“Level Shift Estimation in the Presence of Non-stationary Volatility with an Application to the Unit Root Testing Problem”

“David Harris, Hsein Kew and A.M. Robert Taylor”
Level Shift Estimation in the Presence of Non-stationary Volatility with an Application to the Unit Root Testing Problem

David Harris\textsuperscript{a}, Hsein Kew\textsuperscript{b} and A.M. Robert Taylor\textsuperscript{c}

\textsuperscript{a} Department of Economics, University of Melbourne
\textsuperscript{b} Department of Econometrics and Business Statistics, Monash University
\textsuperscript{c} Essex Business School, University of Essex

Abstract

In this paper we contribute to two separate literatures. Our principal contribution is made to the literature on break fraction estimation. Here we investigate the properties of a class of weighted residual sum of squares estimators for the location of a level break in time series whose shocks display non-stationary volatility (permanent changes in unconditional volatility). This class contains the ordinary least squares (OLS) and weighted least squares (WLS) estimators, the latter based on the true volatility process. For fixed magnitude breaks we show that the estimator attains the same consistency rate under non-stationary volatility as under homoskedasticity. We also provide local limiting distribution theory for the estimator when the break magnitude is either local-to-zero at some rate in the sample size or exactly zero. The former includes the Pitman drift rate which is shown via Monte Carlo experiments to predict well the key features of the finite sample behaviour of the OLS estimator and a feasible version of the WLS estimator based on an adaptive estimate of the volatility path of the shocks. The simulations highlight the importance of the break location, break magnitude, and the form of non-stationary volatility for the finite sample performance of these estimators, and show that the feasible WLS estimator can deliver significant improvements over the OLS estimator in certain heteroskedastic environments. We also contribute to the unit root testing literature. We demonstrate how the results in the first part of the paper can be applied, by using level break fraction estimators on the first differences of the data, when testing for a unit root in the presence of trend breaks and/or non-stationary volatility. In practice it will be unknown whether a trend break is present and so we also discuss methods to select between the break and no break cases, considering both standard information criteria and feasible weighted information criteria based on our adaptive volatility estimator. Simulation evidence suggests that the use of these feasible weighted estimators and information criteria can deliver unit root tests with significantly improved finite sample behaviour under heteroskedasticity relative to their unweighted counterparts.

Keywords: Level break fraction, non-stationary volatility, adaptive estimation, feasible weighted estimator, information criteria, unit root tests and trend breaks.

JEL: C12, C22.

\textsuperscript{*}We thank Giuseppe Cavaliere and participants at the International Conference in Celebration of the 65th Birthday of Maxwell King for useful comments. Taylor gratefully acknowledges financial support provided by the Economic and Social Research Council of the United Kingdom under research grant ES/M01147X/1. Correspondence to: Robert Taylor, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom. Email: rtaylor@essex.ac.uk.
1 Introduction

Breaks in the deterministic trend function appear prevalent in macroeconomic series; see, *inter alia*, Stock and Watson (1996, 1999, 2005) and Perron and Zhu (2005). The impact of these on standard unit root tests has been well known since the seminal work of Perron (1989). In his original work, Perron (1989) treated the location of potential breaks as known. Subsequent approaches have focused on the case where the break date is unknown and is replaced by a break fraction estimator; see, *inter alia*, Perron (1997) and Perron and Rodríguez (2003). More recently, Harris et al. (2009) and Carrion-i-Silvestre et al. (2009), among others, extend these approaches to allow for the case where there is uncertainty as to whether a trend break has occurred by incorporating pre-test procedures for the presence of a trend break. Central to these approaches is an estimate of the break fraction and, in the case of the latter two, some form of pre-test for the presence of a break. Under a fixed magnitude trend break the break fraction estimator needs to be consistent at a rate faster than \( T^{-1/2} \), \( T \) denoting the sample size, for the resulting unit root test to be asymptotically valid in the case where a trend break occurs. As a result, the ordinary least squares [OLS] level break estimator of Bai (1994) has tended to be applied to the first differences of the series because it is consistent at rate \( T^{-1} \) for the true break fraction where a break occurs, while the corresponding OLS-based estimator of the trend break fraction in the levels is only consistent at rate \( T^{-1/2} \).

The aforementioned procedures, while allowing for the possibility of breaks in the deterministic trend function, do not allow for time-varying behaviour in the unconditional volatility (often referred to as non-stationary volatility in the literature) of the driving shocks. This is an important practical drawback given that a large number of empirical studies have reported a substantial decline, often referred to as the Great Moderation, in the unconditional volatility of the shocks driving macroeconomic series in the twenty years or so leading up to the Great Recession that started in late 2007, with a subsequent sharp increase again in volatility observed after 2007; see, *inter alia*, McConnell and Perez-Quiros (2000), Clark (2009), Stock and Watson (2012), and the references therein. Cavaliere et al. (2011) refine the approach of Harris et al. (2009) to use wild bootstrap unit root tests. However, their procedures are still based around the applying the OLS level break fraction estimator of Bai (1994) to the first differences and trend break pre-test, each of which were developed for homoskedastic innovations. While they show that both of these (and, indeed, the resulting unit root procedures) are asymptotically robust to time-varying volatility, their finite sample efficacy will clearly have important forward implications for the behaviour of the resulting unit root tests.

Our principal aim in this paper is to explore the properties of the OLS level break estimator of Bai (1994), which is central to the unit root testing procedures discussed above, in cases where the shocks can display non-stationary volatility and to develop and explore the properties of a corresponding feasible weighted least squares [WLS] level break estimator based around the use of data which have been weighted by a non-parametric estimate of the volatility path. We will consider a set-up for the volatility process of a very general form, taken to be unknown to the practitioner, which allows, for example, single and multiple abrupt variance breaks, smooth transition variance breaks, and trending variances. We will also allow for the presence of conditional heteroskedas-
ticity in the shocks. The feasible WLS estimator we propose uses adaptive methods to estimate the volatility path of the shocks. Adaptive methods have been successfully employed in a number of areas of the literature including inference on the parameters of finite-order unconditionally heteroskedastic autoregressive models by Phillips and Xu (2006) and Xu and Phillips (2008), testing for ARCH effects in unconditionally heteroskedastic autoregressive models by Patilea and Raïssi (2014), testing for long memory in unconditionally heteroskedastic ARFIMA models by Harris and Kew (2017) and adaptive testing for autocorrelation in Harris and Kew (2014), and for adaptive estimation of VAR models in Patilea and Raïssi (2012, 2013). Of most relevance to this paper, Boswijk and Zu (2017) propose an adaptive estimator of the unconditional variance process in the context of testing for a unit root in an autoregressive model driven by heteroskedastic errors, but where no allowance is made for the possibility of a trend break.

We establish the large sample properties of the OLS and feasible WLS break fraction estimators under a variety of assumptions on the magnitude of the level shift. For level shifts of either fixed (non-zero) magnitude or where the level shift magnitude is local-to-zero at a rate slower than the Pitman rate of $T^{-1/2}$, we demonstrate the consistency of these estimators, and indeed those from a generic class of residual sums of squares [RSS] based break fraction estimators. This rate of consistency is shown to be the same for the OLS and feasible WLS estimators and to be unaffected by the location of the break or by time variation in the volatility process. We also derive the asymptotic distributions of these estimators in the case where the magnitude of the level shift lies within a Pitman neighbourhood of zero. Elliott and Müller (2007) argue that the finite sample behaviour of break fraction estimators such as those considered in this paper is likely to be far better approximated for the sort of break magnitudes typically encountered in practice by asymptotic theory based on a Pitman localisation rather than a fixed magnitude break. The results we present in this paper entirely accord with this viewpoint. Under Pitman drift the limiting distributions of the OLS and feasible WLS estimators are shown to differ and to depend on the location and (local drift) magnitude of the level break and, to differing extents, on the time variation in the volatility process. Monte Carlo simulations are used to investigate and compare the finite sample behaviour of the estimators and these are shown to agree closely with the qualitative predictions that we draw from the limiting distributions under Pitman drift. In particular, they highlight the importance of the break location, break magnitude, and the time path of the volatility process for the finite sample performance of these estimators. They also show that the feasible WLS estimator can deliver significant improvements over the OLS estimator in certain heteroskedastic environments, most notably in cases where the level break occurs in a low volatility regime.

We then investigate to what extent the improved behaviour we observe for the feasible WLS level break fraction estimator relative to the standard OLS estimator of Bai (1994) when non-stationary volatility is present can effect improvements in the finite sample behaviour of the unit root tests discussed above designed for the case where a trend break (might) occur in the data and where the shocks are heteroskedastic. Here we also propose model selection methods based on the familiar Schwarz (1978) criterion to select between the trend break and no trend break models in the practically relevant case where it is unknown whether a trend break is present in the data or not. We discuss such information criteria based on standard OLS model estimation
and on feasible WLS model estimation, the latter based on the same adaptive estimator of the unconditional variance process as used in constructing the feasible WLS break fraction estimator. Simulation evidence suggests that the use of these feasible weighted estimators and feasible weighted information criteria can deliver unit root tests with significantly improved finite sample behaviour in the presence of non-stationary volatility relative to using their unweighted counterparts.

The paper is organised as follows. Our reference heteroskedastic level break model is outlined in section 2. Section 3 outlines a generic RSS level break fraction estimator which contains the standard OLS estimator of Bai (1994) and the infeasible WLS estimator as special cases. Here we also show how a feasible version of the WLS estimator can be constructed, using an adaptive estimator of the volatility path of the innovations. The large sample properties of these estimators are compared for both fixed, local and zero magnitude level shifts. The relative finite sample properties of these estimators in both homoskedastic and a variety of heteroskedastic environments are explored in section 4. In section 5 the application of level break estimation methods to the problem of unit root testing when breaks in trend and/or volatility may be present is discussed. Section 6 concludes. All proofs are collected in a mathematical appendix. Further simulation results are included in a supplementary appendix.

In what follows, ‘⌊·⌋’ denotes the integer part, ‘1(·)’ denotes the indicator function, and ‘x := y’ (‘x =: y’) indicates that x is defined by y (y is defined by x). The symbols ‘d’ and ‘p’ are respectively used to denote convergence in distribution and probability respectively, as $T \to \infty$. The maximum and minimum of a and b are denoted $a \lor b$ and $a \land b$, respectively. Finally, $D := D[0, 1]$ denotes the space of right continuous with left limit (càdlàg) processes on $[0, 1]$.

## 2 The Heteroskedastic Level Break Model

We consider the time series process $\{y_t\}$ generated according to the following level break model,

\[
y_t = \mu + \delta_T \cdot 1_{t > \lfloor \tau_0 T \rfloor} + e_t \tag{2.1}
\]

\[
e_t = \sigma_t \varepsilon_t. \tag{2.2}
\]

Equation (2.1) is seen to comprise the sum of a constant and level shift at time $\lfloor \tau_0 T \rfloor$, together with a stochastic component $e_t$, which we detail further below. As is standard, for the purposes of the large sample results which follow, we assume that the break date depends on the sample size such that the break occurs at a fixed fraction of the sample size; that is, we parameterise the breakpoint in terms of the break fraction $\tau_0$ where $0 < \tau_L \leq \tau_0 \leq \tau_U < 1$.

In order to allow for level breaks which are either of fixed magnitude or are local-to-zero at some rate in the sample size, $T$, we follow Elliott and Müller (2007) and parameterise the break magnitude parameter as $\delta_T := \delta T^{-d}$ with $\delta$ a fixed constant and $d \geq 0$. It can be seen that for a given fixed value of $T$ a level break exists in $y_t$ of (2.1) only if $\delta \neq 0$. No break occurs when $\delta = 0$, regardless of the value of $d$, while a level break of fixed magnitude $\delta$ occurs when $d = 0$ and $\delta \neq 0$. In the unconditionally homoskedastic case, where $\sigma_t = \sigma$ for all $t$, Bai (1997), shows that when $\delta \neq 0$, then $\tau_0$ is consistently estimated by least squares for any $0 \leq d < 1/2$. In particular, although
the magnitude of the level break shrinks here as the sample size increases, the level break is still sufficiently large for the location of the break, \( \tau_0 \), to be consistently estimated and for consistent tests to exist for testing for a level break. Again in the unconditionally homoskedastic case, setting \( d = 1/2 \) gives the Pitman drift rate for this problem such that \( \tau_0 \) cannot be consistently estimated nor can a consistent test for the presence of a level break be obtained. We will show that these consistency rates in \( d \) also hold in the heteroskedastic case we focus on in this paper.

To complete the specification of (2.1)-(2.2) the following conditions, collectively labelled Assumption \( \mathcal{A} \), will be assumed to hold.

**Assumption \( \mathcal{A} \).**

\( \mathcal{A}_1. \) The innovations \( \{\varepsilon_t\} \) form a martingale difference sequence with respect to the filtration \( \mathcal{F}_t \), where \( \mathcal{F}_{t-1} \subseteq \mathcal{F}_t \) for \( t = ..., -1, 0, 1, 2, ..., \) satisfying: (i) the global homoskedasticity condition:

\[
\frac{1}{T} \sum_{t=1}^{T} E(\varepsilon_t^2 | \mathcal{F}_{t-1}) \overset{p}{\to} 1,
\]

and (ii) \( E|\varepsilon_t|^r < K < \infty \) for some \( r \geq 4; \)

\( \mathcal{A}_2. \) The volatility term \( \sigma_t \) satisfies \( \sigma_t = \sigma(t/T) \), where \( \sigma(\cdot) \in \mathcal{D} \) is non-stochastic, bounded above and below as \( 0 < \underline{\sigma} \leq \sigma(s) \leq \bar{\sigma} < \infty \) for all \( s \), and satisfies a Lipschitz condition.

**Remark 2.1.** The process \( \{e_t\} \) in (2.2) is formed as the product of two components, \( \{\varepsilon_t\} \) and \( \{\sigma_t\} \). The former is assumed to satisfy a relatively weak globally stationary martingale difference assumption which allows for certain forms of conditional heteroskedasticity, such as that arising from stationary GARCH models, in the errors; see Davidson (1994, pp.454-455) for further discussion. Notice that, conditionally on \( \sigma_t \), the error term \( e_t \) has mean zero and time-varying variance \( \sigma_t^2 \).

**Remark 2.2.** Assumption \( \mathcal{A}_2 \) casts the dynamics of the disturbance variance in a quite general framework, requiring it only to be non-stochastic, bounded and to be smooth in between a countable number of jumps. A detailed discussion of the class of variance processes allowed under \( \mathcal{A}_2 \) is given in Cavaliere and Taylor (2007). They show that this includes variance processes displaying multiple volatility shifts, polynomially (possibly piecewise) trending volatility and smooth transition variance breaks, among other things. In the case where volatility displays jumps, these are not constrained to be located at the same point in the sample as the putative level shift, nor indeed are they required to lie within the search set, \( \Lambda \). The conventional unconditionally homoskedastic assumption that \( \sigma_t = \sigma \) for all \( t \), also satisfies Assumption \( \mathcal{A}_2 \), since here \( \sigma(s) = \sigma \) for all \( s \).

**Remark 2.3.** In order to focus our attention on the impact of non-stationary volatility of the form considered in Assumption \( \mathcal{A}_2 \) on level break estimation, we have omitted autocorrelation in the model for the disturbance \( e_t \). We will, however, discuss generalisations to allow for this at relevant points in the text.

### 3 Level Break Fraction Estimation

In this section we will consider level break fraction estimation based on the minimisation of a RSS criterion from a weighted regression model. In this general framework setting the weights to unity yields the OLS break fraction estimator while setting the weighting factor at time \( t \) to be equal to occurs at the same break fraction, \( \tau_0 \), as the level break.
the inverse of the volatility process at time $t$ yields the corresponding weighted least squares [WLS] estimator. A generic form for this estimator is first given in section 3.1. In section 3.2, under the assumption of non-stochastic weights, we establish the large sample properties of this estimator under “large” magnitude breaks (where the magnitude is either fixed and non-zero or local-to-zero at a slower rate than the Pitman drift), demonstrating the consistency of the estimator in such cases. In section 3.3 we analyse, again for non-stochastic weights, the large sample behaviour of the estimator under “small” magnitude breaks (where the break magnitude is either exactly zero or local-to-zero at the Pitman rate or faster). In practice the volatility process is unknown and so a feasible version of the WLS estimator will require an estimate of the volatility process. In section 3.4 we show how this can be done using an adaptive estimator and demonstrate that the resulting feasible WLS break fraction estimator is asymptotically equivalent to the infeasible WLS estimator based on the true volatility process.

3.1 Residual Sum of Squares Break Fraction Estimator

In what follows we define a generic RSS-based level break fraction estimator which contains weighted and unweighted break fraction estimators as special cases. To that end, define the weights $x_t$, $t = 1, ..., T$, and a generic RSS-based estimator

$$\hat{\tau} := \arg \min_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} \hat{e}_{\tau,t}^2$$

(3.1)

where, for any $\tau \in [\tau_L, \tau_U] \subset [0, 1]$, the residuals $\hat{e}_{\tau,t}^*$ are obtained from the OLS regression

$$y_t^* = \hat{\mu}_\tau x_t + \hat{\delta}_\tau (1_{t>T} \cdot x_t) + \hat{e}_{\tau,t}^*$$

(3.2)

where $y_t^* := x_t y_t$. Setting $x_t := 1$ for $t = 1, ..., T$, in (3.2) yields the usual OLS estimator of $\tau_0$ considered in Bai (1994), while setting $x_t := 1/\sigma_t$, $t = 1, ..., T$, yields the infeasible WLS estimator that obtains with knowledge of $\sigma_t$. In what follows, where we wish to make reference to the OLS and WLS estimators specifically, rather than the generic RSS-based estimator in (3.1), we will use the notation $\hat{\tau}_{OLS}$ and $\hat{\tau}_{WLS}$, respectively. The assumption of non-stochastic weights will be relaxed in section 3.4 when we detail our feasible WLS estimator of $\tau_0$ based on adaptive estimation of $\sigma_t$.

3.2 Asymptotic Behaviour of $\hat{\tau}$ under Large Breaks

We first analyse the large sample behaviour of $\hat{\tau}$ in the case where the trend break magnitude is “large” in that it can be either non-zero and fixed or is such that it is local-to-zero but at a rate slower than the Pitman drift rate of $T^{-1/2}$. We will show that the standard OLS estimator of $\tau_0$ retains the consistency property established under unconditional homoskedasticity in Bai (1997) and that the rate also holds for the corresponding WLS estimator, and indeed for any of a wide class of weights. These results are now formally stated in Theorem 1.

\footnote{The form of estimated coefficients $\hat{\mu}_\tau$ and $\hat{\delta}_\tau$ obviously depend on the choice of $x_t$ but this is omitted from the notation for brevity.}
Theorem 1. Let $y_t$ be generated according to (2.1) with $\delta_T := \delta T^{-d}$ and let Assumption $A$ hold. Moreover let the non-stochastic weights, $x_t = x(t/T)$, $t = 1, ..., T$, used in constructing $\hat{\tau}$ of (3.1) be such that $x(.)$ satisfies the same conditions as $\sigma(.)$ given in Assumption $A_2$. Then if $\delta \neq 0$ and $0 \leq d < 1/2$, it holds that $\hat{\tau} \overset{p}{\to} \tau_0$. Moreover, if $\delta \neq 0$ and $0 < d < 1/2$ then

$$T \delta_T^2 \frac{x(\tau_0)^4}{\sigma(\tau_0)^2} (\hat{\tau} - \tau_0) \overset{d}{\to} \arg \max_{s \in (-\infty, \infty)} Z(s),$$

where

$$Z(s) := \begin{cases} W_1(-s) - \frac{|s|}{T}, & s \leq 0 \\ \sqrt{\phi} W_2(s) - \frac{|s|}{T}, & s > 0 \end{cases}$$

in which $W_1$ and $W_2$ are independent standard Brownian motions each on $[0, \infty)$, and

$$\phi := \frac{\sigma(\tau_0)^2 x(\tau_0)^2}{\sigma(\tau_0)^2 \bar{x}(\tau_0)^2}, \quad \xi := \left( \frac{\bar{x}(\tau_0)}{\bar{Z}(\tau_0)} \right)^2,$$

where $\sigma(\tau_0) := \lim_{\tau \downarrow \tau_0} \sigma(\tau)$, $\sigma(\tau_0) := \lim_{\tau \uparrow \tau_0} \sigma(\tau)$, $\bar{x}(\tau_0) := \lim_{\tau \downarrow \tau_0} x(\tau)$ and $\bar{Z}(\tau_0) := \lim_{\tau \uparrow \tau_0} Z(\tau)$.

Remark 3.1. From Theorem 1 we obtain that $\hat{\tau}$ is a consistent estimator for $\hat{\tau}$ at rate $O_p(T^{-1} \delta_T^{-2})$ and, hence, is a consistent estimator regardless of the value of $d$, provided $0 < d < 1/2$, and irrespective of any conditional or unconditional heteroskedasticity present in $\sigma_t$ satisfying Assumption $A$, or the form of the weights, $x_t$, used in its construction. For the fixed break magnitude case where $d = 0$ and $\delta \neq 0$, $(\hat{\tau} - \tau_0)$ is of $O_p(T^{-1})$ under the conditions of Assumption $A$; see Cavaliere et al. (2011).

Remark 3.2. The results in Theorem 1 extend the results of Bai (1994) for the unweighted OLS estimator to cover both weighted and unweighted level break estimators and to allow for the general form of unconditional heteroskedasticity permitted in $\sigma_t$ under Assumption $A_2$. Bai (1997) establishes the same $O_p(T^{-1} \delta_T^{-2})$ rate in regression models (including (3.2)) allowing for weak dependence and conditional heteroskedasticity in the shocks, the latter of a similar form to that allowed under Assumption $A_1$. We specify martingale difference disturbances here in to order to focus attention on the role of non-stationary volatility in this model, but it can also be shown that the $O_p(T^{-1} \delta_T^{-2})$ consistency rate given in Theorem 1 continues to hold when $e_t$ is autocorrelated. For example, if $e_t = \sigma_t u_t$ where, as in equation (2) of Bai (1994), $u_t$ is generated by a linear process $u_t = C(L)\epsilon_t$, where $C(L) := \sum_{j=0}^{\infty} c_j L^j$ satisfies the standard summability condition (assumption B of Bai, 1994) $\sum_{j=0}^{\infty} j |c_j| < \infty$, and $\epsilon_t$ and $\sigma_t$ again satisfy the conditions in Assumption $A$, then the short run variance, $\sigma(\tau_0)^2$, in (3.3) would simply need to be replaced by the corresponding long run variance, $\sigma(\tau_0)^2 C(1)^2$.

Remark 3.3. A comparison of the results in Theorem 1 with the corresponding results in Proposition 3 of Bai (1997) shows that the presence of non-stationary volatility in $e_t$ has no effect on the large sample properties of $\hat{\tau}$, with the exception of the situation where a jump in the variance occurs at $\tau_0$ (a possibility also allowed for in the set-up considered by Bai, 1997) inducing the presence of the terms $\phi$ and $\xi$ in $Z(s)$. Other than this, $\sigma_t$ does not affect the limiting distribution of $\hat{\tau}$. The weighting factor $x_t$ is also seen to make no difference to the asymptotic distribution of

6
\( \hat{\tau} \) (again with the one exception where a break \( x(s) \) occurs at \( \tau_0 \)). Indeed, it can be seen that it is only the variance of the shocks at \( \tau_0 \) (or on either side of \( \tau_0 \) if the variance also changes at \( \tau_0 \)) that features in this limit; that is, the prediction from the large sample result in Theorem 1 is that where a level break occurs, the efficacy of \( \hat{\tau} \) is a function of the level break magnitude relative to the volatility of the shocks at that point in the sample only. As we will see in section 3.3, this contrasts with the “small” breaks asymptotics which predicts that the efficacy of \( \hat{\tau} \) is a function of the level break magnitude relative to (a function of) the entire sample path of the volatility function, and, moreover, depends on the weighting factor used in constructing \( \hat{\tau} \). The intuition behind these features is that the asymptotic distribution of \( \hat{\tau} \) when the break magnitude is “large” in Theorem 1 is derived from a functional central limit theorem [FCLT] applied only to observations within a shrinking neighbourhood of \( \tau_0 \). The càdlàg assumption on \( \sigma(\cdot) \) therefore implies that all such observations will have (asymptotically) the same variance. \( \square \)

### 3.3 Asymptotic Behaviour of \( \hat{\tau} \) under Small Breaks

Elliott and Müller (2007) argue that the asymptotic behaviour of break fraction estimators such as \( \hat{\tau} \) in (3.1) under “large” breaks is likely to provide a poor approximation to the finite sample properties of the estimator for the sort of break magnitudes typically encountered in practice. They argue that asymptotic theory based on the Pitman rate, \( T^{-1/2} \), is likely to provide more accurate predictions for the behaviour of \( \hat{\tau} \) in finite samples. They suggest that the asymptotics for \( d = 1/2 \) provides a continuous bridge between the no break case, \( \delta = 0 \), and the fixed magnitude break case considered in section 3.2. Accordingly, we now explore the asymptotic distribution theory for \( \hat{\tau} \) in cases where the break magnitude can be “small” (i.e. \( d \geq 1/2 \)) or, indeed, exactly zero (\( \delta = 0 \)). In particular, we will see that under such small breaks \( \hat{\tau} \) is no longer consistent for the true break location, \( \tau_0 \), but instead has a well-defined limiting distribution which depends on \( \tau_L \) and \( \tau_U \), the upper and lower limits, respectively, of the search set, and on the form of unconditional heteroskedasticity present through the function \( \sigma(\cdot) \), and the weight function \( x(\cdot) \) used in constructing the weighted estimator. For break magnitudes which are local-to-zero at the Pitman rate, \( d = 1/2 \), these limiting distributions additionally depend on \( \tau_0 \) and on the local break magnitude (i.e. the local drift parameter). The results we provide here will also cover the large break case considered previously. These results are now formally collected in Theorem 2.

**Theorem 2.** Let the conditions of Theorem 1 hold. Then for \( d \geq 0 \),

\[
\hat{\tau} \quad \overset{d}{\to} \quad \max_{\tau \in [\tau_L, \tau_U]} Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)
\]

where

\[
Q(\tau; x(\cdot), \sigma(\cdot), \delta, d) := \left( 1_{0 \leq d \leq 1/2} \frac{\delta}{\omega} (\chi(\tau_0)(1 - \chi(\tau_0)))^{1/2} \left( \chi(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)} \right) \right.
\]

\[
- 1_{d \geq 1/2} \frac{B_\eta(\tau) - \chi(\tau)B_\eta(1)}{(\chi(\tau)(1 - \chi(\tau)))^{1/2}} \left( \chi_1(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)} \right) \left( \chi(\tau_0)(1 - \chi(\tau_0)))^{1/2} \right) \left( \chi(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)} \right)
\]

(3.5)
with \( \omega^2 := (\int_0^1 x(s)^2 ds)^{-2} (\int_0^1 x(s)^4 \sigma(s)^2 ds) \),
\[
\chi(\tau) := \frac{\int_0^1 x(s)^2 ds}{\int_0^1 x(s)^2 ds}, \quad \eta(\tau) := \frac{\int_0^1 x(s)^4 \sigma(s)^2 ds}{\int_0^1 x(s)^4 \sigma(s)^2 ds},
\]
and
\[
\chi_1(\tau; \tau_0) := \left( \frac{\chi(\tau)/(1 - \chi(\tau))}{\chi(\tau)/((1 - \chi(\tau))} \right)^{1/2}
\]
and where \( B_0(\tau) = B(\eta(\tau)) \), with \( B(\cdot) \) a standard Brownian motion, is a variance-transformed Brownian motion; see, for example, Davidson (1994).

**Remark 3.4.** Theorem 2 establishes that \( \hat{\tau} \) has a well-defined asymptotic distribution with support \( \Lambda := [\tau_L, \tau_U] \) with its form depending on the increasing functions \( \chi(\cdot) : [0, 1] \to [0, 1] \) and \( \eta(\cdot) : [0, 1] \to [0, 1] \). The function \( \chi(\tau) \) is the cumulative weighting function associated with the weighted regression (3.2). As regards \( \eta(\tau) \), where \( x_t = 1 \), for all \( t \), this function is the generalisation to weighted estimation of the variance profile, \( (\int_0^1 \sigma(r)^2 dr)^{-1} \int_0^\tau \sigma(r)^2 dr \), of Cavaliere and Taylor (2007). The constant \( \omega^2 \) appearing in the right-hand component of (3.5) is an asymptotic measure of the scaled disturbance variance in the weighted regression (3.2) and relates to the average level of the volatility in the weighted data. For \( x_t = 1 \) (the unweighted OLS estimator) this quantity simplifies to \( \omega^2 := \int_0^1 \sigma(r)^2 dr \) which, by Assumption \( A_2 \), equals the limit of \( T^{-1} \sum_{t=1}^T \sigma_t^2 \), and may therefore be interpreted as the (asymptotic) average innovation variance. For \( x_t = 1/\sigma_t \) (the infeasible WLS estimator), \( \eta(\tau) = \omega^2 \int_0^\tau \sigma(r)^2 dr \) and \( \omega^2 = \left( \int_0^1 \sigma(r)^2 dr \right)^{-1} \).

Notice that, for any given volatility process \( \sigma(\cdot) \), the arithmetic/harmonic mean inequality implies that \( \omega^2 \) is strictly greater for the OLS estimator than it is for the WLS estimator, with the exception of the case where \( \sigma(s) = \sigma \) for all \( s \), as holds under homoskedasticity, where they are equal. \( \square \)

**Remark 3.5.** In the case of the OLS estimator, \( \hat{\tau}_{OLS} \), and under the Pitman drift rate, \( T^{-1/2} \), the general result in Theorem 2 coincides under homoskedasticity with the expression given for \( \hat{\tau}_{OLS} \) in Theorem 3 of Harvey et al. (2012, p.154). Notice also that the limiting function \( Q(\tau; x(\cdot), \sigma(\cdot), \delta, d) \) appearing in Theorem 2 does not depend on any nuisance parameters arising from conditional heteroskedasticity in \( e_t \) satisfying the conditions in Assumption \( A_1 \). As a consequence, the result in Theorem 3 of Harvey et al. (2012) which they derived under the assumption of IID innovations is therefore also valid under conditional heteroskedasticity. \( \square \)

**Remark 3.6.** As discussed in Remark 3.2, it is straightforward to extend the DGP to allow for autocorrelation in \( e_t \). In that case the disturbances \( e_t = \sigma_t u_t \) satisfy a heteroskedastic FCLT as usual, and \( \omega^2 \) in Theorem 2 would become \( \omega^2 = (\int_0^1 x(s)^2 ds)^{-2} (\int_0^1 x(s)^4 \sigma(s)^2 ds) C(1)^2 \). The implications of Theorem 2 are therefore qualitatively unchanged. \( \square \)

Inspection of (3.5) shows that there are two components to the limiting \( Q(\tau; x(\cdot), \sigma(\cdot), \delta, d) \) function. The first is non-stochastic and involves the true break fraction, \( \tau_0 \), the ratio of the break magnitude parameter \( \delta \) to \( \omega \), and the cumulative weighting function \( \chi(\cdot) \). The second is stochastic and depends on the variance transformed Brownian motion \( B_{\eta}(\cdot) \), and the cumulative weighting function, but not on either \( \tau_0 \) or \( \delta \). Heuristically one may view these components as, respectively, the “signal” and the “noise” with respect to the estimation of \( \tau_0 \). Notice that the non-stochastic
component does not directly depend on the sample path of the volatility process, $\sigma(\cdot)$, but rather on $\omega$ the summary measure of the average level of volatility in the weighted data. In contrast, the stochastic component depends on the sample path of the volatility process through $B_\eta(\cdot)$. The relative importance of the two components of $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ depends on the localisation rate, $d$, and the break magnitude parameter, $\delta$. We will now outline the four possible cases of interest:

**Case 1:** $d > 1/2$, $\delta \neq 0$. Here the level break magnitude shrinks to zero as the sample size increases at a faster rate than the Pitman rate ($d = 1/2$), and as a result the signal disappears from $Q(\tau; x_t, \sigma_t, \delta, d)$ such that there is no asymptotic information in $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ regarding $\tau_0$ and, as a consequence, the true break fraction cannot be consistently estimated. Here then the limiting distribution of $\hat{\tau}$ takes exactly the same form as it does in the case where no level break occurs, $\delta = 0$, which we consider in further detail in Case 4 below.

**Case 2:** $0 \leq d < 1/2$, $\delta \neq 0$. Here the break is sufficiently large such that the signal asymptotically dominates the noise. In this case the result in Theorem 2 reduces to

$$\hat{\tau} \xrightarrow{d} \arg \max_{\tau \in [\tau_L, \tau_U]} \left( \chi_1(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)} \right) = \tau_0,$$

and so the result agrees with the consistency result for $\hat{\tau}$ given in Theorem 1 for $0 \leq d < 1/2$.

**Case 3:** $d = 1/2$, $\delta \neq 0$. The most interesting case is where the Pitman drift rate, $d = 1/2$, holds. Here the signal and noise components of the limiting $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ function are seen to have equivalent orders of magnitude provided $\delta \neq 0$ (i.e. such that a local break occurs) and the $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ function captures the trade-off between the two; here $\tau_0$ cannot be consistently estimated, precisely because the signal does not dominate the noise, even asymptotically. It is of course this trade-off between the two components that makes the Pitman-based local asymptotics useful for predicting the finite sample behaviour of $\hat{\tau}$. Now, because

$$\max_{\tau} \left( \chi_1(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)} \right) = \chi(\tau_0; \tau_0) = 1$$

we may consider the scaling on the “signal” relative to the “noise” as being determined by the constant $\frac{d}{2}(\chi(\tau_0)(1 - \chi(\tau_0)))^{1/2}$. Considering this term by term, evidently a larger break magnitude $\delta$ increases the signal, other things equal, and makes the break fraction more accurately estimatable. Similarly the signal is larger, other things equal, the smaller is the constant $\omega$; recall that $\omega$ is a measure of the average level of volatility in the weighted data. Notice that in contrast to the “large” break asymptotics discussed in section 3.2, the “small” break asymptotics therefore predicts that the the efficacy of $\hat{\tau}$ is related to both the weighting scheme used in constructing $\hat{\tau}$, and to the magnitude of the level break, $\delta$, relative to a measure of the average volatility across the whole sample not just the level of volatility at the level break location; cf. Remark 3.3. The constant $\chi(\tau_0)(1 - \chi(\tau_0))$ is maximised for $\tau_0$ satisfying $\chi(\tau_0) = \frac{1}{2}$, showing that the signal for estimating $\tau_0$ is not necessarily highest at $\tau_0 = 0.5$ (as it is for the unweighted estimator) once weighting is applied. Rather it is maximised at the value of $\tau_0$ where the cumulative weighting reaches 0.5, i.e. $\int_0^{\tau_0} x(s)^2 ds = \frac{1}{2} \int_0^1 x(s)^2 ds$. We now provide two examples to illustrate.
Example 1: Consider the case where the weighted estimator is formed on the assumption that the variance follows the linear trend path \( \sigma_t^2 = 1 + t/T \). The corresponding weighted estimator obtains setting \( x_t = 1/(1 + t/T)^{1/2} \), and hence \( x(s) = (1 + s)^{-1/2} \) and \( \chi(\tau) = \log(1 + \tau) \). Then \( \chi(\tau_0) = 1/2 \) gives \( \tau_0 = \sqrt{2} - 1 \approx 0.414 \). Consequently, when weighting is used appropriate for a linear trend in the variance the position of a break fraction that maximises the asymptotic signal in \( Q(\tau; \sigma(\cdot), \sigma(\cdot), \delta, d) \) is \( \tau_0 \approx 0.414 \), rather than \( \tau_0 = 0.5 \). Notice that this result obtains regardless of whether this weighting leads to the true WLS estimator; that is, the result holds regardless of the true variance process, \( \sigma_t \).

Example 2: As a second example, suppose that it is assumed that there is one-time change in variance at time \( [T\lambda] \); that is, under the assumption that \( \sigma_t = 1 + \kappa 1_{\tau > \lambda} \). Here \( x_t = 1/\sigma_t \), so that

\[
x(s) = \begin{cases} 
1, & s \leq \lambda \\
(1 + \kappa)^{-1}, & s > \lambda
\end{cases}
\]

and, hence,

\[
\chi(\tau) = \begin{cases} 
\frac{\tau}{\lambda + (1 + \kappa)^{-2}(1 - \lambda)}, & \tau \leq \lambda \\
\frac{\lambda + (1 + \kappa)^{-2}(\tau - \lambda)}{\lambda + (1 + \kappa)^{-2}(1 - \lambda)}, & \tau > \lambda.
\end{cases}
\]

In the homoskedastic case, where no break in variance occurs, such that \( \kappa = 0 \), the weighted and unweighted estimators coincide. Solving here, \( \chi(\tau_0) = 1/2 \) gives \( \tau_0 = \sqrt{2} \), as expected; that is, with homoskedastic data the asymptotic signal in \( Q(\tau; \sigma(\cdot), \sigma(\cdot), \delta, d) \) is maximised for a break occurring in the middle of the sample. However, in the case where \( \lambda = 0.3 \) and \( \kappa = 2 \), such that the volatility increases threefold 30% of the way into the sample, then solving \( \chi(\tau_0) = 1/2 \) yields \( \tau_0 \approx 0.19 \). Using the weights appropriate to this form of variance step function therefore results in the largest “signal” for a break occurring at \( \tau_0 \approx 0.19 \). In contrast if \( \lambda = 0.7 \) and \( \kappa = 2 \), such that the volatility increases threefold 70% of the way into the sample, then solving \( \chi(\tau_0) = 1/2 \) yields \( \tau_0 \approx 0.40 \). We therefore see, again noting that these results obtain regardless of whether or not these weightings lead to the true WLS estimator in each case, that in these two examples the weighting based on either the assumption of an early or late increase in variance results in the largest “signal” for a break occurring in the lower variance regime of the sample, as seems intuitively reasonable. □

Case 4: \( \delta = 0 \). Consider finally the case where no trend break occurs, \( \delta = 0 \). Here the result in Theorem 2 implies (irrespective of the value of \( d \)) that

\[
\hat{\tau} d \rightarrow \max \limits_{\tau \in [\tau_L, \tau_U]} Q(\tau; x(\cdot), \sigma(\cdot), 0)
\]

\[
= \max \limits_{\tau \in [\tau_L, \tau_U]} \frac{(B_\eta(\tau) - \chi(\tau)B_\eta(1))^2}{(\chi(\tau)(1 - \chi(\tau)))}
\]

\[
= \max \limits_{\tau \in [\tau_L, \tau_U]} \frac{B_\eta(\tau)^2}{\chi(\tau)} + \frac{B_\eta(1) - B_\eta(\tau))^2}{1 - \chi(\tau)}.
\]

The result in (3.7) coincides with the form of the distribution in part 1(a) of Theorem 3.1 of Nunes et al. (1995) specialised to the case of a level shift and generalised to allow for heteroskedasticity. The latter is also in the general form reported in Proposition 1 of Elliott and M"{u}ller (2007).

The OLS estimator, \( \hat{\tau}_{OLS} \), applies equal weighting \( (x_t = 1) \) to the observations, implying \( \chi(\tau) = \tau \). Under homoskedasticity \( (\sigma_t = \sigma) \) we have \( \eta(\tau) = \tau \), in which case \( Q(\tau; 1, \sigma(\cdot), 0) \) reduces to the square of a standard Brownian Bridge \( B(\tau) - \tau B(1) \) divided by its standard deviation process, \( (\tau(1-
This scaled Brownian Bridge has a marginal standard normal distribution for each \( \tau \). In contrast, where unconditional heteroskedasticity is present, the limit \( Q(\tau; 1, \sigma(\cdot), 0) \) in (3.6) involves the square of \( (\tau(1 - \tau))^{-1/2}(B_{\eta}(\tau) - \tau B_{\eta}(1)) \) where \( \eta(\tau) = \int_0^\tau \sigma(s)^2 ds / \int_0^1 \sigma(s)^2 ds \) now differs from \( \tau \). Heuristically, this dependence suggests that the distribution of \( \hat{\tau}_{OLS} \) will be significantly affected by the presence of unconditional heteroskedasticity. The WLS estimator, \( \hat{\tau}_{WLS} \), applies weighting of the form \( x_t = 1/\sigma_t \), implying that \( \chi(\tau) = \eta(\tau) = \int_0^\tau \sigma(s)^{-2} ds / \int_0^1 \sigma(s)^{-2} ds \), and, hence, that \( Q(\tau; 1/\sigma(\cdot), \sigma(\cdot), 0) \) is a function of the variance transformed Brownian Bridge \( B_{\eta}(\tau) - \eta(\tau)B_{\eta}(1) \) divided by its standard deviation process, \((\eta(\tau)(1 - \eta(\tau)))^{1/2}\). As in the homoskedastic case, this latter scaled process has a marginal standard normal distribution for each \( \tau \). Although formally the asymptotic distribution of \( \hat{\tau}_{WLS} \) depends on the joint distribution of \( Q(\cdot; 1/\sigma(\cdot), \sigma(\cdot), 0) \) on \([\tau_L, \tau_U]\) and, hence, will depend on \( \sigma(\cdot) \) in some form, the marginal properties of the scaled process are suggestive that \( \hat{\tau}_{WLS} \) will be less affected by any unconditional heteroskedasticity present in \( e_t \) than \( \hat{\tau}_{OLS} \). This conjecture is supported by the simulation evidence reported in section 4.

### 3.4 A Feasible WLS Break Fraction Estimator

The WLS estimator, \( \hat{\tau}_{WLS} \), outlined in section 3.1 is infeasible in practice because it requires knowledge of the volatility process, \( \sigma_t \), \( t = 1, ..., T \). It can, however, be made operational by replacing \( \sigma_t \) in the formulation of \( \hat{\tau}_{WLS} \) by an estimate of \( \sigma_t \). In practice the volatility process could be estimated either parametrically or non-parametrically. The former could be useful where the practitioner wishes to specify a particular model for the volatility process but of course has the drawback that an incorrectly specified model will likely give a very poor estimate of the volatility path. Given our focus in this paper is on setting up general assumptions on the heteroskedasticity present in the shocks without assuming a parametric model for the volatility process, it is more natural for us to consider a two-step approach based on a non-parametric (adaptive) estimator of the volatility process. In this approach the volatility, \( \sigma_t \), is first estimated using the residuals from estimating the level break model as in (3.2) by standard OLS (i.e. treating the shocks as homoskedastic) and then substituting \( \sigma_t \) in the expression for \( \hat{\tau}_{WLS} \) by the the resulting estimator, \( \hat{\sigma}_t \). Our proposed estimator of \( \sigma_t \) is based on the approach developed in Hansen (1995) and Xu and Phillips (2008), which has recently been adapted to the unit root testing context by Boswijk and Zu (2017). We will demonstrate that the large sample behaviour of the resulting feasible weighted estimator coincides with that of the infeasible WLS estimator.

To that end, let \( \hat{e}_{\tau,t} := y_t - \hat{\mu}_t - \delta_{\tau} \mathbb{1}_{T>\lfloor \tau T \rfloor} \), \( t = 1, ..., T \), denote the standard OLS residuals which obtain from estimating (2.1) under the assumption that \( e_t \) is homoskedastic. In doing so an initial estimate of the level break location is needed. This could be provided by any form of the generic estimator \( \hat{\tau} \) given in (3.1) such that the consistency result in Theorem 1 holds and a natural choice would be the simple OLS estimator, \( \hat{\tau}_{OLS} \). Next let \( K(\cdot) \) be a kernel function, and let \( K_h(t) := K(t/h) \) with \( h > 0 \) a bandwidth. Then, given the residuals \( \hat{e}_{\tau,t} \), and \( K_h(t) \), a kernel smoothing estimator for \( \sigma_t^2 \) can be defined as

\[
\hat{\sigma}_t^2 := \frac{\sum_{i=1}^T K_h \left( \frac{\tau_{\tau,i}}{h} \right) \hat{e}_{\tau,i}^2}{\sum_{i=1}^T K_h \left( \frac{\tau_{\tau,i}}{h} \right)}.
\]
By choosing different kernel functions one can obtain either one-sided or two-sided smoothing. We will follow Xu and Phillips (2008) and set $K_h(0) = 0$, and also avoid the need for boundary value adjustments to (3.8) of the type discussed in Hansen (1995) by assuming the use of two-sided smoothing in what follows. In particular, we will assume that $K(\cdot)$ is a bounded non-negative function defined on the real line and is such that $\int_{-\infty}^{\infty} K(x) dx = 1$ and $0 < \int_{0}^{\infty} K(x) dx < 1$. The bandwidth, $h := h(T)$, is assumed to satisfy the (standard) rate condition that $h + (Th)^{-1} \rightarrow 0$ as $T \rightarrow \infty$. The practical implementation of the estimator $\hat{\sigma}_t^2$ depends on the choice of kernel function, $K(\cdot)$, and the bandwidth, $h$. Commonly used kernels which satisfy the stated conditions include the uniform, Epanechnikov, biweight and Gaussian functions. The bandwidth condition implies that $h \rightarrow 0$ but at a slower rate than $T^{-1/2}$. In practice bandwidth selection can be crucial to performance, and cross-validation and plug-in rules can be defined for $h$. The latter is used in the simulations in section 4 below.

If $\sigma(s)$ was continuous in $s \in [0, 1]$, then it would be possible to establish that $\hat{\sigma}_t^2$ in (3.8) was a uniformly consistent estimator for $\sigma_t^2$. However, we do not want to impose continuity on $\sigma(s)$ and we will show below that even without doing so the resulting feasible weighted break fraction estimator will have the same large sample properties as the infeasible estimator under the conditions stated above for the kernel function and bandwidth.

Based on the adaptive estimate $\hat{\sigma}_t^2$ we can then define the corresponding feasible version of the WLS estimator

$$\hat{\tau}_{FWLS} := \arg\min_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} e_{t, t}^2,$$

where $e_{t, t}$, $t = 1, ..., T$, are the OLS residuals from the weighted regression

$$\frac{y_t}{\hat{\sigma}_t} = \hat{\mu}_\tau + \hat{\delta}_\tau \left( 1_{t > [\tau T]} \cdot \frac{1}{\hat{\sigma}_t} \right) + \hat{\epsilon}_{t, t}.$$  \hspace{1cm} (3.9)

We now detail the large sample properties of the feasible WLS estimator, $\hat{\tau}_{FWLS}$. As in Xu and Phillips (2008), in order to do so we need to appropriately strengthen the moment condition in part (ii) of Assumption $A_1$.

**Theorem 3.** Let the conditions of Theorem 1 hold with Assumption $A_1$ (ii) replaced by $\sup_t E(\epsilon_t^3) < \infty$. If the kernel function $K(\cdot)$ and bandwidth $h$ satisfy the conditions stated below equation (3.8), then $\hat{\tau}_{FWLS} - \hat{\tau}_{WLS} \xrightarrow{p} 0$.

**Remark 3.7.** The result in Theorem 3 demonstrates that the feasible WLS level break estimator, $\hat{\tau}_{FWLS}$, based on the adaptive estimation of $\sigma_t$ is asymptotically equivalent to the infeasible WLS estimator $\hat{\tau}_{WLS}$. Consequently, all of the results given previously for the large sample properties of $\hat{\tau}_{WLS}$ apply equally to $\hat{\tau}_{FWLS}$.

**Remark 3.8.** It is straightforward to show that the adaptive estimator for $\sigma_t$ remains consistent (except, as usual, at the points of discontinuity of $\sigma(s)$) in the presence of serial correlation in $e_t$ of the form mentioned in Remark 3.2. The result in Theorem 3 will continue to hold in such cases. Boswijk and Zu (2017) also discuss the kernel estimation of variances in the presence of autocorrelation in a related unit root testing context.
4 Numerical Results

We now provide a detailed Monte Carlo comparison of the finite sample behaviour of the OLS and feasible WLS break fraction estimators, \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \) respectively, from section 3 under both homoskedasticity and a variety of heteroskedastic environments. Here we will also explore how useful the large sample results from the previous section are in predicting the finite sample behaviour of these estimators.

All of the simulation results reported in this paper were performed for the following settings. All experiments were based on 10,000 Monte Carlo replications programmed in Gauss 15 using the \texttt{rndn} random number generator. For both \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \) we set \( \tau_L = 0.2 \) and \( \tau_U = 0.8 \) in (3.1), thereby defining the set of possible breakpoints to be searched over for a given value of \( T \) as \( \{T/5, \ldots, 4T/5\} \). For the kernel variance estimator for \( \hat{\tau}_{FWLS} \) we used a QS kernel and plug-in bandwidth \( h = sT^{-0.2} \) where \( s \) is the sample deviation of the regressor \( 1, \ldots, T \) (see section 2.2.1 of Li and Racine, 2007), and the results were found to be quite insensitive to reasonable variations of this choice.

The Monte Carlo simulations reported in this section are based on the level break DGP:

\[
y_t = \mu + \delta \cdot 1_{t>T0} + \sigma_t \varepsilon_t, \quad t = 1, \ldots, T, \quad \text{with} \quad \varepsilon_t \sim \text{i.i.d.} N(0,1). \tag{4.1}
\]

Data were generated from this DGP allowing for both the no break case, \( \delta = 0 \), and for level breaks occurring at \( \tau_0 \in \{0.3, 0.5, 0.7\} \). The volatility process, \( \sigma_t \), was varied among the following models:

- **SD0**: \( \sigma_t = 1, t = 1, \ldots, T \)
- **SD1**: \( \sigma_t = 1 + \kappa \cdot 1_{t>T0} \)
- **SD2**: \( \sigma_t = 1 + \kappa \cdot 1_{t<T0} \), with \( \kappa \in \{1, 2\} \) and \( \lambda_0 \in \{0.3, 0.5, 0.7\} \)
- **SD3**: \( \sigma_t = 1 + \kappa \cdot (1_{t<T0} + 1_{t>T(1-\lambda_0)}) \), with \( \lambda_0 = 0.3 \) and \( \kappa \in \{1, 2\} \)
- **SD4**: \( \sigma_t = 1 + \kappa \cdot t/T \), with \( \kappa \in \{1, 2\} \).

SD0 is the case of unconditional homoskedasticity. SD1 (SD2) allows for an increase (decrease) in volatility at break fraction \( \lambda_0 \) from 1 to \( (1 + \kappa) \) \((1 + \kappa) \) to 1). SD3 allows for a double change in volatility from \((1 + \kappa) \) to 1 at break fraction \( \lambda_0 \) reverting back to \((1 + \kappa) \) at \((1 - \lambda_0) \). Finally SD4 generates a volatility process which follows a positive linear trend between 1 at the start of the sample and \((1 + \kappa) \) at the end of the sample.

While the full set of results are reported in Tables A.1-A.12 and Figures A.1-A.4 in the supplementary appendix, in the main text we report a representative selection of these results. In particular in Tables 1-4 we report the empirical mean and standard deviation, together with the root mean squared error [RMSE] from the empirical distributions of \( \hat{\tau}_{OLS} \) (Panel A in each table) and \( \hat{\tau}_{FWLS} \) (Panel B in each table) for samples of size 100 and 300 and for level break magnitudes \( \delta \in \{0,0.5,1\} \). Table 1 reports results for the case where no level break occurs \( (\delta = 0) \), for SD1 and SD2 each with \( \kappa = 2 \) and \( \lambda_0 \in \{0.3,0.5,0.7\} \), along with the results for SD0, SD3 and SD4. Tables 2, 3 and 4 report results for the case where a level break occurs. Table 2 reports results for SD0 for \( \kappa = 2 \) and \( \lambda_0 \in \{0.3,0.5,0.7\} \) for models with a level break occurring at each of \( \tau_0 = 0.3, 0.5 \) and 0.7 for break magnitudes \( \delta \in \{0.5,1\} \). Table 3 mirrors Table 2 but for SD2 for \( \kappa = 2 \) and \( \lambda_0 \in \{0.3,0.5,0.7\} \). Corresponding results for SD3 and SD4 are contained in Table
4. Figures 1-4 report corresponding plots of the empirical density functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ for samples of size 100, 200 and 300 and break magnitudes $\delta \in \{0, 0.5\}$, organised so that Figure 1 presents all of the results for the no level break case, while Figures 2, 3 and 4 present results for the case where a level break occurs at $\tau_0 = 0.3, 0.5$ and 0.7, respectively.

The behaviour of the $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ estimators and the role of the heteroskedasticity and its interplay with the location of the level break location in their finite sample properties can best be revealed by considering the results in Tables 1-4 and Figures 1-4 in tandem. A consideration of all of the results presented in the tables and figures is suggestive of the basic conclusion that while both $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ will, other things being equal, be drawn towards the location of a level break where it happens, and increasingly so as the sample size, $T$, and/or the break magnitude, $\delta$, increase, the $\hat{\tau}_{OLS}$ estimator is at the same time drawn towards periods of high volatility in the series, both where a level break occurs in the data and where it does not. This phenomenon is not evident in the weighted estimator, $\hat{\tau}_{FWLS}$, because, by construction, $\hat{\tau}_{FWLS}$ down-weights the data in periods of high volatility, thereby ameliorating the tendency of the unweighted OLS break estimator to be drawn towards high volatility periods.

To illustrate these effects, consider first the results in Table 1 and Figure 1 for the case where no level break occurs, $\delta = 0$. Here we see that for the homoskedastic case $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ behave almost identically with a relatively uniform empirical density across the search interval with slight pile-up effects at the ends of the search set, $\tau_L = 0.2$ and $\tau_U = 0.8$. Both have an empirical mean of about 0.5. However, when heteroskedasticity is present the two estimators behave quite differently. While the behaviour of $\hat{\tau}_{FWLS}$ is seen to be relatively unchanged from the homoskedastic case in all of the heteroskedastic cases considered, the behaviour of $\hat{\tau}_{OLS}$ varies considerably across the different non-constant volatility cases. In particular we see that the mass of the distribution of the estimator is redistributed towards high volatility periods vis-à-vis the homoskedastic case. This phenomenon is most obviously seen in Figure 1(g) which relates to the case where the volatility increases by a factor of 3 at $\lambda_0 = 0.7$. Here we see that a large bulk of the mass of the empirical density of $\hat{\tau}_{OLS}$ is now spread out across the high volatility period in the data, with the empirical mean of $\hat{\tau}_{OLS}$ now very close to 0.8, the upper limit of the search set. In contrast, the empirical density of $\hat{\tau}_{FWLS}$ in Figure 1(h) is seen to be almost unchanged from the homoskedastic case. This is of course to be expected as, by construction, $\hat{\tau}_{FWLS}$ down-weights the data in periods of high volatility, thereby reducing the tendency of the break estimator to be drawn towards such periods.

The tendency of $\hat{\tau}_{OLS}$ to be drawn towards high volatility periods within the data when no level break occurs has clear repercussions on the finite sample performance of $\hat{\tau}_{OLS}$ when a level break does occur in the data. The weighting of the data inherent in $\hat{\tau}_{FWLS}$ is also important for its efficacy where a level break occurs. In particular, the weighting works best in cases where the level break occurs in a low volatility period of the data, because here it down-weights the high volatility periods in the sample again reducing its tendency, relative to the OLS break estimator, to be drawn towards those periods. To illustrate this, consider Figures 2e and 2f relative to Figures 2a and 2b - in each case a level break of magnitude $\delta = 0.5$ occurs at $\tau_0 = 0.3$. In Figures 2a and 2b, where the volatility is constant, both $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ are centred on $\tau_0$ with the estimated densities becoming increasingly concentrated around $\tau_0$ as the sample size increases. However, in Figures
2e and 2f where the volatility increases threefold at $\lambda_0 = 0.5$, although the density of $\hat{\tau}_{FWLS}$ is almost identical to that seen in Figure 2b, the density of $\hat{\tau}_{OLS}$ is radically altered. A relative peak still exists at $\tau_0$, at least for the larger sample sizes considered, but it can be observed that, as also happens when no level break is present (see Figure 1e), a large mass of the density has shifted into the high volatility region with a relative peak seen at $\tau_U = 0.8$. Notice also that the performance of the $\hat{\tau}_{OLS}$ estimator is little improved between $T = 100$ and $T = 300$ here. Further illustration of these effects can also be seen from the associated results in Table 2, where the empirical mean of $\hat{\tau}_{OLS}$ is seen to be as high as 0.678 for $T = 100$ in this case.

The results also show that the weighted estimator is not a panacea and can in some cases display apparently inferior finite sample performance to $\hat{\tau}_{OLS}$. This can occur in cases where the level break occurs in a high volatility period of the data, and especially so where the period of high volatility is short-lived. Where the level break occurs within an extended period of high volatility, weighting is relatively innocuous and there is little difference seen between $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. This phenomenon occurs because here, as we have already observed, some of the mass of the unweighted $\hat{\tau}_{OLS}$ estimator is attracted to the high volatility regime, regardless of whether a level break occurs or not. In contrast, $\hat{\tau}_{FWLS}$ down-weights the high volatility period and, as a result, where a level break occurs within the high volatility regime $\hat{\tau}_{FWLS}$ will have less mass in the vicinity of the level break than the $\hat{\tau}_{OLS}$ estimator. However, for $\hat{\tau}_{OLS}$ this mass will be spread across the high volatility regime and so one will still see reduced performance relative to the homoskedastic case (even where the level and volatility break locations coincide) and increasingly so the longer the duration of the high volatility period. A good illustration of this phenomenon is seen in Figures 4a-4h relating to the case where a level break occurs at $\tau_0 = 0.7$. In the homoskedastic case, $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ perform similarly well. However, in cases where the volatility increases by a factor 3 at $\lambda_0$ we see that the performance of both estimators deteriorates. For $\hat{\tau}_{FWLS}$ the performance is roughly similar regardless of where in the sample the volatility break occurs. For $\hat{\tau}_{OLS}$ the pile up of mass in the high volatility region is evident (see also Figures 1c, 1e and 1f) and so it has more mass in the vicinity of the level break - increasingly so as $\lambda_0$ increases, such that the duration of the high volatility region decreases. Indeed, for the case of the longest period of high volatility where this regime starts at $\lambda_0 = 0.3$ the empirical densities of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ are relatively similar.

We can also use the results in Figures 1-4 and Tables 1-4 to explore further how well the finite sample behaviour of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ conform to the predictions of the large sample theory given in Theorem 1 for level breaks of fixed magnitude and Theorem 2 for level breaks whose magnitude is local-to-zero at the Pitman rate, $d = 1/2$. To that end, we first recall that Theorem 1 predicts that both $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ will be consistent for $\tau_0$ regardless of the pattern of heteroskedasticity present. Looking at the results for the constant volatility case in Table 2 and Figures 2-4 we indeed see this prediction being borne for both $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ with each of the empirical bias, standard deviation and RMSE of the estimators decreasing, other things equal, the larger the sample size, $T$, for a fixed break magnitude, $\delta$. These quantities also all decrease as the break magnitude increases while keeping the sample size constant, as anticipated by the result in Theorem 2 when $d = 1/2$. Moreover, as anticipated from our discussion in section 3.3, these quantities are all at their smallest when the level break occurs in the middle of the sample, i.e. $\tau_0 = 0.5$. In the heteroskedastic cases
considered, the reported results are still in general consistent with these predictions from Theorems 1 and 2 but much less obviously so.

A key prediction from Theorem 2 is that for level breaks whose magnitude is local-to-zero at the Pitman rate, the asymptotic distributions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ will differ from one another, and that the form of these limiting distributions will differ for both estimators according to the pattern of unconditional heteroskedasticity present. Moreover, this large sample result also predicts that the efficacy of the two estimators will depend on the break magnitude, $\delta$, considered relative to the parameter $\omega$. We recall from the discussion in section 3.3 that $\omega$ provides a measure of the average volatility in the weighted data and is a function of the volatility path $\sigma(\cdot)$ and of the weighting function used (and therefore differs between $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$). In contrast, Theorem 1 predicts that the two estimators will be identically behaved and that it is only the volatility in the neighbourhood of the level break that matters for the efficacy of the estimators. The results in Figures 1-4 and Tables 1-4 clearly demonstrate the superiority of the foregoing predictions from Theorem 2. That the finite sample behaviour of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ differ significantly from one another according to the form of heteroskedasticity present has already been discussed in some detail above. To illustrate the role of $\omega$, consider Figures 2m-2p together with Table 4 which relate to the case where a level break occurs at $\tau_0 = 0.3$ and the volatility displays an upward linear trend through the sample (SD4). We can see that relative to the homoskedastic case (see Figures 2a and 2b and Table 1) the efficacy of both $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ is considerably reduced when a trend in volatility is present, and increasingly so as the magnitude of the linear trend, $\kappa$, is increased. It is also seen that the peaks in the empirical densities at $\tau_0$ are somewhat smaller for $\hat{\tau}_{OLS}$ than for $\hat{\tau}_{FWLS}$. Noting that $\omega$ increases as the magnitude of the linear trend increases and is higher for $\hat{\tau}_{OLS}$ than for $\hat{\tau}_{FWLS}$ and that the level break occurs near the start of the series (where the volatility at that point is relatively small compared to the average volatility), we clearly see that the efficacy of the estimators in finite samples is related to the average volatility across the whole sample rather than just to the volatility level near the level break, and to the weighting function used in constructing the level break fraction estimator, in each case as Theorem 2 predicts.

To illustrate further the usefulness of the asymptotic approximation provided by Theorem 2, Figure 5 provides simulations of the distribution of $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ with comparisons to the finite sample distributions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$ from the same DGPs. Figure 5a shows, in the broken lines, the simulated sampling distributions of $\hat{\tau}_{OLS}$ for $T = 100, 200, 300$ from a DGP with no level shift ($\delta = 0$) and heteroskedasticity of the form SD2 with $\kappa = 2$ and $\lambda_0 = 0.7$. The solid line shows the asymptotic approximation for this same DGP, obtained using a 2000 step discretisation. Clearly in this case the distribution of $\hat{\tau}_{OLS}$ is seen to be essentially the same across these sample sizes. Figure 5b shows the same information for $\hat{\tau}_{FWLS}$. The asymptotic approximation remains very accurate in this case, other than a minor divergence around the time of the break in variance ($\lambda_0 = 0.7$) arising from the differences of the finite sample properties of the kernel variance estimator used for finite $T$ and the true variance process that is used in $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$. These two figures illustrate

\[ \text{In particular, in this example the parameter } \omega^2 = 1 \text{ when } \kappa = 0 \text{ (the homoskedastic case) for both } \hat{\tau}_{OLS} \text{ and } \hat{\tau}_{FWLS}, \text{ but for } \hat{\tau}_{OLS}, \omega^2 = 2 \frac{1}{3} \text{ when } \kappa = 1 \text{ and } \omega^2 = 4 \frac{1}{3} \text{ when } \kappa = 2, \text{ while for } \hat{\tau}_{FWLS}, \omega^2 = 2 \text{ when } \kappa = 1 \text{ and } \omega^2 = 3 \text{ when } \kappa = 2. \]
the applicability of the stochastic component of $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ for predicting the finite sample behaviour of the estimators when no level shift occurs.

Figures 5c and 5d graph the simulated finite sample and asymptotic distributions when a level shift of magnitude $\delta_T = \delta T^{-1/2}$ at $\tau_0 = 0.5$ is included in the DGP. Both figures show that the approximation provided by the asymptotic distribution given in Theorem 2 is very accurate in cases where both a level shift and unconditional heteroskedasticity are present in the DGP. The level shift magnitude in the previous simulations was held fixed, while in this case it becomes smaller as the sample size increases. Figure 3g and 3h show the finite sample distributions with fixed level shift magnitude of 0.5, and the asymptotic approximations given in Figures 5c and 5d evidently match well with this for $T = 300$ in particular, since for $T = 300$ the implied level shift magnitude $\delta_T = 8T^{-1/2} = 0.46$ is close to 0.5.

The heuristic discussion of Theorem 2 given in Example 2 of Case 3 above can also be illustrated numerically. An additional simulation was carried out based on the DGP (4.1) with

$$
\sigma_t = \frac{(\lambda_0(1 + \kappa)^2 + (1 - \lambda_0))^{1/2}}{1 + \kappa} (1 + \kappa \cdot 1_{t > \lfloor \lambda_0 T \rfloor}).
$$

The additional scaling of $\sigma_t$ relative to that in Example 2 of Case 3 above is used so that $\omega^2 = 1$ for this standard deviation process for any values of $\lambda_0$ and $\kappa$, which allows meaningful comparisons to be drawn across these two parameters. In particular, simulation can be used to obtain approximations to the true value of the break fraction $\tau_0$ that can be more accurately estimated from a finite sample data, based on different values of the standard deviation parameters. The analysis of the signal given in Example 2 of Case 3 above suggested that the multiplier on the deterministic signal component of $Q(\tau; x(\cdot), \sigma(\cdot), \delta, 1/2)$ would be maximised at $\tau_0 = 1/2$ for $\kappa = 0$ (homoskedasticity), at $\tau_0 \approx -0.19$ for $\kappa = 2$ and $\lambda_0 = 0.3$, and at $\tau_0 \approx -0.40$ for $\kappa = 2$ and $\lambda_0 = 0.7$. These calculations do not constitute formal proof that these values of $\tau_0$ are those that can be most efficiently estimated under these variance patterns. However, the simulation results summarised in Figure 6 show that they provide a good approximation in these cases at least. Figure 6 gives plots, one for each of the three variance processes discussed here, of the simulated RMSEs of $\hat{\tau}_{FWLS}$ for estimating each of the indicated values of $\tau_0$ the horizontal axis, based on samples of size $T = 200$ and with break size $\delta_T = 8T^{-1/2}$ (i.e. the same break size considered in Figure 5 for the purposes of comparison). The values of $\tau_0$ that returned minimum RMSE of $\hat{\tau}_{FWLS}$ in each case were respectively $\tau_0 = 0.49$ ($\kappa = 0$), $\tau_0 = 0.18$ ($\kappa = 2$, $\lambda_0 = 0.3$) and $\tau_0 = 0.40$ ($\kappa = 2$, $\lambda_0 = 0.7$). Again we see that, even when used heuristically, the asymptotic approximation provided by $Q(\tau; x(\cdot), \sigma(\cdot), \delta, 1/2)$ in Theorem 2 provides a very useful guide to finite sample properties.

5 An Application to the Unit Root Testing Problem

We have so far demonstrated how non-stationary volatility can affect the asymptotic and finite sample properties of the OLS and (feasible) WLS estimators of the timing of a level break in time series data. However, such estimation is rarely the ultimate goal of the analysis of the time series; rather, the level break estimator is used as an input into subsequent inference. In this section we illustrate the relevance of these findings in one such important case, where the estimated level break
is used to date a trend break in a time series prior to running a unit root test. We will consider two possible scenarios which have been considered in the literature. The first, detailed in section 5.1, relates to the scenario where a trend break is \textit{known} to have occurred in the data but its location is unknown to the practitioner who must therefore employ some estimate of the unknown trend break location. The second, detailed in section 5.2, relates to the empirically more relevant scenario where the practitioner neither knows whether a trend break has occurred, nor the location where it might occur. Both cases require a break date estimator, while the latter also requires some form of model selection procedure for assessing whether a break has occurred or not which can again be based on standard OLS regression estimation or on feasible WLS estimation. A Monte Carlo comparison of the relative finite sample properties of the unit root procedures based on weighted and unweighted trend break pre-tests and break fraction estimators will then be provided in section 5.3.

The underlying DGP is common to both the case where a trend break is known to have happened and where it is unknown as to whether a trend break has occurred and so we outline this first before moving to the two separate cases. To that end, consider the time series process \( y_t \) generated according to the following DGP,

\[
\begin{align*}
y_t = \begin{cases} 
\mu_{0,0} + \mu_{1,0}t + z_t, & t = 1, \ldots, \lfloor \tau_0 T \rfloor \\
\mu_{0,1} + \mu_{1,1}t + z_t, & t = \lfloor \tau_0 T \rfloor + 1, \ldots, T
\end{cases}
\end{align*}
\] (5.1)

where \( z_t = \phi_T z_{t-1} + e_t \), (5.2) and where \( e_t \) is generated according to (2.2) and is again taken to satisfy the conditions of Assumption \( A \). As is common in this literature, we also assume that the initial condition satisfies \( T^{-1/2} z_0 \overset{P}{\to} 0 \). This latter condition can be weakened, but at the expense of additional complexity. In (5.2) we will follow the convention in the unit root testing literature and focus on the near-integrated autoregressive model, \( H_c : \phi_T := 1 + c/T \) with \(-\infty < c \leq 0\). We will therefore be concerned with testing the unit root null hypothesis, \( H_0 : c = 0 \), against local alternatives, \( H_c \) where \( c < 0 \).

In the context of the observation equation in (5.1), \( y_t \) has a linear trend with break in both intercept and slope occurring at time \( \lfloor \tau_0 T \rfloor \). Following Harris \textit{et al.} (2009) and Cavaliere \textit{et al.} (2011), among others, we will focus on the situation where the trend function is restricted to be continuous at the break point, so that the coefficients satisfy \( \mu_{0,0} + \mu_{1,0} \lfloor \tau_0 T \rfloor = \mu_{0,1} + \mu_{1,1} \lfloor \tau_0 T \rfloor \).

In this case the trend specification can be written as

\[
y_t = \alpha + \mu t + \delta_T 1_{t > \lfloor \tau_0 T \rfloor} (t - \lfloor \tau_0 T \rfloor) + z_t
\] (5.3)

with \( \alpha := \mu_{0,0}, \mu := \mu_{1,0} \) and \( \delta_T := \mu_{1,1} - \mu_{1,0} \) (allowing for the magnitude of the break to depend on \( T \) as the previous sections). Taking first differences we obtain

\[
\Delta y_t = \mu + \delta_T 1_{t > \lfloor \tau_0 T \rfloor} + \Delta z_t,
\] (5.4)

\footnote{The imposition of continuity on the trend function makes the connection to the level shift results clear and simple. The restriction is not compulsory, however, as without it the equation corresponding to (5.4) would be given by \( \Delta y_t = \mu + \lambda 1_{t=\lfloor \tau_0 T \rfloor} + \gamma 1_{t > \lfloor \tau_0 T \rfloor} + \Delta z_t \), and the effect of the additional impulse dummy variable \( 1_{t=b_0} \) is asymptotically negligible.}
where $\Delta := (1 - L)$ denotes the first difference operator. Under the unit root null hypothesis, $H_0$, (5.4) can be seen to coincide with (2.1) on replacing $y_t$ by $\Delta y_t$ in the latter. Consequently, the results obtained in section 3 relating to the estimation of the level break location therefore continue to apply in this context, so that we estimate the trend break location via level break estimation applied to the first differences of the data.

### 5.1 The Case Where a Trend Break is Known to Have Occurred

In this first scenario we consider, the practitioner knows that the trend break magnitude $\delta$ is non-zero, but does not know the true location, $\tau_0$, of the trend break.

We will base our unit root test on Dickey-Fuller [DF] type statistics which model the trend break. These statistics are based on a two step procedure whereby the data are de-trended in the first step and in the second step a standard DF test is applied to the de-trended data. We will follow the recent literature and use the quasi-difference [QD] de-trending approach of Elliott et al. (1996) in what follows although OLS de-trending could alternatively be used. For a generic trend break location, $\tau$, the QD de-trended data are given by $\hat{\varepsilon}_{\tau,t} := y_t - X_t(\tau)'\hat{\theta}_c$, where $X_t(\tau) := \{(1, t, (t - [T\tau]) \cdot 1_{t \geq [T\tau]}\}'$ and $\hat{\theta}_c$ the vector of OLS parameter estimates from the regression of $y_{c,t}$ on $X_{c,t}(\tau)$, with $y_{c,1} := y_1$, $y_{c,t} := y_t - \hat{\phi}_T y_{t-1}$, $t = 2, ..., T$; $X_{c,1}(\tau) := X_1(\tau)$, $X_{c,t}(\tau) := X_t(\tau) - \hat{\phi}_T X_{t-1}(\tau)$, $t = 2, ..., T$, and where $\hat{\phi}_T := 1 + \bar{c}/T$, where $\bar{c}$ is the QD parameter. The QD de-trended data $\hat{\varepsilon}_{\tau,t}$ can then be used to estimate the DF regression

$$\hat{\varepsilon}_{\tau,t} = \hat{\phi}_{\tau} \hat{\varepsilon}_{\tau,t-1} + \hat{\epsilon}_{\tau,t}$$

and hence to obtain the usual DF $t$-statistic

$$t_{\tau} := \frac{\hat{\phi}_{\tau} - 1}{\text{s.e}(\hat{\phi}_{\tau})}.$$  

Theorem 4 provides the limiting distribution under the local alternative $H_c$ of $t_{\tau}$ when evaluated at the true break fraction $\tau = \tau_0$. The theorem also shows that this limit is unchanged when $\tau_0$ is replaced by either the OLS break fraction estimate, $\hat{\tau}_{OLS}$, or the corresponding feasible WLS estimate, $\hat{\tau}_{FWLS}$. We will use the simplified notation $t_{OLS}$ and $t_{FWLS}$ for the DF tests based on $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$, respectively, in what follows.

**Theorem 4.** Let $y_t$ be generated according to (5.1)-(5.2) and with $e_t$ generated according to (2.2), and let the conditions of Assumption $\mathcal{A}$ hold. Let $\delta_T = \delta T^{-d}$, $d \geq 0$. Then, under $H_c$:

(i) For any $d \geq 0$, and regardless of whether $\delta = 0$ or $\delta \neq 0$,

$$t_{\tau_0} \overset{d}{\rightarrow} \frac{1}{2}(Z(1; \tau_0, c, \bar{c}, \eta)^2 - 1) \left(\int_0^1 Z(s; \tau_0, c, \eta)^2 ds\right)^{1/2} := \xi(\tau_0, c, \bar{c}, \eta)$$

where

$$Z(s; \tau, c, \bar{c}, \eta) := B^c_\eta(s) - X(s; \tau)' \left(\int_0^1 X_c(s; \tau)X_c(s; \tau)' ds\right)^{-1} \int_0^1 X_c(s; \tau) dB^c_\eta(s; \bar{c})$$
and 
\[ B^c_\eta(s) := \int_0^s \exp(c(s-r))dB_\eta(r), \quad B^c_\eta(s; \tilde{c}) := B^c_\eta(s) - \tilde{c}\int_0^s B^c_\eta(r)dr, \]
with \( B^c_\eta(\cdot) \) as defined in Theorem 2, and
\[ X(s; \tau) := \left( \begin{array}{c} s \\ (s - \tau) \wedge 0 \end{array} \right), \quad X_c(s; \tau) := \left( \begin{array}{c} 1 - \tilde{c}s \\ 1 - \tilde{c}((s - \tau) \wedge 0) \end{array} \right). \]

(ii) For \( 0 \leq d < 1/2 \), and provided \( \delta \neq 0 \), it holds that: (a) \( t_{OLS} - t_{\tau_0} \overset{p}{\to} 0 \), and (b) provided the additional conditions of Theorem 3 hold, \( t_{FWLS} - t_{\tau_0} \overset{p}{\to} 0 \).

**Remark 5.1.** The results in Theorem 4 and in what follows will continue to hold in the case where \( e_t \) admits serial correlation of the form outlined in Remark 3.2. Under the standard invertibility condition on \( C(L) \) that \( C(z) \neq 0 \) for all \( |z| \leq 1 \), the serial correlation in \( e_t \) can be captured in the usual way using an augmented Dickey-Fuller [ADF] statistic, whereby (5.5) is augmented by the addition of the lagged dependent variables, \( \{\Delta y_{t-j}\}_{j=1}^p \), where \( p \) satisfies the usual rate condition that \( 1/p + p^3/T \to 0 \), as \( T \to \infty \).

**Remark 5.2.** The results in part (ii) of Theorem 4 might appear to contradict with Proposition 3 of Kim and Perron (2009,p.12) where it is shown that for some generic break fraction estimator, \( \hat{\tau} \), the break fraction, \( \tau_0 \) must be consistently estimated at some rate greater than \( T^{1/2} \) in order for a DF test based on \( \hat{\tau}, \hat{\tau}_t \) say, and \( t_{\tau_0} \) to be asymptotically equivalent. However, the result in Kim and Perron (2009) relates only to the case where the trend break magnitude \( \delta_T \) is fixed and non-zero (see their Assumption 1 on page 3), and therefore corresponds to the specific case of \( d = 0 \) and \( \delta \neq 0 \) in Theorem 4. In this case we know from Theorem 1 that both \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \) are consistent at rate \( O_p(T^{-1}) \), which certainly satisfies the condition in Proposition 3 of Kim and Perron (2009, p.4). In the more general set-up we consider here, the trend break magnitude and convergence rate of \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \) change together; as the break magnitude slows, so commensurately does the convergence rate of \( \hat{\tau}_{OLS}, \hat{\tau}_{FWLS} \). In particular, where the trend break magnitude is of order \( T^{-d}, \) \( d \geq 0 \), then, as shown in Theorems 1 and 3, respectively, \( (\hat{\tau}_{OLS} - \tau_0) \) and \( (\hat{\tau}_{FWLS} - \tau_0) \) are both of \( O_p(T^{2d-1}) \). However, this rate of consistency is still sufficiently fast for the asymptotic equivalence results in part (ii) of Theorem 4 to hold, precisely because the magnitude of the trend break is shrinking commensurately with the reduced consistency rate.

**Remark 5.3.** The result in Theorem 4 relates to the “large” break case of section 3.2 where \( 0 \leq d < 1/2 \) in the localisation of the trend break magnitude, such that the trend break location \( \tau_0 \) can be consistently estimated. Localisations which converge to zero at a faster rate, as considered in section 3.3, including the Pitman drift rate where \( d = 1/2 \), are excluded. Our aim in this section is not to provide a comprehensive treatment of the large sample properties of unit root tests in the present setting but rather to show how weighted trend break estimators can improve the finite sample properties of unit root tests relative to standard OLS estimation. However, the results could be extended to cover the case of \( d \geq 1/2 \). For \( d = 1/2 \), results comparable to those given in section 5 of Harvey et al. (2012), but generalised by the non-stationary volatility allowed for under Assumption \( A_2 \), would be obtained. For \( d > 1/2 \), as discussed in Case 1 in section 3.3, the
magnitude of the trend break would be such that it would lead to trend break estimators which behave asymptotically the same as in the no break case. As a result, both $t_{OLS}$ and $t_{FWLS}$ would converge to limiting distributions of a similar form to that given in (5.7) evaluated not at $\tau_0$ but at the random outcomes of $t_{OLS}$ and $t_{FWLS}$ within the search set $[\tau_L, \tau_U]$. As we will subsequently discuss in section 5.2, the model selection criteria discussed there would select the no trend break model rather than the trend break model here and, hence, one would construct a standard DF test statistic, which we denote $t_0$ in what follows, allowing only for a constant and linear trend in the QD detrending, by replacing $X_t(\tau)$ with $X_t := (1, t)'$ in the de-trending step.

**Remark 5.4.** Theorem 4 replicates the result given for the $t(\bar{\tau})$ statistic (which is based around the OLS break fraction estimator, $\hat{\tau}_{OLS}$) in part (ii) of Theorem 1 of Cavaliere et al. (2011, p.966) which pertains to the case of a trend break of fixed magnitude, $\delta \neq 0$ and $d = 0$. Theorem 4 shows that the result also holds for $t_{FWLS}$, the DF statistic based on the feasible WLS break fraction estimator, and that it continues to hold (for both break fraction estimators) for breaks which are local to zero, provided the localisation rate is smaller than the Pitman rate, $d = 1/2$. 

**Remark 5.5.** The (common) limiting null distribution of $t_{OLS}$ and $t_{FWLS}$, which obtains on setting $c = 0$ in (5.7), is seen to depend on the volatility process $\sigma(\cdot)$ through the presence of the heteroskedastic Brownian motion $B_\eta(\cdot)$. This is also the case for the corresponding no-break statistic, $t_0$, defined in Remark 5.3, whose limiting distribution under $H_c$ coincides with that given for the $MZ_t$ statistic in Theorem 1 of Cavaliere and Taylor (2008, pp.49-50). Consequently for pivotal inference on the unit root null hypothesis, $H_0$, we will need to base these tests on either the simulated critical value approach outlined in section 4.2 of Cavaliere and Taylor (2007) or a wild bootstrap approach, the latter outlined for the $t_0$ statistic in section 4.1 of Cavaliere and Taylor (2008), and for the trend break case in Algorithm 1 of Cavaliere et al. (2011, p.971). In the no trend break case, $\delta = 0$, the asymptotic validity of the simulated critical value and wild bootstrap methods is established in Cavaliere and Taylor (2007) and Cavaliere and Taylor (2008), respectively. In the case where a trend break occurs, $\delta \neq 0$, both approaches can be shown to deliver asymptotically pivotal inference for $t_{OLS}$ and $t_{FWLS}$ under the conditions of Theorem 4. For the wild bootstrap approach, the proof of asymptotic validity follows directly from Cavaliere et al. (2011), noting the asymptotic equivalence of $t_{OLS}$ and $t_{FWLS}$. The asymptotic validity of the simulated critical value approach follows using the same arguments as are given in section 4 of Cavaliere and Taylor (2007).

### 5.2 The Case Where it is Unknown if a Trend Break Has Occurred

In practice it is unlikely to be known with certainty whether or not a trend break has occurred. Allowing for a trend break which is not there (and, as a result, estimating a phantom break date) results in a unit root test with a different limiting distribution from when a trend break is present; see Remark 5.3. As a result, the correct asymptotic critical value for the unit root test that allows for a trend break fitted at this estimated date differs between the break and no break cases. In order to control asymptotic size the test must be based on the no break asymptotic critical value; using a critical value based on the estimated break fraction leads to an over-sized test when no trend break
occurs. However, doing so leads to a loss in test power, even asymptotically, both where a break occurs because a conservative critical value is being used, and where a break does not occur because the inclusion of a redundant trend break regressor leads to a considerable power loss relative to the corresponding unit root test that does not allow for trend break; see, for example, the numerical results presented in section 5 of Harris et al. (2009) and sections 3.2 and 5 of Cavaliere et al. (2011). In this section we show that a modified version of the usual Schwarz (1978) criterion [SC] can be used to select between the trend break and no trend break versions of the unit root tests. We also demonstrate that break selection, like break dating, can benefit from weighting for unconditional heteroskedasticity and can therefore result in unit root test procedures with improved properties.

We calculate the SC for break selection based on the representation for \( y_t \) provided by equations (5.2) and (5.3). For the calculation excluding the break, define the OLS residuals \( \tilde{e}_{0,t} \) from an OLS regression of \( y_t \) on an intercept, trend \( (t) \) and \( y_{t-1} \), with associated residual variance \( s^2_0 := T^{-1} \sum_{t=2}^{T} \tilde{e}_{0,t}^2 \). The SC for the model excluding the trend break is therefore

\[
SC_0 := T \log(s^2_0) + 3 \log T
\]

the “3” appearing in the penalty function derives from the estimation of the coefficients on the intercept, trend and \( y_{t-1} \) regressors. Similarly the calculation for the model including a trend break at break fraction \( \tau \) involves the residuals \( \tilde{e}_{\tau,t} \) from an OLS regression of \( y_t \) on an intercept, trend \( (t) \), \( y_{t-1} \) and also the break regressors \( 1_{t>[\tau T]} \) and \( 1_{t>[\tau T]}(t - [\tau T]) \), giving residual variance \( s^2_\tau := T^{-1} \sum_{t=2}^{T} \tilde{e}_{\tau,t}^2 \), and SC

\[
SC_\tau := T \log(s^2_\tau) + 6 \log T.
\]

The penalty of 6 presumes that the break fraction \( \tau \) is an estimated parameter, as it will be in our applications. If a fixed \( \tau \) were used then the penalty would become 5. The SC decision rule is to include a trend break at time \( t = [\tau T] \) if \( SC_\tau < SC_0 \), and to exclude the trend break otherwise.

We evaluate below an implementation of this decision rule with \( \tau \) replaced by the OLS estimator \( \hat{\tau}_{OLS} \), taking no account for heteroskedasticity.

The evidence of section 3 suggests that \( \hat{\tau}_{FWLS} \) can be substantially superior to \( \hat{\tau}_{OLS} \) under certain forms of non-stationary volatility, and so we also consider its use in the SC. In addition in this case, since weighting for heteroskedasticity was found to be effective for break point estimation, we also consider its effectiveness for break selection by including weighting in the SC calculation. The weighted residuals \( \tilde{e}^*_{0,t} \) are calculated from a regression of \( y_t/\hat{\sigma}_t \) on \( 1/\hat{\sigma}_t \), \( t/\hat{\sigma}_t \) and \( y_{t-1}/\hat{\sigma}_t \), where \( \hat{\sigma}^2_t \) is defined in (3.8). Similarly the residuals \( \tilde{e}^*_{FWLS,t} \) are calculated from a regression of \( y_t/\hat{\sigma}_t \) on \( 1/\hat{\sigma}_t \), \( t/\hat{\sigma}_t \), \( y_{t-1}/\hat{\sigma}_t \), \( 1_{t>[\hat{\tau}_{FWLS} T]}/\hat{\sigma}_t \) and \( 1_{t>[\hat{\tau}_{FWLS} T]}(t - [\hat{\tau}_{FWLS} T]) \). The weighted SC analogues of (5.2) and (5.2) are then given by

\[
SC^*_0 := T \log(s^*_0) + 3 \log T, \quad SC^*_{FWLS} := T \log(s^*_T) + 6 \log T,
\]

where \( s^*_0 := T^{-1} \sum_{t=2}^{T} \tilde{e}^2_{0,t} \) and \( s^2_{FWLS} := T^{-1} \sum_{t=2}^{T} \tilde{e}^2_{FWLS,t} \).

We will use the unweighted and weighted SC decision rules outlined above to choose whether or not to include a trend break in the de-trending regression used in the first step of computing the unit root statistics outlined above. Our proposed weighted and unweighted SC-based DF test
statistics are then defined as,
\[ t_{SC} := \begin{cases} t_0 & \text{if } SC_0 < SC^*_{\tau_{OLS}} \\ t_{OLS} & \text{if } SC_0 \geq SC^*_{\tau_{OLS}} \end{cases} \]
and
\[ t_{WSC} := \begin{cases} t_0 & \text{if } SC^*_0 < SC^*_{\tau_{FWLS}} \\ t_{FWLS} & \text{if } SC^*_0 \geq SC^*_{\tau_{FWLS}} \end{cases} \]  
(5.8)
respectively, where we recall that \( t_0 \) denotes the DF test that obtains when allowing only for a constant and linear trend, so that \( X_t := (1, t)' \), in the QD detrending step.

In Theorem 5 we now establish the large sample properties of the weighted and unweighted SC-based procedures.

**Theorem 5.** Let \( y_t \) be generated according to (5.1)-(5.2) with \( e_t \) generated according to (2.2), and let the conditions of Assumption A hold. Let \( \delta_T = \delta T^{-d}, 0 \leq d < 1/2 \). Then, under \( H_c \), and in each case as \( T \to \infty \):

(a) For the unweighted SC-based procedure:
(i) if \( \delta \neq 0 \), then \( \Pr( SC^*_{\tau_{OLS}} \leq SC^*_0 ) \to 1 \); (ii) if \( \delta = 0 \), then \( \Pr( SC^*_{\tau_{OLS}} > SC^*_0 ) \to 1 \).

(b) For the weighted SC-based procedure, and provided the additional conditions of Theorem 3 hold:
(i) if \( \delta \neq 0 \), \( \Pr( SC^*_{\tau_{FWLS}} \leq SC^*_0 ) \to 1 \); (ii) if \( \delta = 0 \), \( \Pr( SC^*_{\tau_{FWLS}} > SC^*_0 ) \to 1 \).

**Remark 5.6.** Theorem 5 shows that for both the unweighted and weighted SC procedures the probability of selecting the trend break model converges to one (zero) when a trend break is (is not) present in the DGP, in each case as the sample size diverges. As a result, as the sample size diverges, both \( t_{SC} \) and \( t_{WSC} \) converge (in probability) to \( t_0 \) when no trend break is present, while when a trend break is present \( t_{SC} \) converges to \( t_{OLS} \) and \( t_{WSC} \) converges to \( t_{FWLS} \), both of which coincide with the known \( \tau_0 \) limiting distribution in (5.7). Consequently, the tests from both SC procedures will be asymptotically correctly sized when using the appropriate asymptotic critical value, obtained as outlined in Remark 5.5 using either the simulated critical value approach of Cavaliere and Taylor (2007) or the wild bootstrap approach in Algorithm 1 of Cavaliere et al. (2011), regardless of whether a trend break occurs or not. Moreover, the asymptotic local power of the SC tests will be identical to that of the (size-adjusted) infeasible test which assumes knowledge of whether a break has occurred or not, together with knowledge of the true break fraction, \( \tau_0 \), in the former case.

**Remark 5.7.** Observe that the unweighted SC decision rule can be equivalently expressed in terms of the (pseudo) likelihood ratio test based decision rule to include the trend break if \( T(\log(s^2_0) - \log(s^2_{\tau_{OLS}})) \geq 3 \log T \) and, similarly, for the weighted SC rule if \( T(\log(s^2_0) - \log(s^2_{\tau_{FWLS}})) \geq 3 \log T \). This is therefore seen to be analogous to testing for the presence of a trend break at the random fraction \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \) for the unweighted and weighted SC rules, respectively, and as such is related to a sup-LR type statistic in the spirit of Andrews (1993), but where the decision rule is based not on a fixed critical value but on a Schwarz-type penalty. These are then essentially pre-tests for the presence of a trend break both of which, by design, have size which shrinks to zero as the sample size diverges; the same requirement is needed on the trend break pre-tests used in the univariate testing analogue of the problem considered here in Harris et al. (2009) and Carrion-i-Silvestre et al. (2009). Because the weighted SC approach corresponds to a test based on feasible
WLS estimation, is it anticipated that it will be more efficacious than the unweighted SC approach in selecting between the trend break and no trend break models in the presence of non-stationary volatility.

**Remark 5.8.** Although the test procedures defined in (5.8) are based on the use of the Schwarz information criterion to choose between the trend break and no trend break models, analogous procedures based on any consistent information criterion, such as the HannanQuinn [HQ] information criterion, would have the same asymptotic properties as we report for the SC-based procedures in this paper. It is also worth noting that both the unweighted and weighted SC penalties given above assign a penalty of 1 to the unknown breakpoint parameter. Theoretical results provided in Zhang and Siegmund (2007), Kurozumi and Tuvaandorj (2011) and Kim (2012) suggest that a stricter penalty of 2 might be appropriate for this parameter.

### 5.3 Finite Sample Simulations

In this section we use Monte Carlo simulation methods to investigate whether the superior finite sample behaviour observed for the feasible weighted break fraction estimator, \( \hat{\tau}_{FWLS} \), over the unweighted estimator, \( \hat{\tau}_{OLS} \), seen in the simulation results in section 3, carries over to the unit root test procedures based on \( \hat{\tau}_{FWLS} \) and the feasible weighted model selection criteria outlined in sections 5.1 and 5.2, relative to unit root tests based on the corresponding unweighted quantities.

The results reported in this section are based on the DGP:

\[
y_t = \alpha + \mu t + \delta(t - \lfloor \tau_0 T \rfloor) \cdot 1_{t > \lfloor \tau_0 T \rfloor} + z_t \tag{5.9}
\]

\[
z_t = \phi_T z_{t-1} + \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.} \mathcal{N}(0,1). \tag{5.10}
\]

We set \( \alpha = \mu = 0 \) in our experiments because all of the unit root tests considered are exact invariant to these parameters. For the volatility process, \( \sigma_t \), we considered the same set of models as outlined in section 4. Again we report only a representative selection in the main text. The full set of results are available from the authors on request. In particular, Figures 7 and 8 for \( T = 100 \) and \( T = 200 \), respectively, report results for the homoskedastic case \( \kappa = 0 \), and for a one-time break in volatility occurring at \( \lfloor \lambda_0 T \rfloor \) for \( \lambda_0 \in \{0.3, 0.5, 0.7\} \). In each case results are reported for the no trend break case, \( \delta = 0 \), and where a trend break of magnitude \( \delta = 0.5 \) occurs at \( \lfloor \tau_0 T \rfloor \) for \( \tau_0 \in \{0.3, 0.5, 0.7\} \).

Figures 7 and 8 compare the empirical rejection frequencies, for \( \phi_T := (1 + c/T) \) with \( c \in \{0, -1, -2, ..., -50\} \), of the \( t_{SC} \) and \( t_{WSC} \) SC-based unit root test procedures of (5.8), comparing each with a number of benchmark tests that are also required in the definition of \( t_{SC} \) and \( t_{WSC} \). First \( t_0 \), the DF test which does not allow for a trend break in the de-trending step and where we used \( \bar{c} = -13.5 \) in the QD de-trending procedure. Second, in cases where a trend break occurs in the DGP, \( t_{\tau_0} \) the infeasible DF test based on including a trend break in the de-trending step at the true break fraction \( \tau_0 \). Finally, we also report \( t_{OLS} \) and \( t_{FWLS} \), the DF tests which always including a trend break located at \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \), respectively, in the de-trending step. For all of the tests which include a trend break we set \( \bar{c} \) in the QD de-trending procedure according to the relevant entry from Table 1 from Cavaliere et al. (2011, p.964). In all cases the tests were run at the nominal 5% level using the Gaussian wild bootstrap with 499 bootstrap replications. For the \( t_{SC} \)
and \( t_{WSC} \) procedures the SC rule with the penalties stated in section 5.2 is used. Also reported in the graphs under the labels ‘SC’ and ‘WSC’, respectively, are the empirical frequencies with which the unweighted and weighted SC decision rules select the model which allows for a trend break.

The finite sample properties of \( t_{SC} \) relative to \( t_{WSC} \), and of \( t_{OLS} \) relative to \( t_{FWLS} \) generally mirror the corresponding differences seen between the unweighted and weighted break fraction estimators, \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \), seen in the results for these models for \( \sigma_t \) in section 3. In all of the Figures relating to a trend break, it is seen that the differences between the weighted and unweighted SC decision rules and tests are generally rather smaller, other things equal, for \( T = 200 \) than for \( T = 100 \). This is to be expected, given that both approaches are consistent and \( \delta \) is fixed and non-zero.

Consider first the homoskedastic cases in Figures 7a, 7e, 7i and 7m and Figures 8a, 8e, 8i and 8m. Here we see no discernible differences between the finite sample behaviour of \( t_{SC} \) and \( t_{WSC} \) and between \( t_{OLS} \) and \( t_{FWLS} \), even for \( T = 100 \). Where no trend break is present (Figures 7a and 8a), both the weighted and unweighted SC decision rules select the no trend break model with high probability and, as a result, both \( t_{SC} \) and \( t_{WSC} \) lie very close to the (near-) efficient \( t_0 \) test. Notice that a degree of over-sizing is seen here for both \( t_{OLS} \) and \( t_{FWLS} \) and, as a consequence, also for \( t_{SC} \) and \( t_{WSC} \), although this is reduced for \( T = 200 \) vis-à-vis \( T = 100 \). The power gains from using the SC-based \( t_{SC} \) and \( t_{WSC} \) tests, relative to the \( t_{OLS} \) and \( t_{FWLS} \) tests which always include a trend break (at the fitted break fractions \( \hat{\tau}_{OLS} \) and \( \hat{\tau}_{FWLS} \), respectively), when no break occurs can also clearly be seen both for \( T = 100 \) and \( T = 200 \). Where a trend break is present (Figures 7e, 7i and 7m and Figures 8e, 8i and 8m) the power of the \( t_0 \) test is effectively zero, regardless of the value of \( c \). Consequently, we want the \( t_{SC} \) and \( t_{WSC} \) procedures to select the no break case, and hence \( t_0 \), as infrequently as possible. It is seen from the results that both the weighted and unweighted SC rules perform well in this regard, with \( t_{SC} \) and \( t_{WSC} \) generally lying reasonably close to \( t_{OLS} \) and \( t_{FWLS} \) respectively, the more so the later in the sample the trend break occurs, which in turn lie close to the infeasible efficient benchmark \( t_{\tau_0} \) test. In these cases where a trend break is present, an interesting feature seen for both SC decision rules is that their efficacy to select the trend break model improves the further the AR parameter \( \phi_T \) lies into the stationarity region (i.e. the bigger is \( c \)). This phenomenon is clearly beneficial to the finite sample performance of the \( t_{SC} \) and \( t_{WSC} \) procedures, and is to be expected given that it is well known that a trend break is more easily detected in stationary noise than it is in noise which contains a unit root; see, for example, Harvey et al. (2009).

Consider next the cases where \( \sigma_t \) is heteroskedastic. Where no trend break occurs (\( \delta = 0 \)), it is seen in Figures 7b, 7c and 7d and Figures 8b, 8c and 8d that although the weighted SC decision rule is marginally more efficacious in selecting the no trend break model than the unweighted SC rule, and increasingly so as \( \lambda_0 \) increases, in selecting the no break model, there is almost nothing to choose between the resulting \( t_{SC} \) and \( t_{WSC} \) procedures, each of which again performs well lying very close to the \( t_0 \) test, as in the homoskedastic case. However, where a trend break occurs (\( \delta \neq 0 \)) this picture changes considerably. The most dramatic differences between the weighted and unweighted tests are seen for precisely those cases where \( \hat{\tau}_{FWLS} \) was observed in the simulations in section 3 to be significantly more efficacious than \( \hat{\tau}_{OLS} \). These are the cases where the trend
break occurs in a low volatility regime and correspond with Figures 7g, 7h and 7l and Figures 8g, 8h and 8l. In these cases the superior finite sample performance of $t_{FWLS}$ over $t_{OLS}$ is clearly seen with the former lying very close to the infeasible efficient benchmark $t_{\tau_0}$ test, while the latter lies some considerable distance from this benchmark. In these examples the weighted SC decision rule is also considerably more efficacious than the unweighted SC decision rule in (correctly) selecting the trend break model for the de-trending step. This is crucial to explaining the differences in the behaviour of $t_{SC}$ relative to $t_{WSC}$. Too often in these cases, the unweighted SC rule is wrongly selecting the no trend break model and hence selecting the inappropriate no break $t_{\tau_0}$ test and as such is heavily compromised. The superior performance of both the weighted SC decision rule and the DF test based on the weighted break fraction estimator translate into very significant power gains for $t_{WSC}$ over $t_{SC}$ in these cases, especially so for $T = 100$. As an example, in Figure 7g the empirical power of $t_{WSC}$ for $T = 100$ is around 90% for $c = -40$ while that of $t_{SC}$ is only about 35%.

Interestingly, the weighted SC decision rule often outperforms the unweighted SC rule, and $t_{WSC}$ accordingly outperforms $t_{SC}$, even in cases where $\hat{\tau}_{FWLS}$ was seen to be no more efficacious than $\hat{\tau}_{FWLS}$ in the simulations in section 3. Examples of this can be seen in Figures 7f and 7k and 8f and 8k where the location of the trend and volatility breaks coincides. In these examples $t_{FWLS}$ also performs better than $t_{OLS}$. Finally, in those cases where $\hat{\tau}_{FWLS}$ performed least well relative to $\hat{\tau}_{OLS}$, which are the cases where the trend break lies in a high volatility regime (see Figures 7j, 7n and 7o and Figures 8j, 8n and 8o) the unweighted SC decision rule is seen to perform slightly better than the weighted SC rule. In these examples $t_{OLS}$ correspondingly also performs slightly better than $t_{FWLS}$ as does $t_{SC}$ over $t_{WSC}$.

In unreported simulations we also explored corresponding procedures based on the HQ information criterion, and procedures using the stricter double penalty on the estimated break fraction discussed in Remark 5.8. These govern the strength of the penalty (the SC penalty is stricter than the HQ penalty) imposed on including the trend break. Clearly the weaker the penalty, the higher the frequency with which the trend break will be retained in the de-trending step, other things being equal. As we have seen, this break retention frequency affects the finite sample size and power properties of the resulting unit root tests. We found that the stricter the penalty used the better the finite sample size control of the information criteria based test procedures (so that, for example, using the SC with a double penalty on the trend break reduced the over-sizing in $t_{SC}$ over $t_{WSC}$ relative to that seen in Figures 7 and 8), but came at the expense of lower finite sample power where a trend break is present. However, the qualitative conclusions drawn above regarding the relative finite sample performance of the unweighted and weighted information criteria and associated unit root tests were unaltered between these different possible penalties.

6 Conclusions

We have investigated the properties of RSS-based estimators, including OLS and feasible WLS estimators, the latter formed using a non-parametric kernel-based estimate of the volatility process, for the location of a level break in series driven by shocks displaying non-stationary volatility.
Consistency rates were derived for these estimators against breaks of fixed magnitude and shown to coincide with those obtained under homoskedasticity. Distribution theory for these estimators was also derived for cases where the break magnitude was either local-to-zero or exactly zero. Under Pitman drift these limiting distributions were shown to depend on nuisance parameters deriving from the non-stationary volatility as well as on the location and magnitude of the level break and the bounds of the search set. Monte Carlo simulation evidence was provided which demonstrated that these Pitman limits closely predict the finite sample behaviour of both the OLS and feasible WLS estimators, and demonstrated the potential for the feasible WLS estimator to deliver significant improvements over the OLS estimator in certain heteroskedastic environments. We then showed how the feasible WLS level break fraction estimator can be used in the context of the problem of unit root testing when trend and/or volatility breaks may be present in the data by applying it to the first differences of the data. This was shown to have the potential to deliver significant improvements in the finite sample properties of the resulting unit root tests relative to the case where the standard OLS break fraction estimator is used. We also discussed the use of feasible weighted information criteria, based on the same estimate of the volatility process, to select between the trend break and no trend break models in cases where it is unknown whether a trend break is present. Again these were shown to have the potential to deliver unit root tests with considerably improved finite sample behaviour under heteroskedasticity relative to those based on standard information criteria.

Although our focus in this paper has been on a single possible level break, the ideas we have presented naturally extend to the case of multiple level breaks in the deterministic trend function and to structural breaks in the parameters of more general time series regression settings. Moreover, the procedures we develop here should extend to the multivariate case and so would be anticipated to improve inference on determining the co-integration rank in the case of multiple time series potentially subject to breaks in both trend and volatility. These issues are currently being investigated by the authors.

References


A Appendix

This appendix contains proofs for the theorems stated in the main text. Section A.1 contains some preliminary results. Proofs of the results in Theorems 1-5 are then provided in Section A.2.

A.1 Preliminary Results

The following Lemmas are useful for the proofs of Theorems 1 and 2. The first provides a general representation of the weighted least squares criterion, and the second the heteroskedastic FCLT results required for the theorems.

The estimator of \( \tau_0 \) defined in (3.1) minimises the sum of squared residuals from regressions (3.2) estimated over the range of \( \tau \), which can equivalently be expressed as regressions of \( y_t^* := y_t x_t \) on \( x_t \) for \( \tau \in [\tau_L, \tau_U] \). For any \( \tau \) such a regression can be represented

\[
y_t^* = \hat{\mu}_1,\tau(x_t \text{ for } t \leq \lfloor \tau T \rfloor) + \hat{\mu}_2,\tau(x_t \text{ for } t > \lfloor \tau T \rfloor) + \tilde{e}_{\tau,t},
\]

where

\[
\hat{\mu}_1,\tau := \frac{\sum_{t=1}^{\lfloor \tau T \rfloor} y_t x_t}{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2}, \quad \hat{\mu}_2,\tau := \frac{\sum_{t=\lfloor \tau T \rfloor + 1}^{T} y_t x_t}{\sum_{t=\lfloor \tau T \rfloor + 1}^{T} x_t^2}
\]

are the OLS coefficient estimates.

**Lemma A.1.** For any weights \( x_t \), \( \hat{\tau} \) defined in (3.1) can be represented as

\[
\hat{\tau} = \arg \max_{\tau \in [\tau_L, \tau_U]} Q_T(\tau),
\]

where

\[
Q_T(\tau) := \frac{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2 \sum_{t=\lfloor \tau T \rfloor + 1}^{T} x_t^2}{\sum_{t=1}^{T} x_t^2} (\hat{\mu}_2,\tau - \hat{\mu}_1,\tau)^2,
\]

and where

\[
\hat{\mu}_2,\tau - \hat{\mu}_1,\tau = \frac{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2 \sum_{t=\lfloor \tau T \rfloor + 1}^{T} x_t^2}{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2} \left( \sum_{t=1}^{\lfloor \tau T \rfloor} e_t x_t^2 - \sum_{t=\lfloor \tau T \rfloor + 1}^{T} e_t x_t^2 \right) - \frac{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2}{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2} \left( \sum_{t=1}^{\lfloor \tau T \rfloor} e_t x_t^2 \right). \quad (A.1)
\]

**Proof of Lemma A.1**

For any \( \tau \), standard least squares algebra gives the sum of squared residuals

\[
\sum_{t=1}^{T} \tilde{e}_{\tau,t}^2 = \sum_{t=1}^{T} y_t^2 - \left( \hat{\mu}_1,\tau \sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2 + \hat{\mu}_2,\tau \sum_{t=\lfloor \tau T \rfloor + 1}^{T} x_t^2 \right). \quad (A.2)
\]

Similar algebra for a regression of \( y_t^* \) on \( x_t \) alone gives the sum of squared residuals

\[
\sum_{t=1}^{T} \tilde{e}_t^2 = \sum_{t=1}^{T} y_t^2 - \hat{\mu}^2 \sum_{t=1}^{T} x_t^2 \quad (A.3)
\]
where $\hat{\mu} := T^{-1} \sum_{t=1}^T x_t$. Subtracting (A.3) from (A.2) gives

$$\sum_{t=1}^T \hat{e}_{t,\tau}^2 = \sum_{t=1}^T \hat{e}_t^2 + \frac{(\sum_{t=1}^T x_t y_t^*)^2}{\sum_{t=1}^T x_t^2} - \frac{(\sum_{t=1}^{[\tau T]} x_t y_t^*)^2}{\sum_{t=1}^{[\tau T]} x_t^2} - \frac{(\sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t y_t^*)^2}{\sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t^2}.$$  

Substituting $\sum_{t=1}^T x_t y_t^* = \sum_{t=1}^{[\tau T]} x_t y_t^* + \sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t y_t^*$, expanding and rearranging gives

$$\sum_{t=1}^T \hat{e}_{t,\tau}^2 = \sum_{t=1}^T \hat{e}_t^2 - \frac{\sum_{t=1}^{[\tau T]} x_t^2 \sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t^2}{\sum_{t=1}^{[\tau T]} x_t^2} (\hat{\mu}_{2,\tau} - \hat{\mu}_{1,\tau})^2$$

so that arg min $\sum_{t=1}^T \hat{e}_{t,\tau}^2$ is identical to arg max $Q_T(\tau)$.

Next consider $\hat{\mu}_{2,\tau} - \hat{\mu}_{1,\tau}$. Using

$$y_t^* = \mu x_t + \delta_T (x_t \cdot 1_{t \geq |\gamma_0 T|}) + x_t e_t$$

in the expressions for $\hat{\mu}_{1,\tau}$ and $\hat{\mu}_{2,\tau}$ gives

$$\hat{\mu}_{1,\tau} = \mu + \delta_T \frac{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2}{\sum_{t=1}^{[\tau T]} x_t^2} + \frac{\sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t}{\sum_{t=1}^{[\tau T]} x_t^2}$$

( thanked second term being zero for $|\gamma T| \leq |\gamma_0 T|$) and

$$\hat{\mu}_{2,\tau} = \mu + \delta_T \frac{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2}{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2} + \frac{\sum_{t=\max(\tau T,|\tau T|\leq|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t}{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2}$$

To calculate $\hat{\mu}_{2,\tau} - \hat{\mu}_{1,\tau}$, note that

$$\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t - \frac{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t}{\sum_{t=1}^{[\tau T]} x_t^2} = \begin{cases} \sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2, & \text{if } |\tau T| \leq |\gamma_0 T| \\ \sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2, & \text{if } |\tau T| > |\gamma_0 T| \end{cases}$$

and that straightforward rearrangements give

$$\frac{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t}{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2} = \frac{\sum_{t=1}^{[\tau T]} x_t^2 \hat{e}_t}{\sum_{t=1}^{[\tau T]} x_t^2} - \frac{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2 \hat{e}_t}{\sum_{t=\min(\tau T,|\gamma_0 T|)+1}^T x_t^2}$$

and together these imply the result in (A.1).
Lemma A.2. Let \( e_t \) be generated by (2.2) under the conditions of Assumption A, and let \( x_t = x(t/T) \) be non-stochastic weights such that the function \( x(\cdot) \) satisfies the same conditions as \( \sigma(\cdot) \) in Assumption A2. Then,

\[
\omega_{xe}^{-1} T^{-1/2} \sum_{t=1}^{\lfloor T \rfloor} x_t^2 e_t d \to B_\eta(\cdot) \tag{A.4}
\]

where \( B_\eta(s) := B(\eta(s)) \), with \( B(\cdot) \) a standard Brownian motion, and where \( \eta(s) := \omega_{xe}^{-2} \int_0^s x(r)^4 \sigma(r)^2 dr \) with \( \omega_{xe}^2 := \int_0^1 x(r)^4 \sigma(r)^2 dr \). Moreover,

\[
\sup_{\tau \in [\tau_L, \tau_U]} T^{-1/2} \sum_{t=1}^{\lfloor \tau T \rfloor} e_t x_t^2 = O_p(1) \tag{A.5}
\]

Proof of Lemma A.2

Equation (A.4) follows directly from Lemma 4 of Cavaliere and Taylor (2007). The maximal inequality in equation (A.5) follows by noting that

\[
\sup_{\tau \in [\tau_L, \tau_U]} T^{-1/2} \sum_{t=1}^{\lfloor \tau T \rfloor} e_t x_t^2 \leq \max_{b \in [1, T]} T^{-1/2} \sum_{t=1}^{b} e_t x_t^2,
\]

and then, since \( \sum_{t=1}^{b} e_t x_t^2 \) is a martingale indexed by \( b \), Doob’s inequality (result 15.15 of Davidson, 1994) implies that

\[
E \left( \max_{b \in [1, T]} \left( T^{-1/2} \left| \sum_{t=1}^{b} e_t x_t^2 \right| \right)^2 \right) \leq 4 T^{-1} E \left( \sum_{t=1}^{T} e_t x_t^2 \right)^2 = 4 T^{-1} \sum_{t=1}^{T} \sigma_t^2 x_t^4 = O(1).
\]

\( \blacksquare \)

A.2 Main Proofs

Proof of Theorem 1

The argument for consistency follows Bai (1994, 1997) and especially Theorem 3.1 of Nunes et al. (1995). The assumptions on the weights \( x_t \) imply that

\[
\frac{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2}{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2} \to \frac{\int_0^\tau x(s)^2 ds}{\int_0^1 x(s)^2 ds} =: \chi(\tau),
\]

with \( \chi : [0, 1] \to [0, 1] \) being an increasing function (exactly analogous to the variance-profile \( \eta \) constructed from the variance sequence \( \sigma_t^2 \)). Thus, for example,

\[
\frac{\sum_{t=1}^{\lfloor \tau T \rfloor} x_t^2 \sum_{t=\lfloor \tau T \rfloor + 1}^{T} x_t^2}{\sum_{t=1}^{T} x_t^2} \to \chi(\tau)(1 - \chi(\tau)).
\]

and

\[
\frac{\sum_{t=1}^{\lfloor \tau_0 T \rfloor} x_t^2}{\sum_{t=1}^{\lfloor \tau_0 T \rfloor} x_t^2} \wedge \frac{\sum_{t=\lfloor \tau_0 T \rfloor + 1}^{T} x_t^2}{\sum_{t=\lfloor \tau_0 T \rfloor + 1}^{T} x_t^2} \to \frac{\chi(\tau_0)}{\chi(\tau)} \wedge \frac{1 - \chi(\tau_0)}{1 - \chi(\tau)}.
\]

A.3
Applying \((A.1)\), it follows that for \(\delta \neq 0\) and \(0 \leq d < 1/2\)

\[ T^d (\hat{\mu}_{2,\tau} - \hat{\mu}_{1,\tau}) = \delta \left( \frac{\sum_{t=1}^{[\tau T]} x_t^2}{\sum_{t=1}^T x_t^2} \wedge \frac{\sum_{t=[\tau T]+1} T x_t^2}{\sum_{t=[\tau T]+1}^T x_t^2} \right) \]

\[ - T^{-d/2} \frac{T \sum_{t=1} T x_t^2}{\sum_{t=1}^T x_t^2 \sum_{t=[\tau T]+1}^T x_t^2} \left( T^{-\frac{1}{2}} \sum_{t=1}^{[\tau T]} \epsilon_t x_t^2 - \frac{\sum_{t=1} T x_t^2}{\sum_{t=1}^T x_t^2} T^{-\frac{1}{2}} \sum_{t=1}^T \epsilon_t x_t^2 \right) \]

\[ B_\delta \left( \frac{\chi(\tau_0)}{\chi(\tau)} \wedge \frac{1 - \chi(\tau_0)}{1 - \chi(\tau)} \right), \]

uniformly in \(\tau\) on \([\tau_L, \tau_U]\), and, hence,

\[ \frac{T^{2d}}{\sum_{t=1}^T x_t^2} Q_T(\tau) \overset{P}{\to} Q(\tau) := \delta^2 \chi(\tau)(1 - \chi(\tau)) \left( \frac{\chi(\tau_0)}{\chi(\tau)} \wedge \frac{1 - \chi(\tau_0)}{1 - \chi(\tau)} \right)^2. \quad (A.6) \]

For \(\tau \in [\tau_L, \tau_0]\)

\[ Q(\tau) = \delta^2 (1 - \chi(\tau_0))^2 \frac{\chi(\tau)}{1 - \chi(\tau)} \]

is an increasing function with maximum value at \(\tau = \tau_0\) of \(\frac{Q(\tau_0) = \chi(\tau_0)(1 - \chi(\tau_0))}{\chi(\tau)}\). For \(\tau \in [\tau_0, \tau_U]\)

\[ Q(\tau) = \delta^2 \chi(\tau_0)^2 \frac{1 - \chi(\tau)}{\chi(\tau)} \]

is a decreasing function with maximum value of \(Q(\tau_0)\) at \(\tau = \tau_0\). So \(Q(\tau)\) has a unique maximum at \(\tau = \tau_0\) and, since the convergence in \((A.6)\) is uniform in \(\tau\), it therefore follows that

\[ \hat{\tau} = \arg \max_{\tau \in [\tau_L, \tau_U]} \frac{T^{2d}}{\sum_{t=1}^T x_t^2} Q_T(\tau) \overset{P}{\to} \arg \max_{\tau \in [\tau_L, \tau_U]} Q(\tau) = \tau_0. \]

The asymptotic distribution of \(\hat{\tau}\) follows by the same arguments as Bai (1997), except that it is necessary to check for any effect that unconditional heteroskedasticity satisfying Assumption \(A_2\) has on the form of the distribution. In particular, using Bai’s notation in the proof of his Proposition 3, we will verify that the following two convergence results from Bai (1997) still hold under our assumptions:

(a) \(v_T \sum_{t=k_0+1}^{k_0+d} z_t \epsilon_t \overset{d}{\to} B_1(s)\) and

(b) \(v_T^2 \sum_{t=k_0+1}^{k_0+d} z_t' \epsilon_t \overset{d}{\to} sQ_1\),

where \(B_1(s)\) is Brownian motion on \([0, \infty)\) with variance \(s\Omega_1\), and \(k\) is defined as a function of \(s\) to be \(k := k_0 + [s\delta_T^{-2}]\). Note that both limits, being simple functions of \(s\), reflect the unconditional homoskedasticity assumed in Bai (1997). Under the form of heteroskedasticity assumed in this paper, it might have been expected that the corresponding \(B_1\) would be a heteroskedastic Brownian Motion (as in, for example, Cavaliere and Taylor, 2007), but we will show that this is not the case.

To translate the notation of Bai (1997) notation into our own, his \(z_t\) (the regressor(s) whose coefficient(s) break) is the same as our \(x_t\), \(\epsilon_t = \epsilon_t\), \(v_t = \delta_T\) and \(k = [\tau_0 T] + [s\delta_T^{-2}]\). This specification of \(k\) is relevant for the scaled criterion function for values of \(\tau\) above the true value \(\tau_0\). In our equivalent to (a), as \(T \to \infty\) the term

\[ \delta_T \sum_{t=[\tau_0 T]+1}^{[\tau_0 T]+[s\delta_T^{-2}]} x_t^2 \epsilon_t \]

\(A.4\).
can straightforwardly be shown to satisfy the heteroskedastic FCLT in Lemma 4 of Cavaliere and Taylor (2007), indexed by $s$, but as in their proof we need to explicitly derive the form of the marginal distribution for given $s$. In particular the variance is

$$\text{var} \left( \delta_T \sum_{t=\lceil T \delta_T \rceil + 1}^{\lfloor T \delta_T \rfloor} x_t^2 e_t \right)$$

$$= \delta_T^2 \sum_{i=1}^{\lfloor s \delta_T^2 \rfloor} x_{\lfloor T \delta_T \rfloor + i}^4 + \sigma_{\lfloor T \delta_T \rfloor + i}^2$$

$$= \delta_T^2 \sum_{i=1}^{\lfloor s \delta_T^2 \rfloor} \int_{i/(s \delta_T^2)}^{(i+1)/(s \delta_T^2)} x \left( \frac{\tau_0 T + |r| s \delta_T^2}{T} \right)^4 \sigma \left( \frac{\tau_0 T + |r| s \delta_T^2}{T} \right)^2 dr$$

$$= \int_{1/(s \delta_T^2)}^{1+1/(s \delta_T^2)} x \left( \frac{\tau_0 T + |r| s \delta_T^2}{T} \right)^4 \sigma \left( \frac{\tau_0 T + |r| s \delta_T^2}{T} \right)^2 dr$$

$$\rightarrow s \int_0^1 \bar{x}(\tau_0)^4 \bar{\sigma}(\tau_0)^2 dr = s \bar{x}(\tau_0)^4 \bar{\sigma}(\tau_0)^2,$$

with the convergence in the second last step following because $\delta_T^{-2}/T \rightarrow 0$ for $0 < d < 1/2$ and from the assumed càdlàg property of both the $x(\cdot)$ and $\sigma(\cdot)$ functions. The limits from above $\bar{x}(\tau_0) := \lim_{\tau \uparrow \tau_0} x(\tau)$ and $\bar{\sigma}(\tau_0) := \lim_{\tau \uparrow \tau_0} \sigma(\tau)$ may differ from $x(\tau_0)$ and $\sigma(\tau_0)$ if either function has a jump at $\tau = \tau_0$. Similarly for our equivalent to (b) above,

$$\delta_T^2 \sum_{t=\lceil T \delta_T \rceil + 1}^{\lfloor T \delta_T \rfloor} x_t^4 \rightarrow s \bar{x}(\tau_0)^4.$$

Similarly for values of $\tau$ below $\tau_0$ we find

$$\text{var} \left( \delta_T \sum_{t=\lceil T \delta_T \rceil - \lfloor s \delta_T^2 \rfloor}^{\lfloor T \delta_T \rfloor - 1} x_t^2 e_t \right) \rightarrow s \bar{x}(\tau_0)^4 \bar{\sigma}(\tau_0)^2$$

and

$$\delta_T^2 \sum_{t=\lceil T \delta_T \rceil - \lfloor s \delta_T^2 \rfloor}^{\lfloor T \delta_T \rfloor - 1} x_t^4 \rightarrow s \bar{x}(\tau_0)^4.$$

where $\bar{x}(\tau_0) := \lim_{\tau \uparrow \tau_0} x(\tau)$ and $\bar{\sigma}(\tau_0) := \lim_{\tau \uparrow \tau_0} \sigma(\tau)$.

The rate of $\delta_T^{-2}$ when $0 < d < 1/2$, being less than $T$, is such that the effect of the weighting function $x(\tau)$ and the volatility process $\sigma(\tau)$ drop out of these limits for all values of $\tau$ other than $\tau_0$. For a “large” break magnitude (one for which the break fraction can be consistently estimated) the asymptotic distribution of the break fraction is determined by the behaviour of the criterion function in a diminishing interval around $\tau_0$, such that in the limit it is only the variance properties at $\tau_0$ that enter the FCLT results.

With these expressions in hand, the proof follows each step of Bai’s (1997) Proposition 3 to reach the analogous result in this theorem. The equivalences in the notation are Bai’s $Q_1$ and $Q_2$ with $\bar{x}(\tau_0)^2$ and $\bar{x}(\tau_0)^2$ here, and Bai’s $\Omega_1$ and $\Omega_2$ with $\lim_{\tau \uparrow \tau_0} \text{var}(x_t e_t) = \bar{x}(\tau_0)^2 \bar{\sigma}(\tau_0)^2$ and $\lim_{\tau \uparrow \tau_0} \text{var}(x_t e_t) = \bar{x}(\tau_0)^2 \bar{\sigma}(\tau_0)^2$, respectively. ■
Proof of Theorem 2

For $\delta \neq 0$ and $d = 1/2$, the representation in (A.1) can be written

$$T^{1/2}(\tilde{\mu}_{2,\tau} - \tilde{\mu}_{1,\tau})$$

$$= T^{1/2}d \left( \frac{\sum_{t=1}^{[\gamma T]} x_t^2 \wedge \sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{T} x_t^2} \right) - \frac{\omega{x}{e}}{T-1} \frac{\sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{[\gamma T]} x_t^2} \left( \omega^{-1} T^{-1/2} \sum_{t=1}^{[\gamma T]} x_t^2 - \sum_{t=1}^{T} x_t^2 \right)$$

$$d \delta \left( \frac{\delta}{\omega} \frac{\chi(\tau) - \chi(\tau)}{\chi(\tau)(1 - \chi(\tau))} \right)^{1/2} \left( \frac{\chi_1(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)}}{\chi(\tau)(1 - \chi(\tau))} - \frac{B_\eta(\tau - \chi(\tau) B_\eta(1)}{\chi(\tau)(1 - \chi(\tau))} \right)$$

where

$$\omega^2 = \frac{\omega^2 x e}{(\int_0^1 x(s)^2 ds)^2} = \int_0^1 x(s)^2 ds \int_0^1 x(s)^2 ds.$$ 

Hence,

$$\frac{T}{\omega^2 \sum_{t=1}^{T} x_t^2} Q_T(\tau) = \sum_{t=1}^{[\gamma T]} x_t^2 \sum_{t=1}^{T} x_t^2 \sum_{t=1}^{T} x_t^2 T(\tilde{\mu}_{2,\tau} - \tilde{\mu}_{1,\tau})^2$$

$$d \left( \frac{\delta}{\omega} \frac{\chi(\tau) - \chi(\tau)}{\chi(\tau)(1 - \chi(\tau))} \right)^{1/2} \left( \frac{\chi_1(\tau; \tau_0) \wedge \frac{1}{\chi_1(\tau; \tau_0)}}{\chi(\tau)(1 - \chi(\tau))} - \frac{B_\eta(\tau - \chi(\tau) B_\eta(1)}{\chi(\tau)(1 - \chi(\tau))} \right)^2 = Q(\tau; x(\cdot), \sigma(\cdot), \delta, \frac{1}{2}).$$

If $d > 1/2$ then the representation in (A.1) is

$$T^{1/2}(\tilde{\mu}_{2,\tau} - \tilde{\mu}_{1,\tau})$$

$$= T^{1/2}d \left( \frac{\sum_{t=1}^{[\gamma T]} x_t^2 \wedge \sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{T} x_t^2} \right) - \frac{\omega{x}{e}}{T-1} \frac{\sum_{t=1}^{T} x_t^2}{\sum_{t=1}^{[\gamma T]} x_t^2} \left( \omega^{-1} T^{-1/2} \sum_{t=1}^{[\gamma T]} x_t^2 - \sum_{t=1}^{T} x_t^2 \right)$$

$$d \omega \frac{B_\eta(\tau) - \chi(\tau) B_\eta(1)}{\chi(\tau)(1 - \chi(\tau))}$$

so that

$$\hat{\tau} = \arg \max_{\tau \in [\tau_1, \tau_0]} \frac{T}{\omega^2 \sum_{t=1}^{T} x_t^2} Q_T(\tau) \quad \text{and} \quad \hat{\tau} = \arg \max_{\tau \in [\tau_1, \tau_0]} \frac{B_\eta(\tau) - \chi(\tau) B_\eta(1)}{\chi(\tau)(1 - \chi(\tau))}$$

which is the arg max of $Q(\tau; x(\cdot), \sigma(\cdot), \delta, d)$ with $d > 1/2$ as required. Clearly this latter results also holds when $\delta = 0$ (in which case the value of $d$ is irrelevant).

For $d < 1/2$ (and $\delta \neq 0$), the result follows immediately as an implication of equation (A.6) established in the proof of Theorem 1.
Proof of Theorem 3
For clarity we write $\hat{d}_t^2 := \sum_{i=1}^T w_{it} \hat{e}_{i,t}^2$, where $w_{it} := K_h \left( \frac{t-i}{T} \right) / \sum_{i=1}^T K_h \left( \frac{t-i}{T} \right)$, instead of $\hat{d}_t^2$ and $Q_T (\tau; x_t)$ instead of $Q_T (\tau)$ defined in Lemma A.1 in the rest of the proof of this Theorem. We prove the result separately for Case 1: $d = 1/2$ and Case 2: $0 \leq d < 1/2$.

Proof of Theorem 3, Case 1: $d = 1/2$
By Theorem 2,
$$\hat{\tau}_{WLS} = \arg \max_{\tau \in [\tau, \tau_\nu]} \frac{T}{\omega^2 \sum_{t=1}^T \sigma_t^2} Q_T (\tau; \sigma_t^{-1}) \xrightarrow{d} \arg \max_{\tau \in [\tau, \tau_\nu]} Q (\tau; 1/\sigma (\cdot), \sigma (\cdot), \delta, d)$$
and
$$\hat{\tau}_{FWLS} = \arg \max_{\tau \in [\tau, \tau_\nu]} \frac{T}{\omega^2 \sum_{t=1}^T \sigma_t^2} Q_T (\tau; \hat{\sigma}_t^{-1}).$$
Let
$$R_T (\tau; \hat{\sigma}_t, \sigma_t) := \frac{T}{\omega^2 \sum_{t=1}^T \sigma_t^2} \left( Q_T (\tau; \hat{\sigma}_t^{-1}) - Q_T (\tau; \sigma_t^{-1}) \right).$$
Theorem 3 can be established by showing that
$$\hat{\tau}_{FWLS} = \arg \max_{\tau \in [\tau, \tau_\nu]} \frac{T}{\omega^2 \sum_{t=1}^T \sigma_t^2} Q_T (\tau; \hat{\sigma}_t^{-1}) + R_T (\tau; \hat{\sigma}_t, \sigma_t) \xrightarrow{d} \arg \max_{\tau \in [\tau, \tau_\nu]} Q (\tau; 1/\sigma (\cdot), \sigma (\cdot), \delta, d),$$
(by appealing to the continuous mapping theorem for argmax functionals; see Kim and Pollard, 1990, Theorem 2.7), where
$$\sup_{\tau \in [\tau, \tau_\nu]} |R_T (\tau; \hat{\sigma}_t, \sigma_t)| \xrightarrow{p} 0.$$  (A.7)
Because $T^{-1} \sum_{t=1}^T \sigma_t^{-2} \to \int_0^1 \sigma (s)^{-2} \, ds$, to show (A.7) we need to show that the following two results hold:
$$\sup_{\tau \in [\tau, \nu, 1]} \left| T^{-1/2} \sum_{t=1}^{\lfloor \tau T \rfloor} e_t \hat{\sigma}_{t, \tau}^2 - e_t \sigma_t^2 \right| \xrightarrow{p} 0.$$  (A.8)
and
$$\sup_{\tau \in [\tau, \nu, 1]} \left| T^{-1} \sum_{t=1}^{\lfloor \tau T \rfloor} \hat{\sigma}_{t, \tau}^2 - \sigma_t^2 \right| \xrightarrow{p} 0.$$  (A.9)
Similar arguments apply for the $0 \leq d < 1/2$ case (see the proof of Theorem 1), which also requires the above two results as will be done below in Case 2.

To prove (A.8), observe first that
$$\sup_{\tau \in [\tau, \nu, 1]} \left| T^{-1/2} \sum_{t=1}^{\lfloor \tau T \rfloor} e_t \hat{\sigma}_{t, \tau}^2 - e_t \sigma_t^2 \right| \leq T^{-1/2} \sum_{t=1}^T \left| e_t \left( \hat{\sigma}_{t, \tau}^2 - \sigma_t^2 \right) \right|.$$  (A.10)
Define $\hat{\sigma}_t^2 := \sum_{i=1}^T w_{it} e_{i,t}^2$. Then, following Robinson (1987), to prove that the right hand side of (A.10) is $o_p(1)$, it is sufficient, by virtue of results in the proof of Theorem 2 of Xu and Phillips.
(2008) (specifically their equation (20) and part (a) on page 276) and in Lemma A of Xu and Phillips (2008), for us to show that
\[
\sum_{t=1}^{T} |\hat{\sigma}_{\tau,t}^2 - \tilde{\sigma}_{i}^2|^2 \overset{p}{\to} 0. \tag{A.11}
\]
Similarly, to prove (A.9), it is sufficient, by virtue of the results in part (d) of the proof of Theorem 2 of Xu and Phillips (2008, p.277), for us to show that
\[
\max_{1 \leq t \leq T} |\hat{\sigma}_{\tau,t}^2 - \tilde{\sigma}_{i}^2| \overset{p}{\to} 0. \tag{A.12}
\]
We will now establish the validity of the results in (A.11) and (A.12). To show (A.11), write
\[
\sum_{t=1}^{T} |\hat{\sigma}_{\tau,t}^2 - \tilde{\sigma}_{i}^2|^2 \leq \sup_{\tau \in [\tau, \tau_0]} \sum_{t=1}^{T} \left( e_{\tau,t}^2 - e_{t}^2 \right) + \sum_{i=1}^{[\tau_0]} w_i \left( e_{\tau,t}^2 - e_{t}^2 \right) + \sum_{i=[\tau]+1}^{T} w_i \left( e_{\tau,t}^2 - e_{t}^2 \right) =: \sup_{\tau \in [\tau, \tau_0]} \sum_{t=1}^{T} (A_{\tau,t} + B_{\tau,t} + C_{\tau,t})^2. \tag{A.13}
\]
Throughout this proof, in the interest of brevity, we will only discuss the case where \(|\tau T| > |\tau_0 T|\), as the results for the case where \(|\tau T| \leq |\tau_0 T|\) follow along exactly the same lines.

For the \(A_{\tau,t}\) term in (A.13), in which \(i = 1, ..., [\tau_0 T]\),
\[
\hat{e}_{\tau,i} = y_i - \hat{\mu}_{1,\tau} = \mu + e_i - [\tau T]^{-1} \sum_{t=1}^{[\tau T]} (\mu + \delta T \cdot 1_{t>[\tau_0 T]} + e_t) = e_i + \alpha_{T,\tau} + \hat{e}_{\tau} \tag{A.14}
\]
where \(\alpha_{T,\tau} := \delta T^{-d} T^{-1} (\tau - \tau_0)\) and \(\hat{e}_{\tau} := [\tau T]^{-1} \sum_{t=1}^{[\tau T]} e_t\).

For the \(B_{\tau,t}\) term in (A.13), with \(i = [\tau_0 T] + 1, ..., [\tau T]\),
\[
\hat{e}_{\tau,i} = y_i - \hat{\mu}_{1,\tau} = \mu + \delta T + e_i - [\tau T]^{-1} \sum_{t=1}^{[\tau T]} (\mu + \delta T \cdot 1_{t>[\tau_0 T]} + e_t) = e_i + \beta_{T,\tau} + \hat{e}_{\tau} \tag{A.15}
\]
where \(\beta_{T,\tau} := \delta T^{-d} T^{-1} \tau_0\).

For the \(C_{\tau,t}\) term in (A.13), in which \(i = [\tau T] + 1, ..., T\),
\[
\hat{e}_{\tau,i} = y_i - \hat{\mu}_{2,\tau} = \mu + \delta T + e_i - \frac{1}{T-[\tau T]} \sum_{t=1}^{[\tau T]} (\mu + \delta T \cdot 1_{t>[\tau_0 T]} + e_t) = e_i - \hat{e}_{(T-[\tau T])} \tag{A.16}
\]
where \(\hat{e}_{(T-[\tau T])} := (T-[\tau T])^{-1} \sum_{t=[\tau T]+1}^{T} e_t\).

In what follows, we will only deal with the \(A_{\tau,t}\) and \(B_{\tau,t}\) terms. The expressions for \(\hat{e}_{\tau,i}\) above imply that the arguments when dealing with the \(C_{\tau,t}\) term follow similarly because it does not feature \(\alpha_{T,\tau}\) and \(\beta_{T,\tau}\), and because \(\hat{e}_{(T-[\tau T])} = \hat{e}_{1} - \hat{e}_{\tau} \).

A.8
For the second term in (A.17), we have

\[
\sup_{\tau \in [\tau_0, T]} \left| T^{1/2} \tilde{e}_\tau \right| = \sup_{\tau \in [\tau_0, T]} \left| \tau^{-1} T^{-1/2} \sum_{t=1}^{\lfloor r T \rfloor} e_t \right| \leq \tau_L^{-1} \sup_{\tau \in [\tau_0, T]} \left| T^{-1/2} \sum_{t=1}^{\lfloor r T \rfloor} e_t \right| = O_p(1)
\]

by Lemma A.2: (ii) \( \sup_{\tau \in [\tau_0, T]} |\alpha_{T, \tau}| = \sup_{\tau \in [\tau_0, T]} |\delta T^{-d-1} \tau^{-1} (\tau - \tau_0)| = O(T^{-d}) \), which holds because \( \sup_{\tau \in [\tau_0, T]} |\alpha_{T, \tau}| \leq \delta T^{-d-1} \sup_{\tau \in [\tau_0, T]} |\tau - \tau_0| \leq c T^{-d} \) for some \( c > 0 \); and (iii)

\[
\sup_{\tau \in [\tau_0, T]} |\beta_{T, \tau}| = \sup_{\tau \in [\tau_0, T]} |\delta T^{-d-1} \tau_0| = O(T^{-d}). \tag{A.16}
\]

We will now show that \( \sup_{\tau \in [\tau_0, T]} \sum_{t=1}^{\lfloor r T \rfloor} |A_{t, \tau, T}| = o_p(1) \) in (A.13). Using (A.14), we have that

\[
\sum_{t=1}^{\lfloor r T \rfloor} |A_{t, \tau, T}| = \sum_{t=1}^{\lfloor r T \rfloor} w_{ti} \left( \alpha_{T, \tau}^2 + \tilde{e}_\tau^2 - 2 \epsilon_i \alpha_{T, \tau} - 2 \epsilon_i \tilde{e}_\tau - 2 \alpha_{T, \tau} \tilde{e}_\tau \right)^2. \tag{A.17}
\]

Consider the first term in (A.17). We have (because \( d = 1/2 \))

\[
\sup_{\tau \in [\tau_0, T]} \sum_{t=1}^{\lfloor r T \rfloor} w_{ti} \alpha_{T, \tau}^2 \leq \sup_{\tau \in [\tau_0, T]} |\alpha_{T, \tau}|^4 \sum_{t=1}^{\lfloor r T \rfloor} w_{ti}^2 \leq \sup_{\tau \in [\tau_0, T]} |\alpha_{T, \tau}|^4 \sum_{t=1}^{\lfloor r T \rfloor} w_{ti}^2 = O_p(T^{-1}).
\]

For the second term in (A.17), we have

\[
\sup_{\tau \in [\tau_0, T]} \sum_{t=1}^{\lfloor r T \rfloor} w_{ti} \tilde{e}_\tau^2 \leq \sup_{\tau \in [\tau_0, T]} \left| T^{1/2} \tilde{e}_\tau \right|^4 T^{-2} \sum_{t=1}^{\lfloor r T \rfloor} w_{ti}^2 \leq \sup_{\tau \in [\tau_0, T]} \left| T^{1/2} \tilde{e}_\tau \right|^4 T^{-2} \sum_{t=1}^{\lfloor r T \rfloor} w_{ti}^2 = O_p(T^{-1}).
\]

Before we consider the third term, note that

\[
T^{-1} \sum_{t=1}^{\lfloor r T \rfloor} w_{ti} e_i = O_p \left( \frac{1}{T h} \right). \tag{A.18}
\]

because

\[
T^{-1} \sum_{t=1}^{\lfloor r T \rfloor} E \left( \sum_{i=1}^{\lfloor r T \rfloor} w_{ti} e_i \right)^2 = T^{-1} \sum_{t=1}^{\lfloor r T \rfloor} \sum_{i=1}^{\lfloor r T \rfloor} w_{ti}^2 \sigma_i^2 \leq \sigma^2 T^{-1} \sum_{t=1}^{\lfloor r T \rfloor} \sum_{i=1}^{\lfloor r T \rfloor} w_{ti}^2 \leq \sigma^2 \left( \max_{t, i} w_{ti} \right) T^{-1} \sum_{t=1}^{\lfloor r T \rfloor} \sum_{i=1}^{\lfloor r T \rfloor} w_{ti} = O \left( \frac{1}{T h} \right),
\]

A.9
since $\max_{t,i} w_{ti} = O(1/Th)$; see Lemma A(d) of Xu and Phillips, 2008. Consequently, for the third term in (A.17), we have that
\[
\sup_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} e_{i,\alpha_{T,\tau}} \right)^2 = T \sup_{\tau \in [\tau_L, \tau_U]} \alpha_{T,\tau}^2 \cdot T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} e_{i} \right)^2 = O_p \left( \frac{1}{Th} \right).
\]
For the fourth term in (A.17), we have that
\[
\sup_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} e_{i,\epsilon_{\tau}} \right)^2 \leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \epsilon_{\tau}|^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} e_{i} \right)^2 = O_p \left( \frac{1}{Th} \right).
\]
Finally, for the fifth term in (A.17), we have that
\[
\sup_{\tau \in [\tau_L, \tau_U]} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} \alpha_{T,\tau} \epsilon_{\tau} \right)^2 \leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \epsilon_{\tau}|^2 \sup_{\tau \in [\tau_L, \tau_U]} |\alpha_{T,\tau}|^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{[\tau_T]} w_{ti} \epsilon_{i} \right)^2 = O_p \left( T^{-1} \right).
\]
The cross-product terms in (A.17) are all of $o_p(1)$ uniformly in $\tau \in [\tau_L, \tau_U]$ by the application of the Cauchy-Schwarz inequality.

The term $B_{\tau,t}$ in (A.13) follows similarly since $\sup_{\tau \in [\tau_L, \tau_U]} |\beta_{T,\tau}| = O \left( T^{-1/2} \right)$; see (A.16). All of the cross-product terms are again of $o_p(1)$ by the Cauchy-Schwarz inequality. This completes the proof for (A.11).

We next prove (A.12). We have
\[
\max_{1 \leq t \leq T} |\hat{\sigma}_{T,\tau,t}^2 - \sigma_t^2| \leq \max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} |\hat{\sigma}_{T,\tau,t}^2 - \sigma_t^2| \leq \max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} |A_{\tau,t} + B_{\tau,t} + C_{\tau,t}|.
\]
We will only show that $\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} |A_{\tau,t}| = o_p(1)$ since the term $B_{\tau,t}$ follows similarly as just noted. We have
\[
|A_{\tau,t}| = \sum_{i=1}^{[\tau_T]} w_{ti} \left( \alpha_{T,\tau}^2 + \hat{\epsilon}_{\tau}^2 - 2 \alpha_{T,\tau} \hat{\epsilon}_{\tau} - 2 \alpha_{T,\tau} \epsilon_{\tau} \right).
\]
(A.19)

For the first term in (A.19), we have that
\[
\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} \sum_{i=1}^{[\tau_T]} w_{ti} \alpha_{T,\tau}^2 \leq \sup_{\tau \in [\tau_L, \tau_U]} |\alpha_{T,\tau}|^2 \max_{1 \leq t \leq T} \sum_{i=1}^{[\tau_T]} w_{ti} \leq \sup_{\tau \in [\tau_L, \tau_U]} |\alpha_{T,\tau}|^2 \max_{1 \leq t \leq T} \sum_{i=1}^{[\tau_T]} w_{ti} = O_p \left( T^{-1} \right).
\]
For the second term in (A.19),
\[
\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} \sum_{i=1}^{[\tau_T]} w_{ti} \hat{\epsilon}_{\tau}^2 \leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \hat{\epsilon}_{\tau}|^2 \max_{1 \leq t \leq T} \sum_{i=1}^{[\tau_T]} w_{ti} \leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \hat{\epsilon}_{\tau}|^2 \max_{1 \leq t \leq T} \sum_{i=1}^{[\tau_T]} w_{ti} = O_p \left( T^{-1} \right).
\]
A.10
Before turning to the third term, we note that $T^{-1} \sum_{t=1}^{T} |e_t| = O_p(1)$. Using this result, for the third term in (A.19) we have that,

$$
\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} \left| \sum_{i=1}^{[\tau T]} w_t e_i \alpha_{T, \tau} \right| \leq \sup_{\tau \in [\tau_L, \tau_U]} \left| \alpha_{T, \tau} \right| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right|

\leq T \sup_{\tau \in [\tau_L, \tau_U]} \left| \alpha_{T, \tau} \right| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right|

\leq T \sup_{\tau \in [\tau_L, \tau_U]} \left| \alpha_{T, \tau} \right| \left( \max_{t,i} w_{ti} \right) T^{-1} \sum_{i=1}^{[\tau T]} |e_i|

= O_p \left( \frac{1}{\sqrt{Th}} \right)

Next, for the fourth term in (A.19) we have that,

$$
\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} \left| \sum_{i=1}^{[\tau T]} w_t e_i \bar{e}_{T} \right| \leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \bar{e}_{T} \right| \left( T^{-1/2} \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right| \right)

\leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \bar{e}_{T} \right| \left( T^{-1/2} \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right| \right)

\leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \bar{e}_{T} \right| \left( T^{-1/2} \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right| \right)

= O_p \left( \frac{1}{\sqrt{Th}} \right)

Finally, for the fifth term in (A.19) we have that,

$$
\max_{1 \leq t \leq T} \sup_{\tau \in [\tau_L, \tau_U]} \left| \sum_{i=1}^{[\tau T]} w_t \alpha_{T, \tau} \bar{e}_{T} \right| \leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \bar{e}_{T} \right| \sup_{\tau \in [\tau_L, \tau_U]} \left| \alpha_{T, \tau} \right| \left( T^{-1/2} \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right| \right)

\leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \bar{e}_{T} \right| \sup_{\tau \in [\tau_L, \tau_U]} \left| \alpha_{T, \tau} \right| \left( T^{-1/2} \max_{1 \leq t \leq T} \left| \sum_{i=1}^{[\tau T]} w_t e_i \right| \right)

= O_p \left( T^{-1} \right)

This completes the proof of the theorem for the $d = 1/2$ case.

**Proof of Theorem 3, Case 2 : $0 \leq d < 1/2$**

We now consider $0 \leq d < 1/2$ and prove that (A.11) and (A.12) also hold here. To ease notation we will simply use $\hat{\tau}$ to denote $\hat{\tau}_{DLS}$ in the remainder of this proof.

To show (A.11), we write, for any $\tau$,

$$
\sum_{t=1}^{T} |\hat{\sigma}_{T,t}^2 - \sigma_t^2|^2 = \sum_{t=1}^{T} |A_{\tau,t} + B_{\tau,t} + C_{\tau,t}|^2.
$$

(A.20)
We now evaluate each term in (A.20) at \( \tau = \hat{\tau} \) where \( \hat{\tau} - \tau_0 = O_p \left( T^{-1+2d} \right) \). To that end, we first note that
\[
|\alpha_{T,\hat{\tau}}| = |\delta T^{-d} \hat{\tau}^{-1} (\hat{\tau} - \tau_0)| = O_p \left( T^{-1+d} \right)
\]
because \( |\alpha_{T,\hat{\tau}}| = \delta T^{-d} |\hat{\tau}^{-1} (\hat{\tau} - \tau_0)| \leq \delta \tau_0^{-1} T^{-d} |\hat{\tau} - \tau_0| \). Moreover, \( |\beta_{T,\hat{\tau}}| = \delta T^{-d} |\hat{\gamma}_0| = O_p \left( T^{-d} \right) \).

Let us first deal with the \( A_{\tau,t} \) term in (A.20). To that end, recall first that
\[
\sum_{t=1}^{T} A_{\tau,t}^2 = \sum_{t=1}^{T} \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \left( \alpha_{T,t}^2 + \bar{\varepsilon}_t^2 - 2e_t \alpha_{T,t} - 2e_t \bar{\varepsilon}_t - 2\alpha_{T,t} \bar{\varepsilon}_t \right)^2 \tag{A.21}
\]
We now evaluate every term in (A.21) at \( \tau = \hat{\tau} \). For the first term,
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \alpha_{T,t}^2 \right)^2 \leq |\alpha_{T,\hat{\tau}}|^4 \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \right)^2 = O_p \left( T^{4d-3} \right) = o_p \left( 1 \right).
\]
For the second term,
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \bar{\varepsilon}_t^2 \right)^2 \leq \sup \limits_{\tau \in \lfloor \tau_{L,\tau_U} \rfloor} T^{1/2} \bar{\varepsilon}_\tau^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \right)^2 = O_p \left( T^{-1} \right).
\]
For the third term, recalling equation (A.18),
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} e_t \alpha_{T,\hat{\tau}} \right)^2 = T |\alpha_{T,\hat{\tau}}|^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} e_t \right)^2 = O_p \left( T^{-2} \right) = o_p \left( 1 \right).
\]
For the fourth term,
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} e_t \bar{\varepsilon}_t \right)^2 \leq \sup \limits_{\tau \in \lfloor \tau_{L,\tau_U} \rfloor} T^{1/2} \bar{\varepsilon}_\tau^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} e_t \right)^2 = O_p \left( \frac{1}{T h} \right).
\]
Finally, for the fifth term,
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \alpha_{T,t} \bar{\varepsilon}_t \right)^2 \leq \sup \limits_{\tau \in \lfloor \tau_{L,\tau_U} \rfloor} T^{1/2} \bar{\varepsilon}_\tau^2 |\alpha_{T,\hat{\tau}}|^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \right)^2 = O_p \left( T^{-2d-2} \right) = o_p \left( 1 \right).
\]

We next deal with \( B_{\tau,t} \) term in (A.20). Using (A.13) and (A.15), we have that
\[
\sum_{t=1}^{T} B_{\tau,t}^2 = \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \left( \beta_{T,t}^2 + \bar{\varepsilon}_t^2 - 2e_t \beta_{T,t} - 2e_t \bar{\varepsilon}_t - 2\beta_{T,t} \bar{\varepsilon}_t \right) \right)^2 \tag{A.22}
\]
We will need to evaluate each term at \( \tau = \hat{\tau} \). For the first term we have that
\[
\sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \beta_{T,t}^2 \right)^2 = |\beta_{T,\hat{\tau}}|^4 \sum_{t=1}^{T} \left( \sum_{i=1}^{\lfloor T/2 \rfloor} w_{ti} \right)^2 \leq |\beta_{T,\hat{\tau}}|^4 \left( \max \limits_{t,i} w_{ti} \right)^2 T (T (\hat{\tau} - \tau_0))^2 = O_p \left( \frac{1}{T h} \right) = o_p \left( 1 \right).
\]

A.12
For the second term
\[
\sum_{t=1}^{T} \left( \sum_{i=\lfloor \hat{\tau}T \rfloor + 1}^{\lfloor \hat{\tau}T \rfloor} w_{ti} \tilde{e}_i^2 \right)^2 = |\tilde{e}_T|^4 \sum_{t=1}^{T} \left( \sum_{i=\lfloor \tau_0T \rfloor + 1}^{\lfloor \hat{\tau}T \rfloor} w_{ti} \right)^2 \\
\leq \sup_{\tau \in [\hat{\tau}, \tau_U]} \left| T^{1/2} \tilde{e}_\tau \right|^4 \left( \max_{i,t} w_{ti} \right)^2 T^{-1} (T (\hat{\tau} - \tau))^2 \\
= O_p \left( \frac{1}{T^{3-4d_h^2}} \right) = o_p(1).
\]

To deal with the third term in (A.22), define
\[
V_{[\tau T],t} := \sum_{i=\lfloor \tau T \rfloor + 1}^{\lfloor \hat{\tau}T \rfloor} w_{ti} \beta_i T^{1/2} 
\]
for \( t = 1, \ldots, T \). Since \( \hat{\tau} - \tau_0 = O_p \left( T^{-1} \delta_T^{-2} \right) \) we consider values of \( \tau \) in the set
\[
K_T(M) := \{ \tau : \lfloor \tau_0T \rfloor + 1 \leq \lfloor \tau T \rfloor \leq \lfloor \tau_0T \rfloor + v\delta_T^{-2} \text{ for all } 0 \leq v \leq M \},
\]
for \( M < \infty \), which is an interval above \( \tau_0 \). (The corresponding interval below \( \tau_0 \) is handled identically, see the proof of Theorem 1 of Bai (1994) for the same approach.) Since \( V_{[\tau T],t} \) is a martingale indexed by \( \lfloor \tau T \rfloor \), Doob’s inequality implies
\[
E \left( \sup_{\tau \in K_T(M)} \left| \sum_{i=1}^{\lfloor \tau T \rfloor - \lfloor \tau_0T \rfloor} w_{(i+\lfloor \tau_0T \rfloor)} \beta_i \right|^2 \right) \leq 4E \left( \sum_{i=1}^{\lfloor \tau T \rfloor - \lfloor \tau_0T \rfloor} w_{(i+\lfloor \tau_0T \rfloor)} \beta_i \right)^2 \\
\leq 4 \sum_{i=1}^{\lfloor \tau T \rfloor - \lfloor \tau_0T \rfloor} w_{(i+\lfloor \tau_0T \rfloor)}^2 \sigma_i^2 \\
\leq 4 \sigma_i^2 v \left( \max_{i,t} w_{ti} \right)^2 \delta_T^{-2} \\
= O \left( \frac{1}{T^{2-2d_h^2}} \right),
\]
uniformly in \( t \) (in view of the uniform boundedness of \( w_{ti} \) over \( t \)). Hence, uniformly in \( t \),
\[
|V_{[\tau T],t}| \leq \sup_{\tau \in K_T(M)} |V_{[\tau T],t}| = O_p \left( \frac{1}{T^{1-d_h}} \right). \tag{A.23}
\]

Therefore the third term in (A.22) is such that
\[
\sum_{t=1}^{T} \left( \sum_{i=\lfloor \tau_0T \rfloor + 1}^{\lfloor \hat{\tau}T \rfloor} w_{ti} \beta_i \right)^2 = T |\beta_{T,\hat{T}}|^2 T^{-1} \sum_{t=1}^{T} \left( \sum_{i=\lfloor \tau_0T \rfloor + 1}^{\lfloor \hat{\tau}T \rfloor} w_{ti} \right)^2 \\
= O_p \left( \frac{1}{Th^2} \right).
\]
For the second term in (A.25) we have that,
\[
\sum_{t=1}^{T} \left( \sum_{i=[T_{0}]+1}^{[T]} w_{ti} e_{i} \right)^{2} = T|\bar{e}_{\hat{\tau}}|^{2} T^{-1} \sum_{t=1}^{T} \left( \sum_{i=[T_{0}]+1}^{[T]} w_{ti} e_{i} \right)^{2} \leq \sup_{\tau \in [\tau_{L}, \tau_{U}]} \left| T^{1/2} \bar{e}_{\tau} \right|^{2} T^{-1} \sum_{t=1}^{T} \left( \sum_{i=[T_{0}]+1}^{[T]} w_{ti} e_{i} \right)^{2} \leq O_p \left( \frac{1}{T^{2-2d\eta^2}} \right).
\]

Finally, for the fifth term in (A.22) we have that,
\[
\sum_{t=1}^{T} \left( \sum_{i=[T_{0}]+1}^{[T]} w_{ti} \beta_{T,\hat{\tau}} \bar{e}_{\hat{\tau}} \right)^{2} \leq \sup_{\tau \in [\tau_{L}, \tau_{U}]} \left| T^{1/2} \bar{e}_{\tau} \right|^{2} T^{-1} \left| \beta_{T,\hat{\tau}} \right|^{2} \sum_{t=1}^{T} \left( \sum_{i=[T_{0}]+1}^{[T]} w_{ti} \right)^{2} \leq \sup_{\tau \in [\tau_{L}, \tau_{U}]} \left| T^{1/2} \bar{e}_{\tau} \right|^{2} T^{-1} \left( \max_{t,i} w_{ti} \right)^{2} T (T (\hat{\tau} - \bar{\tau}_{0}))^{2} = O_p \left( \frac{1}{T^{2-2d\eta^2}} \right).
\]

Next, to show (A.12), we write, for any $\tau$,
\[
\max_{1 \leq t \leq T} \left| \hat{\sigma}_{\tau,t}^{2} - \hat{\sigma}_{t}^{2} \right| \leq \max_{1 \leq t \leq T} \left| A_{\tau,t} + B_{\tau,t} + C_{\tau,t} \right|.
\]

(A.24)

For $A_{\tau,t}$, we recall that
\[
|A_{\tau,t}| = \left| \sum_{i=1}^{[T_{0}]} w_{ti} \left( \alpha_{T,\tau}^{2} + \bar{e}_{\tau}^{2} - 2e_{i} \alpha_{T,\tau} - 2\bar{e}_{\tau} \bar{e}_{\tau} - 2\alpha_{T,\tau} \bar{\bar{e}}_{\tau} \right) \right|.
\]

(A.25)

We therefore need to evaluate each term in (A.25) at $\tau = \hat{\tau}$. For the first term, we have that
\[
\max_{1 \leq t \leq T} \left| \sum_{i=1}^{[T_{0}]} w_{ti} \alpha_{T,\hat{\tau}}^{2} \right| \leq \left| \alpha_{T,\hat{\tau}} \right|^{2} \max_{1 \leq t \leq T} \sum_{i=1}^{[T_{0}]} w_{ti} \leq \left| \alpha_{T,\hat{\tau}} \right|^{2} \sum_{i=1}^{T} w_{ti} = O_p \left( T^{-2+2d} \right) = O_p (1).
\]

For the second term in (A.25) we have that
\[
\max_{1 \leq t \leq T} \left| \sum_{i=1}^{[T_{0}]} w_{ti} \bar{e}_{\tau}^{2} \right| \leq \left| \bar{e}_{\tau} \right|^{2} \max_{1 \leq t \leq T} \sum_{i=1}^{[T_{0}]} w_{ti} \leq \sup_{\tau \in [\tau_{L}, \tau_{U}]} \left| T^{1/2} \bar{e}_{\tau} \right|^{2} T^{-1} \max_{1 \leq t \leq T} \sum_{i=1}^{T} w_{ti} = O_p (T^{-1}).
\]

A.14
For the third term in (A.25) we have that
\[
\max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t e_i \alpha_{T,i} \right| \leq |\alpha_{T,i}| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t e_i \right|
\]
\[
\leq T |\alpha_{T,i}| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t |e_i| \right|
\]
\[
\leq T |\alpha_{T,i}| \left( \max_{t,i} w_t \right) T^{-1} \left( \sum_{i=1}^{\lfloor \tau T \rfloor} |e_i| \right)
\]
\[
= O_p \left( \frac{1}{T^{1-d_h}} \right) = o_p(1).
\]

For the fourth term in (A.25) we have that
\[
\max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t e_i \bar{e}_{T,i} \right| \leq |\bar{e}_{T,i}| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t |e_i| \right|
\]
\[
\leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \bar{e}_{T,i}| T^{-1/2} \max_{1 \leq t \leq T} \left( \sum_{i=1}^{\lfloor \tau T \rfloor} w_t |e_i| \right)
\]
\[
\leq \sup_{\tau \in [\tau_L, \tau_U]} T^{1/2} \bar{e}_{T,i} T^{1/2} \left( \max_{t,i} w_t \right) T^{-1} \left( \sum_{i=1}^{\lfloor \tau T \rfloor} |e_i| \right)
\]
\[
= O_p \left( \frac{1}{\sqrt{h}} \right).
\]

Finally, for the fifth term in (A.25) we have that
\[
\max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t \alpha_{T,i} \bar{e}_{T,i} \right| \leq |\bar{e}_{T,i}| \max_{1 \leq t \leq T} \left| \sum_{i=1}^{\lfloor \tau T \rfloor} w_t \right|
\]
\[
\leq \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \bar{e}_{T,i}| \frac{1}{\sqrt{T}} |\alpha_{T,i}| \max_{1 \leq t \leq T} \left( \sum_{i=1}^{\lfloor \tau T \rfloor} w_t \right)
\]
\[
\leq |\alpha_{T,i}| \sup_{\tau \in [\tau_L, \tau_U]} |T^{1/2} \bar{e}_{T,i}| \frac{1}{T^{1/2}} \frac{1}{\sqrt{T}} \max_{1 \leq t \leq T} \left( \sum_{i=1}^{\lfloor \tau T \rfloor} w_t \right) = O_p \left( T^{d-3/2} \right).
\]

Turning next to the \( B_{\tau,i} \) term in (A.24), we have that
\[
|B_{\tau,i}| = \left| \sum_{i=\lfloor \tau T \rfloor+1}^{\lfloor \tau T \rfloor} w_t (\beta_{T,i}^2 + \bar{e}_{T,i}^2 - 2 \epsilon_i \beta_{T,i} - 2 \epsilon_i \bar{e}_{T,i} - 2 \beta_{T,i} \bar{e}_{T,i}) \right|
\]
and we will again need to evaluate each term at \( \tau = \hat{\tau} \). The first term is of \( o_p(1) \) because
\[
\max_{1 \leq t \leq T} \left| \sum_{i=\lfloor \tau T \rfloor+1}^{\lfloor \tau T \rfloor} w_t \beta_{T,i}^2 \right| \leq |\beta_{T,i}|^2 \max_{1 \leq t \leq T} \left| \sum_{i=\lfloor \tau T \rfloor+1}^{\lfloor \tau T \rfloor} w_t \right|
\]
\[
\leq |\beta_{T,i}|^2 \left( \max_{t,i} w_t \right) \left| T (\hat{\tau} - \tau_0) \right| = O_p \left( \frac{1}{T h} \right).
\]

A.15
The second term is of $o_p(1)$ because
\[
\max_{1 \leq t \leq T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} \hat{e}_t^2 \right| \leq |\hat{e}_t|^2 \max_{1 \leq t \leq T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} \right| \\
\leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \hat{e}_t \right|^2 T^{-1} \left( \max_{t,i} w_{ti} \right) T (\hat{\tau} - \tau_0) \\
= O_p \left( \frac{1}{T^{2 - 2d_h}} \right).
\]

For the third term, we note first that from part (d) of the proof of Theorem 2 of Xu and Phillips (2008, p.277), as $X_{t-1} = 1$ in their notation, (A.24) becomes $T^{-1} \sum_{t=1}^{T} |B_{\tau,t}|$. Using (A.23), we therefore have that
\[
T^{-1} \sum_{t=1}^{T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} e_i \beta_{T,\hat{\tau}} \right| \leq |\beta_{T,\hat{\tau}}| T^{-1} \sum_{t=1}^{T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} e_i \right| = O_p \left( \frac{1}{Th} \right).
\]

Similarly for the fourth term,
\[
T^{-1} \sum_{t=1}^{T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} e_i \hat{e}_t \right| \leq |\hat{e}_t| T^{-1} \sum_{t=1}^{T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} e_i \right| \\
\leq T^{-1/2} \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \hat{e}_t \right| T^{-1} \sum_{t=1}^{T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} e_i \right| \\
= O_p \left( \frac{1}{T^{3/2 - d_h}} \right).
\]

Finally, the fifth term is also of $o_p(1)$ because
\[
\max_{1 \leq t \leq T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} \beta_{T,\hat{\tau}} \right| \leq |\hat{e}_t| |\beta_{T,\hat{\tau}}| \max_{1 \leq t \leq T} \left| \sum_{i=[\tau T]+1}^{[\tau T]} w_{ti} \right| \\
\leq \sup_{\tau \in [\tau_L, \tau_U]} \left| T^{1/2} \hat{e}_t \right| T^{-1/2} |\beta_{T,\hat{\tau}}| \left( \max_{t,i} w_{ti} \right) T (\hat{\tau} - \tau_0) \\
= O_p \left( \frac{1}{T^{3/2 - d_h}} \right).
\]

This completes the proof for the $0 \leq d < 1/2$ case.

**Proof of Theorem 4**

For any $\tau$, define $X_t(\tau) := (1, t, 1_{t>\lfloor \tau T \rfloor}(t - \lfloor \tau T \rfloor))'$. The DGP in (2.1) can then be expressed as
\[
y_t = X_t(\tau)' \theta + z_t,
\]
\[
y_t = X_t(\tau)' \theta - d_t(\tau) \delta_T + z_t
\]

where $\theta := (\alpha, \mu, \delta_T)'$, and $d_t(\tau) := 1_{t>\lfloor \tau T \rfloor}(t - \lfloor \tau T \rfloor) - 1_{t>\lfloor \tau_0 T \rfloor}(t - \lfloor \tau_0 T \rfloor)$. Applying QD to (A.26) for any $\hat{e}$ gives
\[
y_{\hat{e},t} = X_{\hat{e},t}(\tau)' \theta - d_{\hat{e},t} \delta_T + z_{\hat{e},t},
\]

(A.27)

A.16
where \( z_{\ell,t} := \Delta z_t - \bar{c}T^{-1}z_{t-1} \),

\[
X_{\ell,t}(\tau) := \begin{pmatrix} 1_{t=1} + 1_{t>1}(-\bar{c}/T) \\ 1 - \bar{c}(t-1)/T \\ 1_{t>\lfloor \tau T \rfloor}(1 - \bar{c}(t - \lfloor \tau T \rfloor - 1)/T) \end{pmatrix},
\]

and

\[
d_{\ell,t}(\tau) := 1_{t>\lfloor \tau T \rfloor} \left( 1 - \bar{c} \frac{t - \lfloor \tau T \rfloor - 1}{T} \right) - 1_{t>\lfloor \tau_0 T \rfloor} \left( 1 - \bar{c} \frac{t - \lfloor \tau_0 T \rfloor - 1}{T} \right). \tag{A.28}
\]

Therefore, using equation (A.27), the estimated QD coefficient vector satisfies

\[
\hat{\theta}_e(\tau) = \left( \sum_{t=1}^{T} X_{\ell,t}(\tau)X_{\ell,t}(\tau)' \right)^{-1} \sum_{t=1}^{T} X_{\ell,t}(\tau)y_{\ell,t} = \theta + \left( \sum_{t=1}^{T} X_{\ell,t}(\tau)X_{\ell,t}(\tau)' \right)^{-1} \sum_{t=1}^{T} X_{\ell,t}(\tau)(z_{\ell,t} - d_{\ell,t}\delta_T),
\]

and the QD residuals are

\[
\hat{z}_{\ell,t}(\tau) := y_t - X_t(\tau)'\hat{\theta}_e(\tau) = z_t - d_t(\tau)\delta_T - X_t(\tau)'(\hat{\theta}_e(\tau) - \theta) = z_t - X_t(\tau)' \left( \sum_{t=1}^{T} X_{\ell,t}(\tau)X_{\ell,t}(\tau)' \right)^{-1} \sum_{t=1}^{T} X_{\ell,t}(\tau)z_{\ell,t}
\]

\[= - \left[ d_t(\tau)\delta_T - X_t(\tau)' \left( \sum_{t=1}^{T} X_{\ell,t}(\tau)X_{\ell,t}(\tau)' \right)^{-1} \sum_{t=1}^{T} X_{\ell,t}(\tau)d_{\ell,t}\delta_T \right]. \tag{A.29}
\]

The asymptotic distribution of the DF \( t \)-statistic calculated using \( \hat{z}_{\ell,t}(\hat{\tau}) \) for either \( \hat{\tau} = \hat{\tau}_{OLS} \) or \( \hat{\tau} = \hat{\tau}_{FWLS} \) follows from applying the FCLT results in (A.29) and showing that (A.30) (which captures the estimation effect of \( \hat{\tau} \) for \( \tau_0 \)) is asymptotically negligible.

In (A.29) we make use of the heteroskedastic FCLT for \( z_t = \phi T z_{t-1} + e_t \) with \( \phi := 1 + c/T \), such that

\[
\omega_e^{-1}T^{-1/2}z_{sT} = \omega_e^{-1}T^{-1/2} \sum_{t=1}^{sT} e_t \xrightarrow{d} B^c_{\eta}(s) := \int_0^s \exp(c(s - r))dB_{\eta}(r)
\]

where \( \omega_e^2 := \int_0^1 \sigma(r)^2dr \). It also then follows that

\[
\omega_e^{-1}T^{-1/2} \sum_{t=1}^{sT} z_{\ell,t} = \omega_e^{-1}T^{-1/2} \sum_{t=1}^{sT} (\Delta z_t - \bar{c}T^{-1}z_{t-1}) \xrightarrow{d} B^c_{\eta}(s) - \bar{c} \int_0^s B^c_{\eta}(r)dr =: B^c_{\eta}(s; \bar{c}).
\]

Defining \( D_T := \text{diag}(1, T^{-1/2}, T^{-1/2}) \), standard calculations (cf. Perron and Rodríguez, 2003, and Cavaliere et al., 2011) give the following limits, in each case uniformly in \( \tau \in [\tau_L, \tau_U] \),

\[
D_T^{-1}T^{-1/2}X_{sT}(\tau) \xrightarrow{d} \begin{pmatrix} 0 \\ s \\ (s - \tau) \lor 0 \end{pmatrix}
\]
Defining
\[
\omega^{-1} \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) X_{\tilde{e},t}(\tau)' D_T^{-1} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & f_1^T(1-\bar{c}s)^2 ds & f_1^T(1-\bar{c}s)(1-\bar{c}(s-\tau)) ds \\ 0 & f_1^T(1-\bar{c}s)(1-\bar{c}(s-\tau)) ds & f_1^T(1-\bar{c}(s-\tau))^2 ds \end{pmatrix}
\]

where
\[
\omega^{-1} \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) z_{\tilde{e},t} \rightarrow \begin{pmatrix} \bar{z}_{\tilde{e},1} \\ f_1^T(1-\bar{c}s) dB^e_{\eta}(s; \bar{c}) \\ f_1^T(1-\bar{c}(s-\tau)) dB^e_{\eta}(s; \bar{c}) \end{pmatrix}.
\]

Defining
\[
X(s; \tau) := \begin{pmatrix} s \\ (s - \tau) \vee 0 \end{pmatrix} \quad \text{and} \quad X_{\tilde{e}}(s; \tau) := \begin{pmatrix} 1 - \bar{c}s \\ 1 - \bar{c}((s - \tau) \vee 0) \end{pmatrix}
\]

the FCLT applied to (A.29) can be expressed
\[
\omega^{-1} T^{-1/2} z_{[sT]} - T^{-1/2} X_{[sT]}(\tau)'D_T^{-1} \left( \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) X_{\tilde{e},t}(\tau)' D_T^{-1} \right)^{-1} \omega^{-1} \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) z_{\tilde{e},t}
\]

\[
d \rightarrow B^e_{0}(s) - X(s; \tau)' \left( \int_{0}^{1} X_{\tilde{e}}(s; \tau) X_{\tilde{e}}(s; \tau)' ds \right)^{-1} \int_{0}^{1} X_{\tilde{e}}(s; \tau) dB^e_{0}(s; \bar{c})
\]

\[
=: Z(s; \tau, c, \bar{c}, \eta).
\]

(A.31)

If \( \tau \) is evaluated at an estimator \( \hat{\tau} \) that is consistent for \( \tau_0 \) then the continuous mapping theorem applies to \( \omega^{-1} T^{-1/2} z_{[s\hat{\tau}]}(\hat{\tau}) \) to deduce that the limiting process is \( Z(s; \tau_0, c, \bar{c}, \eta) \). The derivation of the asymptotic distribution of the DF \( t \)-statistic from the derived behaviour of this partial sum process is then entirely standard.

Now consider (A.30), scaled by \( T^{-1/2} \) commensurately with (A.31), written as
\[
r_{[sT]}(\tau) = T^{-1/2} d_{[sT]}(\tau) \delta_T - a_{[sT]}(\tau)' b_T(\tau)
\]

where
\[
a_{[sT]}(\tau) := T^{-1/2} D_T^{-1} X_{[sT]}(\tau)
\]
\[
b_T(\tau) := \left( \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) X_{\tilde{e},t}(\tau)' D_T^{-1} \right)^{-1} \sum_{t=1}^{T} D_T^{-1} X_{\tilde{e},t}(\tau) d_{\tilde{e},t} \delta_T.
\]

Then
\[
\max_{s \in [0,1]} |r_{[sT]}(\tau)| \leq \max_{s \in [0,1]} T^{-1/2} |\delta_T d_{[sT]}(\tau)| + \max_{s \in [0,1]} \left( a_{[sT]}(\tau)' a_{[sT]}(\tau) \right)^{1/2} (b_T(\tau)' b_T(\tau))^{1/2}
\]

(A.32)

The first term includes
\[
d_{t}(\tau) = \begin{cases} 0 & \text{if } t \leq [T\tau_0], t \leq [T\tau] \\ -(t - [T\tau_0]) & \text{if } [T\tau_0] < t \leq [T\tau] \\ t - [T\tau] & \text{if } [T\tau] < t \leq [T\tau_0] \\ -([T\tau] - [T\tau_0]) & \text{if } t > [T\tau_0], t > [T\tau] \end{cases}
\]

A.18
so that \( |d_t(\tau)| \leq ||T\tau| - |T\tau_0|| \) for every \( t \), and hence

\[
\max_{s \in [0,1]} T^{-1/2}|\delta_T d_{[sT]}(\tau)| \leq T^{1/2}\delta_T \left| \frac{|\tau T|}{T} - \frac{|\tau_0 T|}{T} \right|. \tag{A.33}
\]

For the second term in (A.32) we have

\[
a_{[sT]}(\tau) = T^{-1/2}D_T^{-1}X_{[sT]}(\tau) = \begin{pmatrix} T^{-1/2} \\ \frac{|sT|}{T} \end{pmatrix} \leq \begin{pmatrix} 1 \\ 1 \end{pmatrix},
\]

so

\[
\max_{s \in [0,1]} \left( a_{[sT]}(\tau)' a_{[sT]}(\tau) \right)^{1/2} \leq \sqrt{3}.
\]

For the last term in (A.32), we have shown above that \( \left( \sum_{t=1}^T D_T^{-1}X_{\epsilon,t}(\tau)X_{\epsilon,t}(\tau)'D_T^{-1} \right)^{-1} \) is a well-behaved bounded matrix, and therefore we focus on \( \sum_{t=1}^T D_T^{-1}X_{\epsilon,t}(\tau)d_{\epsilon,t}(\tau)\delta_T \). The structure of \( d_{\epsilon,t}(\tau) \) can be seen from (A.28) to be

\[
d_{\epsilon,t}(\tau) = \begin{cases} 0, & \text{if } t \leq |\tau T|, t \leq |\tau_0 T| \\ 1 - e^{-t - |\tau_0 T| - 1}, & \text{if } |\tau_0 T| < t \leq |\tau T| \\ 1 - e^{-t - |\tau_0 T| - 1}, & \text{if } |\tau_0 T| < t \leq |\tau T| \\ e^{-t - |\tau_0 T| - 1}, & \text{if } t > |\tau T|, t > |\tau_0 T|. \end{cases}
\]

Taking \( \tau_0 \leq \tau \) (the reverse follows similarly and gives the same inequality)

\[
T^{-1} \sum_{t=1}^T |d_{\epsilon,t}(\tau)| = T^{-1} \sum_{t=|\tau_0 T|+1}^{|\tau T|} \left| 1 - \frac{t - |\tau_0 T| - 1}{\tau T} \right| + T^{-1} \sum_{t=|\tau T|+1}^{|\tau_0 T|+1} |\epsilon| \left| \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right|
\]

\[
\leq (1 + |\epsilon|) \left| \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right| + |\epsilon|T^{-2} \sum_{t=1}^{|\tau_0 T|+1} t
\]

\[
= (1 + |\epsilon|) \left| \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right| + \frac{|\epsilon|}{2} \left( \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right) \left( \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} - \frac{1}{T} \right)
\]

\[
\leq \left| \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right| \left( 1 + \frac{3|\epsilon|}{2} \right).
\]

Considering

\[
\sum_{t=1}^T D_T^{-1}X_{\epsilon,t}(\tau)d_{\epsilon,t}(\tau)\delta_T = \begin{pmatrix} \delta T^{d-2} \sum_{t=2}^T (-\frac{\epsilon}{T}) d_{\epsilon,t}(\tau) \\ \delta T^{-d-1/2} \sum_{t=1}^T (1 - \frac{t - |\tau_0 T| - 1}{|\tau T|}) d_{\epsilon,t}(\tau) \\ \delta T^{-d-1/2} \sum_{t=|\tau T|+1}^{|\tau_0 T|+1} (1 - \frac{t - |\tau_0 T| - 1}{|\tau T|}) d_{\epsilon,t}(\tau) \end{pmatrix}. \tag{A.34}
\]

it follows that

\[
\left| T^{-d} \sum_{t=2}^T \left(-\frac{\epsilon}{T}\right) d_{\epsilon,t}(\tau) \right| \leq |\epsilon|T^{-d-1} \sum_{t=1}^T |d_{\epsilon,t}(\tau)| \leq |\epsilon| \left( 1 + \frac{3|\epsilon|}{2} \right) T^{-d} \left| \frac{|\tau T|}{\tau T} - \frac{|\tau_0 T|}{\tau T} \right|
\]

A.19
and
\[
T^{-d-1/2} \sum_{t=1}^{T} \left(1 - \tilde{c} \frac{t-1}{T}\right) d_{\tilde{c}, t}(\tau) \leq (1 + |\tilde{c}|) T^{-d+1/2} T^{-1} \sum_{t=1}^{T} |d_{\tilde{c}, t}(\tau)| \leq (1 + |\tilde{c}|) \left(1 + \frac{3|\tilde{c}|}{2}\right) T^{-d+1/2} \left|\frac{\tau T}{T} - \frac{\tau_0 T}{T}\right|
\]
and a bound of the same order applies to the third term. Putting these together gives
\[
(\delta_T(\tau)'b_T(\tau))^{1/2} \leq k_2 T^{-d+1/2} \left|\frac{\tau T}{T} - \frac{\tau_0 T}{T}\right|
\]
for some finite constant \(k\) (depending on the various constants involving \(\tilde{c}\) above), and combining this with (A.33) in (A.32) gives
\[
\max_{s \in [0,1]} |\rho_{(sT)}(\tau)| \leq k T^{-d+1/2} \left|\frac{\tau T}{T} - \frac{\tau_0 T}{T}\right|
\]
for another finite constant \(k\).

If we now consider the evaluation of this at \(\tau = \hat{\tau}\), where \(\hat{\tau} = \hat{\tau}_{OLS}\) or \(\hat{\tau} = \hat{\tau}_{FWLS}\), the rate of consistency \(\hat{\tau} - \tau_0 = O_p(T^{2d-1})\) implies that
\[
\left|\frac{\hat{\tau} T}{T} - \frac{\tau_0 T}{T}\right| = O_p(T^{2d-1}),
\]
and, hence, that
\[
\max_{s \in [0,1]} |\rho_{(sT)}(\hat{\tau})| \leq k T^{-d+1/2} \left|\frac{\hat{\tau} T}{T} - \frac{\tau_0 T}{T}\right| = O_p(T^{d-1/2}).
\]
This is therefore of \(o_p(1)\) for all \(d < 1/2\).  \(\blacksquare\)

**Proof of Theorem 5**
The two equations of the DGP
\[
y_t = \alpha + \mu t + \delta T 1_{t > \lfloor \tau_0 T \rfloor} (t - \lfloor \tau_0 T \rfloor) + z_t
\]
\[
z_t = \phi T z_{t-1} + e_t
\]
can be combined to give
\[
y_t = X_{1,t}' \beta_{1,T} + X_{2,t}(\tau)' \beta_{2,T} + \epsilon_t
\]
\[
= X_{1,t}' \beta_{1,T} + X_{2,t}(\tau)' \beta_{2,T} + \epsilon_{T,t}(\tau)
\]
where
\[
X_{1,t} := \begin{pmatrix} 1 \\ t \\ y_{t-1} \end{pmatrix}, \quad \beta_{1,T} := \begin{pmatrix} \mu \phi T + \alpha (1 - \phi T) \\ \mu (1 - \phi T) \\ \phi \end{pmatrix}
\]
and
\[
X_{2,t}(\tau) := \begin{pmatrix} 1_{t > \lfloor \tau_0 T \rfloor} \\ 1_{t > \lfloor \tau_0 T \rfloor} (t - \lfloor \tau_0 T \rfloor) \end{pmatrix}, \quad \beta_{2,T} := \delta_T \begin{pmatrix} \phi T \\ 1 - \phi T \end{pmatrix}
\]
\[ e_{T,t}(\tau) := e_t - (X_{2,t}(\tau) - X_{2,t}(\tau_0))'\beta_{2,T}. \]

The compound disturbance \( e_{T,t}(\tau) \) includes the DGP disturbance term \( e_t \) and also \((X_{2,t}(\tau) - X_{2,t}(\tau_0))'\beta_{2,T}\), which captures the effect of \( \tau \) differing from \( \tau_0 \), if applicable. When \( \tau \) is replaced by a consistent estimator, e.g. \( \hat{\tau}_{OLS} \) or \( \hat{\tau}_{FWLS} \), the asymptotic negligibility of this latter term follows by similar arguments to those used to the same effect in the proof of Theorem 4. It is convenient to stack the observations as

\[
y := \begin{pmatrix} y_2 \\ \vdots \\ y_T \end{pmatrix}, \quad X_1 := \begin{pmatrix} X'_{1,2} \\ \vdots \\ X'_{1,T} \end{pmatrix}, \quad X_{2,T} := \begin{pmatrix} X_{2,2}(\tau)' \\ \vdots \\ X_{2,T}(\tau)' \end{pmatrix},
\]

and to define the orthogonal projection matrix \( \bar{P}_1 := I_{T-1} - X_1(X_1'X_1)^{-1}X_1' \), so that

\[
\begin{align*}
s_0^2 &= T^{-1}y'\bar{P}_1y \\
s_\tau^2 &= T^{-1}(y'\bar{P}_1y - y'\bar{P}_1X_{2,\tau}(X_{2,\tau}'\bar{P}_1X_{2,\tau})^{-1}X_{2,\tau}'\bar{P}_1y).
\end{align*}
\]

The difference between the resulting SC statistics can be written

\[
SC_\tau - SC_0 = T \log \left( \frac{s_\tau^2}{s_0^2} \right) + 2 \log T \\
= T \log \left( 1 - \frac{y'\bar{P}_1X_{2,\tau}(X_{2,\tau}'\bar{P}_1X_{2,\tau})^{-1}X_{2,\tau}'\bar{P}_1y}{y'\bar{P}_1y} \right) + 2 \log T \\
= T \log \left( 1 - \frac{\hat{\beta}_2(\tau)'X_{2,\tau}'\bar{P}_1X_{2,\tau}\hat{\beta}_2(\tau)}{y'\bar{P}_1y} \right) + 2 \log T. \quad (A.35)
\]

When a trend break is present in the DGP (\( \delta \neq 0 \)) and \( 0 \leq d < 1/2 \), standard but tedious least squares derivations show that with \( \tau = \tau_0 \)

\[
\hat{\beta}_2(\tau_0) = T^{-d}b_{2,T} + o_p(T^{-d})
\]

with \( b_{2,T} := (1, -c/T)' \). This is essentially the consistency of the OLS estimator of the regression, which can be shown to hold in the presence of heteroskedasticity of the form allowed in Assumption A. Hence

\[
T \log \left( 1 - \frac{\hat{\beta}_2(\tau_0)'X_{2,\tau_0}'\bar{P}_1X_{2,\tau_0}\hat{\beta}_2(\tau_0)}{y'\bar{P}_1y} \right) \approx -T^{1-2d}b_{2,T}^2X_{2,\tau_0}'\bar{P}_1X_{2,\tau_0}b_{2,T} \frac{y'\bar{P}_1y}{y'P_1y}
\]

which diverges to \( -\infty \) at rate \( T^{1-2d} \) for \( 0 \leq d < 1/2 \), because the ratio of quadratic forms in this expression is \( O_p(1) \). This in turn implies that the terms in \( SC_{\tau_0} - SC_0 \) have orders \(-O_p(T^{1-2d}) + O(\log T)\), and hence that \( SC_{\tau_0} - SC_0 \) also diverges to \( -\infty \). The conclusion is that the probability that the model with the break (i.e. including \( X_{2,\tau}(\tau_0) \)) is chosen converges to one when \( \delta \neq 0 \) and \( 0 \leq d < 1/2 \). The same results follow similarly when \( \tau_0 \) is replaced by a consistent estimator, such as \( \hat{\tau}_{OLS} \) or \( \hat{\tau}_{FWLS} \). It also turns out that the same rates of convergence, and hence consistent detection of a break, can also be found for sequences of \( \tau \) not converging to \( \tau_0 \), although in that
case the expression for $b_{2,T}$ will differ and the finite sample properties would presumably be inferior to when a consistent estimator is used for $\tau_0$.

When a trend break is absent in the DGP ($\delta = 0$), the estimator $\hat{\beta}_2(\tau_0)$ converges to zero; i.e., it is a consistent estimator of $\beta_{2,T} = 0$ in this case. Moreover it can be shown to be $O_p(T^{-1/2})$ in the usual way, so that the first term in (A.35) satisfies

$$T \log \left( 1 - \frac{\hat{\beta}_2(\tau_0)'X_{2,\tau_0}'\bar{P}_1X_{2,\tau_0}\hat{\beta}_2(\tau_0)}{y'y} \right) \approx T \log(1 - O_p(T^{-1})) \approx O_p(1),$$

from which it follows that $SC_T - SC_0 \approx O_p(1) + 2 \log T \rightarrow +\infty$. Thus when a break is absent, the probability that it is excluded by the SC comparison converges to 1.

The preceding arguments apply without substantial change when the regressions used for the SC calculations are weighted to allow for unconditional heteroskedasticity of the form given in $A_2$. ■
Table 1: Finite Sample Properties of Break Fraction Estimators. No Level Break.
Volatility Models SD0-SD4.

<table>
<thead>
<tr>
<th></th>
<th>( T )</th>
<th>( \text{Mean} )</th>
<th>( \text{SD} )</th>
<th>( \text{Mean} )</th>
<th>( \text{SD} )</th>
<th>( \text{Mean} )</th>
<th>( \text{SD} )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: ( \hat{\tau}_{OLS} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD1 : ( \kappa = 2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.667</td>
<td>0.187</td>
<td>0.730</td>
<td>0.148</td>
<td>0.782</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.669</td>
<td>0.187</td>
<td>0.733</td>
<td>0.143</td>
<td>0.788</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD2 : ( \kappa = 2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.229</td>
<td>0.148</td>
<td>0.277</td>
<td>0.146</td>
<td>0.336</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.215</td>
<td>0.134</td>
<td>0.270</td>
<td>0.146</td>
<td>0.332</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD3 SD4 SD0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \kappa = 2, \lambda_0 = 0.3 )</td>
<td>( \gamma = 2 )</td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \gamma = 2 )</td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \gamma = 2 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.359</td>
<td>0.281</td>
<td>0.699</td>
<td>0.201</td>
<td>0.509</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.341</td>
<td>0.276</td>
<td>0.707</td>
<td>0.196</td>
<td>0.502</td>
<td>0.275</td>
</tr>
<tr>
<td><strong>Panel B: ( \hat{\tau}_{FWLS} )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD1 : ( \kappa = 2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.533</td>
<td>0.280</td>
<td>0.515</td>
<td>0.271</td>
<td>0.537</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.490</td>
<td>0.290</td>
<td>0.479</td>
<td>0.269</td>
<td>0.510</td>
<td>0.258</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD2 : ( \kappa = 2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
<td>( \lambda_0 = 0.7 )</td>
<td>( \lambda_0 = 0.5 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.478</td>
<td>0.263</td>
<td>0.505</td>
<td>0.270</td>
<td>0.493</td>
<td>0.282</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.493</td>
<td>0.258</td>
<td>0.525</td>
<td>0.268</td>
<td>0.519</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD3 SD4 SD0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( \kappa = 2, \lambda_0 = 0.3 )</td>
<td>( \kappa = 2 )</td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \kappa = 2 )</td>
<td>( \lambda_0 = 0.3 )</td>
<td>( \kappa = 2 )</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.486</td>
<td>0.261</td>
<td>0.549</td>
<td>0.269</td>
<td>0.510</td>
<td>0.272</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.492</td>
<td>0.251</td>
<td>0.523</td>
<td>0.275</td>
<td>0.504</td>
<td>0.275</td>
</tr>
</tbody>
</table>
Table 2: Finite Sample Properties of Break Fraction Estimators. Break size $\delta$, Break fraction $\tau_0$. Volatility Models SD0 and SD1.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>T</th>
<th>Mean SD RMSE</th>
<th>Mean SD RMSE</th>
<th>Mean SD RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_0 = 0.3$</td>
<td></td>
<td>$\tau_0 = 0.5$</td>
<td>$\tau_0 = 0.7$</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.375 0.195 0.209</td>
<td>0.502 0.166 0.166</td>
<td>0.633 0.188 0.200</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.312 0.094 0.094</td>
<td>0.500 0.084 0.084</td>
<td>0.687 0.094 0.095</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.309 0.074 0.074</td>
<td>0.500 0.064 0.064</td>
<td>0.694 0.067 0.067</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.300 0.019 0.019</td>
<td>0.500 0.019 0.019</td>
<td>0.700 0.020 0.020</td>
</tr>
<tr>
<td></td>
<td>SD0</td>
<td></td>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.3$</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>0.614 0.209 0.378</td>
<td>0.631 0.188 0.229</td>
<td>0.669 0.178 0.181</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.536 0.213 0.318</td>
<td>0.585 0.174 0.193</td>
<td>0.670 0.161 0.164</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.507 0.210 0.295</td>
<td>0.569 0.166 0.180</td>
<td>0.672 0.153 0.156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.391 0.144 0.170</td>
<td>0.521 0.115 0.117</td>
<td>0.684 0.114 0.115</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.678 0.193 0.425</td>
<td>0.694 0.155 0.248</td>
<td>0.722 0.140 0.142</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.590 0.230 0.370</td>
<td>0.646 0.146 0.207</td>
<td>0.715 0.122 0.123</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.543 0.235 0.338</td>
<td>0.626 0.141 0.189</td>
<td>0.709 0.118 0.119</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.371 0.164 0.179</td>
<td>0.563 0.097 0.116</td>
<td>0.702 0.086 0.086</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>0.695 0.226 0.455</td>
<td>0.729 0.172 0.286</td>
<td>0.774 0.127 0.147</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.566 0.264 0.375</td>
<td>0.663 0.177 0.241</td>
<td>0.768 0.101 0.122</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.503 0.257 0.327</td>
<td>0.626 0.171 0.212</td>
<td>0.761 0.095 0.113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.334 0.130 0.135</td>
<td>0.531 0.098 0.103</td>
<td>0.742 0.061 0.074</td>
</tr>
<tr>
<td></td>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>0.491 0.260 0.323</td>
<td>0.525 0.254 0.255</td>
<td>0.555 0.264 0.302</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.414 0.219 0.247</td>
<td>0.503 0.221 0.221</td>
<td>0.568 0.253 0.285</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.420 0.205 0.238</td>
<td>0.517 0.196 0.197</td>
<td>0.609 0.222 0.240</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.347 0.109 0.119</td>
<td>0.505 0.124 0.124</td>
<td>0.657 0.159 0.164</td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>0.430 0.237 0.270</td>
<td>0.529 0.242 0.244</td>
<td>0.549 0.257 0.298</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.326 0.131 0.134</td>
<td>0.521 0.200 0.201</td>
<td>0.569 0.239 0.273</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.335 0.140 0.144</td>
<td>0.545 0.179 0.184</td>
<td>0.611 0.214 0.232</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.300 0.023 0.023</td>
<td>0.534 0.097 0.103</td>
<td>0.668 0.140 0.144</td>
</tr>
<tr>
<td></td>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td>0.416 0.222 0.250</td>
<td>0.522 0.199 0.200</td>
<td>0.574 0.250 0.280</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.318 0.106 0.108</td>
<td>0.496 0.107 0.107</td>
<td>0.603 0.226 0.246</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.320 0.105 0.107</td>
<td>0.503 0.101 0.101</td>
<td>0.642 0.204 0.212</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.300 0.019 0.019</td>
<td>0.499 0.021 0.021</td>
<td>0.700 0.117 0.117</td>
</tr>
</tbody>
</table>

Panel B: $\hat{\tau}_{FWLS}$
Table 3: Finite Sample Properties of Break Fraction Estimators. Break size $\delta$, Break fraction $\tau_0$. Volatility Model SD2.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$T$</th>
<th>Mean $\tau_0 = 0.3$</th>
<th>SD $\tau_0 = 0.3$</th>
<th>RMSE $\tau_0 = 0.3$</th>
<th>Mean $\tau_0 = 0.5$</th>
<th>SD $\tau_0 = 0.5$</th>
<th>RMSE $\tau_0 = 0.5$</th>
<th>Mean $\tau_0 = 0.7$</th>
<th>SD $\tau_0 = 0.7$</th>
<th>RMSE $\tau_0 = 0.7$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.233</td>
<td>0.132</td>
<td>0.148</td>
<td>0.277</td>
<td>0.170</td>
<td>0.281</td>
<td>0.314</td>
<td>0.228</td>
<td>0.448</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.230</td>
<td>0.101</td>
<td>0.123</td>
<td>0.337</td>
<td>0.177</td>
<td>0.241</td>
<td>0.437</td>
<td>0.264</td>
<td>0.372</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.244</td>
<td>0.104</td>
<td>0.118</td>
<td>0.384</td>
<td>0.168</td>
<td>0.205</td>
<td>0.510</td>
<td>0.252</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.256</td>
<td>0.062</td>
<td>0.076</td>
<td>0.469</td>
<td>0.098</td>
<td>0.103</td>
<td>0.664</td>
<td>0.134</td>
<td>0.139</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.283</td>
<td>0.136</td>
<td>0.137</td>
<td>0.309</td>
<td>0.150</td>
<td>0.243</td>
<td>0.328</td>
<td>0.193</td>
<td>0.419</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.282</td>
<td>0.120</td>
<td>0.122</td>
<td>0.352</td>
<td>0.146</td>
<td>0.208</td>
<td>0.411</td>
<td>0.231</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.292</td>
<td>0.115</td>
<td>0.115</td>
<td>0.373</td>
<td>0.139</td>
<td>0.188</td>
<td>0.467</td>
<td>0.234</td>
<td>0.330</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.296</td>
<td>0.087</td>
<td>0.087</td>
<td>0.434</td>
<td>0.098</td>
<td>0.118</td>
<td>0.627</td>
<td>0.168</td>
<td>0.183</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.335</td>
<td>0.173</td>
<td>0.177</td>
<td>0.371</td>
<td>0.183</td>
<td>0.224</td>
<td>0.387</td>
<td>0.204</td>
<td>0.374</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.326</td>
<td>0.161</td>
<td>0.163</td>
<td>0.415</td>
<td>0.173</td>
<td>0.193</td>
<td>0.462</td>
<td>0.213</td>
<td>0.319</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.330</td>
<td>0.152</td>
<td>0.155</td>
<td>0.432</td>
<td>0.162</td>
<td>0.176</td>
<td>0.493</td>
<td>0.206</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.313</td>
<td>0.113</td>
<td>0.114</td>
<td>0.477</td>
<td>0.114</td>
<td>0.117</td>
<td>0.607</td>
<td>0.146</td>
<td>0.173</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.440</td>
<td>0.252</td>
<td>0.288</td>
<td>0.489</td>
<td>0.195</td>
<td>0.195</td>
<td>0.592</td>
<td>0.215</td>
<td>0.241</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.399</td>
<td>0.227</td>
<td>0.247</td>
<td>0.503</td>
<td>0.107</td>
<td>0.108</td>
<td>0.680</td>
<td>0.108</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.364</td>
<td>0.208</td>
<td>0.218</td>
<td>0.499</td>
<td>0.097</td>
<td>0.097</td>
<td>0.684</td>
<td>0.096</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.299</td>
<td>0.120</td>
<td>0.120</td>
<td>0.501</td>
<td>0.023</td>
<td>0.023</td>
<td>0.700</td>
<td>0.021</td>
<td>0.021</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.468</td>
<td>0.258</td>
<td>0.308</td>
<td>0.487</td>
<td>0.241</td>
<td>0.241</td>
<td>0.586</td>
<td>0.229</td>
<td>0.256</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.432</td>
<td>0.239</td>
<td>0.273</td>
<td>0.479</td>
<td>0.200</td>
<td>0.201</td>
<td>0.671</td>
<td>0.135</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.397</td>
<td>0.217</td>
<td>0.238</td>
<td>0.463</td>
<td>0.177</td>
<td>0.181</td>
<td>0.673</td>
<td>0.127</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.332</td>
<td>0.143</td>
<td>0.146</td>
<td>0.465</td>
<td>0.099</td>
<td>0.105</td>
<td>0.700</td>
<td>0.023</td>
<td>0.023</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.463</td>
<td>0.268</td>
<td>0.314</td>
<td>0.490</td>
<td>0.254</td>
<td>0.254</td>
<td>0.526</td>
<td>0.258</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.429</td>
<td>0.252</td>
<td>0.283</td>
<td>0.500</td>
<td>0.221</td>
<td>0.221</td>
<td>0.584</td>
<td>0.219</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.403</td>
<td>0.227</td>
<td>0.249</td>
<td>0.488</td>
<td>0.194</td>
<td>0.194</td>
<td>0.585</td>
<td>0.201</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.341</td>
<td>0.160</td>
<td>0.165</td>
<td>0.497</td>
<td>0.126</td>
<td>0.126</td>
<td>0.652</td>
<td>0.111</td>
<td>0.121</td>
</tr>
</tbody>
</table>
Table 4: Finite Sample Properties of Break Fraction Estimators. Break size $\delta$, Break fraction $\tau_0$. Volatility Models SD3 and SD4.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$T$</th>
<th>Mean $\hat{\tau}_{OLS}$</th>
<th>SD $\hat{\tau}_{OLS}$</th>
<th>RMSE $\hat{\tau}_{OLS}$</th>
<th>Mean $\hat{\tau}_{FWLS}$</th>
<th>SD $\hat{\tau}_{FWLS}$</th>
<th>RMSE $\hat{\tau}_{FWLS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>RMSE</td>
<td>Mean</td>
<td>SD</td>
<td>RMSE</td>
</tr>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.339</td>
<td>0.257</td>
<td>0.260</td>
<td>0.379</td>
<td>0.267</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.294</td>
<td>0.207</td>
<td>0.207</td>
<td>0.395</td>
<td>0.235</td>
<td>0.257</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.300</td>
<td>0.198</td>
<td>0.198</td>
<td>0.427</td>
<td>0.214</td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.265</td>
<td>0.093</td>
<td>0.099</td>
<td>0.473</td>
<td>0.112</td>
<td>0.115</td>
</tr>
</tbody>
</table>

Panel A: $\hat{\tau}_{OLS}$

SD3: $\kappa = 2$, $\lambda_0 = 0.3$

SD4: $\kappa = 1$

SD4: $\kappa = 2$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Mean $\hat{\tau}_{FWLS}$</th>
<th>SD $\hat{\tau}_{FWLS}$</th>
<th>RMSE $\hat{\tau}_{FWLS}$</th>
<th>Mean $\hat{\tau}_{FWLS}$</th>
<th>SD $\hat{\tau}_{FWLS}$</th>
<th>RMSE $\hat{\tau}_{FWLS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>100</td>
<td>0.493</td>
<td>0.210</td>
<td>0.210</td>
<td>0.547</td>
<td>0.250</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.402</td>
<td>0.224</td>
<td>0.246</td>
<td>0.501</td>
<td>0.121</td>
<td>0.120</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>0.372</td>
<td>0.213</td>
<td>0.225</td>
<td>0.501</td>
<td>0.121</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.301</td>
<td>0.122</td>
<td>0.122</td>
<td>0.500</td>
<td>0.026</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Panel B: $\hat{\tau}_{FWLS}$

SD3: $\kappa = 2$, $\lambda_0 = 0.3$

SD4: $\kappa = 1$

SD4: $\kappa = 2$
Figure 1: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. No Level Break.

(a) $\hat{\tau}_{OLS}$: $\delta = 0$, SD0, $\kappa = 0$, $\lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD0, $\kappa = 0$, $\lambda_0 = 0$

(c) $\hat{\tau}_{OLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD1, $\kappa = 2$, $\lambda_0 = 0.7$
(i) $\hat{\tau}_{OLS} : \delta = 0, SD3, \kappa = 1, \lambda_0 = 0.3$

(j) $\hat{\tau}_{FWLS} : \delta = 0, SD3, \kappa = 1, \lambda_0 = 0.3$

(k) $\hat{\tau}_{OLS} : \delta = 0, SD3, \kappa = 2, \lambda_0 = 0.3$

(l) $\hat{\tau}_{FWLS} : \delta = 0, SD3, \kappa = 2, \lambda_0 = 0.3$

(m) $\hat{\tau}_{OLS} : \delta = 0, SD4, \kappa = 1, \lambda_0 = 0$

(n) $\hat{\tau}_{FWLS} : \delta = 0, SD4, \kappa = 1, \lambda_0 = 0$

(o) $\hat{\tau}_{OLS} : \delta = 0, SD4, \kappa = 2, \lambda_0 = 0$

(p) $\hat{\tau}_{FWLS} : \delta = 0, SD4, \kappa = 2, \lambda_0 = 0$
Figure 2: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.3$.

(a) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD0, \kappa = 0, \lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD0, \kappa = 0, \lambda_0 = 0$

(c) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.7$
Figure 2: continued ...

(i) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}3, \kappa = 1, \lambda_0 = 0.3$

(j) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}3, \kappa = 1, \lambda_0 = 0.3$

(k) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}3, \kappa = 2, \lambda_0 = 0.3$

(l) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}3, \kappa = 2, \lambda_0 = 0.3$

(m) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}4, \kappa = 1, \lambda_0 = 0.0$

(n) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}4, \kappa = 1, \lambda_0 = 0.0$

(o) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}4, \kappa = 2, \lambda_0 = 0.0$

(p) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.3; \text{SD}4, \kappa = 2, \lambda_0 = 0.0$
Figure 3: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.5$.

(a) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(c) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 2$, $\lambda_0 = 0.7$
Figure 3: continued ...

(i) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD3}, \kappa = 1, \lambda_0 = 0.3$

(ii) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD3}, \kappa = 1, \lambda_0 = 0.3$

(iii) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD4}, \kappa = 1, \lambda_0 = 0.3$

(iv) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD4}, \kappa = 1, \lambda_0 = 0.3$
Figure 4: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.7$.

(a) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD0, \kappa = 0, \lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD0, \kappa = 0, \lambda_0 = 0$

(c) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.7$
Figure 4: continued ...

(i) $\hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.7; SD3, \kappa = 1, \lambda_0 = 0.3$

(j) $\hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.7; SD3, \kappa = 1, \lambda_0 = 0.3$

(k) $\hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.7; SD3, \kappa = 2, \lambda_0 = 0.3$

(l) $\hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.7; SD3, \kappa = 2, \lambda_0 = 0.3$

(m) $\hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.7; SD4, \kappa = 1, \lambda_0 = 0$

(n) $\hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.7; SD4, \kappa = 1, \lambda_0 = 0$

(o) $\hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.7; SD4, \kappa = 2, \lambda_0 = 0$

(p) $\hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.7; SD4, \kappa = 2, \lambda_0 = 0$
Figure 5: Simulated sampling density functions with $\delta_T = \delta T^{-1/2}$

(a) $\hat{\tau}_{OLS}: \delta = 0, SD1, \kappa = 2, \lambda_0 = 0.7$

(b) $\hat{\tau}_{FWLS}: \delta = 0, SD1, \kappa = 2, \lambda_0 = 0.7$

(c) $\hat{\tau}_{OLS}: \delta = 8, SD1, \kappa = 2, \lambda_0 = 0.7$

(d) $\hat{\tau}_{FWLS}: \delta = 8, SD1, \kappa = 2, \lambda_0 = 0.7$
Figure 6: Simulated RMSEs of $\hat{\tau}_{FWLS}$. $T = 200$, $\delta_T = 8T^{-1/2}$.
Figure 7: Finite Sample Local Power Comparisons, $T = 100$

(a) $\delta = 0; \kappa = 0$

(b) $\delta = 0; \kappa = 2, \lambda_0 = 0.3$

(c) $\delta = 0; \kappa = 2, \lambda_0 = 0.5$

(d) $\delta = 0; \kappa = 2, \lambda_0 = 0.7$

(e) $\delta = 0.5, \tau_0 = 0.3; \kappa = 0$

(f) $\delta = 0.5, \tau_0 = 0.3; \kappa = 2, \lambda_0 = 0.3$

Legend:
- $t_0$
- $t_{OLS}$
- $t_{FWLS}$
- $t_{WSC}$
- $t_{SC}$
- WSC
Figure 7: continued ...

(g) $\delta = 0.5$, $\tau_0 = 0.3$; $\kappa = 2$, $\lambda_0 = 0.5$

(h) $\delta = 0.5$, $\tau_0 = 0.3$; $\kappa = 2$, $\lambda_0 = 0.7$

(i) $\delta = 0.5$, $\tau_0 = 0.5$; $\kappa = 0$

(j) $\delta = 0.5$, $\tau_0 = 0.5$; $\kappa = 2$, $\lambda_0 = 0.3$

(k) $\delta = 0.5$, $\tau_0 = 0.5$; $\kappa = 2$, $\lambda_0 = 0.5$

(l) $\delta = 0.5$, $\tau_0 = 0.5$; $\kappa = 2$, $\lambda_0 = 0.7
Figure 7: continued ...

(m) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 0$

(n) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.3$

(o) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.5$

(p) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.7$
Figure 8: Finite Sample Local Power Comparisons, $T = 200$

(a) $\delta = 0; \kappa = 0$

(b) $\delta = 0; \kappa = 2, \lambda_0 = 0.3$

(c) $\delta = 0; \kappa = 2, \lambda_0 = 0.5$

(d) $\delta = 0; \kappa = 2, \lambda_0 = 0.7$

(e) $\delta = 0.5, \tau_0 = 0.3; \kappa = 0$

(f) $\delta = 0.5, \tau_0 = 0.3; \kappa = 2, \lambda_0 = 0.3$

Legend:
- $t_0$
- $t_{OLS}$
- $t_{FWLS}$
- $t_{WSC}$
- $t_{SC}$
- WSC
Figure 8: continued ...

(g) $\delta = 0.5, \tau_0 = 0.3; \kappa = 2, \lambda_0 = 0.5$

(h) $\delta = 0.5, \tau_0 = 0.3; \kappa = 2, \lambda_0 = 0.7$

(i) $\delta = 0.5, \tau_0 = 0.5; \kappa = 0$

(j) $\delta = 0.5, \tau_0 = 0.5; \kappa = 2, \lambda_0 = 0.3$

(k) $\delta = 0.5, \tau_0 = 0.5; \kappa = 2, \lambda_0 = 0.5$

(l) $\delta = 0.5, \tau_0 = 0.5; \kappa = 2, \lambda_0 = 0.7$

$t_0$, $t_{OLS}$, $t_{FWLS}$, $t_{WSC}$, SC, WSC
Figure 8: continued ...

(m) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 0$

(n) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.3$

(o) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.5$

(p) $\delta = 0.5$, $\tau_0 = 0.7$; $\kappa = 2$, $\lambda_0 = 0.7$
Supplementary Appendix

to

Level Shift Estimation in the Presence of Non-stationary Volatility with an Application to the Unit Root Testing Problem

by

D. Harris, H. Kew and A.M.R. Taylor
Table A.1: Finite sample properties of $\hat{\tau}_{OLS} : \delta = 0$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A.2: Finite sample properties of $\hat{\tau}_{\text{FWLS}} : \delta = 0$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD1</td>
<td>$\kappa = 1, \lambda_0 = 0.3$</td>
<td>$\kappa = 1, \lambda_0 = 0.5$</td>
<td>$\kappa = 1, \lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.532</td>
<td>0.277</td>
<td>0.521</td>
<td>0.271</td>
<td>0.538</td>
<td>0.265</td>
</tr>
<tr>
<td>200</td>
<td>0.511</td>
<td>0.283</td>
<td>0.502</td>
<td>0.273</td>
<td>0.520</td>
<td>0.264</td>
</tr>
<tr>
<td>300</td>
<td>0.502</td>
<td>0.286</td>
<td>0.490</td>
<td>0.273</td>
<td>0.511</td>
<td>0.264</td>
</tr>
<tr>
<td>400</td>
<td>0.495</td>
<td>0.290</td>
<td>0.486</td>
<td>0.275</td>
<td>0.507</td>
<td>0.266</td>
</tr>
<tr>
<td>$\kappa = 2, \lambda_0 = 0.3$</td>
<td>$\kappa = 2, \lambda_0 = 0.5$</td>
<td>$\kappa = 2, \lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.533</td>
<td>0.280</td>
<td>0.515</td>
<td>0.271</td>
<td>0.537</td>
<td>0.261</td>
</tr>
<tr>
<td>200</td>
<td>0.505</td>
<td>0.288</td>
<td>0.493</td>
<td>0.271</td>
<td>0.521</td>
<td>0.259</td>
</tr>
<tr>
<td>300</td>
<td>0.490</td>
<td>0.290</td>
<td>0.479</td>
<td>0.269</td>
<td>0.510</td>
<td>0.258</td>
</tr>
<tr>
<td>400</td>
<td>0.483</td>
<td>0.294</td>
<td>0.475</td>
<td>0.270</td>
<td>0.509</td>
<td>0.260</td>
</tr>
<tr>
<td>SD2</td>
<td>$\kappa = 1, \lambda_0 = 0.3$</td>
<td>$\kappa = 1, \lambda_0 = 0.5$</td>
<td>$\kappa = 1, \lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.478</td>
<td>0.266</td>
<td>0.498</td>
<td>0.271</td>
<td>0.491</td>
<td>0.278</td>
</tr>
<tr>
<td>200</td>
<td>0.490</td>
<td>0.264</td>
<td>0.512</td>
<td>0.273</td>
<td>0.498</td>
<td>0.284</td>
</tr>
<tr>
<td>300</td>
<td>0.491</td>
<td>0.264</td>
<td>0.515</td>
<td>0.272</td>
<td>0.506</td>
<td>0.286</td>
</tr>
<tr>
<td>400</td>
<td>0.497</td>
<td>0.265</td>
<td>0.517</td>
<td>0.272</td>
<td>0.507</td>
<td>0.288</td>
</tr>
<tr>
<td>$\kappa = 2, \lambda_0 = 0.3$</td>
<td>$\kappa = 2, \lambda_0 = 0.5$</td>
<td>$\kappa = 2, \lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.478</td>
<td>0.263</td>
<td>0.505</td>
<td>0.270</td>
<td>0.493</td>
<td>0.282</td>
</tr>
<tr>
<td>200</td>
<td>0.490</td>
<td>0.259</td>
<td>0.521</td>
<td>0.269</td>
<td>0.507</td>
<td>0.289</td>
</tr>
<tr>
<td>300</td>
<td>0.493</td>
<td>0.258</td>
<td>0.525</td>
<td>0.268</td>
<td>0.519</td>
<td>0.290</td>
</tr>
<tr>
<td>400</td>
<td>0.495</td>
<td>0.259</td>
<td>0.528</td>
<td>0.268</td>
<td>0.520</td>
<td>0.292</td>
</tr>
<tr>
<td>SD3</td>
<td>$\kappa = 1, \lambda_0 = 0.3$</td>
<td>$\kappa = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.509</td>
<td>0.266</td>
<td>0.544</td>
<td>0.270</td>
<td>0.510</td>
<td>0.272</td>
</tr>
<tr>
<td>200</td>
<td>0.505</td>
<td>0.259</td>
<td>0.529</td>
<td>0.273</td>
<td>0.504</td>
<td>0.273</td>
</tr>
<tr>
<td>300</td>
<td>0.502</td>
<td>0.257</td>
<td>0.521</td>
<td>0.275</td>
<td>0.504</td>
<td>0.275</td>
</tr>
<tr>
<td>400</td>
<td>0.504</td>
<td>0.257</td>
<td>0.516</td>
<td>0.278</td>
<td>0.503</td>
<td>0.278</td>
</tr>
<tr>
<td>SD4</td>
<td>$\kappa = 2, \lambda_0 = 0.3$</td>
<td>$\kappa = 2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.486</td>
<td>0.261</td>
<td>0.549</td>
<td>0.269</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.492</td>
<td>0.253</td>
<td>0.533</td>
<td>0.273</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.492</td>
<td>0.251</td>
<td>0.523</td>
<td>0.275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.494</td>
<td>0.250</td>
<td>0.518</td>
<td>0.277</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table A.3: Finite sample properties of $\hat{\tau}_{\text{OLS}}$ : $\delta > 0$, Homoskedasticity (SD0)

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.375</td>
<td>0.195</td>
<td>0.209</td>
<td>0.502</td>
<td>0.166</td>
<td>0.166</td>
<td>0.633</td>
<td>0.188</td>
<td>0.200</td>
</tr>
<tr>
<td>200</td>
<td>0.327</td>
<td>0.130</td>
<td>0.133</td>
<td>0.500</td>
<td>0.115</td>
<td>0.115</td>
<td>0.671</td>
<td>0.133</td>
<td>0.137</td>
</tr>
<tr>
<td>300</td>
<td>0.312</td>
<td>0.094</td>
<td>0.094</td>
<td>0.500</td>
<td>0.084</td>
<td>0.084</td>
<td>0.687</td>
<td>0.094</td>
<td>0.095</td>
</tr>
<tr>
<td>400</td>
<td>0.307</td>
<td>0.071</td>
<td>0.071</td>
<td>0.500</td>
<td>0.065</td>
<td>0.065</td>
<td>0.693</td>
<td>0.072</td>
<td>0.073</td>
</tr>
</tbody>
</table>

### Table A.4: Finite sample properties of $\hat{\tau}_{\text{FWLS}}$ : $\delta > 0$, Homoskedasticity (SD0)

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.376</td>
<td>0.196</td>
<td>0.210</td>
<td>0.504</td>
<td>0.168</td>
<td>0.168</td>
<td>0.630</td>
<td>0.192</td>
<td>0.204</td>
</tr>
<tr>
<td>200</td>
<td>0.328</td>
<td>0.133</td>
<td>0.136</td>
<td>0.500</td>
<td>0.116</td>
<td>0.116</td>
<td>0.671</td>
<td>0.135</td>
<td>0.138</td>
</tr>
<tr>
<td>300</td>
<td>0.313</td>
<td>0.094</td>
<td>0.095</td>
<td>0.499</td>
<td>0.084</td>
<td>0.084</td>
<td>0.686</td>
<td>0.096</td>
<td>0.097</td>
</tr>
<tr>
<td>400</td>
<td>0.308</td>
<td>0.072</td>
<td>0.072</td>
<td>0.500</td>
<td>0.065</td>
<td>0.065</td>
<td>0.692</td>
<td>0.074</td>
<td>0.075</td>
</tr>
</tbody>
</table>

|     | $\delta = 1$, $\tau_0 = 0.3$ | $\delta = 1$, $\tau_0 = 0.5$ | $\delta = 1$, $\tau_0 = 0.7$ |
| 100 | 0.309 | 0.074 | 0.074 | 0.500 | 0.064 | 0.064 | 0.694 | 0.067 | 0.067 |
| 200 | 0.301 | 0.031 | 0.031 | 0.500 | 0.030 | 0.030 | 0.699 | 0.031 | 0.031 |
| 300 | 0.300 | 0.019 | 0.019 | 0.500 | 0.019 | 0.019 | 0.700 | 0.020 | 0.020 |
| 400 | 0.300 | 0.014 | 0.014 | 0.500 | 0.014 | 0.014 | 0.700 | 0.014 | 0.014 |
Table A.5: Finite sample properties of $\hat{\tau}_{\text{OLS}}$: $\delta > 0$, Heteroskedasticity

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 1$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.545</td>
<td>0.226</td>
<td>0.334</td>
<td>0.589</td>
<td>0.190</td>
<td>0.210</td>
<td>0.656</td>
<td>0.183</td>
<td>0.188</td>
</tr>
<tr>
<td>200</td>
<td>0.480</td>
<td>0.214</td>
<td>0.279</td>
<td>0.557</td>
<td>0.169</td>
<td>0.178</td>
<td>0.665</td>
<td>0.164</td>
<td>0.168</td>
</tr>
<tr>
<td>300</td>
<td>0.439</td>
<td>0.192</td>
<td>0.237</td>
<td>0.541</td>
<td>0.150</td>
<td>0.156</td>
<td>0.670</td>
<td>0.147</td>
<td>0.150</td>
</tr>
<tr>
<td>400</td>
<td>0.412</td>
<td>0.173</td>
<td>0.206</td>
<td>0.530</td>
<td>0.134</td>
<td>0.138</td>
<td>0.677</td>
<td>0.134</td>
<td>0.136</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.614</td>
<td>0.209</td>
<td>0.378</td>
<td>0.631</td>
<td>0.188</td>
<td>0.229</td>
<td>0.669</td>
<td>0.178</td>
<td>0.181</td>
</tr>
<tr>
<td>200</td>
<td>0.572</td>
<td>0.215</td>
<td>0.347</td>
<td>0.607</td>
<td>0.183</td>
<td>0.212</td>
<td>0.670</td>
<td>0.171</td>
<td>0.174</td>
</tr>
<tr>
<td>300</td>
<td>0.536</td>
<td>0.213</td>
<td>0.318</td>
<td>0.585</td>
<td>0.174</td>
<td>0.193</td>
<td>0.670</td>
<td>0.161</td>
<td>0.164</td>
</tr>
<tr>
<td>400</td>
<td>0.508</td>
<td>0.209</td>
<td>0.295</td>
<td>0.573</td>
<td>0.167</td>
<td>0.182</td>
<td>0.675</td>
<td>0.155</td>
<td>0.157</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 1$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.577</td>
<td>0.241</td>
<td>0.367</td>
<td>0.635</td>
<td>0.178</td>
<td>0.223</td>
<td>0.693</td>
<td>0.162</td>
<td>0.162</td>
</tr>
<tr>
<td>200</td>
<td>0.497</td>
<td>0.239</td>
<td>0.310</td>
<td>0.604</td>
<td>0.157</td>
<td>0.188</td>
<td>0.698</td>
<td>0.139</td>
<td>0.139</td>
</tr>
<tr>
<td>300</td>
<td>0.442</td>
<td>0.220</td>
<td>0.262</td>
<td>0.583</td>
<td>0.136</td>
<td>0.159</td>
<td>0.698</td>
<td>0.120</td>
<td>0.120</td>
</tr>
<tr>
<td>400</td>
<td>0.407</td>
<td>0.199</td>
<td>0.226</td>
<td>0.569</td>
<td>0.120</td>
<td>0.139</td>
<td>0.698</td>
<td>0.109</td>
<td>0.109</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.678</td>
<td>0.193</td>
<td>0.425</td>
<td>0.694</td>
<td>0.155</td>
<td>0.248</td>
<td>0.722</td>
<td>0.140</td>
<td>0.142</td>
</tr>
<tr>
<td>200</td>
<td>0.636</td>
<td>0.217</td>
<td>0.400</td>
<td>0.670</td>
<td>0.153</td>
<td>0.228</td>
<td>0.720</td>
<td>0.131</td>
<td>0.132</td>
</tr>
<tr>
<td>300</td>
<td>0.590</td>
<td>0.230</td>
<td>0.370</td>
<td>0.646</td>
<td>0.146</td>
<td>0.207</td>
<td>0.715</td>
<td>0.122</td>
<td>0.123</td>
</tr>
<tr>
<td>400</td>
<td>0.556</td>
<td>0.234</td>
<td>0.347</td>
<td>0.631</td>
<td>0.142</td>
<td>0.193</td>
<td>0.713</td>
<td>0.117</td>
<td>0.118</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 1$, $\lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.543</td>
<td>0.235</td>
<td>0.338</td>
<td>0.626</td>
<td>0.141</td>
<td>0.189</td>
<td>0.709</td>
<td>0.118</td>
<td>0.119</td>
</tr>
<tr>
<td>200</td>
<td>0.490</td>
<td>0.248</td>
<td>0.287</td>
<td>0.584</td>
<td>0.115</td>
<td>0.143</td>
<td>0.704</td>
<td>0.101</td>
<td>0.101</td>
</tr>
<tr>
<td>300</td>
<td>0.431</td>
<td>0.228</td>
<td>0.237</td>
<td>0.563</td>
<td>0.097</td>
<td>0.116</td>
<td>0.702</td>
<td>0.086</td>
<td>0.086</td>
</tr>
<tr>
<td>400</td>
<td>0.341</td>
<td>0.164</td>
<td>0.091</td>
<td>0.550</td>
<td>0.082</td>
<td>0.096</td>
<td>0.701</td>
<td>0.078</td>
<td>0.078</td>
</tr>
</tbody>
</table>
Table A.6: Finite sample properties of $\hat{\delta}_{\text{OLS}}$: $\delta > 0$, Heteroskedasticity (ctd)

<table>
<thead>
<tr>
<th>T</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5, \tau_0 = 0.3$</td>
<td>$\delta = 0.5, \tau_0 = 0.5$</td>
<td>$\delta = 0.5, \tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.563</td>
<td>0.265</td>
<td>0.374</td>
<td>0.638</td>
<td>0.205</td>
<td>0.247</td>
<td>0.724</td>
<td>0.170</td>
<td>0.171</td>
</tr>
<tr>
<td>200</td>
<td>0.465</td>
<td>0.250</td>
<td>0.300</td>
<td>0.595</td>
<td>0.182</td>
<td>0.206</td>
<td>0.732</td>
<td>0.139</td>
<td>0.143</td>
</tr>
<tr>
<td>300</td>
<td>0.407</td>
<td>0.217</td>
<td>0.242</td>
<td>0.563</td>
<td>0.155</td>
<td>0.168</td>
<td>0.733</td>
<td>0.114</td>
<td>0.118</td>
</tr>
<tr>
<td>400</td>
<td>0.374</td>
<td>0.187</td>
<td>0.201</td>
<td>0.547</td>
<td>0.133</td>
<td>0.141</td>
<td>0.733</td>
<td>0.099</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1, \tau_0 = 0.3$</td>
<td>$\delta = 1, \tau_0 = 0.5$</td>
<td>$\delta = 1, \tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.370</td>
<td>0.182</td>
<td>0.195</td>
<td>0.544</td>
<td>0.131</td>
<td>0.139</td>
<td>0.731</td>
<td>0.095</td>
<td>0.100</td>
</tr>
<tr>
<td>200</td>
<td>0.312</td>
<td>0.081</td>
<td>0.082</td>
<td>0.512</td>
<td>0.069</td>
<td>0.070</td>
<td>0.726</td>
<td>0.062</td>
<td>0.067</td>
</tr>
<tr>
<td>300</td>
<td>0.303</td>
<td>0.038</td>
<td>0.038</td>
<td>0.503</td>
<td>0.038</td>
<td>0.039</td>
<td>0.721</td>
<td>0.044</td>
<td>0.048</td>
</tr>
<tr>
<td>400</td>
<td>0.301</td>
<td>0.022</td>
<td>0.022</td>
<td>0.501</td>
<td>0.022</td>
<td>0.022</td>
<td>0.717</td>
<td>0.035</td>
<td>0.039</td>
</tr>
</tbody>
</table>

|    | $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ |
| 100 | 0.695  | 0.226 | 0.455 | 0.729  | 0.172 | 0.286 | 0.774  | 0.127 | 0.147 |
| 200 | 0.626  | 0.257 | 0.415 | 0.695  | 0.178 | 0.264 | 0.773  | 0.110 | 0.132 |
| 300 | 0.566  | 0.264 | 0.375 | 0.663  | 0.177 | 0.241 | 0.768  | 0.101 | 0.122 |
| 400 | 0.518  | 0.260 | 0.340 | 0.640  | 0.172 | 0.222 | 0.765  | 0.094 | 0.115 |
|    | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$ |
| 100 | 0.503  | 0.257 | 0.327 | 0.626  | 0.171 | 0.212 | 0.761  | 0.095 | 0.113 |
| 200 | 0.380  | 0.191 | 0.207 | 0.561  | 0.133 | 0.147 | 0.750  | 0.076 | 0.091 |
| 300 | 0.334  | 0.130 | 0.135 | 0.531  | 0.098 | 0.103 | 0.742  | 0.061 | 0.074 |
| 400 | 0.315  | 0.088 | 0.090 | 0.516  | 0.073 | 0.075 | 0.736  | 0.054 | 0.065 |

|    | $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ |
| 100 | 0.288  | 0.181 | 0.181 | 0.375  | 0.206 | 0.241 | 0.451  | 0.265 | 0.364 |
| 200 | 0.270  | 0.140 | 0.143 | 0.410  | 0.180 | 0.201 | 0.534  | 0.248 | 0.299 |
| 300 | 0.266  | 0.115 | 0.120 | 0.436  | 0.156 | 0.169 | 0.590  | 0.218 | 0.244 |
| 400 | 0.266  | 0.096 | 0.102 | 0.452  | 0.134 | 0.142 | 0.627  | 0.186 | 0.200 |
|    | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$ |
| 100 | 0.274  | 0.104 | 0.107 | 0.465  | 0.126 | 0.131 | 0.638  | 0.172 | 0.183 |
| 200 | 0.273  | 0.060 | 0.066 | 0.489  | 0.068 | 0.069 | 0.687  | 0.082 | 0.083 |
| 300 | 0.278  | 0.045 | 0.050 | 0.497  | 0.037 | 0.037 | 0.698  | 0.038 | 0.038 |
| 400 | 0.282  | 0.036 | 0.040 | 0.499  | 0.023 | 0.023 | 0.699  | 0.024 | 0.024 |

<p>|    | $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ |
| 100 | 0.233  | 0.132 | 0.148 | 0.277  | 0.170 | 0.281 | 0.314  | 0.228 | 0.448 |
| 200 | 0.228  | 0.110 | 0.132 | 0.306  | 0.176 | 0.262 | 0.376  | 0.254 | 0.412 |
| 300 | 0.230  | 0.101 | 0.123 | 0.337  | 0.177 | 0.241 | 0.437  | 0.264 | 0.372 |
| 400 | 0.234  | 0.093 | 0.114 | 0.362  | 0.172 | 0.221 | 0.486  | 0.260 | 0.337 |
|    | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$ |
| 100 | 0.244  | 0.104 | 0.118 | 0.384  | 0.168 | 0.205 | 0.510  | 0.252 | 0.316 |
| 200 | 0.249  | 0.073 | 0.089 | 0.441  | 0.131 | 0.144 | 0.619  | 0.191 | 0.207 |
| 300 | 0.256  | 0.062 | 0.076 | 0.469  | 0.098 | 0.103 | 0.664  | 0.134 | 0.139 |
| 400 | 0.262  | 0.055 | 0.067 | 0.484  | 0.073 | 0.074 | 0.684  | 0.092 | 0.093 |</p>
<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SD2 : $\kappa = 1$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td>SD2 : $\kappa = 2$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.314 0.165 0.166</td>
<td>0.369 0.178 0.221</td>
<td>0.431 0.240 0.361</td>
<td>0.302 0.136 0.136</td>
<td>0.395 0.155 0.188</td>
<td>0.501 0.239 0.311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.302 0.136 0.136</td>
<td>0.415 0.137 0.161</td>
<td>0.555 0.221 0.264</td>
<td>0.301 0.121 0.121</td>
<td>0.427 0.122 0.142</td>
<td>0.592 0.200 0.227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.300 0.107 0.107</td>
<td>0.435 0.121 0.137</td>
<td>0.604 0.190 0.213</td>
<td>0.300 0.060 0.060</td>
<td>0.472 0.059 0.066</td>
<td>0.690 0.065 0.065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.301 0.051 0.051</td>
<td>0.480 0.045 0.049</td>
<td>0.697 0.040 0.040</td>
<td>0.286 0.115 0.116</td>
<td>0.365 0.140 0.194</td>
<td>0.444 0.235 0.348</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.297 0.115 0.115</td>
<td>0.373 0.139 0.188</td>
<td>0.467 0.234 0.330</td>
<td>0.292 0.115 0.115</td>
<td>0.373 0.139 0.188</td>
<td>0.467 0.234 0.330</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.295 0.097 0.097</td>
<td>0.411 0.117 0.147</td>
<td>0.569 0.208 0.245</td>
<td>0.295 0.097 0.097</td>
<td>0.411 0.117 0.147</td>
<td>0.569 0.208 0.245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.296 0.087 0.087</td>
<td>0.434 0.098 0.118</td>
<td>0.627 0.168 0.183</td>
<td>0.296 0.087 0.087</td>
<td>0.434 0.098 0.118</td>
<td>0.627 0.168 0.183</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.299 0.078 0.078</td>
<td>0.448 0.083 0.099</td>
<td>0.658 0.132 0.139</td>
<td>0.298 0.078 0.078</td>
<td>0.448 0.083 0.099</td>
<td>0.658 0.132 0.139</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SD2 : $\kappa = 1$, $\lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td>SD2 : $\kappa = 2$, $\lambda_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.350 0.183 0.190</td>
<td>0.414 0.188 0.207</td>
<td>0.458 0.223 0.329</td>
<td>0.327 0.133 0.136</td>
<td>0.471 0.132 0.135</td>
<td>0.590 0.170 0.202</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.334 0.160 0.163</td>
<td>0.439 0.167 0.178</td>
<td>0.516 0.213 0.281</td>
<td>0.311 0.095 0.096</td>
<td>0.489 0.093 0.094</td>
<td>0.642 0.115 0.129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.327 0.147 0.149</td>
<td>0.459 0.150 0.155</td>
<td>0.556 0.194 0.242</td>
<td>0.305 0.072 0.073</td>
<td>0.495 0.070 0.070</td>
<td>0.666 0.079 0.086</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.322 0.134 0.136</td>
<td>0.468 0.135 0.138</td>
<td>0.585 0.175 0.210</td>
<td>0.304 0.060 0.060</td>
<td>0.498 0.056 0.056</td>
<td>0.677 0.056 0.061</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.327 0.133 0.136</td>
<td>0.471 0.132 0.135</td>
<td>0.590 0.170 0.202</td>
<td>0.327 0.133 0.136</td>
<td>0.471 0.132 0.135</td>
<td>0.590 0.170 0.202</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>0.311 0.095 0.096</td>
<td>0.489 0.093 0.094</td>
<td>0.642 0.115 0.129</td>
<td>0.311 0.095 0.096</td>
<td>0.489 0.093 0.094</td>
<td>0.642 0.115 0.129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.305 0.072 0.073</td>
<td>0.495 0.070 0.070</td>
<td>0.666 0.079 0.086</td>
<td>0.305 0.072 0.073</td>
<td>0.495 0.070 0.070</td>
<td>0.666 0.079 0.086</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td>0.304 0.060 0.060</td>
<td>0.498 0.056 0.056</td>
<td>0.677 0.056 0.061</td>
<td>0.304 0.060 0.060</td>
<td>0.498 0.056 0.056</td>
<td>0.677 0.056 0.061</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A.7: Finite sample properties of $\hat{\delta}_{OLS}$, $\delta > 0$, Heteroskedasticity (ctd)
Table A.8: Finite sample properties of \( \hat{\delta} \) : \( \delta > 0 \), Heteroskedasticity (ctd)

<table>
<thead>
<tr>
<th>( T )</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SD3 : ( \kappa = 1, \lambda_0 = 0.3 )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta = 0.5, \tau_0 = 0.3 )</td>
<td>0.448</td>
<td>0.288</td>
<td>0.324</td>
<td>0.511</td>
<td>0.272</td>
<td>0.272</td>
<td>0.575</td>
<td>0.283</td>
<td>0.310</td>
</tr>
<tr>
<td>( 100 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 200 )</td>
<td>0.386</td>
<td>0.259</td>
<td>0.273</td>
<td>0.503</td>
<td>0.243</td>
<td>0.243</td>
<td>0.616</td>
<td>0.257</td>
<td>0.270</td>
</tr>
<tr>
<td>( 300 )</td>
<td>0.352</td>
<td>0.227</td>
<td>0.233</td>
<td>0.500</td>
<td>0.212</td>
<td>0.212</td>
<td>0.647</td>
<td>0.228</td>
<td>0.234</td>
</tr>
<tr>
<td>( 400 )</td>
<td>0.331</td>
<td>0.203</td>
<td>0.206</td>
<td>0.497</td>
<td>0.188</td>
<td>0.188</td>
<td>0.668</td>
<td>0.203</td>
<td>0.205</td>
</tr>
<tr>
<td>( \delta = 1, \tau_0 = 0.3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 100 )</td>
<td>0.339</td>
<td>0.203</td>
<td>0.207</td>
<td>0.507</td>
<td>0.177</td>
<td>0.177</td>
<td>0.675</td>
<td>0.190</td>
<td>0.192</td>
</tr>
<tr>
<td>( 200 )</td>
<td>0.291</td>
<td>0.114</td>
<td>0.114</td>
<td>0.501</td>
<td>0.101</td>
<td>0.101</td>
<td>0.709</td>
<td>0.114</td>
<td>0.114</td>
</tr>
<tr>
<td>( 300 )</td>
<td>0.283</td>
<td>0.068</td>
<td>0.070</td>
<td>0.500</td>
<td>0.058</td>
<td>0.058</td>
<td>0.716</td>
<td>0.067</td>
<td>0.069</td>
</tr>
<tr>
<td>( 400 )</td>
<td>0.284</td>
<td>0.048</td>
<td>0.050</td>
<td>0.499</td>
<td>0.035</td>
<td>0.035</td>
<td>0.715</td>
<td>0.047</td>
<td>0.049</td>
</tr>
<tr>
<td><strong>SD4 : ( \kappa = 2, \lambda_0 = 0.3 )</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta = 0.5, \tau_0 = 0.3 )</td>
<td>0.339</td>
<td>0.257</td>
<td>0.260</td>
<td>0.379</td>
<td>0.267</td>
<td>0.293</td>
<td>0.427</td>
<td>0.293</td>
<td>0.401</td>
</tr>
<tr>
<td>( 100 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 200 )</td>
<td>0.311</td>
<td>0.230</td>
<td>0.231</td>
<td>0.385</td>
<td>0.251</td>
<td>0.276</td>
<td>0.470</td>
<td>0.294</td>
<td>0.373</td>
</tr>
<tr>
<td>( 300 )</td>
<td>0.294</td>
<td>0.207</td>
<td>0.207</td>
<td>0.395</td>
<td>0.235</td>
<td>0.257</td>
<td>0.512</td>
<td>0.287</td>
<td>0.343</td>
</tr>
<tr>
<td>( 400 )</td>
<td>0.287</td>
<td>0.190</td>
<td>0.190</td>
<td>0.410</td>
<td>0.222</td>
<td>0.240</td>
<td>0.545</td>
<td>0.276</td>
<td>0.317</td>
</tr>
<tr>
<td>( \delta = 1, \tau_0 = 0.3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( 100 )</td>
<td>0.300</td>
<td>0.198</td>
<td>0.198</td>
<td>0.427</td>
<td>0.214</td>
<td>0.226</td>
<td>0.563</td>
<td>0.266</td>
<td>0.299</td>
</tr>
<tr>
<td>( 200 )</td>
<td>0.271</td>
<td>0.132</td>
<td>0.135</td>
<td>0.454</td>
<td>0.156</td>
<td>0.163</td>
<td>0.644</td>
<td>0.205</td>
<td>0.213</td>
</tr>
<tr>
<td>( 300 )</td>
<td>0.265</td>
<td>0.093</td>
<td>0.099</td>
<td>0.473</td>
<td>0.112</td>
<td>0.115</td>
<td>0.680</td>
<td>0.150</td>
<td>0.152</td>
</tr>
<tr>
<td>( 400 )</td>
<td>0.266</td>
<td>0.072</td>
<td>0.080</td>
<td>0.484</td>
<td>0.081</td>
<td>0.082</td>
<td>0.697</td>
<td>0.108</td>
<td>0.108</td>
</tr>
</tbody>
</table>

**SD4 : \( \kappa = 1 \)**

| \( \delta = 0.5, \tau_0 = 0.3 \) | 0.527 | 0.252 | 0.339 | 0.592 | 0.204 | 0.224 | 0.670 | 0.193 | 0.195 |
| \( 100 \) |      |     |      |       |     |      |       |     |      |
| \( 200 \) | 0.448 | 0.231 | 0.275 | 0.561 | 0.175 | 0.186 | 0.680 | 0.166 | 0.167 |
| \( 300 \) | 0.403 | 0.203 | 0.227 | 0.543 | 0.150 | 0.156 | 0.689 | 0.140 | 0.140 |
| \( 400 \) | 0.374 | 0.175 | 0.190 | 0.534 | 0.129 | 0.133 | 0.693 | 0.124 | 0.124 |
| \( \delta = 1, \tau_0 = 0.3 \) |      |     |      |       |     |      |       |     |      |
| \( 100 \) | 0.377 | 0.178 | 0.194 | 0.532 | 0.130 | 0.134 | 0.691 | 0.121 | 0.122 |
| \( 200 \) | 0.321 | 0.094 | 0.096 | 0.513 | 0.078 | 0.079 | 0.699 | 0.080 | 0.080 |
| \( 300 \) | 0.307 | 0.053 | 0.054 | 0.505 | 0.052 | 0.052 | 0.701 | 0.057 | 0.057 |
| \( 400 \) | 0.304 | 0.035 | 0.035 | 0.503 | 0.038 | 0.038 | 0.701 | 0.043 | 0.043 |

**SD4 : \( \kappa = 2 \)**

| \( \delta = 0.5, \tau_0 = 0.3 \) | 0.622 | 0.240 | 0.402 | 0.653 | 0.200 | 0.252 | 0.702 | 0.181 | 0.181 |
| \( 100 \) |      |     |      |       |     |      |       |     |      |
| \( 200 \) | 0.563 | 0.249 | 0.363 | 0.626 | 0.192 | 0.230 | 0.703 | 0.168 | 0.168 |
| \( 300 \) | 0.514 | 0.246 | 0.326 | 0.601 | 0.180 | 0.206 | 0.704 | 0.153 | 0.153 |
| \( 400 \) | 0.482 | 0.238 | 0.299 | 0.588 | 0.168 | 0.190 | 0.706 | 0.143 | 0.143 |
| \( \delta = 1, \tau_0 = 0.3 \) |      |     |      |       |     |      |       |     |      |
| \( 100 \) | 0.474 | 0.235 | 0.292 | 0.581 | 0.168 | 0.187 | 0.701 | 0.143 | 0.143 |
| \( 200 \) | 0.381 | 0.178 | 0.196 | 0.542 | 0.126 | 0.132 | 0.704 | 0.111 | 0.111 |
| \( 300 \) | 0.343 | 0.132 | 0.138 | 0.525 | 0.097 | 0.100 | 0.705 | 0.088 | 0.088 |
| \( 400 \) | 0.325 | 0.098 | 0.101 | 0.517 | 0.076 | 0.078 | 0.706 | 0.073 | 0.073 |
Table A.9: Finite sample properties of $\hat{\tau}_{FWLS}$: $\delta > 0$, Heteroskedasticity

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD1 : $\kappa = 1$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.458</td>
<td>0.243</td>
<td>0.290</td>
<td>0.522</td>
<td>0.229</td>
<td>0.230</td>
<td>0.576</td>
<td>0.248</td>
<td>0.277</td>
</tr>
<tr>
<td>200</td>
<td>0.398</td>
<td>0.207</td>
<td>0.229</td>
<td>0.503</td>
<td>0.201</td>
<td>0.201</td>
<td>0.598</td>
<td>0.230</td>
<td>0.251</td>
</tr>
<tr>
<td>300</td>
<td>0.375</td>
<td>0.177</td>
<td>0.192</td>
<td>0.504</td>
<td>0.175</td>
<td>0.175</td>
<td>0.620</td>
<td>0.209</td>
<td>0.223</td>
</tr>
<tr>
<td>400</td>
<td>0.355</td>
<td>0.150</td>
<td>0.160</td>
<td>0.501</td>
<td>0.154</td>
<td>0.154</td>
<td>0.638</td>
<td>0.189</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.370</td>
<td>0.158</td>
<td>0.173</td>
<td>0.510</td>
<td>0.149</td>
<td>0.149</td>
<td>0.647</td>
<td>0.175</td>
<td>0.183</td>
</tr>
<tr>
<td>200</td>
<td>0.331</td>
<td>0.091</td>
<td>0.096</td>
<td>0.502</td>
<td>0.100</td>
<td>0.100</td>
<td>0.677</td>
<td>0.123</td>
<td>0.125</td>
</tr>
<tr>
<td>300</td>
<td>0.319</td>
<td>0.059</td>
<td>0.061</td>
<td>0.499</td>
<td>0.074</td>
<td>0.074</td>
<td>0.687</td>
<td>0.089</td>
<td>0.090</td>
</tr>
<tr>
<td>400</td>
<td>0.314</td>
<td>0.041</td>
<td>0.044</td>
<td>0.500</td>
<td>0.058</td>
<td>0.058</td>
<td>0.693</td>
<td>0.069</td>
<td>0.069</td>
</tr>
<tr>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.491</td>
<td>0.260</td>
<td>0.323</td>
<td>0.525</td>
<td>0.254</td>
<td>0.255</td>
<td>0.555</td>
<td>0.264</td>
<td>0.302</td>
</tr>
<tr>
<td>200</td>
<td>0.438</td>
<td>0.240</td>
<td>0.276</td>
<td>0.504</td>
<td>0.239</td>
<td>0.239</td>
<td>0.556</td>
<td>0.263</td>
<td>0.300</td>
</tr>
<tr>
<td>300</td>
<td>0.414</td>
<td>0.219</td>
<td>0.247</td>
<td>0.503</td>
<td>0.221</td>
<td>0.221</td>
<td>0.568</td>
<td>0.253</td>
<td>0.285</td>
</tr>
<tr>
<td>400</td>
<td>0.395</td>
<td>0.200</td>
<td>0.221</td>
<td>0.499</td>
<td>0.207</td>
<td>0.207</td>
<td>0.583</td>
<td>0.244</td>
<td>0.271</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.420</td>
<td>0.205</td>
<td>0.238</td>
<td>0.517</td>
<td>0.196</td>
<td>0.197</td>
<td>0.609</td>
<td>0.222</td>
<td>0.240</td>
</tr>
<tr>
<td>200</td>
<td>0.367</td>
<td>0.147</td>
<td>0.161</td>
<td>0.506</td>
<td>0.153</td>
<td>0.153</td>
<td>0.639</td>
<td>0.189</td>
<td>0.198</td>
</tr>
<tr>
<td>300</td>
<td>0.347</td>
<td>0.109</td>
<td>0.119</td>
<td>0.505</td>
<td>0.124</td>
<td>0.124</td>
<td>0.657</td>
<td>0.159</td>
<td>0.164</td>
</tr>
<tr>
<td>400</td>
<td>0.336</td>
<td>0.086</td>
<td>0.093</td>
<td>0.502</td>
<td>0.103</td>
<td>0.103</td>
<td>0.672</td>
<td>0.133</td>
<td>0.136</td>
</tr>
<tr>
<td>SD1 : $\kappa = 1$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.419</td>
<td>0.229</td>
<td>0.258</td>
<td>0.532</td>
<td>0.219</td>
<td>0.221</td>
<td>0.578</td>
<td>0.242</td>
<td>0.271</td>
</tr>
<tr>
<td>200</td>
<td>0.348</td>
<td>0.167</td>
<td>0.173</td>
<td>0.524</td>
<td>0.183</td>
<td>0.185</td>
<td>0.600</td>
<td>0.222</td>
<td>0.243</td>
</tr>
<tr>
<td>300</td>
<td>0.323</td>
<td>0.123</td>
<td>0.125</td>
<td>0.524</td>
<td>0.153</td>
<td>0.155</td>
<td>0.626</td>
<td>0.197</td>
<td>0.211</td>
</tr>
<tr>
<td>400</td>
<td>0.311</td>
<td>0.091</td>
<td>0.092</td>
<td>0.524</td>
<td>0.131</td>
<td>0.133</td>
<td>0.643</td>
<td>0.178</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.327</td>
<td>0.122</td>
<td>0.124</td>
<td>0.535</td>
<td>0.131</td>
<td>0.135</td>
<td>0.656</td>
<td>0.162</td>
<td>0.168</td>
</tr>
<tr>
<td>200</td>
<td>0.302</td>
<td>0.043</td>
<td>0.043</td>
<td>0.523</td>
<td>0.080</td>
<td>0.083</td>
<td>0.683</td>
<td>0.108</td>
<td>0.109</td>
</tr>
<tr>
<td>300</td>
<td>0.300</td>
<td>0.022</td>
<td>0.022</td>
<td>0.516</td>
<td>0.053</td>
<td>0.056</td>
<td>0.691</td>
<td>0.077</td>
<td>0.078</td>
</tr>
<tr>
<td>400</td>
<td>0.300</td>
<td>0.015</td>
<td>0.015</td>
<td>0.513</td>
<td>0.040</td>
<td>0.042</td>
<td>0.695</td>
<td>0.059</td>
<td>0.059</td>
</tr>
<tr>
<td>SD1 : $\kappa = 2$, $\lambda_0 = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.430</td>
<td>0.237</td>
<td>0.270</td>
<td>0.529</td>
<td>0.242</td>
<td>0.244</td>
<td>0.549</td>
<td>0.257</td>
<td>0.298</td>
</tr>
<tr>
<td>200</td>
<td>0.353</td>
<td>0.176</td>
<td>0.184</td>
<td>0.518</td>
<td>0.221</td>
<td>0.222</td>
<td>0.553</td>
<td>0.250</td>
<td>0.290</td>
</tr>
<tr>
<td>300</td>
<td>0.326</td>
<td>0.131</td>
<td>0.134</td>
<td>0.521</td>
<td>0.200</td>
<td>0.201</td>
<td>0.569</td>
<td>0.239</td>
<td>0.273</td>
</tr>
<tr>
<td>400</td>
<td>0.313</td>
<td>0.101</td>
<td>0.102</td>
<td>0.524</td>
<td>0.183</td>
<td>0.185</td>
<td>0.585</td>
<td>0.229</td>
<td>0.257</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1$, $\tau_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.335</td>
<td>0.140</td>
<td>0.144</td>
<td>0.545</td>
<td>0.179</td>
<td>0.184</td>
<td>0.611</td>
<td>0.214</td>
<td>0.232</td>
</tr>
<tr>
<td>200</td>
<td>0.302</td>
<td>0.049</td>
<td>0.049</td>
<td>0.538</td>
<td>0.128</td>
<td>0.134</td>
<td>0.646</td>
<td>0.177</td>
<td>0.185</td>
</tr>
<tr>
<td>300</td>
<td>0.300</td>
<td>0.023</td>
<td>0.023</td>
<td>0.534</td>
<td>0.097</td>
<td>0.103</td>
<td>0.668</td>
<td>0.140</td>
<td>0.144</td>
</tr>
<tr>
<td>400</td>
<td>0.300</td>
<td>0.016</td>
<td>0.016</td>
<td>0.529</td>
<td>0.076</td>
<td>0.081</td>
<td>0.681</td>
<td>0.115</td>
<td>0.116</td>
</tr>
</tbody>
</table>
Table A.10: Finite sample properties of $\hat{\tau}_{FWLS}$ : $\delta > 0$, Heteroskedasticity (ctd)

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.407</td>
<td>0.218</td>
<td>0.243</td>
<td>0.521</td>
<td>0.191</td>
<td>0.192</td>
<td>0.601</td>
<td>0.235</td>
<td>0.255</td>
</tr>
<tr>
<td>200</td>
<td>0.340</td>
<td>0.151</td>
<td>0.156</td>
<td>0.502</td>
<td>0.138</td>
<td>0.138</td>
<td>0.627</td>
<td>0.210</td>
<td>0.222</td>
</tr>
<tr>
<td>300</td>
<td>0.316</td>
<td>0.104</td>
<td>0.106</td>
<td>0.499</td>
<td>0.100</td>
<td>0.100</td>
<td>0.653</td>
<td>0.183</td>
<td>0.189</td>
</tr>
<tr>
<td>400</td>
<td>0.310</td>
<td>0.079</td>
<td>0.079</td>
<td>0.499</td>
<td>0.077</td>
<td>0.077</td>
<td>0.670</td>
<td>0.160</td>
<td>0.163</td>
</tr>
</tbody>
</table>

**SD1 : $\kappa = 1$, $\lambda_0 = 0.7$**

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.318</td>
<td>0.101</td>
<td>0.103</td>
<td>0.504</td>
<td>0.089</td>
<td>0.089</td>
<td>0.684</td>
<td>0.146</td>
<td>0.147</td>
</tr>
<tr>
<td>200</td>
<td>0.302</td>
<td>0.036</td>
<td>0.036</td>
<td>0.500</td>
<td>0.037</td>
<td>0.037</td>
<td>0.706</td>
<td>0.086</td>
<td>0.086</td>
</tr>
<tr>
<td>300</td>
<td>0.300</td>
<td>0.019</td>
<td>0.019</td>
<td>0.499</td>
<td>0.020</td>
<td>0.020</td>
<td>0.710</td>
<td>0.056</td>
<td>0.056</td>
</tr>
<tr>
<td>400</td>
<td>0.300</td>
<td>0.015</td>
<td>0.015</td>
<td>0.500</td>
<td>0.015</td>
<td>0.015</td>
<td>0.710</td>
<td>0.039</td>
<td>0.040</td>
</tr>
</tbody>
</table>

**SD1 : $\kappa = 2$, $\lambda_0 = 0.7$**

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.320</td>
<td>0.105</td>
<td>0.107</td>
<td>0.503</td>
<td>0.101</td>
<td>0.101</td>
<td>0.642</td>
<td>0.204</td>
<td>0.212</td>
</tr>
<tr>
<td>200</td>
<td>0.301</td>
<td>0.036</td>
<td>0.036</td>
<td>0.498</td>
<td>0.042</td>
<td>0.042</td>
<td>0.680</td>
<td>0.160</td>
<td>0.161</td>
</tr>
<tr>
<td>300</td>
<td>0.300</td>
<td>0.019</td>
<td>0.019</td>
<td>0.499</td>
<td>0.021</td>
<td>0.021</td>
<td>0.700</td>
<td>0.117</td>
<td>0.117</td>
</tr>
<tr>
<td>400</td>
<td>0.300</td>
<td>0.014</td>
<td>0.014</td>
<td>0.499</td>
<td>0.015</td>
<td>0.015</td>
<td>0.708</td>
<td>0.092</td>
<td>0.093</td>
</tr>
</tbody>
</table>

**SD2 : $\kappa = 1$, $\lambda_0 = 0.3$**

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.410</td>
<td>0.238</td>
<td>0.262</td>
<td>0.490</td>
<td>0.189</td>
<td>0.190</td>
<td>0.597</td>
<td>0.215</td>
<td>0.239</td>
</tr>
<tr>
<td>200</td>
<td>0.371</td>
<td>0.208</td>
<td>0.220</td>
<td>0.498</td>
<td>0.137</td>
<td>0.137</td>
<td>0.658</td>
<td>0.152</td>
<td>0.158</td>
</tr>
<tr>
<td>300</td>
<td>0.346</td>
<td>0.183</td>
<td>0.188</td>
<td>0.501</td>
<td>0.100</td>
<td>0.100</td>
<td>0.681</td>
<td>0.107</td>
<td>0.109</td>
</tr>
<tr>
<td>400</td>
<td>0.328</td>
<td>0.160</td>
<td>0.162</td>
<td>0.502</td>
<td>0.077</td>
<td>0.077</td>
<td>0.690</td>
<td>0.080</td>
<td>0.080</td>
</tr>
</tbody>
</table>

**SD2 : $\kappa = 2$, $\lambda_0 = 0.3$**

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.440</td>
<td>0.252</td>
<td>0.288</td>
<td>0.489</td>
<td>0.195</td>
<td>0.195</td>
<td>0.592</td>
<td>0.215</td>
<td>0.241</td>
</tr>
<tr>
<td>200</td>
<td>0.417</td>
<td>0.239</td>
<td>0.266</td>
<td>0.500</td>
<td>0.144</td>
<td>0.144</td>
<td>0.657</td>
<td>0.153</td>
<td>0.159</td>
</tr>
<tr>
<td>300</td>
<td>0.399</td>
<td>0.227</td>
<td>0.247</td>
<td>0.503</td>
<td>0.107</td>
<td>0.108</td>
<td>0.680</td>
<td>0.108</td>
<td>0.110</td>
</tr>
<tr>
<td>400</td>
<td>0.378</td>
<td>0.215</td>
<td>0.229</td>
<td>0.505</td>
<td>0.083</td>
<td>0.083</td>
<td>0.691</td>
<td>0.080</td>
<td>0.081</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.364</td>
<td>0.208</td>
<td>0.218</td>
<td>0.499</td>
<td>0.097</td>
<td>0.097</td>
<td>0.684</td>
<td>0.096</td>
<td>0.097</td>
</tr>
<tr>
<td>200</td>
<td>0.318</td>
<td>0.156</td>
<td>0.157</td>
<td>0.502</td>
<td>0.042</td>
<td>0.042</td>
<td>0.698</td>
<td>0.037</td>
<td>0.037</td>
</tr>
<tr>
<td>300</td>
<td>0.299</td>
<td>0.120</td>
<td>0.120</td>
<td>0.501</td>
<td>0.023</td>
<td>0.023</td>
<td>0.700</td>
<td>0.021</td>
<td>0.021</td>
</tr>
<tr>
<td>400</td>
<td>0.290</td>
<td>0.089</td>
<td>0.090</td>
<td>0.500</td>
<td>0.015</td>
<td>0.015</td>
<td>0.700</td>
<td>0.014</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Table A.11: Finite sample properties of $\hat{\delta}_{FWLS}$ : $\delta > 0$, Heteroskedasticity (ctd)

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
</table>
|     |  $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$
| 100 | 0.433  | 0.244  | 0.277  | 0.477  | 0.218  | 0.219  | 0.591  | 0.224  | 0.249  |
| 200 | 0.397  | 0.221  | 0.242  | 0.474  | 0.183  | 0.185  | 0.652  | 0.167  | 0.174  |
| 300 | 0.374  | 0.198  | 0.211  | 0.477  | 0.155  | 0.157  | 0.676  | 0.123  | 0.125  |
| 400 | 0.355  | 0.178  | 0.186  | 0.475  | 0.133  | 0.135  | 0.688  | 0.094  | 0.095  |

SD2 : $\kappa = 1, \lambda_0 = 0.5$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
</table>
|     |  $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$
| 100 | 0.439  | 0.166  | 0.173  | 0.470  | 0.131  | 0.134  | 0.678  | 0.113  | 0.115  |
| 200 | 0.315  | 0.076  | 0.076  | 0.483  | 0.054  | 0.057  | 0.700  | 0.022  | 0.022  |
| 300 | 0.305  | 0.060  | 0.061  | 0.487  | 0.041  | 0.043  | 0.700  | 0.016  | 0.016  |

SD2 : $\kappa = 2, \lambda_0 = 0.5$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
</table>
|     |  $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$
| 100 | 0.468  | 0.258  | 0.308  | 0.487  | 0.241  | 0.241  | 0.586  | 0.229  | 0.256  |
| 200 | 0.449  | 0.249  | 0.290  | 0.482  | 0.219  | 0.220  | 0.647  | 0.176  | 0.183  |
| 300 | 0.432  | 0.239  | 0.273  | 0.479  | 0.200  | 0.201  | 0.671  | 0.135  | 0.138  |
| 400 | 0.414  | 0.229  | 0.256  | 0.473  | 0.183  | 0.185  | 0.686  | 0.104  | 0.105  |

SD2 : $\kappa = 1, \lambda_0 = 0.7$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
</table>
|     |  $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$
| 100 | 0.434  | 0.251  | 0.285  | 0.489  | 0.230  | 0.230  | 0.554  | 0.240  | 0.281  |
| 200 | 0.401  | 0.230  | 0.251  | 0.493  | 0.201  | 0.201  | 0.599  | 0.207  | 0.230  |
| 300 | 0.379  | 0.209  | 0.224  | 0.497  | 0.175  | 0.175  | 0.624  | 0.178  | 0.194  |
| 400 | 0.361  | 0.190  | 0.199  | 0.496  | 0.155  | 0.155  | 0.642  | 0.153  | 0.163  |

SD2 : $\kappa = 2, \lambda_0 = 0.7$

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
</table>
|     |  $\delta = 0.5, \tau_0 = 0.3$ | $\delta = 0.5, \tau_0 = 0.5$ | $\delta = 0.5, \tau_0 = 0.7$ | $\delta = 1, \tau_0 = 0.3$ | $\delta = 1, \tau_0 = 0.5$ | $\delta = 1, \tau_0 = 0.7$
| 100 | 0.463  | 0.268  | 0.314  | 0.490  | 0.254  | 0.254  | 0.526  | 0.258  | 0.311  |
| 200 | 0.447  | 0.263  | 0.302  | 0.494  | 0.239  | 0.239  | 0.563  | 0.240  | 0.277  |
| 300 | 0.429  | 0.252  | 0.283  | 0.500  | 0.221  | 0.221  | 0.584  | 0.219  | 0.248  |
| 400 | 0.417  | 0.244  | 0.270  | 0.498  | 0.208  | 0.208  | 0.599  | 0.203  | 0.226  |
Table A.12: Finite sample properties of $\hat{\tau}_{\text{FWLS}}$ : $\delta > 0$, Heteroskedasticity (ctd)

<table>
<thead>
<tr>
<th>$T$</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD3 : $\kappa = 1$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.441</td>
<td>0.249</td>
<td>0.286</td>
<td>0.506</td>
<td>0.208</td>
<td>0.208</td>
<td>0.570</td>
<td>0.247</td>
<td>0.279</td>
</tr>
<tr>
<td>200</td>
<td>0.388</td>
<td>0.218</td>
<td>0.235</td>
<td>0.500</td>
<td>0.156</td>
<td>0.156</td>
<td>0.612</td>
<td>0.218</td>
<td>0.235</td>
</tr>
<tr>
<td>300</td>
<td>0.357</td>
<td>0.190</td>
<td>0.198</td>
<td>0.500</td>
<td>0.115</td>
<td>0.115</td>
<td>0.642</td>
<td>0.191</td>
<td>0.200</td>
</tr>
<tr>
<td>400</td>
<td>0.336</td>
<td>0.168</td>
<td>0.172</td>
<td>0.500</td>
<td>0.090</td>
<td>0.090</td>
<td>0.661</td>
<td>0.169</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD3 : $\kappa = 2$, $\lambda_0 = 0.3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.343</td>
<td>0.179</td>
<td>0.184</td>
<td>0.503</td>
<td>0.114</td>
<td>0.114</td>
<td>0.665</td>
<td>0.170</td>
<td>0.174</td>
</tr>
<tr>
<td>200</td>
<td>0.298</td>
<td>0.098</td>
<td>0.098</td>
<td>0.500</td>
<td>0.049</td>
<td>0.049</td>
<td>0.703</td>
<td>0.098</td>
<td>0.098</td>
</tr>
<tr>
<td>300</td>
<td>0.290</td>
<td>0.061</td>
<td>0.062</td>
<td>0.500</td>
<td>0.024</td>
<td>0.024</td>
<td>0.709</td>
<td>0.060</td>
<td>0.061</td>
</tr>
<tr>
<td>400</td>
<td>0.290</td>
<td>0.042</td>
<td>0.044</td>
<td>0.500</td>
<td>0.016</td>
<td>0.016</td>
<td>0.709</td>
<td>0.042</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD4 : $\kappa = 1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.449</td>
<td>0.252</td>
<td>0.293</td>
<td>0.493</td>
<td>0.210</td>
<td>0.210</td>
<td>0.547</td>
<td>0.250</td>
<td>0.293</td>
</tr>
<tr>
<td>200</td>
<td>0.422</td>
<td>0.237</td>
<td>0.266</td>
<td>0.497</td>
<td>0.159</td>
<td>0.159</td>
<td>0.598</td>
<td>0.222</td>
<td>0.244</td>
</tr>
<tr>
<td>300</td>
<td>0.402</td>
<td>0.224</td>
<td>0.246</td>
<td>0.501</td>
<td>0.121</td>
<td>0.121</td>
<td>0.633</td>
<td>0.195</td>
<td>0.206</td>
</tr>
<tr>
<td>400</td>
<td>0.382</td>
<td>0.213</td>
<td>0.228</td>
<td>0.502</td>
<td>0.095</td>
<td>0.095</td>
<td>0.655</td>
<td>0.173</td>
<td>0.178</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD4 : $\kappa = 2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta = 0.5$, $\tau_0 = 0.3$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.5$</td>
<td>$\delta = 0.5$, $\tau_0 = 0.7$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.372</td>
<td>0.213</td>
<td>0.225</td>
<td>0.501</td>
<td>0.121</td>
<td>0.121</td>
<td>0.647</td>
<td>0.187</td>
<td>0.194</td>
</tr>
<tr>
<td>200</td>
<td>0.322</td>
<td>0.159</td>
<td>0.160</td>
<td>0.501</td>
<td>0.052</td>
<td>0.052</td>
<td>0.696</td>
<td>0.110</td>
<td>0.110</td>
</tr>
<tr>
<td>300</td>
<td>0.301</td>
<td>0.122</td>
<td>0.122</td>
<td>0.500</td>
<td>0.026</td>
<td>0.026</td>
<td>0.708</td>
<td>0.064</td>
<td>0.064</td>
</tr>
<tr>
<td>400</td>
<td>0.290</td>
<td>0.089</td>
<td>0.090</td>
<td>0.500</td>
<td>0.016</td>
<td>0.016</td>
<td>0.709</td>
<td>0.043</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure A.1: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. No Level Break.

(a) $\hat{\tau}_{OLS}: \delta = 0, SD0, \kappa = 0, \lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}: \delta = 0, SD0, \kappa = 0, \lambda_0 = 0$

(c) $\hat{\tau}_{OLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}: \delta = 0, SD1, \kappa = 1, \lambda_0 = 0.7$
Figure A.1: continued ...

(i) \( \hat{\tau}_{\text{OLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.3 \)

(j) \( \hat{\tau}_{\text{FWLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.3 \)

(k) \( \hat{\tau}_{\text{OLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.5 \)

(l) \( \hat{\tau}_{\text{FWLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.5 \)

(m) \( \hat{\tau}_{\text{OLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.7 \)

(n) \( \hat{\tau}_{\text{FWLS}} : \delta = 0, \ SD1, \ \kappa = 2, \ \lambda_0 = 0.7 \)

(o) \( \hat{\tau}_{\text{OLS}} : \delta = 0, \ SD2, \ \kappa = 1, \ \lambda_0 = 0.3 \)

(p) \( \hat{\tau}_{\text{FWLS}} : \delta = 0, \ SD2, \ \kappa = 1, \ \lambda_0 = 0.3 \)
Figure A.1: continued ...

(q) $\hat{\tau}_{OLS}$: $\delta = 0$, SD2, $\kappa = 1$, $\lambda_0 = 0.5$

(r) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD2, $\kappa = 1$, $\lambda_0 = 0.5$

(s) $\hat{\tau}_{OLS}$: $\delta = 0$, SD2, $\kappa = 1$, $\lambda_0 = 0.7$

(t) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD2, $\kappa = 1$, $\lambda_0 = 0.7$

(u) $\hat{\tau}_{OLS}$: $\delta = 0$, SD2, $\kappa = 2$, $\lambda_0 = 0.3$

(v) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD2, $\kappa = 2$, $\lambda_0 = 0.3$

(w) $\hat{\tau}_{OLS}$: $\delta = 0$, SD2, $\kappa = 2$, $\lambda_0 = 0.5$

(x) $\hat{\tau}_{FWLS}$: $\delta = 0$, SD2, $\kappa = 2$, $\lambda_0 = 0.5$
Figure A.1: continued ...

(y) $\hat{\tau}_{OLS} : \delta = 0, SD2, \kappa = 2, \lambda_0 = 0.7$

(aa) $\hat{\tau}_{OLS} : \delta = 0, SD3, \kappa = 1, \lambda_0 = 0.3$

(ac) $\hat{\tau}_{OLS} : \delta = 0, SD3, \kappa = 2, \lambda_0 = 0.3$

(ae) $\hat{\tau}_{OLS} : \delta = 0, SD4, \kappa = 1, \lambda_0 = 0$

(z) $\hat{\tau}_{FWLS} : \delta = 0, SD2, \kappa = 2, \lambda_0 = 0.7$

(ab) $\hat{\tau}_{FWLS} : \delta = 0, SD3, \kappa = 1, \lambda_0 = 0.3$

(ad) $\hat{\tau}_{FWLS} : \delta = 0, SD3, \kappa = 2, \lambda_0 = 0.3$

(af) $\hat{\tau}_{FWLS} : \delta = 0, SD4, \kappa = 1, \lambda_0 = 0$
Figure A.1: continued ...

(a) \( \hat{\tau}_{\text{OLS}} : \delta = 0, \text{SD4}, \kappa = 2, \lambda_0 = 0 \)

(b) \( \hat{\tau}_{\text{FWLS}} : \delta = 0, \text{SD4}, \kappa = 2, \lambda_0 = 0 \)
Figure A.2: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.3$.

(a) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(c) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.3$; SD1, $\kappa = 1$, $\lambda_0 = 0.7$
Figure A.2: continued ...

(i) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.3$

(j) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.3$

(k) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.5$

(l) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.5$

(m) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.7$

(n) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD1, \kappa = 2, \lambda_0 = 0.7$

(o) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.3$

(p) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.3$
Figure A.2: continued ...

(q) \( \hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.5 \)

(r) \( \hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.5 \)

(s) \( \hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.7 \)

(t) \( \hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 1, \lambda_0 = 0.7 \)

(u) \( \hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.3 \)

(v) \( \hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.3 \)

(w) \( \hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.5 \)

(x) \( \hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.5 \)
(y) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.7$

(z) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD2, \kappa = 2, \lambda_0 = 0.7$

(aa) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD3, \kappa = 1, \lambda_0 = 0.3$

(ab) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD3, \kappa = 1, \lambda_0 = 0.3$

(ac) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD3, \kappa = 2, \lambda_0 = 0.3$

(ad) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD3, \kappa = 2, \lambda_0 = 0.3$

(ee) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; SD4, \kappa = 1, \lambda_0 = 0$

(af) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; SD4, \kappa = 1, \lambda_0 = 0$
Figure A.2: continued ...

(af) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.3; \text{SD4}, \kappa = 2, \lambda_0 = 0$

(ah) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.3; \text{SD4}, \kappa = 2, \lambda_0 = 0$
Figure A.3: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.5$.

(a) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD0, $\kappa = 0$, $\lambda_0 = 0$

(c) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}$: $\delta = 0.5$, $\tau_0 = 0.5$; SD1, $\kappa = 1$, $\lambda_0 = 0.7$
Figure A.3: continued ...

(i) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.3$

(j) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.3$

(k) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.5$

(l) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.5$

(m) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.7$

(n) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD1, \kappa = 2, \lambda_0 = 0.7$

(o) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD2, \kappa = 1, \lambda_0 = 0.3$

(p) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD2, \kappa = 1, \lambda_0 = 0.3$
Figure A.3: continued ...

(q) \( \hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 1, \lambda_0 = 0.5 \)

(r) \( \hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 1, \lambda_0 = 0.5 \)

(s) \( \hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 1, \lambda_0 = 0.7 \)

(t) \( \hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 1, \lambda_0 = 0.7 \)

(u) \( \hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 2, \lambda_0 = 0.3 \)

(v) \( \hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 2, \lambda_0 = 0.3 \)

(w) \( \hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 2, \lambda_0 = 0.5 \)

(x) \( \hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD2}, \kappa = 2, \lambda_0 = 0.5 \)
Figure A.3: continued ...

(y) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD2, \kappa = 2, \lambda_0 = 0.7$

(z) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD2, \kappa = 2, \lambda_0 = 0.7$

(aa) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD3, \kappa = 1, \lambda_0 = 0.3$

(ab) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD3, \kappa = 1, \lambda_0 = 0.3$

(ac) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD3, \kappa = 2, \lambda_0 = 0.3$

(ad) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD3, \kappa = 2, \lambda_0 = 0.3$

(ae) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.5; SD4, \kappa = 1, \lambda_0 = 0$

(af) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.5; SD4, \kappa = 1, \lambda_0 = 0$
Figure A.3: continued ...

(ag) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD4, } \kappa = 2, \lambda_0 = 0$

(ah) $\hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.5; \text{SD4, } \kappa = 2, \lambda_0 = 0$
Figure A.4: Simulated Sampling Density Functions of $\hat{\tau}_{OLS}$ and $\hat{\tau}_{FWLS}$. Level Break at $\tau_0 = 0.7$.

(a) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD0, \kappa = 0, \lambda_0 = 0$

(b) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD0, \kappa = 0, \lambda_0 = 0$

(c) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.3$

(d) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.3$

(e) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.5$

(f) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.5$

(g) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.7$

(h) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 1, \lambda_0 = 0.7$
(i) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.3$

(j) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.3$

(k) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.5$

(l) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.5$

(m) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.7$

(n) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD1, \kappa = 2, \lambda_0 = 0.7$

(o) $\hat{\tau}_{OLS}: \delta = 0.5, \tau_0 = 0.7; SD2, \kappa = 1, \lambda_0 = 0.3$

(p) $\hat{\tau}_{FWLS}: \delta = 0.5, \tau_0 = 0.7; SD2, \kappa = 1, \lambda_0 = 0.3$
Figure A.4: continued ...

(q) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 1, \lambda_0 = 0.5$

(r) $\hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 1, \lambda_0 = 0.5$

(s) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 1, \lambda_0 = 0.7$

(t) $\hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 1, \lambda_0 = 0.7$

(u) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 2, \lambda_0 = 0.3$

(v) $\hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 2, \lambda_0 = 0.3$

(w) $\hat{\tau}_{\text{OLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 2, \lambda_0 = 0.5$

(x) $\hat{\tau}_{\text{FWLS}} : \delta = 0.5, \tau_0 = 0.7; \text{SD2}, \kappa = 2, \lambda_0 = 0.5$
(y) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}2, \kappa = 2, \lambda_0 = 0.7$

(z) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}2, \kappa = 2, \lambda_0 = 0.7$

(aa) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}3, \kappa = 1, \lambda_0 = 0.3$

(ab) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}3, \kappa = 1, \lambda_0 = 0.3$

(ac) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}3, \kappa = 2, \lambda_0 = 0.3$

(ad) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}3, \kappa = 2, \lambda_0 = 0.3$

(ae) $\hat{\tau}_{\text{OLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}4, \kappa = 1, \lambda_0 = 0$

#af) $\hat{\tau}_{\text{FWLS}}: \delta = 0.5, \tau_0 = 0.7; \text{SD}4, \kappa = 1, \lambda_0 = 0$
Figure A.4: continued ...

(ag) \( \hat{\tau}_{OLS} : \delta = 0.5, \tau_0 = 0.7; \text{SD4, } \kappa = 2, \lambda_0 = 0 \)

(ah) \( \hat{\tau}_{FWLS} : \delta = 0.5, \tau_0 = 0.7; \text{SD4, } \kappa = 2, \lambda_0 = 0 \)