}^{T-1} k(l/M)\, c_{x(\delta_1)z(\delta,\theta)}(l). \tag{11}$$

By identical arguments to those in the Appendix, it is simple to show that the fixed-b limiting distribution of $T^{\delta_2-\delta_1}(\overline{\nu}(\delta,\theta) - \nu)$ depends on b, $k(\cdot)$, $B_2(r, \delta_2 - \delta_1)$ and $B_{1.2}(r) = B_1(r) - \Omega_{12}\Omega_{22}^{-1}B_2(r)$. Then, noting that $B_2(r, \delta_2 - \delta_1)$ and $B_{1.2}(r)$ are independent processes, the fixed-b limiting distribution of $T^{\delta_2-\delta_1}(\overline{\nu}(\delta,\theta) - \nu)$ can be easily shown to be mixed-Gaussian. This is a crucial result which as in, e.g., Hualde and Robinson (2010), can be exploited to construct a Wald test statistic to test for values of ν with a χ^2 fixed-b null limiting distribution. Additionally, in the present setting, the fixed-b theory enjoys other attractive features. First, the type of correction employed to remove the endogeneity is very simple (the inclusion of $\theta' x_t(\delta_2)$ in (10)). When considering small-b asymptotics instead, this type of correction would work if η_t is a white noise process, but in more general settings, in view of the results for "zero-frequency" estimators of Hualde and Robinson (2010), it would only be adequate if the rate of growth of M is appropriately restricted. Second, the fixed-b limit is identical irrespective of the type of cointegration which characterizes the data. This is relevant, because it is well known in the fractional cointegration literature that, typically, different results apply under strong (with $\delta_2 - \delta_1 > 1/2$) or weak (with $\delta_2 - \delta_1 < 1/2$) cointegration, the borderline case $\delta_2 - \delta_1 = 1/2$ being relatively unexplored. Therefore, while the small-b limit of (11) depends crucially on the gap $\delta_2 - \delta_1$, the fixed-b limit is valid for any $\delta_2 - \delta_1 > 0$. Finally, given that in practice δ, θ are unknown, $\overline{\nu}(\delta, \theta)$ is unfeasible: however, δ, θ can be easily estimated (say by $\widehat{\delta}$, $\widehat{\theta}$), which prompts consideration of the feasible estimator $\overline{\nu}(\widehat{\delta}, \widehat{\theta})$. Then, by very similar techniques to those in, e.g., Hualde and Robinson (2010), it can be shown that if $\widehat{\delta} - \delta = O_p(T^{-\kappa})$, for any $\kappa > 0$, and $\widehat{\theta} \rightarrow_p \theta$, then $T^{\delta_2 - \delta_1}(\overline{\nu}(\delta, \theta) - \nu)$ and $T^{\delta_2 - \delta_1}\left(\overline{\nu}(\widehat{\delta}, \widehat{\theta}) - \nu\right)$ have identical fixed-b limiting distributions. This is a very strong result, because the conditions on the estimators of the nuisance parameters are very mild and contrast heavily with the standard theory, where faster rates of convergence on the estimators of δ, θ are required the smaller the cointegrating gap $(\delta_2 - \delta_1)$ is. Note however that the feasibility issue can be considered from a more attractive (and complex) perspective. Given that θ would typically be estimated by WC estimators, instead of relying on consistency arguments, one could consider a "complete" fixed-b theory, where fixed-b arguments apply also to $\widehat{\theta}$. This type of analysis, proposing also an appropriate Wald statistic with a pivotal fixed-b limit, appears to be a doable but very challenging task.

Remark 7. The frequency domain representations of the WC estimators lead to an interesting connection between fixed-b theory for WC and fixed bandwidth approach to NBLS. For two generic sequences ξ_t, ζ_t, let $I^*_{\xi\zeta}(\lambda) = (2\pi)^{-1}\sum_{|l|<T} c^*_{\xi\zeta}(l)\, e^{-il\lambda}$, where i is the complex operator, and then let $I_{\xi\zeta}(\lambda) = I^*_{\xi\zeta}(\lambda)$ when $c^*_{\xi\zeta}(l) = c_{\xi\zeta}(l)$ is used, whereas let $\widetilde{I}_{\xi\zeta}(\lambda) = I^*_{\xi\zeta}(\lambda)$ when $c^*_{\xi\zeta}(l) = \widetilde{c}_{\xi\zeta}(l)$, so that $I_{\xi\zeta}(\lambda)$ and $\widetilde{I}_{\xi\zeta}(\lambda)$ are (cross-) periodograms. Then, the WC estimators have frequency domain representation

$$\nu^* = \left(\int_{-\pi}^{\pi} K_M(\lambda)\, I^*_{xx}(\lambda)\, d\lambda\right)^{-1} \int_{-\pi}^{\pi} K_M(\lambda)\, I^*_{xy}(\lambda)\, d\lambda, \tag{12}$$

(see, e.g., Brockwell and Davis, 1991, pp.358-360), where $K_M(\lambda) = (2\pi)^{-1}\sum_{|l|<T} k(l/M)\, e^{-il\lambda}$ is the spectral window associated to $k(\cdot)$ (see, e.g., Priestley, 1981, p.436), and ν^* equals $\widehat{\nu}$ or $\widetilde{\nu}$ depending on whether I or $\widetilde{I}$ are used. Focusing just on $\widehat{\nu}$, approximating the integrals in (12) by sums over the Fourier frequencies $\lambda_j = 2\pi j/T$ for $j = 0, \pm 1, ..., \pm\lfloor T/2 \rfloor$ (where $\lfloor \cdot \rfloor$ denotes integer part), yields an alternative estimator of ν, say $\overline{\nu}$. This is particularly interesting if the Daniell kernel is used: this kernel has spectral window $M/(2\pi)$ when $-\pi/M \leq \lambda \leq \pi/M$ and 0 otherwise, so by the symmetry of the periodogram, $\overline{\nu} = \left(\sum_{j=0}^{m} s_j I_{xx}(\lambda_j)\right)^{-1} \sum_{j=0}^{m} s_j \operatorname{Re} I_{xy}(\lambda_j)$, where $s_j = 1$, $j = 0, T/2$, $s_j = 2$, otherwise, and following Brockwell and Davis (1991), pp.359,360, $m = \lfloor T/(2M) \rfloor$. Note that $\overline{\nu}$ is the NBLS estimator, whose limiting properties have been derived by Robinson and Marinucci (2001) under the assumption that $1/m + m/T \rightarrow 0$ as $T \rightarrow \infty$, implying

$M/T\rightarrow 0$. Robinson and Marinucci (2001, p.866) conjectured that the $T^{\delta_2-\delta_1}$ convergence rate could be achievable by the NBLS with m fixed as $T\rightarrow\infty$. Later, Chen and Hurvich (2003) verified this conjecture for a tapered NBLS used on differenced data. By our previous reasoning, analyzing the properties of $\widehat{\nu}$ (or $\widetilde{\nu}$) when $M=bT$, $b>0$ and the Daniell kernel is used, relates closely to discussing the properties of the NBLS when m is fixed, specifically $m=\lfloor 1/(2b)\rfloor$. Providing formal results is beyond the scope of the present paper, but we conjecture that if $M=bT$, $b\in(0,1]$, under standard conditions, $\widehat{\nu}-\overline{\nu}=O_p\left(T^{\delta_2-\delta_1}\right)$, although $\widehat{\nu}$ and $\overline{\nu}$ may have different limiting distributions.

3. FINITE SAMPLE PERFORMANCE

In a simple bivariate case, we compare the sampling distributions of $\widetilde{\nu}$ with its small-b and fixed-b limiting distributions for cases $(\delta_1,\delta_2)=(0,1),(0,1.4),(.8,1.2)$. While the fixed-$b$ limit is given by Theorem 1, the small-b can be straightforwardly derived from Theorem 2 (although stronger moment conditions might be needed). In particular, the small-b limits of $T^{\delta_2-\delta_1}(\widetilde{\nu}-\nu)$ for the bivariate case are

$$\left(\int_0^1 \overline{B}_2(r;1)\,dB_1(r;1)+\frac{1}{2}\sum_{l=-\infty}^{\infty}\gamma_l\right)\Big/\int_0^1 \overline{B}_2^2(r;1)\,dr, \text{ if } \delta_1 = 0,\delta_2=1, \quad (13)$$

$$\int_0^1 \overline{B}_2(r;\delta_2)\,dB_1(r;1)\Big/\int_0^1 \overline{B}_2^2(r;\delta_2)\,dr, \text{ if } \delta_1 = 0,\delta_2>1, \quad (14)$$

$$\int_0^1 \overline{B}_2(r;\delta_2)\,\overline{B}_1(r;\delta_1)\,dr\Big/\int_0^1 \overline{B}_2^2(r;\delta_2)\,dr, \text{ if } \delta_1 > \frac{1}{2},\delta_2\geq 1, \quad (15)$$

where $\overline{B}_j(r;d)=B_j(r;d)-\int_0^1 B_j(r;d)\,dr,\ j=1,2$. Note also that the small-b approximation of $T^{\delta_2-\delta_1}(\widehat{\nu}-\nu)$ is given by corresponding expressions (13), (14), (15), just replacing $\overline{B}_j$ by B_j. Interestingly, note that (13) with $\overline{B}_2$ replaced by B_2 differs from (A.12) of Phillips (1991) (there seems to be a minor typo in the proof in p.433, where the contribution of $k(x)$ is missing) and also from (18) of Marinucci (2000).

We generate ε_t as an iid Gaussian process with $E(\varepsilon_t)=0$, $Var(\varepsilon_{1,t})=Var(\varepsilon_{2,t})=1$, $Cov(\varepsilon_{1,t},\varepsilon_{2,t})=.5$, and also η_t as in Assumption 1 with $A(z)=diag\{1/(1-.5z),(1+.5z)\}$. Fixing $\mu=\alpha=\nu=0$, we generate $(y_t,x_t)'$ using (1) and (2), for the three different (δ_1,δ_2) combinations and compute $\widetilde{\nu}$ for $b=M/T=.1,.25,.5,1$, using the Bartlett kernel. Next, using $R=5000$ replications we computed the empirical cumulative distribution functions (CDF) of $T^{\delta_2-\delta_1}(\widetilde{\nu}-\nu)$, $T=64,256$, using $\widehat{F}(x)=R^{-1}\sum_{i=1}^{R}1\left(T^{\delta_2-\delta_1}(\widetilde{\nu}_i-\nu)<x\right)$, where $\widetilde{\nu}_i$, $i=1,...,R$, are the estimates corresponding to each replication. To evaluate the accuracy of the small-b and fixed-b asymptotic approximations (which are nuisance-parameter dependent), we compare the two sampling

CDFs with corresponding asymptotic CDFs (given in Theorem 1 and (13)-(15)), which, as in Hashimzade and Vogelsang (2007), were simulated approximating the standard Brownian motion by scaled partial sums of iid $N(0,1)$ random variables using 1000 increments and 50000 replications.

Results for $(\delta_1, \delta_2) = (0,1), (0,1.4), (0.8,1.2)$ are given in Figures 1, 2, 3, respectively. As expected, we always find that the limit distribution computed assuming $b = 0.1$ is the closest one to the small-b limit, being these two distributions very close to each other when $(\delta_1, \delta_2) = (0,1), (0.8,1.2)$. On the other hand, the small-b and fixed-b limit distributions differentiate more as b increases. In all cases the fixed-b limit distribution is closer to the empirical distributions of the estimates in small samples, thus providing a better approximation: indeed, the fixed-b limit distribution provides a good approximation of the distributions of the estimates in small samples already for $T = 256$ and, at least for $b = 0.5, 1$, even for $T = 64$. The gains in accuracy achieved by the fixed-b limit are most evident when $(\delta_1, \delta_2) = (0, 1.4)$, the small-$b$ approximation being very inaccurate here.

4. CONCLUSION

We have compared the traditional (small-b) and fixed-b limiting approximations to the sampling distribution of WC estimators. First, we have derived the fixed-b limiting distribution of two WC estimators, filling also some gaps in the small-b theory. Then, by means of a Monte Carlo experiment, we have compared both limiting distributions, concluding that the fixed-b limit is more accurate. Given that these distributions depend in general on nuisance parameters, our results are not of direct use in testing. However, we have proposed an appropriate modification of one of the WC estimators along the lines of second stage estimation of cointegrating parameters (see, e.g., Robinson and Hualde, 2003, Hualde and Robinson, 2010, Nielsen and Frederiksen, 2011), which can be exploited to construct a Wald test statistics with pivotal fixed-b limit. This, in view of the evidence provided by the present paper and the size problems displayed by the Wald test statistics based on second stage estimators, appears to be a promising research avenue which will be explored in future work.

Appendix.

Proof of Theorem 1. We first give the proof for $\widetilde{\nu}$. From (2), $\overline{y} = \alpha + \nu\overline{x} + \overline{e}$ and $y_t - \overline{y} = \nu(x_t - \overline{x}) + e_t - \overline{e}$. Therefore,

$$\begin{aligned} \widetilde{c}_{xy}(l) &= \nu T^{-1} \textstyle\sum_{t=1}^{T-l} (x_t - \overline{x})(x_{t+l} - \overline{x})' + T^{-1} \textstyle\sum_{t=1}^{T-l} (x_t - \overline{x})(e_{t+l} - \overline{e}), \; l \geq 0, \\ &= \nu T^{-1} \textstyle\sum_{t=1-l}^{T} (x_t - \overline{x})(x_{t+l} - \overline{x})' + T^{-1} \textstyle\sum_{t=1-l}^{T} (x_t - \overline{x})(e_{t+l} - \overline{e}), \; l < 0, \end{aligned}$$

so $\widetilde{c}_{xy}(l) = \widetilde{c}_{xx}(l)\,\nu + \widetilde{c}_{xe}(l)$, and

$$\widetilde{\nu} = \nu + \left(\sum_{l=-T+1}^{T-1} k(l/M)\,\widetilde{c}_{xx}(l)\right)^{-1} \sum_{l=-T+1}^{T-1} k(l/M)\,\widetilde{c}_{xe}(l).$$

Next, adapting notation on p.1353 of Kiefer and Vogelsang (2002), let $\widehat{e}_t = e_t - \overline{e}$, $\widehat{x}_t = x_t - \overline{x}$, $\widehat{v}_t = (\widehat{e}_t, \widehat{x}_t')'$, $\widehat{\Gamma}_l = T^{-1}\sum_{t=l+1}^{T} \widehat{v}_t \widehat{v}_{t-l}'$ for $l \geq 0$, $\widehat{\Gamma}_l = \widehat{\Gamma}'_{-l}$ for $l < 0$, $\widehat{\Omega} = \sum_{l=-T+1}^{T-1} k(l/M)\,\widehat{\Gamma}_l$, $\kappa_{h,l} = k((h-l)/(bT))$, $\nabla^2\kappa_{h,l} = (\kappa_{h,l} - \kappa_{h,l+1}) - (\kappa_{h+1,l} - \kappa_{h+1,l+1})$, $\widehat{S}_l = \sum_{t=1}^{l} \widehat{v}_t$. Then, as on p.1365 of Kiefer and Vogelsang (2002),

$$\widehat{\Omega} = T^{-1}\sum_{h=1}^{T-1} T^{-1}\sum_{l=1}^{T-1} T^2\nabla^2\kappa_{h,l} T^{-1/2}\widehat{S}_h T^{-1/2}\widehat{S}_l', \tag{16}$$

so $T^{-(\delta_1+\delta_2)}\sum_{l=-T+1}^{T-1} k(l/M)\,\widetilde{c}_{xe}(l) = T^{-(\delta_1+\delta_2)}\left(\widetilde{0}, I_{p-1}\right)\widehat{\Omega}\left(1, \widetilde{0}'\right)'$ equals

$$T^{-1}\sum_{h=1}^{T-1} T^{-1}\sum_{l=1}^{T-1} T^2\nabla^2\kappa_{h,l}\left(T^{-1/2-\delta_1}\sum_{t=1}^{h}\widehat{e}_t\right)\left(T^{-1/2-\delta_2}\sum_{t=1}^{l}\widehat{x}_t\right).$$

We discuss Type 1 kernels first. Adapting results from Kiefer and Vogelsang (2005), p.1159, and Bunzel and Vogelsang (2005), $T^2\nabla^2\kappa_{h,l} \to -b^{-2}k''(b^{-1}(r-s))$, so, by the FCLT (4) and the continuous mapping theorem,

$$T^{-(\delta_1+\delta_2)}\sum_{l=-T+1}^{T-1} k(l/M)\,\widetilde{c}_{xe}(l) \Rightarrow \widetilde{Q}_{xe}(b;\delta), \tag{17}$$

and, in the same way,

$$T^{-2\delta_2}\sum_{l=-T+1}^{T-1} k(l/M)\,\widetilde{c}_{xx}(l) \Rightarrow \widetilde{Q}_{xx}(b;\delta). \tag{18}$$

The proofs of (17) and (18) for Type 2 kernels and for Type 3 kernel follow again using formulae in Kiefer and Vogelsang (2005) and Bunzel and Vogelsang (2005), but applying the FCLT (4) for fractional processes as in the proof for Type 1 kernels. Finally, (5) follows by the continuous mapping theorem.

The proof for $\widehat{\nu}$ is almost identical, just noting that additional terms arise in the expansion of $\widehat{\Omega}$ (see (16)) because the series are not demeaned, and also that when x_t and/or e_t are truly nonstationary, the contribution of nonzero μ and/or α is of smaller order.

Proof of Theorem 2. First, noting that $M^{-1}\sum_{l=-M}^{M} k(l/M) - 1 = o(1)$, (7) for

$c_{xx}^{*}=c_{xx}$ follows as in Lemma 3 of Marinucci (2000) by showing

$$\begin{aligned}\frac{1}{MT}\sum_{l=1}^{M}k\left(l/M\right)\sum_{t=1}^{l}x_t x_t' &= o_p\left(T^{2\delta_2-1}\right),\\ \frac{1}{MT}\sum_{l=1}^{M}k\left(l/M\right)\sum_{t=l+1}^{T}x_t\left(x_t-x_{t-l}\right)' &= o_p\left(T^{2\delta_2-1}\right),\end{aligned}$$

which, as $M/T^{2\delta_2-1}\rightarrow 0$, can be easily justified as in Robinson and Marinucci (2001). The proof for $c_{xx}^{*}=\widetilde{c}_{xx}$ is almost identical and thus we omit it.

The proof of (8) follows by replicating some of the steps given in the proof of Lemma 4 of Marinucci (2000) and also by results in Robinson and Marinucci (2001). In particular

$$c_{xe}^{*}\left(0\right)-\frac{1}{M}\sum_{l=-M}^{M}\kappa\left(l/M\right)c_{xe}^{*}\left(l\right)=\left(I\right)+\left(II\right)+\left(III\right)+\left(IV\right)+\left(V\right), \qquad (19)$$

where for $c_{xe}^{*}=c_{xe}$ the terms of the right hand side of (19) are defined in p.701 of Marinucci (2000), whereas for $c_{xe}^{*}=\widetilde{c}_{xe}$ the definitions are almost identical, with the only difference that the series are demeaned. Then, for either definition, it can be shown that $\left(II\right)=\left(IV\right)=o_p\left(1\right)$, $\left(V\right)=o_p\left(1\right)$,

$$\begin{aligned}\left(I\right) &= -E\left(\frac{1}{MT}\sum_{l=1}^{M}k\left(l/M\right)\sum_{t=l+1}^{T}\left(x_t-x_{t-l}\right)e_{t-l}\right)+o_p\left(1\right),\\ \left(III\right) &= E\left(\frac{1}{MT}\sum_{l=1}^{M}k\left(l/M\right)\sum_{t=l+1}^{T}\left(x_t-x_{t-l}\right)e_t\right)+o_p\left(1\right),\end{aligned}$$

so (8) follows because $\left(I\right)\rightarrow_p -\frac{1}{2}\sum_{l=1}^{\infty}\gamma_{-l}$, $\left(III\right)\rightarrow_p \frac{1}{2}\sum_{l=0}^{\infty}\gamma_l$. Finally, the proof of (9) follows as in Marinucci (2000) using results from Robinson and Marinucci (2001). With Marinucci's (2000) notation, the requirement $T^{-1}M\log T+T^{1-\delta_1-\delta_2}M\rightarrow 0$ is needed to show that terms Δ_3 and $\left(III\right)$ are $o_p\left(T^{\delta_1+\delta_2-1}\right)$.

REFERENCES

Brockwell, P.J., and R.A. Davis, 1991. *Time Series: Theory and Methods, 2nd Ed.* New York: Springer.

Bunzel, H., 2006. Fixed-b asymptotics in single-equation cointegration models with endogenous regressors, *Econometric Theory* 22, 743-755.

Bunzel, H., and T.J. Vogelsang, 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.

Chen, W.W., and C.M. Hurvich, 2003. Estimating fractional cointegration in the presence of polynomial trends, *Journal of Econometrics* 117, 95-121.

Hashimzade, N., and T.J. Vogelsang, 2007. Fixed-b asymptotic approximation of the sampling behaviour of nonparametric spectral density estimators, *Journal of Time Series Analysis* 29, 142-162.

Hualde, J., and F. Iacone, 2012. First stage estimation of fractional cointegration, *Journal of Time Series Econometrics* 4, Issue 1, Article 2.

Hualde, J., and P.M. Robinson, 2010. Semiparametric inference in multivariate fractionally cointegrated systems, *Journal of Econometrics* 157, 492-511.

Iacone, F., Leybourne, S.J., and A.M. R. Taylor, 2013. On the behaviour of fixed-b trend break tests under fractional integration, *Econometric Theory* 29, 393-418.

Jin, S., Phillips, P.C.B., and Y. Sun, 2006. A new approach to robust inference in cointegration, *Economics Letters* 91, 300-306.

Johansen, S., and M.Ø. Nielsen, 2012a. Likelihood inference for a fractionally cointegrated vector autoregressive model, *Econometrica* 80, 2667–2732.

Johansen, S., and M.Ø. Nielsen, 2012b. A necessary moment condition for the fractional functional central limit theorem, *Econometric Theory* 28, 671-679.

Kiefer, N.M., and T.J. Vogelsang, 2002. Heteroskedasticity-autocorrelation robust testing using bandwidth equal to sample size, *Econometric Theory* 18, 1350-1366.

Kiefer, N.M., and T.J. Vogelsang, 2005. A new asymptotic theory for heteroskedasticity - autocorrelation robust tests. *Econometric Theory* 21, 1130-64.

Marinucci, D., 2000. Spectral regression for cointegrated time series with long-memory innovations, *Journal of Time Series Analysis* 21, 685-705.

Marinucci, D., and P.M. Robinson, 1999. Alternative forms of fractional Brownian motion, *Journal of Statistical Planning and Inference* 80, 111-122.

Marinucci, D., and P.M. Robinson, 2000. Weak convergence of multivariate fractional processes. *Stochastic Processes and their Applications* 86, 103-120.

Neave, H.R., 1970. An improved formula for the asymptotic variance of spectrum estimates, *Annals of Mathematical Statistics* 41, 70-77.

Nielsen, M.Ø., and P. Frederiksen, 2011. Fully modified narrow-band least squares estimation of weak fractional cointegration, *Econometrics Journal* 14, 77-120.

Phillips, P.C.B., 1991. Spectral regression for cointegrated time series. In: Barnett, W.A., Powell, J., Tauchen, G. (Eds.), Nonparametric and Semiparametric Methods in Econometrics and Statistics. Cambridge University Press, Cambridge.

Priestley, M.B., 1981. *Spectral Analysis and Time Series*, Academic Press, London.

Robinson, P.M., and J. Hualde, 2003. Cointegration in fractional systems with unknown integration orders, *Econometrica* 71, 1727-1766.

Robinson, P.M., and D. Marinucci, 2001. Narrow band analysis of nonstationary processes, *Annals of Statistics* 29, 947-86.

Robinson, P.M., and D. Marinucci, 2003. Semiparametric frequency domain analysis of fractional cointegration, in *Time series with long memory* (P.M. Robinson ed.), Oxford: New York: Oxford University press, 334-373.

Vogelsang, T.J., and M. Wagner, 2014. Integrated modified OLS estimation and fixed-b inference for cointegrating regressions, *Journal of Econometrics* 178, 741–760.

Figure 1 (a). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.0, b=0.1

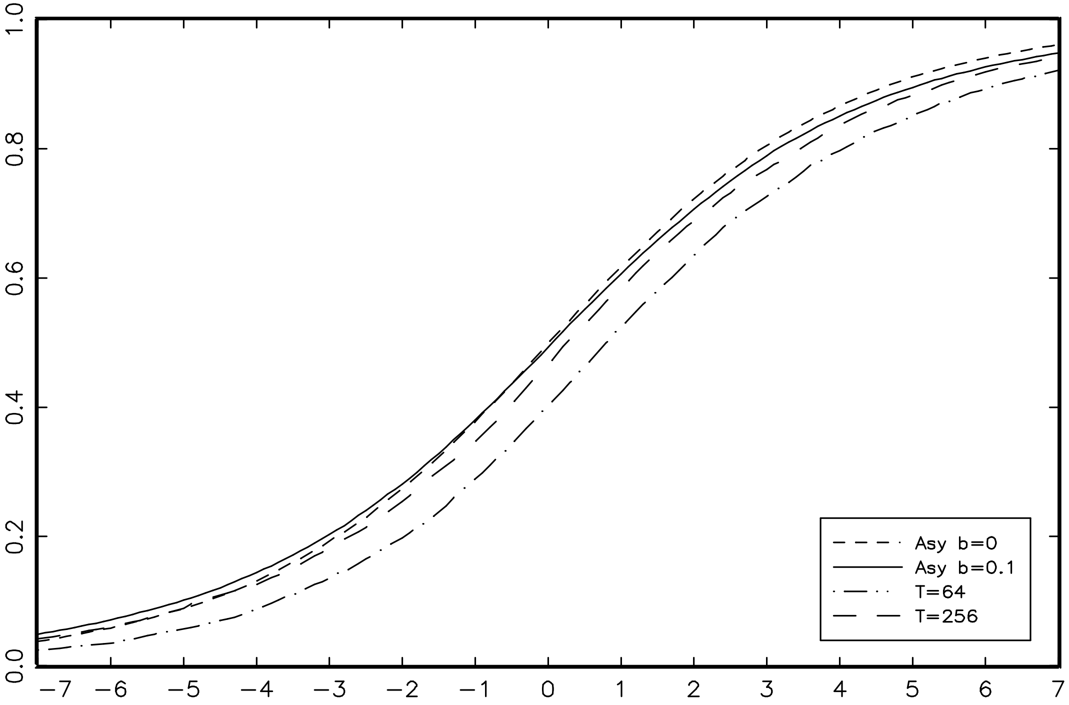

Figure 1 (b). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.0, b=0.25

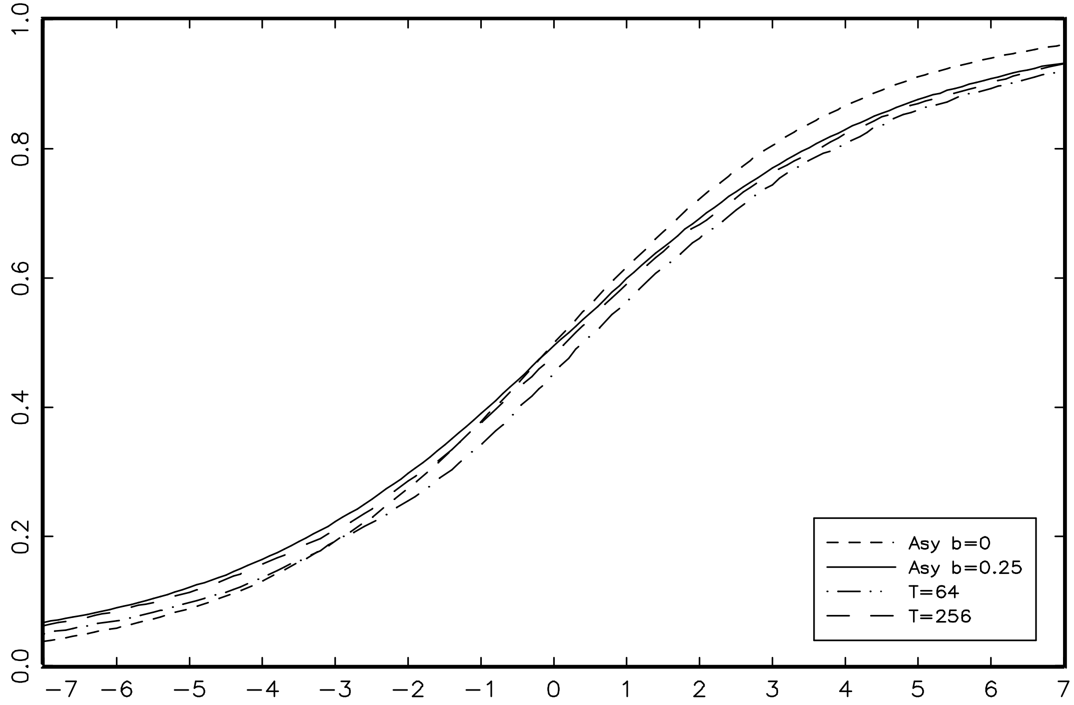

Figure 1 (c). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.0, b=0.5

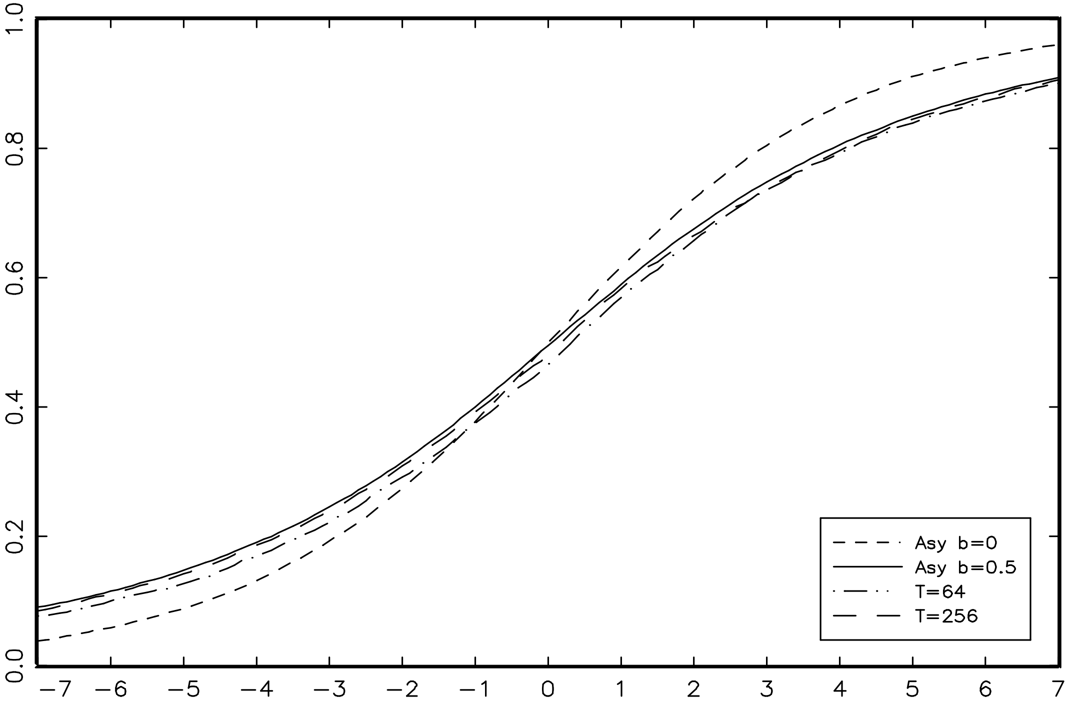

Figure 1 (d). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.0, b=1

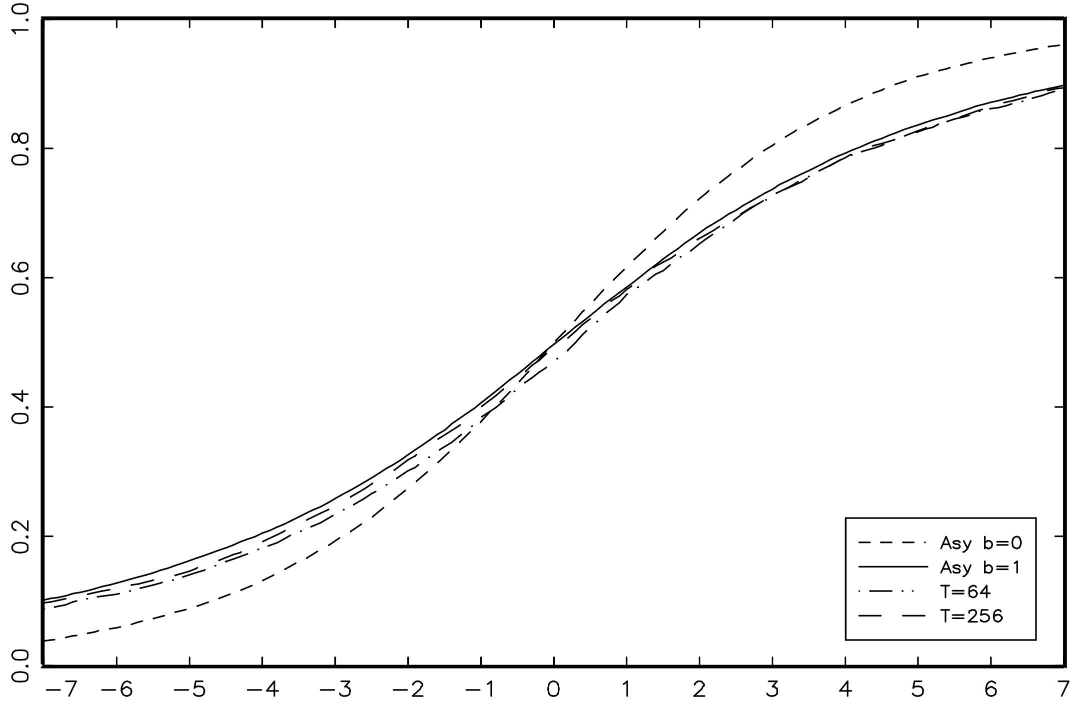

Figure 2 (a). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.4, b=0.1

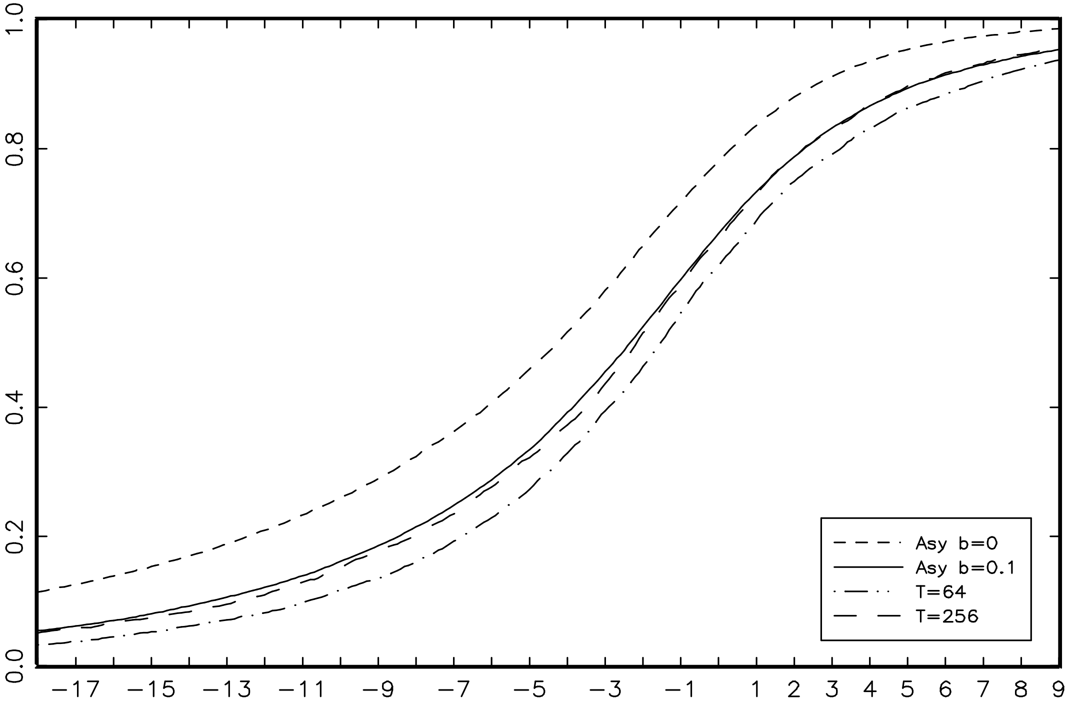

Figure 2 (b). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.4, b=0.25

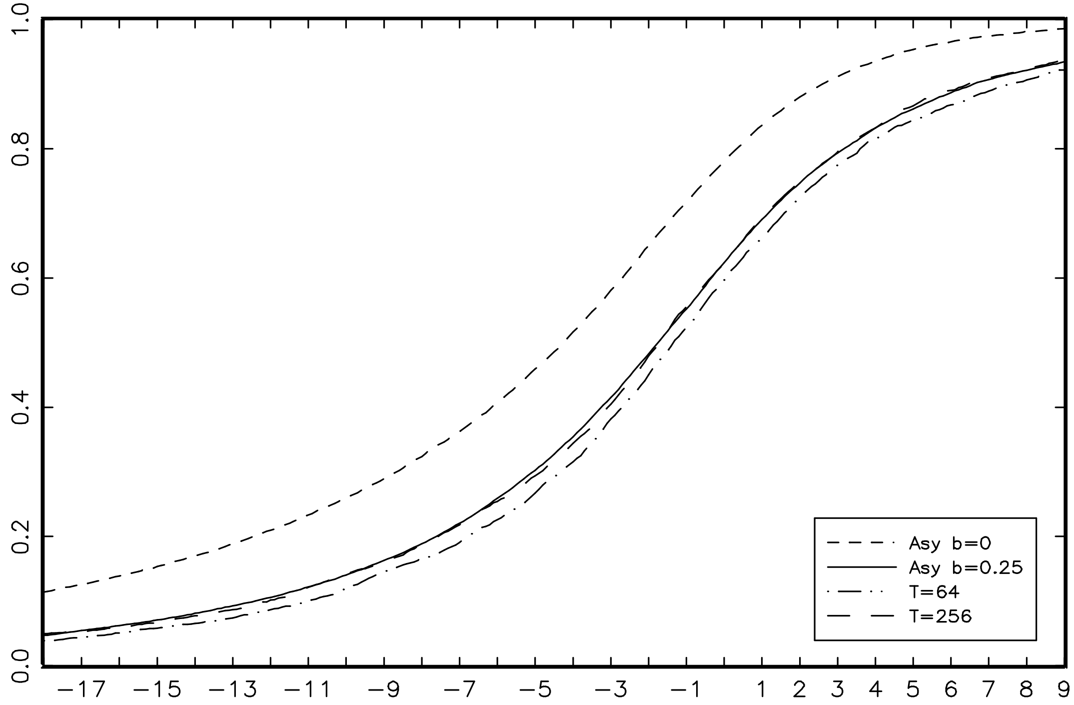

Figure 2 (c). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.4, b=0.5

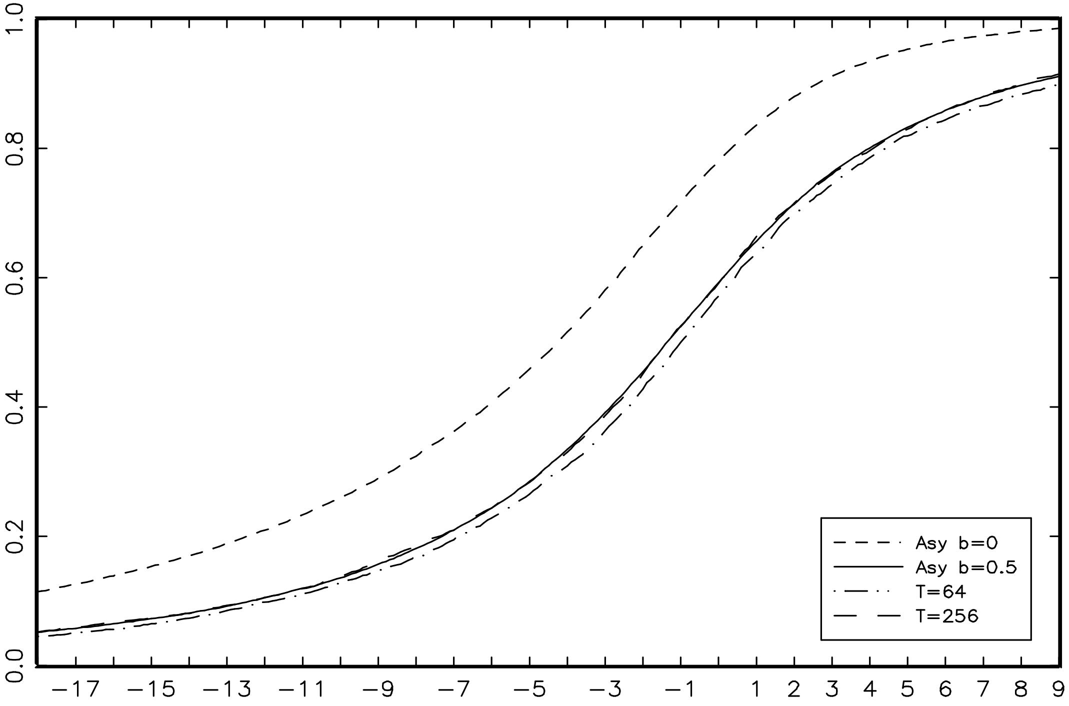

Figure 2 (d). Sampling and Asymptotic CDFs, δ_1=0.0 δ_2=1.4, b=1

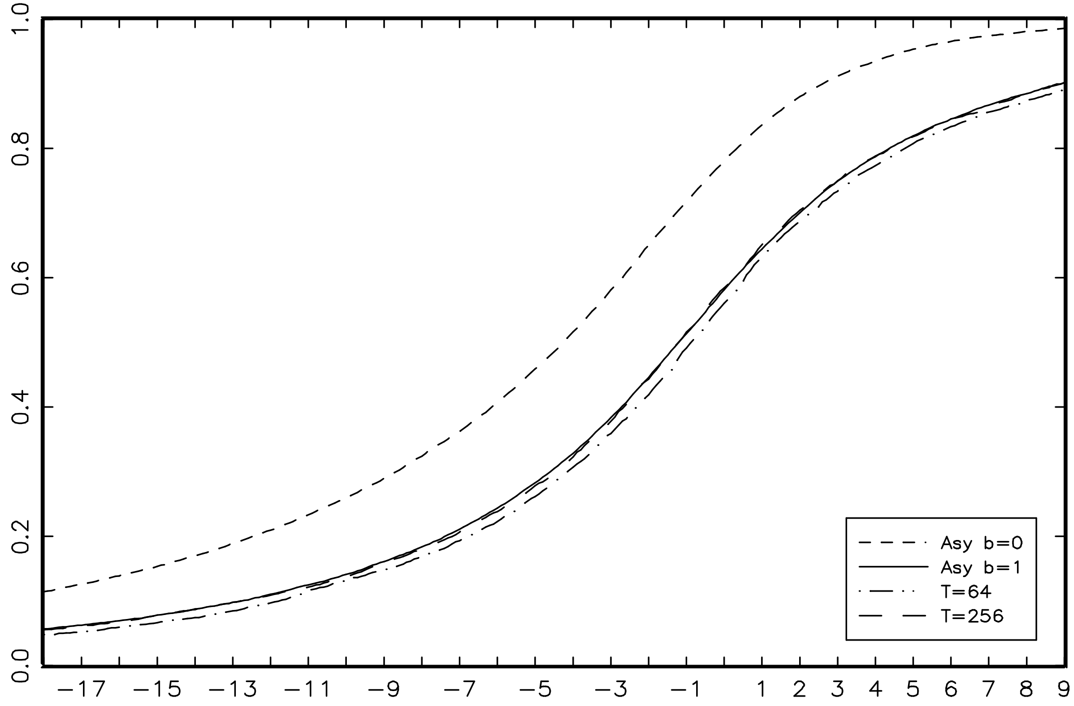

Figure 3 (a). Sampling and Asymptotic CDFs, δ_1=0.8 δ_2=1.2, b=0.1

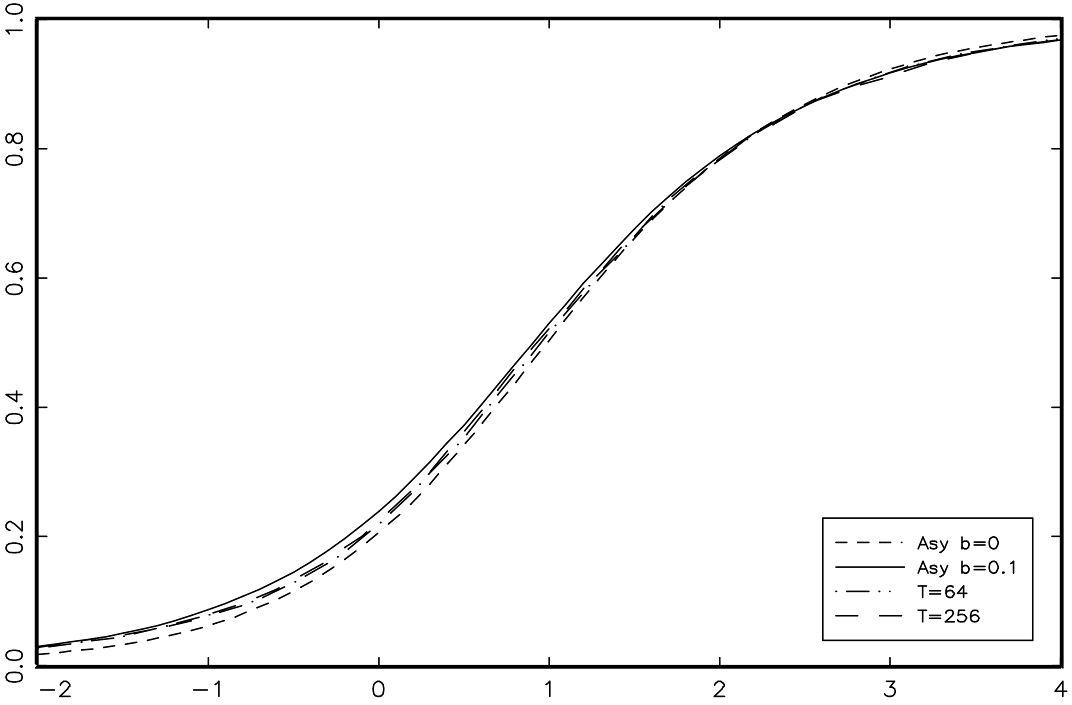

Figure 3 (b). Sampling and Asymptotic CDFs, δ_1=0.8 δ_2=1.2, b=0.25

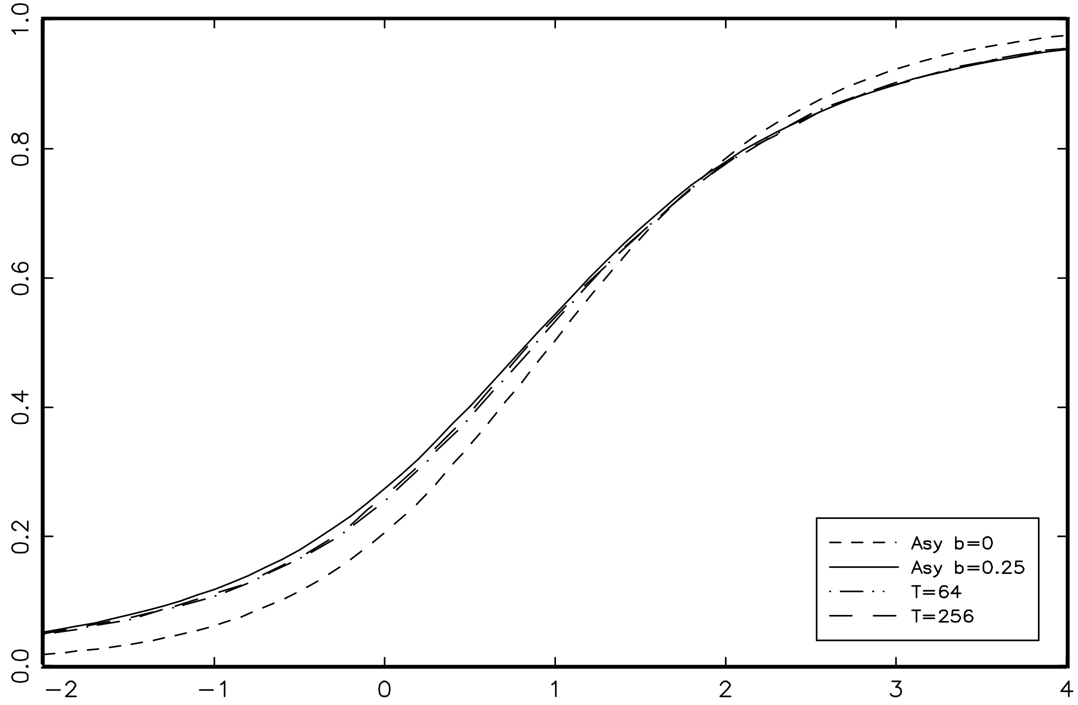

Figure 3 (c). Sampling and Asymptotic CDFs, δ_1=0.8 δ_2=1.2, b=0.5

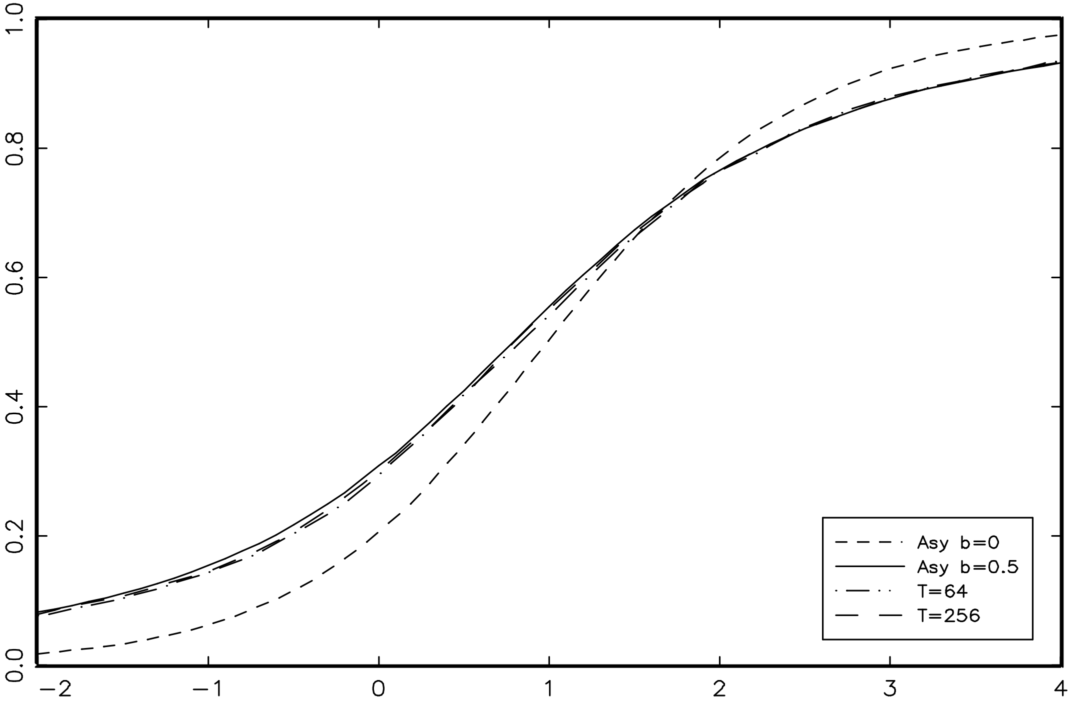

Figure 3 (d). Sampling and Asymptotic CDFs, δ_1=0.8 δ_2=1.2, b=1