Working Paper 94-04
Statistics and Econometrics Series 03
March 1994

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

A FIXED-WIDTH INTERVAL FOR $1/\beta$ IN SIMPLE LINEAR REGRESSION

Daniel A. Coleman*

Abstract

Consider the regression model $y_i = \beta x_i + e_i$ and the problem of constructing a confidence interval for $1/\beta$ with $\beta \in (0,\beta^*)$ where $\beta^* > 0$. Uniformity down to $\beta = 0$ is a major difficulty. In fact any procedure based on a fixed sample size, will have either infinite expected width or zero confidence (Gleser and Hwang 1987), confidence being the infimum of the coverage probability. Sequential sampling is used to construct fixed-width intervals of the form

$$(1/\hat{\beta}_\tau - h,\ 1/\hat{\beta}_\tau + h)$$

where τ is an integer valued stopping time, $\hat{\beta}_\tau$ is the least squares estimator for β based on τ observations and h is the half-length of the interval. Stopping times τ_h are derived so that these intervals have coverage probabilities converging to a set value γ as $h \to 0$. This convergence is uniform down to $\beta = 0$. Furthermore the predictors x_i may be chosen adaptively.

Key Words
Brownian motion; sequential estimation; Strassen.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. Research partially supported by National Science Foundation grant DMS 89-02188.

A Fixed-Width Interval for $1/\beta$ in Simple Linear Regression

Daniel A. Coleman
Dpto. de Estadistica y Econometria, Universidad Carlos III de Madrid
Calle Madrid 126, 28903 Getafe Madrid, Spain

Abstract: Consider the regression model $y_i = \beta x_i + e_i$ and the problem of constructing a confidence interval for $1/\beta$ with $\beta \in (0, \beta^*)$ where $\beta^* > 0$. Uniformity down to $\beta = 0$ is a major difficulty. In fact any procedure based on a fixed sample size, will have either infinite expected width or zero confidence (Gleser and Hwang 1987), confidence being the infimum of the coverage probability. Sequential sampling is used to construct fixed-width intervals of the form

$$(1/\hat{\beta}_\tau - h, 1/\hat{\beta}_\tau + h)$$

where τ is an integer valued stopping time, $\hat{\beta}_\tau$ is the least squares estimator for β based on τ observations and h is the half-length of the interval. Stopping times τ_h are derived so that these intervals have coverage probabilities converging to a set value γ as $h \to 0$. This convergence is uniform down to $\beta = 0$. Furthermore the predictors x_i may be chosen adaptively.

AMS Subject Classification: Primary 6L12; Secondary 60G40, 60F15.

Keywords: Brownian motion; sequential estimation; Strassen.

Research partially supported by National Science Foundation grant DMS 8902188.

1 An interval for $1/\beta$

In this paper, fixed-width, asymptotic confidence intervals are set for $1/\beta$ from the model

$$y_i = x_i\beta + e_i. \tag{1}$$

Intervals for $1/\beta$ are of the form

$$(1/\hat{\beta}_\tau - h, 1/\hat{\beta}_\tau + h), \tag{2}$$

where τ is an integer valued stopping time, $\hat{\beta}_\tau$ is the least squares estimator for β based on τ-observations and h is the half-length. Stopping times τ_a are derived so that these confidence intervals have coverage probabilities converging to a set value $\gamma \in (0,1)$ as $h \to 0$ or as $a \to \infty$ where

$$a = \sqrt{\frac{-\Phi^{-1}(\frac{1-\gamma}{2})}{h}} \tag{3}$$

and Φ is the distribution function for a $N(0,1)$ random variable. This coverage is uniform in $\beta \in (0, \beta_a^*)$ where $\beta_a^* = \beta^* a^{\frac{1}{4}}$ and $\beta^* > 0$.

Furthermore, the predictors x_i may be chosen adaptively. That is x_i may be a function of $(x_{i-1}, y_{i-1}), \ldots, (x_1, y_1)$. In particular, x_i may be a function of $\hat{\beta}_{i-1}$ and hence may implicitly depend on the parameter β.

Sequential methods have previously been used by Lai and Siegmund (1983) to construct fixed-width, asymptotic confidence intervals for the parameter β of an $AR(1)$ model, $y_i = \beta y_{i-1} + \epsilon$, uniformly for $|\beta| \leq 1$. The difficulty in this case is that for $|\beta| = 1$ the least squares estimator is no longer asymptotically normal.

Assume the following assumptions on the errors.

(E) The errors, e_i are assumed to be independent, identically distributed random variables with $\mathbb{E}e_i = 0$, $\mathbb{E}e_i^2 = \sigma^2 > 0$ and for some $p > 1$, $\mathbb{E}|e_i|^{2p} < \infty$.

The estimators for β and σ^2 are

$$\hat{\beta}_n = t_n^{-1} \sum_{i=1}^{n} x_i y_i \text{ and } \hat{\sigma}_n^2 = (n-1)^{-1} \left[\sum_{i=1}^{n} y_i^2 - t_n \hat{\beta}_n^2 \right] + t_n^{-\frac{1}{2}}$$

where

$$\overline{x}_n = n^{-1} \sum_{i=1}^{n} x_i, \ \overline{y}_n = n^{-1} \sum_{i=1}^{n} y_i, \text{ and } t_n = \sum_{i=1}^{n} x_i^2.$$

The least squares estimator for σ^2 is modified by adding $t_n^{-\frac{1}{2}}$ to prevent stopping early.

The stopping time τ is motivated by the following. Assume

$$\frac{\hat{\beta}_n - \beta}{\hat{\sigma}_n / \sqrt{t_n}} \Rightarrow N(0,1).$$

This should hold under mild conditions by the martingale central limit theorem. Then by Slutsky's Theorem

$$\frac{\sqrt{t_n}}{\hat{\sigma}_n} \left(\frac{1}{\hat{\beta}_n} - \frac{1}{\beta} \right) \Rightarrow N\left(0, \frac{1}{\beta^4}\right).$$

Hence

$$\mathbb{P}\left(\left| \frac{1}{\hat{\beta}_n} - \frac{1}{\beta} \right| \leq h \right) \approx 1 - 2\Phi\left(\frac{-h\beta^2\sqrt{t_n}}{\sigma} \right).$$

This coverage should be at least γ, a fixed value. Replace β and σ with their estimators to obtain

$$1 - 2\Phi\left(\frac{-h\hat{\beta}_n^2\sqrt{t_n}}{\hat{\sigma}_n} \right) \geq \gamma$$

and

$$\frac{h\hat{\beta}_n^2\sqrt{t_n}}{\hat{\sigma}_n} \geq -\Phi^{-1}\left(\frac{1-\gamma}{2} \right).$$

Hence

$$\left| \sum_{i=1}^{n} x_i y_i \right| \geq a t_n^{\frac{3}{4}} \sqrt{\hat{\sigma}_n}$$

where a is defined in (3). Based on these calculations it's natural to consider the stopping time

$$\tau_a = \inf\left\{ n | n \geq 2, t_n \geq t^o \text{ and } \sum_{i=1}^{n} x_i y_i \geq a t_n^{\frac{3}{4}} \hat{\sigma}_n^{\frac{1}{2}} \right\} \tag{4}$$

where $t^o > 0$ is a constant set by the experimenter. Theorem 1, below, shows that this stopping time produces fixed-width asymptotic confidence intervals as described in (2).

Let $\lfloor z \rfloor$ be the largest integer less than or equal to z and define $f(a) = O(g(a))$ as

$$\limsup_{a \to \infty} \left| \frac{f(a)}{g(a)} \right| < M$$

for some positive constant M.

Assume the following assumptions on the predictors:

(P1) $x_i = x_i((x_{i-1}, y_{i-1}), ..., (x_1, y_1), v_i)$ where v_i, $i \geq 1$, are independent random variables such that $\{v_i\}$ is independent of $\{e_j\}$,
(P2) $\exists k \geq p$ such that $\sup_{0<\beta<\beta_a^*} \sum_{i=1}^{a} \mathbb{E}|x_i|^{2k} = O(a)$,
(P3) $\exists z_o > 0$ such that $\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{n>a} n t_n^{-1} > z_o\right) = O\left(a^{-\frac{k}{2}}\right)$.

If the predictors are deterministic the assumptions simplify to

(P2) $\exists k \geq p$ such that $\limsup_{n\to\infty} n^{-1} \sum_{i=1}^{n} |x_i|^{2k} < \infty$,
(P3) $\exists z_o > 0$ such that $\limsup_{n\to\infty} n t_n^{-1} \leq z_o$.

Let

$$\mathcal{W}_0 = \sigma\{\emptyset, \Omega\} \text{ and } \mathcal{W}_i = \sigma\{e_{i-1}, ..., e_1, v_i, ..., v_1\} \text{ for } i \geq 1. \tag{5}$$

Assuption (P3) may be replaced by

(P3') $\exists m_x > 0$ such that $\inf_{i\geq 1} \mathbb{E}(x_i^2 | \mathcal{W}_{i-1}) \geq m_x$.

Hence the assumptions are satisfied for predictors,x_i, independent, identitically distributed such that $\{x_i\}$ is independent of $\{e_i\}$, $\mathbb{E}x_1^2 > 0$ and $\mathbb{E}|x_1|^{2k} < \infty$, for some $k \geq p$.

The main result is stated in the following theorem.

Theorem 1 *Assume (E) and (P1) - (P3). Then*

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \left| \mathbb{P}\left(\left| \frac{1}{\hat{\beta}_{\tau_a}} - \frac{1}{\beta} \right| \leq h \right) - \gamma \right| = 0$$

and for $0 < p' < 4kp/(4k + 5p)$,

$$\sup_{0<\beta<\beta_a^*} \mathbb{E}\left(\beta \left| \frac{1}{\hat{\beta}_{\tau_a}} - \frac{1}{\beta} \right| \right)^{2p'} = O\left(a^{-\min\{p, 2p'\}}\right).$$

The proof of Theorem 1 will require some properties of the stopping time. At stopping

$$\sum_{i=1}^{\tau} x_i y_i = \sum_{i=1}^{\tau} x_i e_i + t_\tau \beta \geq a t_\tau^{\frac{3}{4}} \hat{\sigma}_\tau^{\frac{1}{2}}.$$

Setting $\sum_{i=1}^{\tau} x_i e_i = 0$ and solving for t_τ yields

$$t_\tau \approx \frac{a^4}{\beta^4} \hat{\sigma}_\tau^2.$$

Hence uniformity for β down to zero is obtained by sampling untill t_n is sufficiently large. Let

$$t_\tau^* = \frac{\beta^4 t_\tau}{a^4 \sigma^2}.$$

Let $d > 0$ such that

$$d < k^2/(k+2) \text{ for } k \leq 2 \text{ and } d < \min(k/2, p) \text{ for } k > 2. \tag{6}$$

The following theorem is required in the proof of Theorem 1.

Theorem 2 *Assume (E) and (P1)-(P3). Then for* $\epsilon_o > 0$,

$$\lim_{a \to 0} \sup_{0 < \beta < \beta_a^*} \sup_{\epsilon > \epsilon_o} \epsilon^d \mathbb{P}\left(t_\tau^* \geq 1 + \epsilon\right) = 0.$$

For $\epsilon > 0$,

$$\lim_{a \to 0} \sup_{0 < \beta < \beta_a^*} \mathbb{P}\left(t_\tau^* \leq 1 - \epsilon\right) = 0.$$

Furthermore

$$\lim_{a \to 0} \sup_{0 < \beta < \beta_a^*} \mathbb{E}\left|t_\tau^* - 1\right|^d = 0.$$

The rate ϵ^d obtained in the first assertion of Theorem 2 leads directly to the bound for the expectation in the third assertion.

The second assertion of Theorem 2 shows that the probability of stopping early is small. Stopping early means that the process, $\sum_{i=1}^{n} x_i y_i$, exceeds the boundary, $a t_n^{\frac{3}{4}} \hat{\sigma}_n^{\frac{1}{2}}$, for some time $t_n \leq (a/\beta)^4 \sigma^2 (1 - \epsilon)$. The main idea of the proof is to approximate the process $\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} x_i e_i + t_n \beta$ with $W(t_n) + t_n \beta$ where $W(t)$ is a Brownian motion. Then the probability of stopping early is roughly the probability that $W(t) + t\beta$ exceeds the boundary $a t^{\frac{3}{4}} \sigma^{\frac{1}{2}}$ for some time $t \leq (a/\beta)^4 \sigma^2 (1-\epsilon)$. This probability is shown to be small in Keener and Woodroofe (1992). Note that the approximation is uniform for $\beta \in (0, \beta_a^*)$.

The strong approximation to Brownian motion is proved in Section 2. Theorems 1 and 2 are proved in Sections 4 and 3 respectively. In section 5, it is shown that (P3) may be replaced by (P3').

2 Strassen's Strong Approximation Result

The strong approximation result here is almost a special case of a strong approximation result for martingales by Strassen, see Theorem 4.4, Strassen (1965). It requires a Skorohod type embedding for martingales by Jonas, see Theorem 4.3, Strassen (1965).

Theorem 3 *Let Y_i be random variables such that*

$$\mathbb{E}(Y_i|\mathcal{L}_{i-1}) = 0,\ \mathbb{E}(Y_i^2|\mathcal{L}_{i-1}) < \infty \text{ and } \mathbb{E}(Y_i^2|\mathcal{L}_{i-1}) > 0,$$

where $\mathcal{L}_i = \sigma\{Y_i, ..., Y_1\}$. Then, without loss of generality, there exists a Brownian motion $W(t)$ and random variables $\xi_i > 0$ such that

$$\sum_{i=1}^n Y_i = W\left(\sum_{i=1}^n \xi_i\right) \ a.s.,$$

$$\xi_i \text{ is measurable } \mathcal{L}_i' = \sigma\left\{Y_i, ..., Y_1, W(s); 0 \le s \le \sum_{j=1}^i \xi_j\right\},$$

$$W\left(\sum_{i=1}^n \xi_i + s\right) - W\left(\sum_{i=1}^n \xi_i\right) \text{ is independent of } \mathcal{L}_n' \text{ for } s > 0$$

and

$$\mathbb{E}(\xi_i|\mathcal{L}_{i-1}') = \mathbb{E}(Y_i^2|\mathcal{L}_{i-1}).$$

Furthermore if $\mathbb{E}(Y_n^{2k}|\mathcal{L}_{n-1}) < \infty$ for some $k > 1$, then there exists a constant L_k, depending only on k, such that

$$\mathbb{E}(\xi_i^k|\mathcal{L}_{i-1}') \le L_k \mathbb{E}(Y_i^{2k}|\mathcal{L}_{i-1}).$$

Here as in Strassen, the phrase, without loss of generality, means that there exist a probability space with a Brownian motion and random variables equal in distribution to the original random variables such that the relation is satisfied.

Theorem 4 *Let $\Theta \subseteq \mathbb{R}^k$, for k a positive integer, $\theta \in \Theta$ and $\Theta_a \subseteq \Theta$ such that $\Theta_{a'} \subseteq \Theta_a$ for all $a' \le a$. Assume $\{e_i\}$ satisfy (E), $x_i = x_i(\theta)$ are such that x_i is indpendent of e_j for all $j \ge i$ and*

$$\sup_{\Theta} \sum_{i=1}^a \mathbb{E}|x_i|^{2k} = O(a).$$

Then, without loss of generality, there exist Brownian motions $W(t) = W_\theta(t)$ such that for $\gamma > \frac{1}{4}$, $\frac{1}{4} < \gamma' < \gamma$, $\gamma' \le (6k - 2 + p)/4p$ and $\epsilon > 0$,

$$\sup_{\Theta} \epsilon^p \mathbb{P}\left(\sup_{n>a} n^{-\gamma}\left|\sigma^{-1}\sum_{i=1}^n x_i e_i - W(t_n)\right| > \epsilon\right) = O\left(a^{-(2\gamma'-\frac{1}{2})p}\right).$$

The proof of Theorem 4 requires two lemmas. The first lemma is a strong law for martingales. The lemma is adapted from a result by Brunk and Chung (see Chow and Teicher, Corollary 2, pg. 397 and Theorem 3, pg. 345).

Lemma 1 *Let $d_i = d_i(\theta)$ be martingale differences, $S_n = \sum_{i=1}^n d_i, k > 1$, $\phi > \frac{1}{2}$ and*

$$\sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}|d_i|^k = O(a). \tag{7}$$

Then

$$\sup_{\Theta} \mathbb{E} \sup_{1 \le j \le a} |S_j|^k = O\left(a^{\frac{k}{2}}\right)$$

and for $\epsilon > 0$

$$\sup_{\Theta} \epsilon^k \mathbb{P}\left(\sup_{n>a} n^{-\phi}|S_n| > \epsilon\right) = O\left(a^{-(\phi-\frac{1}{2})k}\right).$$

Proof. By Burkholder's inequality (see Chow and Teicher, Corollary 1, pg. 397), Holder's inequality and Jensen's inequality

$$\begin{aligned}
\sup_{\Theta} \mathbb{E} \sup_{1 \le j \le a} |S_j|^k &\le B_k^k \sup_{\Theta} \mathbb{E}\left[\sum_{i=1}^{\lfloor a \rfloor} d_i^2\right]^{\frac{k}{2}} \\
&\le B_k^k \sup_{\Theta} \mathbb{E}\left[\lfloor a \rfloor^{\frac{k-2}{k}} \left(\sum_{i=1}^{\lfloor a \rfloor} |d_i|^k\right)^{\frac{2}{k}}\right]^{\frac{k}{2}} \\
&\le B_k^k \sup_{\Theta} \lfloor a \rfloor^{\frac{k-2}{2}} \sum_{i=1}^{\lfloor a \rfloor} \mathbb{E}|d_i|^k \\
&= O\left(a^{\frac{k}{2}}\right)
\end{aligned}$$

where B_k is a known constant. This proves the first assertion. By Doob's submartingale inequality (see Chow and Tiecher, Theorem 8, pg. 247)

$$\begin{aligned}
&\sup_{\Theta} \epsilon^k \mathbb{P}\left(\sup_{n>a} n^{-\phi}|S_n| \ge \epsilon\right) \\
&\le \lim_{M\to\infty} \sup_{\Theta} \epsilon^k \mathbb{P}\left(\sup_{\lfloor a \rfloor \le n \le M} n^{-\phi k}|S_n|^k \ge \epsilon^k\right) \\
&\le \lim_{M\to\infty} \sup_{\Theta} \left[\lfloor a \rfloor^{-k\phi} \mathbb{E}|S_{\lfloor a \rfloor}|^k + \sum_{n=\lfloor a \rfloor+1}^{M} n^{-\phi k}\left(\mathbb{E}|S_n|^k - \mathbb{E}|S_{n-1}|^k\right)\right] \\
&\le O\left(a^{-(\phi-\frac{1}{2})k}\right) + \lim_{M\to\infty} \sum_{n=\lfloor a \rfloor+2}^{M} \left((n-1)^{-\phi k} - n^{-\phi k}\right) \sup_{\Theta} \mathbb{E}|S_{n-1}|^k \\
&\quad + \lim_{M\to\infty} M^{-\phi k} \sup_{\Theta} \mathbb{E}|S_M|^k \\
&\le O\left(a^{-(\phi-\frac{1}{2})k}\right) + O(1) \sum_{n=\lfloor a \rfloor+2}^{\infty} n^{-\phi k-1}(n-1)^{\frac{k}{2}} \\
&= O\left(a^{-(\phi-\frac{1}{2})k}\right).
\end{aligned}$$

□

Lemma 2

$$\sup_{\Theta} \mathbb{E} t_{\lfloor a \rfloor}^k = O\left(a^k\right).$$

Proof. By Holder's inequality, Jensen's inequality and (P2)

$$\begin{aligned}
\sup_{\Theta} \mathbb{E} t_{\lfloor a \rfloor}^k &= \sup_{\Theta} \mathbb{E}\left(\sum_{i=1}^{\lfloor a \rfloor} x_i^2\right)^k \\
&\leq \sup_{\Theta} \mathbb{E}\left(\lfloor a \rfloor^{\frac{k-1}{k}}\left(\sum_{i=1}^{\lfloor a \rfloor} |x_i|^{2k}\right)^{\frac{1}{k}}\right)^k \\
&= \lfloor a \rfloor^{k-1} \sup_{\Theta} \sum_{i=1}^{\lfloor a \rfloor} \mathbb{E}|x_i|^{2k} \\
&= O\left(a^k\right).
\end{aligned}$$

□

Proof of Theorem 4. For each $\theta \in \Theta$, apply Theorem 3, a Skorohod type embedding, to the random variables $\sigma^{-1}x_i e_i$. Then for each θ, there exists a probability space $(\Omega_\theta, \mathcal{A}_\theta, \mathbb{P}_\theta)$ supporting r.v.s, $e_i(\theta)$ and $x_i(\theta)$, equal in distribution to e_i and x_i, a Brownian motion W_θ and r.v.s $\xi_i(\theta)$ such that

$$\sigma^{-1}\sum_{i=1}^{n} x_i(\theta)e_i(\theta) = W_\theta\left(\sum_{i=1}^{n}\xi_i(\theta)\right) \ a.s.$$

Suppose that the result holds on each of these probability spaces, uniformly in θ, that is

$$\sup_{\Theta} \mathbb{P}_\theta \epsilon^p \left(\sup_{n>a} n^{-\gamma}\left|\sigma^{-1}\sum_{i=1}^{n} x_i(\theta)e_i(\theta) - W_\theta(t_n(\theta))\right| > \epsilon\right) = O\left(a^{-(2\gamma'-\frac{1}{2})p}\right). \tag{8}$$

By Theorem 1A, de Acosta (1982), there exists a new probability space, $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$. This new probability space supports r.v.s, $e_i^*(\theta)$ and $x_i^*(\theta)$, equal in distribution to e_i and v_i, Brownian motions W_θ^* and r.v.s $\xi_i^*(\theta)$, for all $\theta \in \Theta$. In addition, (8) holds with these new random variables and $\mathbb{P}_\theta$ replaced by $\mathbb{P}^*$. The probability space $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$ is the new probability space referred to in the phrase, without loss of generality, in the statement of the theorem. For ease of exposition, assume $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$ is the original probability space and ommit $*$ and θ from the notation.

It's sufficient to show (8) holds. Let $\gamma' \in (\frac{1}{4}, \gamma)$. A preliminary step is to establish

$$\sup_{\Theta} \mathbb{P}\left(\sup_{n>a} n^{-2\gamma'}\left|\sum_{i=1}^{n}\xi_i - t_n\right| > \epsilon\right) = \epsilon^{-p} O\left(a^{-(2\gamma'-\frac{1}{2})p}\right). \tag{9}$$

Let

$$\mathcal{X}_0 = \sigma\{v_1\} \text{ and } \mathcal{X}_i = \sigma\{e_i, ..., e_1, x_{i+1}, ..., x_1\} \text{ for } i \geq 1. \tag{10}$$

By Theorem 3 and smoothing, define

$$\begin{aligned}
\nu_n &= \sum_{i=1}^{n} \mathbb{E}\left(\xi_i|\mathcal{L}'_{i-1}\right) = \sum_{i=1}^{n} \mathbb{E}\left(\sigma^{-2}x_i^2 e_i^2|\mathcal{L}_{i-1}\right) = \sum_{i=1}^{n} \mathbb{E}\left(\mathbb{E}\left(\sigma^{-2}x_i^2 e_i^2|\mathcal{X}_{i-1}\right)|\mathcal{L}_{i-1}\right)\\
&= \sum_{i=1}^{n} \mathbb{E}\left(x_i^2 \mathbb{E}\left(\sigma^{-2}e_i^2|\mathcal{X}_{i-1}\right)|\mathcal{L}_{i-1}\right) = \sum_{i=1}^{n} \mathbb{E}\left(x_i^2|\mathcal{L}_{i-1}\right),
\end{aligned}$$

where $\mathcal{L}_i$ and $\mathcal{L}'_i$ are defined in Theorem 3. Then

$$\sum_{i=1}^{n} \xi_i - \nu_n = \sum_{i=1}^{n} \left[\xi_i - \mathbb{E}(\xi_i|\mathcal{L}'_{i-1})\right]$$

with the filtration $\mathcal{L}'_n$ is a martingale. By Jensen's inequality for conditional expectations

$$|\xi_i - \mathbb{E}(\xi_i|\mathcal{L}'_{i-1})|^p \leq \xi_i^p + \mathbb{E}(\xi_i^p|\mathcal{L}'_{i-1}).$$

Then by Theorem 3

$$\begin{aligned}
\sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}\left|\xi_i - \mathbb{E}(\xi_i|\mathcal{L}'_{i-1})\right|^p &\leq 2 \sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}\xi_i^p\\
&\leq 2L_p \sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}|\sigma^{-1}x_i e_i|^{2p}\\
&= 2L_p \sigma^{-2p}\mathbb{E}|e_1|^{2p} \sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}|x_i|^{2p}\\
&= O(a).
\end{aligned}$$

By Lemma 1,

$$\sup_{\Theta} \mathbb{P}\left(\sup_{n>a} n^{-2\gamma'} \left|\sum_{i=1}^{n} \xi_i - \nu_n\right| > \epsilon\right) = \epsilon^{-p} O\left(a^{-(2\gamma' - \frac{1}{2})p}\right). \tag{11}$$

Similarly,

$$\sigma^{-2}\sum_{i=1}^{n} x_i^2 e_i^2 - \nu_n = \sum_{i=1}^{n} \left[\sigma^{-2}x_i^2 e_i^2 - \mathbb{E}(x_i^2|\mathcal{L}_{i-1})\right]$$

with the filtration $\mathcal{L}_n$ is a martingale and

$$\sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}\left|\sigma^{-2}x_i^2 e_i^2 - \mathbb{E}(x_i^2|\mathcal{L}_{i-1})\right|^p \leq \left(1 + \sigma^{-2p}\mathbb{E}|e_1|^{2p}\right) \sup_{\Theta} \sum_{i=1}^{a} \mathbb{E}|x_i|^{2p} = O(a).$$

By Lemma 1,

$$\sup_{\Theta} \mathbb{P}\left(\sup_{n>a} n^{-2\gamma'} \left|\sigma^{-2}\sum_{i=1}^{n} x_i^2 e_i^2 - \nu_n\right| > \epsilon\right) = \epsilon^{-p} O\left(a^{-(2\gamma' - \frac{1}{2})p}\right). \tag{12}$$

Finally, since

$$t_n - \sigma^{-2}\sum_{i=1}^{n} x_i^2 e_i^2 = \sum_{i=1}^{n} \left[x_i^2(1 - \sigma^{-2}e_i^2)\right]$$

with the filtration $\mathcal{X}_n$ is a martingale,

$$\sup_\Theta \sum_{i=1}^a \mathbb{E}\left|x_i^2(1-\sigma^{-2}e_i^2)\right|^p \le \left(1+\sigma^{-2p}\mathbb{E}|e_1|^{2p}\right)\sup_\Theta \sum_{i=1}^a \mathbb{E}|x_i|^{2p} = O(a).$$

By Lemma 1,

$$\sup_\Theta \mathbb{P}\left(\sup_{n>a} n^{-2\gamma'}\left|t_n - \sigma^{-2}\sum_{i=1}^n x_i^2 e_i^2\right| > \epsilon\right) = \epsilon^{-p} O\left(a^{-(2\gamma'-\frac{1}{2})p}\right). \tag{13}$$

The first preliminary result (9) follows from (11), (12) and (13). By Lemma 2 the second preliminary result is

$$\sup_\Theta \epsilon^k \mathbb{P}\left(\sup_{n>a}\frac{t_n}{n^4} > \epsilon\right) \le \sum_{n>a} n^{-4k}\sup_\Theta \mathbb{E}t_n^k \le \sum_{n>a} O\left(n^{-3k}\right) = O\left(a^{-(3k-1)}\right). \tag{14}$$

Define the set

$$\mathcal{A}_a = \left\{\sup_{n>a}\frac{t_n}{n^4} \le \epsilon, \sup_{n>a} n^{-2\gamma'}\left|\sum_{i=1}^n \xi_i - t_n\right| \le \epsilon\right\}.$$

Since $\gamma' \le (6k-2+p)/4p$ then $(3k-1) \ge (2\gamma' - \frac{1}{2})p$ and by (9) and (14),

$$\sup_\Theta \mathbb{P}\left(\mathcal{A}_a^c\right) = \epsilon^{-p} O\left(a^{-(2\gamma'-\frac{1}{2})p}\right). \tag{15}$$

Hence it is sufficient to consider

$$\begin{aligned}
&\sup_\Theta \mathbb{P}\left(\sup_{n>a} n^{-\gamma}\left|\sigma^{-1}\sum_{i=1}^n x_i e_i - W_\theta(t_n)\right| > \epsilon, \mathcal{A}_a\right) \\
&= \sup_\Theta \mathbb{P}\left(\sup_{n>a} n^{-\gamma}\left|W_\theta\left(\sum_{i=1}^n \xi_i\right) - W_\theta(t_n)\right| > \epsilon, \mathcal{A}_a\right) \\
&\le \sum_{n>a} \mathbb{P}\left(\sup\left\{n^{-\gamma}\left|W(c)-W(t)\right|; 0 \le t \le n^4\epsilon, t \le c \le t + n^{2\gamma'}\epsilon\right\} > \epsilon\right) \\
&\le \sum_{n>a}\sum_{m=1}^{\lceil n^4\epsilon\rceil+1}\left[2\mathbb{P}\left(\sup\left\{n^{-\gamma}\left|W(m)-W(t)\right|; m-1 \le t \le m\right\} > \frac{\epsilon}{2}\right)\right. \\
&\quad + \mathbb{P}\left(\sup\left\{n^{-\gamma}\left|W(c)-W(m)\right|; m-1 \le c \le m\right\} > \frac{\epsilon}{2}\right) \\
&\quad \left. + \mathbb{P}\left(\sup\left\{n^{-\gamma}\left|W(c)-W(m)\right|; m \le c \le m + n^{2\gamma'}\epsilon\right\} > \frac{\epsilon}{2}\right)\right].
\end{aligned} \tag{16}$$

For a sufficiently large, the first and second probabilities are less then the third probability. By Levy's inequality (see Lemma d, pg. 243, Loeve 1977) and Mills' inequality (see Lemma b pg. 241, Loeve 1977) the third probability is

$$\mathbb{P}\left(\sup\left\{n^{-\gamma}\left|W(m)-W(c)\right|; m < c < m + n^{2\gamma'}\epsilon\right\} > \frac{\epsilon}{2}\right)$$

$$
\begin{aligned}
&\leq \; 2\mathbb{P}\left(n^{-\gamma}\left|W(m)-W(m+n^{2\gamma'}\epsilon)\right| > \frac{\epsilon}{2}\right) \\
&\leq \; 2\mathbb{P}\left(|N(0,1)| > \frac{\epsilon^{\frac{1}{2}}n^{\gamma-\gamma'}}{2}\right) \\
&\leq \; \frac{1}{\epsilon^{\frac{1}{2}}n^{\gamma-\gamma'}}\exp\left[-\frac{1}{2}\left(\frac{\epsilon^{\frac{1}{2}}n^{\gamma-\gamma'}}{2}\right)^2\right].
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\sup_{\Theta}\mathbb{P}\left(\sup_{n>a} n^{-\gamma}\left|\sigma^{-1}\sum_{i=1}^{n} x_i e_i - W_\theta(t_n)\right| > \epsilon, \mathcal{A}_a\right) \\
&\leq \; O(1)\sum_{n>a}\sum_{m=1}^{\lceil n^4\epsilon\rceil+1}\frac{1}{\epsilon^{\frac{1}{2}}n^{\gamma-\gamma'}}\exp\left[-\frac{1}{2}\left(\frac{\epsilon^{\frac{1}{2}}n^{\gamma-\gamma'}}{2}\right)^2\right] \\
&\leq \; O(1)\int_{n=\lfloor a\rfloor}^{\infty} n^{4-(\gamma-\gamma')}\epsilon^{\frac{1}{2}}\exp\left(-\frac{\epsilon^{\frac{1}{2}}n^{2(\gamma-\gamma')}}{8}\right)dn. \qquad (17)
\end{aligned}
$$

Integration by parts shows that this bound goes to zero geometrically as a goes to infinity. Then (8) follows by (15) and (17). □

3 Results for the Stopping Time, τ

For the remainder of the paper assume (E) and (P1)-(P3). The following lemma is used frequently.

Lemma 3

$$
\sup_{0<\beta<\beta_a^*}\mathbb{E}\sup_{1\leq n\leq a}\left|\sum_{i=1}^{n}x_i e_i\right|^{2p} = O\left(a^p\right).
$$

For $\phi > \frac{1}{2}$ and $\epsilon > 0$,

$$
\sup_{0<\beta<\beta_a^*}\epsilon^{2p}\mathbb{P}\left(\sup_{n>a} n^{-\phi}\left|\sum_{i=1}^{n}x_i e_i\right| > \epsilon\right) \leq O\left(a^{-(\phi-\frac{1}{2})2p}\right).
$$

Proof. The sum $\sum_{i=1}^{n} x_i e_i$ with the filtration $\mathcal{X}_n$, defined in (10), is a martingale such that

$$
\sup_{0<\beta<\beta_a^*}\sum_{i=1}^{a}\mathbb{E}\left|x_i e_i\right|^{2p} = \mathbb{E}|e_1|^{2p}\sup_{0<\beta<\beta_a^*}\sum_{i=1}^{a}\mathbb{E}|x_i|^{2p} = O(a)
$$

the conditions of Lemma 1 are satisfied and the results follow. □

For d in (6), choose $\delta > 0$ such that

$$
\delta < \min\{1, 2k/(k+2)\},\; d < k\delta/2, \text{ and } d < p\delta.
$$

For $\epsilon > 0$ define the stopping time $n^* = n^*(a,\delta,\beta,\epsilon)$ by

$$
n^* = \inf\left\{n \geq 2 | t_n \geq (a/\beta)^4\sigma^2(1+\epsilon)^\delta\right\}.
$$

Define the set

$$\mathcal{B}_a = \left\{a < n^* \le n^o, (a/\beta)^4\sigma^2(1+\epsilon)^\delta \le t_{n^*} < (a/\beta)^4\sigma^2(1+\epsilon)\right\}$$

where

$$n^o = \lceil 2z_o(a/\beta)^4\sigma^2(1+\epsilon)^\delta \rceil \tag{18}$$

and z_o is defined in (P3). On the set $\mathcal{B}_a$

$$\left\{t_\tau \ge (a/\beta)^4\sigma^2(1+\epsilon)\right\} \subseteq \left\{\sum_{i=1}^{n^*} x_i e_i + t_{n^*}\beta < a t_{n^*}^{\frac{3}{4}} \hat\sigma_{n^*}^{\frac{1}{2}}\right\} \tag{19}$$

Lemma 4 states that $\mathbb{P}(\mathcal{B}_a^c)$ tends to zero, Lemma 5 shows that $\hat\sigma_{n^*}^2$ converges to σ^2 and Lemma 6 uses (19) and Lemmas 4 and 5 to prove the first assertion of Theorem 2.

Lemma 4 *For* $\epsilon > 0$,

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \epsilon^d \mathbb{P}(\mathcal{B}_a^c) = 0.$$

Proof. Note that

$$\mathcal{B}_a^c \subseteq \{a \ge n^*\} \bigcup \{n^o < n^*\} \bigcup \left\{t_{n^*} \ge (a/\beta)^4\sigma^2(1+\epsilon), n^* \le n^o\right\}.$$

Since t_n is nondecreasing,

$$\{a \ge n^*\} \subseteq \left\{t_{\lfloor a \rfloor} \ge t_{n^*}\right\} \subseteq \left\{t_{\lfloor a \rfloor} \ge (a/\beta)^4\sigma^2(1+\epsilon)^\delta\right\}$$

and by Lemma 2 and Markovs inequality

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}(a \ge n^*) \le \sup_{0<\beta<\beta_a^*} \left(\frac{\beta^4}{a^4\sigma^2(1+\epsilon)^\delta}\right)^k \mathbb{E} t_{\lfloor a \rfloor}^k = (1+\epsilon)^{-k\delta} O\left(a^{-2k}\right).$$

Since n^* is a stopping time

$$\{n^* > n^o\} \subseteq \left\{\frac{a^4}{\beta^4}\sigma^2(1+\epsilon)^\delta > t_{n^o}\right\} \subseteq \left\{\frac{n^o}{t_{n^o}} > \frac{\beta^4 n^o}{a^4\sigma^2(1+\epsilon)^\delta}\right\} \subseteq \left\{\frac{n^o}{t_{n^o}} > z_o\right\}$$

and by (P3)

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}(n^* > n^o) \le \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\frac{n^o}{t_{n^o}} > z_o\right) = \sup_{0<\beta<\beta_a^*} O\left((n^o)^{-2p}\right) = (1+\epsilon)^{-\frac{k\delta}{2}} O\left(a^{-\frac{3}{2}k}\right).$$

Since $t_{n^*-1} < (a/\beta)^4\sigma^2(1+\epsilon)^\delta$ then

$$t_{n^*} - (a/\beta)^4\sigma^2(1+\epsilon) \le x_{n^*}^2 - (a/\beta)^4\sigma^2\left[(1+\epsilon) - (1+\epsilon)^\delta\right] \le \sup_{2\le n\le n^o} x_n^2 - (a/\beta)^4\sigma^2(1+\delta)\epsilon$$

and

$$\begin{aligned}\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t_{n^*} \ge \frac{a^4}{\beta^4}\sigma^2(1+\epsilon), n^* \le n^o\right) &\le \sup_{0<\beta<\beta_a^*} \left(\frac{\beta^4}{a^4\sigma^2(1+\delta)\epsilon}\right)^k \sum_{n=2}^{n^o} \mathbb{E}|x_n|^{2k} \\ &= \epsilon^{-(k-\delta)} O\left(a^{-3(k-1)}\right).\end{aligned}$$

Note that $\delta < \min\{1, 2k/(k+2)\}$ implies $k - \delta \ge k\delta/2 \ge d$. □

Lemma 5 *For $\epsilon_o > 0$,*

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \sup_{\epsilon>\epsilon_o} \epsilon^d \mathbb{P}\left(\hat{\sigma}_{n^*}^2 > (1+\epsilon)^\delta \sigma^2, \mathcal{B}_a\right) = 0.$$

For $\epsilon > 0$,

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\inf_{n>a} \hat{\sigma}_n^2 < (1-\epsilon)\sigma^2\right) = 0.$$

Proof. Note that

$$\hat{\sigma}_n^2 = \frac{n\sigma^2}{n-1} + \frac{\sum_{i=1}^n (e_i^2 - \sigma^2)}{n-1} - \frac{\sum_{i=1}^n x_i e_i}{(n-1)t_n} + t_n^{-\frac{1}{2}}.$$

Choose $m = m(\epsilon_o, \delta) > 0$ such that for $\epsilon \geq \epsilon_o$, $(1+\epsilon)^\delta \geq (1+\frac{\epsilon}{2})^\delta + 2m\epsilon^\delta$. On the set $\mathcal{B}_a$, $t_{n^*}^{-\frac{1}{2}} \leq [(a/\beta)^4\sigma^2(1+\epsilon)^\delta]^{-\frac{1}{4}} = O(a^{-\frac{3}{4}})$ then for a sufficiently large, $t_{n^*}^{-\frac{1}{2}} - m\epsilon^\delta\sigma^2 < 0$ and

$$\begin{aligned}
\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\hat{\sigma}_{n^*}^2 > (1+\epsilon)^\delta\sigma^2, \mathcal{B}_a\right) &\leq \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{n>a} \frac{1}{n-1}\left|\sum_{i=1}^n (e_i^2 - \sigma^2)\right| > m\epsilon^\delta\sigma^2\right) \\
&= \epsilon^{-p\delta} O\left(a^{-\frac{p}{2}}\right)
\end{aligned}$$

The bound follows by applying Lemma 1 to the martingale differences $(e_i^2 - \sigma^2)$. For the second assertion note that for a sufficiently large

$$\sup_{n>a}\left[(1-\epsilon)\sigma^2 - \hat{\sigma}_n^2\right] \leq \sup_{n>a}\left[\left(\frac{\sum_{i=1}^n (e_i^2-\sigma^2)}{n-1} - \frac{\epsilon\sigma^2}{4}\right) + \left(\frac{n}{t_n}\frac{(\sum_{i=1}^n x_i e_i)^2}{n(n-1)} - \frac{\epsilon\sigma^2}{4}\right)\right].$$

The result follows by applying (P3) and Lemmas 1 and 3. □

Lemma 6 *For $\epsilon_o > 0$,*

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \sup_{\epsilon>\epsilon_o} \epsilon^d \mathbb{P}\left(t_\tau^* \geq 1+\epsilon\right) = 0.$$

Proof: Define the set

$$\mathcal{C}_a = \left\{\omega | \hat{\sigma}_{n^*}^{\frac{1}{2}} < \left(1+\frac{\epsilon}{2}\right)^{\frac{\delta}{4}} \sigma^{\frac{1}{2}}, \mathcal{B}_a\right\}.$$

By Lemmas 4 and 5,

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \sup_{\epsilon>\epsilon_o} \epsilon^d \mathbb{P}\left(\mathcal{C}_a^c\right) = 0.$$

Choose $M = M(\epsilon_o, \delta) > 0$ such that for $\epsilon > \epsilon_o$,

$$(1+\epsilon)^{\frac{\delta}{4}}\left[\left(1+\frac{\epsilon}{2}\right)^{\frac{\delta}{4}} - (1+\epsilon)^{\frac{\delta}{4}}\right] \leq -M\epsilon^{\frac{\delta}{2}}.$$

On the set $\mathcal{C}_a$, $(a/\beta)^4\sigma^2(1+\epsilon)^\delta \leq t_{n^*} < (a/\beta)^4\sigma^2(1+\epsilon)$ and so

$$a t_{n^*}^{\frac{3}{4}} \hat{\sigma}_{n^*}^{\frac{1}{2}} - t_{n^*}\beta \leq t_{n^*}^{\frac{3}{4}}\left[a\sigma^{\frac{1}{2}}\left(1+\frac{\epsilon}{2}\right)^{\frac{\delta}{4}} - t_{n^*}^{\frac{1}{4}}\beta\right] < 0,$$

which is maximized with $t_{n^*} = (a/\beta)^4\sigma^2(1+\epsilon)^\delta$. Hence

$$\begin{aligned} a t_{n^*}^{\frac{3}{4}} \hat{\sigma}_{n^*}^{\frac{1}{2}} - t_{n^*}\beta &\leq \frac{a^4}{\beta^3}\sigma^2(1+\epsilon)^{\frac{3\delta}{4}}\left[\left(1+\frac{\epsilon}{2}\right)^{\frac{\delta}{4}} - (1+\epsilon)^{\frac{\delta}{4}}\right] \\ &\leq -a^{-\frac{7}{4}} m M (n^o)^{\frac{1}{2}} \epsilon^{\frac{\delta}{2}} \end{aligned}$$

where $m > 0$ is such that $m \leq \inf_{0<\beta<\beta_a^*}(a^{\frac{1}{4}}\sigma/\beta\sqrt{2z_o})$. By (19) and Lemma 3,

$$\begin{aligned} \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t_\tau \geq \frac{a^4}{\beta^4}\sigma^2(1+\epsilon), \mathcal{C}_a\right) &\leq \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sum_{i=1}^{n^*} x_i e_i + t_{n^*}\beta < a t_{n^*}^{\frac{3}{4}} \hat{\sigma}_{n^*}^{\frac{1}{2}}, \mathcal{C}_a\right) \\ &\leq \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{2\leq n\leq n^o} (n^o)^{-\frac{1}{2}} \left|\sum_{i=1}^{n} x_i e_i\right| \geq a^{-\frac{7}{4}} m M \epsilon^{\frac{\delta}{2}}\right) \\ &= \epsilon^{-\frac{p\delta}{2}} O\left(a^{-\frac{7}{2}p}\right). \end{aligned}$$

□

Consider the second assetion of Theorem 2. For a sufficiently large, define the random variable $n_* = n_*(a,\beta,\epsilon)$ as

$$n_* = \sup\left\{n > a | t_n \leq (a/\beta)^4\sigma^2(1-\epsilon)\right\}.$$

Then

$$\left\{t_\tau < (a/\beta)^4\sigma^2(1-\epsilon)\right\} \subseteq \left\{t^o \leq t_\tau \leq t_{\lfloor a \rfloor}\right) \bigcup \left\{t_{\lfloor a \rfloor} < t_\tau \leq t_{n_*}\right\}. \tag{20}$$

Lemma 8 proves the first set on the r.h.s. of (20) tends to zero. Lemma 10 uses Theorem 4, the strong approximation, to rewrite the second set on the r.h.s. of (20) in terms of a stopping time for Brownian motion. This new set is shown to tend to zero in Lemma 9.

Lemma 7 *For* $\epsilon > 0$,

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t^o \leq t_\tau \leq t_{\lfloor a \rfloor}\right) = O\left(a^{-p}\right).$$

Proof. Let $B = (2\beta^*)^{-\frac{8}{3}}$ define the event

$$\mathcal{D}_a = \left\{t_{\lfloor a \rfloor} \leq a^2 B\right\}.$$

By Lemma 2

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}(\mathcal{D}_a^c) = \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t_{\lfloor a \rfloor} > a^2 B\right) \leq \sup_{0<\beta<\beta_a^*} \left(a^2 B\right)^{-k} \mathbb{E} t_{\lfloor a \rfloor}^k = O\left(a^{-k}\right). \tag{21}$$

On the set $\mathcal{D}_a \cap \{t_n \geq t^o\}$ and for $\beta \in (0, \beta_a^*)$,

$$\sup_{2\leq n\leq a} t_n^{\frac{5}{8}}\left(t_n^{\frac{3}{8}}\beta - a\right) \leq \sup_{t^o\leq t\leq a^2 B} t^{\frac{5}{8}}\left(t^{\frac{3}{8}} a^{\frac{1}{4}} \beta^* - a\right) \leq (t^o)^{\frac{5}{8}}\left(\frac{a}{2} - a\right) \leq -(t^o)^{\frac{5}{8}}\frac{a}{2}.$$

Since $\hat{\sigma}_n^2 \geq t_n^{-\frac{1}{2}}$ then $at_n^{\frac{3}{4}}\sqrt{\hat{\sigma}_n} \geq at_n^{\frac{5}{8}}$ and by Lemma 3

$$\begin{aligned}
\sup_{0<\beta<\beta_a^*} \mathbb{P}(t_\tau \leq t_{\lfloor a \rfloor}, \mathcal{D}_a) &\leq \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{2\leq n\leq a} \sum_{i=1}^{n} x_i e_i + t_n\beta - at_n^{\frac{5}{8}} \geq 0, \mathcal{D}_a, t_n \geq t^o\right) \\
&\leq \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{2\leq n\leq a} \sum_{i=1}^{n} x_i e_i \geq (t^o)^{\frac{5}{8}} \frac{a}{2}\right) \\
&= O\left(a^{-p}\right). \qquad (22)
\end{aligned}$$

The result follows from (21) and (22). □

Lemma 8 *Let $W(t)$ be a standard Brownian motion, $c = c(a) > 0$ such that $\lim_{a\to} ac = \infty$,*

$$a' = \frac{ac}{t^o}\left(1 - \frac{1}{\sqrt{act^o}}\right) \text{ and } \tau_W = \inf\left\{t | t \geq t^o \text{ and } W(t) + t\mu \geq act^{\frac{3}{4}}\right\}.$$

Then for $ac > e^4$ and $0 < \mu \leq a'$,

$$\mathbb{P}\left(\tau_W \leq \left(\frac{ac}{\mu} - \sqrt{\frac{1}{\mu}}\right)^4\right) \leq 11\left(1 - \Phi\left(\sqrt{act^o}\right)\right) + 4ac\phi\left(\sqrt{ac} - 1\right),$$

where Φ and ϕ are the distribution and density functions of a $N(0,1)$ random variable.

After rescaling for c this lemma is the second result in Proposition 2.3 of Keener and Woodroofe (1992). Note that the bound tends to zero geometrically as $a \to \infty$.

Lemma 9 *For $\epsilon > 0$,*

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t_{\lfloor a \rfloor} < t_\tau \leq t_{n_*}\right) = 0.$$

Proof. Define the event

$$\mathcal{E}_a = \left\{\inf_{n>a}\left[a\sigma^{-1} t_n^{\frac{3}{4}}\sqrt{\hat{\sigma}_n} - \Delta_n\right] - \left[a\sigma^{-\frac{1}{2}} t_n^{\frac{3}{4}}\left(1 - \frac{\epsilon}{2}\right)^{\frac{1}{4}}\right] \geq 0\right\}$$

where $\Delta_n = \left|\sigma^{-1}\sum_{i=1}^n x_i e_i - W(t_n)\right|$. Let $\epsilon' > 0$ such that $(1 - \frac{\epsilon}{2})^{\frac{1}{4}} \leq (1 - \frac{\epsilon}{4})^{\frac{1}{4}} - \epsilon'$. Then

$$\begin{aligned}
\mathcal{E}_a^c &= \left\{\inf_{n>a} \hat{\sigma}_n^{\frac{1}{2}} - a^{-1}\sigma t_n^{-\frac{3}{4}}\Delta_n < \sigma^{\frac{1}{2}}\left(1 - \frac{\epsilon}{2}\right)^{\frac{1}{4}}\right\} \\
&\subseteq \left\{\inf_{n>a} \hat{\sigma}_n^2 < \sigma^2\left(1 - \frac{\epsilon}{4}\right)\right\} \cup \left\{\sup_{n>a} n^{\frac{3}{4}} t_n^{-\frac{3}{4}} > z_o^{\frac{3}{4}}\right\} \cup \left\{\sup_{n>a} n^{-\frac{3}{4}}\Delta_n > a\sigma^{-\frac{1}{2}}\epsilon' z_o^{-\frac{3}{4}}\right\}
\end{aligned}$$

and by Lemma 5, (P3) and Theorem 4 with $\gamma = \frac{3}{4}$ and $\gamma' = \frac{3}{8}$,

$$\lim_{a\to\infty} \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\mathcal{E}_a^c\right) = 0. \qquad (23)$$

On the set $\mathcal{E}_a$, for $\beta \in (0, \beta_a^*)$ and a sufficiently large, define R as

$$t_{n_*} \le \frac{a^4}{\beta^4}\sigma^2(1-\epsilon) \le \left(\frac{a(1-\frac{\epsilon}{2})^{\frac{1}{4}}\sqrt{\sigma}}{\beta} - \sqrt{\frac{\sigma}{\beta}}\right)^4 = R.$$

Then

$$\begin{aligned}
\left\{t_{\lfloor a\rfloor} < t_\tau \le t_{n_*}, \mathcal{E}_a\right\} &\subseteq \left\{\sum_{i=1}^{n} x_i e_i + t_n\beta \ge a t_n^{\frac{3}{4}}\sqrt{\hat{\sigma}_n}, \text{ for some } a < n \le n^*, \mathcal{E}_a\right\} \\
&\subseteq \left\{W(t_n) + t_n\frac{\beta}{\sigma} \ge a t_n^{\frac{3}{4}}\sigma^{-1}\sqrt{\hat{\sigma}_n} - \Delta_n, \text{ for some } a < n \le n^*, \mathcal{E}_a\right\} \\
&\subseteq \left\{W(t_n) + t_n\frac{\beta}{\sigma} \ge a t_n^{\frac{3}{4}}\sigma^{-\frac{1}{2}}\left(1-\frac{\epsilon}{2}\right)^{\frac{1}{4}}, \text{ for some } a < n \le n^*\right\} \\
&\subseteq \left\{W(t) + t\frac{\beta}{\sigma} \ge a t^{\frac{3}{4}}\sigma^{-\frac{1}{2}}\left(1-\frac{\epsilon}{2}\right)^{\frac{1}{4}}, \text{ for some } t^o \le t \le R\right\} \\
&\subseteq \{\tau_W \le R\}
\end{aligned}$$

where τ_W is the stopping time defined in Lemma 7 with

$$c = \sigma^{-\frac{1}{2}}\left(1-\frac{\epsilon}{2}\right)^{\frac{1}{4}} \text{ and } \mu = \frac{\beta}{\sigma} \le \frac{\beta^* a^{\frac{1}{4}}}{\sigma}.$$

Hence by Lemma 7,

$$\lim_{a\to\infty}\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(t_{\lfloor a\rfloor} < t_\tau \le t_{n_*}, \mathcal{E}_a\right) \le \lim_{a\to\infty}\sup_{0<\beta<\beta_a^*} \mathbb{P}(\tau_W \le R) = 0. \tag{24}$$

The result follows by (23) and (24) . □

Proof of Theorem 2. Lemmas 6, 7 and 9 imply the first and second assertions of Theorem 2. Consider the third assertion of Theorem 2. Choose $\epsilon \in (0,1)$ and d' such that $d' < d$. Then by the first and second assertions of Theorem 2,

$$\begin{aligned}
&\sup_{0<\beta<\beta_a^*} \mathbb{E}\,|t_\tau^* - 1|^{d'} \\
&= \sup_{0<\beta<\beta_a^*} \Big[\mathbb{E}\left(|t_\tau^*-1|^{d'}; 0 \le t_\tau^* \le 1-\epsilon\right) + \mathbb{E}\left(|t_\tau^*-1|^{d'}; |t_\tau^*-1| \le \epsilon\right) \\
&\qquad + \mathbb{E}\left(|t_\tau^*-1|^{d'}; 1+\epsilon \le t_\tau^* \le 2\right) + \mathbb{E}\left((t_\tau^*-1)^{d'}; t_\tau^* \ge 2\right)\Big] \\
&\le \sup_{0<\beta<\beta_a^*}\left[\mathbb{P}(t_\tau^* \le 1-\epsilon) + \epsilon^{d'} + \mathbb{P}(t_\tau^* \ge 1+\epsilon) + \sum_{n=1}^{\infty}\mathbb{P}\left((t_\tau^*-1)^{d'} \ge n\right)\right] \\
&\le \epsilon + \epsilon^{d'} + \epsilon + o(1)\left(\sum_{n=1}^{\infty} n^{-\frac{d}{d'}}\right) \\
&\le 4\max\{\epsilon, \epsilon^{d'}\},
\end{aligned}$$

for a sufficiently large. Since ϵ was arbitrary the result follows.

4 Proof of the Main Result, Theorem 1

In this section let $n^o = \lfloor 4z_o(a/\beta)^4\sigma^2 \rfloor$. It was previously defined slightly differently in (18). For $\epsilon \in (0,1)$, define the set

$$\mathcal{F}_a = \{|t^*_\tau - 1| < \epsilon \,\text{and}\, a < \tau \leq n^o\}.$$

Lemma 10 *For* $\epsilon \in (0,1)$,

$$\lim_{a\to 0} \sup_{0<\beta<\beta^*_a} \mathbb{P}(\mathcal{F}^c_a) = 0.$$

Proof. Consider

$$\mathcal{F}^c_a = \{|t^*_\tau - 1| \geq \epsilon\} \cup \{\tau \leq a,\, t^*_\tau > 1-\epsilon\} \cup \{\tau > n^o,\, t^*_\tau < 1+\epsilon\}.$$

By Theorem 2 the probability of the first set tends to zero uniformly for $0 < \beta < \beta^*_a$. Since

$$\{a \geq \tau, t^*_\tau > 1-\epsilon\} \subseteq \{t_{\lfloor a \rfloor} \geq t_\tau, t^*_\tau > 1-\epsilon\} \subseteq \{t_{\lfloor a \rfloor} \geq (a/\beta)^4\sigma^2(1-\epsilon)\}$$

then by Lemma 2 the probability of the second set is

$$\sup_{0<\beta<\beta^*_a} \mathbb{P}(a \geq \tau, t^*_\tau > 1-\epsilon) \leq \sup_{0<\beta<\beta^*_a} \left(\frac{\beta^4}{a^4\sigma^2(1-\epsilon)}\right)^k \mathbb{E}t^k_{\lfloor a \rfloor} = O\left(a^{-2k}\right).$$

Since

$$\{\tau > n^o, t^*_\tau < 1+\epsilon\} \subseteq \{t_\tau \geq t_{n^o},\, t^*_\tau < 1+\epsilon\} \subseteq \left\{t_{n^o} \leq (a/\beta)^4\sigma^2(1+\epsilon)\right\} \subseteq \left\{n^o t^{-1}_{n^o} > z_o\right\}$$

then by (P3) the probability of the third set is

$$\sup_{0<\beta<\beta^*_a} \mathbb{P}(\tau > n^o, t^*_\tau < 1+\epsilon) = \sup_{0<\beta<\beta^*_a} O\left((n^o)^{-\frac{k}{2}}\right) = O\left(a^{-\frac{3}{2}k}\right).$$

□

Lemma 11 *For* $\epsilon > 0$,

$$\lim_{a\to\infty} \sup_{0<\beta<\beta^*} \mathbb{P}\left(\left|\frac{a^4\sigma^2}{\beta^3 \sum_{i=1}^{\tau} x_i y_i} - 1\right| \geq \epsilon, \mathcal{F}_a\right) = 0.$$

Proof. It's sufficient to consider

$$\begin{aligned}
&\sup_{0<\beta<\beta^*_a} \mathbb{P}\left(\left|\frac{\beta^3 \sum_{i=1}^{\tau} x_i y_i}{a^4\sigma^2} - 1\right| \geq 2\epsilon, \mathcal{F}_a\right) \\
&\leq \sup_{0<\beta<\beta^*_a} \mathbb{P}\left(\left|\frac{\beta^3}{a^4\sigma^2}\sum_{i=1}^{\tau} x_i e_i + t^*_\tau - 1\right| \geq 2\epsilon\right) \\
&\leq \sup_{0<\beta<\beta^*_a} \left[\mathbb{P}\left(\sup_{1\leq n\leq n^o}\left|\sum_{i=1}^{n} x_i e_i\right| \geq \frac{a^4\sigma^2}{\beta^3}\epsilon\right) + \mathbb{P}\left(|t^*_\tau - 1| \geq \epsilon, \mathcal{F}_a\right)\right] \\
&\leq \sup_{0<\beta<\beta^*_a} \left(\frac{2\sqrt{z_o}\beta}{a^2\sigma\epsilon}\right)^{2p} (n^o)^p \\
&= O(a^{-p}).
\end{aligned}$$

□

Lemma 12 *Let $W(t) = W_\beta$ be the Brownian motion given in Theorem 4. Then for $\kappa \in (0,1)$*

$$\lim_{a\to\infty} \sup_{0<\beta<\beta^*} \mathbb{P}\left(\left| a^2\left(\frac{1}{\beta} - \frac{1}{\hat{\beta}_\tau}\right) - \frac{\beta^2}{a^2\sigma} W\left(\frac{a^4\sigma^2}{\beta^4}\right) \right| \geq \kappa, \mathcal{F}_a \right) = 0.$$

Note $\frac{\beta^2}{a^2\sigma^2} W_\beta\left(\frac{a^4\sigma^2}{\beta^4}\right) \sim N(0,1)$.

Proof. Let $\kappa > 0$. Choose $\epsilon > 0$ in the set $\mathcal{F}_a$ such that $\epsilon \leq \kappa^3/8$. Since

$$\begin{aligned}
&\left| a^2\left(\frac{1}{\beta} - \frac{1}{\hat{\beta}_\tau}\right) - \left(\frac{\beta^2}{a^2\sigma}\right)\left(\frac{a^4\sigma^2}{\beta^4}\right) \right| \\
\leq\ & \left| \left(\frac{a^4\sigma^2}{\beta^3 \sum_{i=1}^\tau x_i y_i}\right)\left(\frac{\beta^2}{a^2\sigma}\right)\left(\frac{1}{\sigma}\sum_{i=1}^\tau x_i e_i\right) - \left(\frac{\beta^2}{a^2\sigma}\right) W\left(\frac{a^4\sigma^2}{\beta^4}\right) \right| \\
\leq\ & \left(\frac{\beta^2}{a^2\sigma}\right)\left[\left| \frac{a^4\sigma^2}{\beta^3 \sum_{i=1}^\tau x_i y_i} - 1 \right| \left| \frac{1}{\sigma}\sum_{i=1}^\tau x_i e_i \right| + \left| \frac{1}{\sigma}\sum_{i=1}^\tau x_i e_i - W(t_\tau) \right| + \left| W(t_\tau) - W\left(\frac{a^4\sigma^2}{\beta^4}\right) \right| \right]
\end{aligned}$$

it's sufficient to consider the following three probabilities. By Lemma 11 and the first assertion of Lemma 3, the first probability is

$$\begin{aligned}
&\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\left| \frac{a^4\sigma^2}{\beta^3 \sum_{i=1}^\tau x_i y_i} - 1 \right| \left(\frac{\beta^2}{a^2\sigma}\right) \left| \sigma^{-1}\sum_{i=1}^\tau x_i e_i \right| \geq \kappa, \mathcal{F}_a \right) \\
\leq\ & \sup_{0<\beta<\beta_a^*} \left[\mathbb{P}\left(\sup_{2\leq n\leq n^o} \left| \sum_{i=1}^n x_i e_i \right| \geq \frac{a^2\sigma^2}{\beta^2\kappa} \right) + \mathbb{P}\left(\left| \frac{a^4\sigma^2}{\beta^3 \sum_{i=1}^\tau x_i y_i} - 1 \right| \geq \kappa^2, \mathcal{F}_a \right) \right] \\
\leq\ & \sup_{0<\beta<\beta_a^*} \left(\frac{\beta^2\kappa}{a^2\sigma^2}\right)^{2p} (n^o)^p + \frac{\kappa}{2} \\
\leq\ & \kappa
\end{aligned}$$

for a sufficiently large. Taking $\gamma = \frac{1}{2}$ and $\gamma' = \frac{3}{8}$ in Theorem 4 the second probability is

$$\begin{aligned}
&\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\frac{\beta^2}{a^2\sigma} \left| \sigma^{-1}\sum_{i=1}^\tau x_i e_i - W(t_\tau) \right| \geq \kappa, \mathcal{F}_a \right) \\
\leq\ & \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{a<n\leq n^o} \left| \sigma^{-1}\sum_{i=1}^n x_i e_i - W(t_n) \right| \geq \frac{a^2\sigma\kappa}{\beta^2} \right) \\
\leq\ & \sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{n>a} n^{-\frac{1}{2}} \left| \sigma^{-1}\sum_{i=1}^n x_i e_i - W(t_n) \right| \geq \frac{a^2\sigma\kappa}{\beta^2}(n^o)^{-\frac{1}{2}} \right) \\
\leq\ & O\left(a^{-\frac{p}{4}}\right).
\end{aligned}$$

By Levy's Inequality (see Lemma d, pg. 243, Loeve 1977) the third probability in is

$$\begin{aligned}
&\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\frac{\beta^2}{a^2\sigma} \left| W\left(\frac{a^4\sigma^2}{\beta^4}\right) - W(t_\tau) \right| \geq \kappa, \mathcal{F}_a \right) \\
\leq\ & \sup_{0<\beta<\beta_a^*} 4\mathbb{P}\left(\frac{\beta^2}{a^2\sigma\sqrt{2\epsilon}} \left| W\left(\frac{a^4\sigma^2}{\beta^4}(1-\epsilon)\right) - W\left(\frac{a^4\sigma^2}{\beta^4}(1+\epsilon)\right) \right| \geq \frac{\kappa}{\sqrt{2\epsilon}} \right)
\end{aligned}$$

$$\begin{aligned} &\leq 4\left(\frac{\sqrt{2\epsilon}}{\kappa}\right)^2 \\ &\leq \kappa. \end{aligned}$$

Since κ was arbituary the result follows. □

The first assertion of Theorem 1 follows from Lemmas 10 and 12.

Proof of the second assertion of Theorem 1. Claim for $k' < k$,

$$\sup_{0<\beta<\beta_a^*} \mathbb{E}\left(\left(\frac{\tau}{t_\tau}\right)^{\frac{k'}{4}}; \tau > a\right) = \sup_{0<\beta<\beta_a^*} \mathbb{E}\left(\sup_{n>a}\left(\frac{n}{t_n}\right)^{\frac{k'}{4}}; t_n \geq t^o\right) = O(1). \tag{25}$$

By (P3), for $m \geq 1$ and $z \geq z_o$,

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{a^m<n\leq a^{m+1}} nt_n^{-1} > z, t_n \geq t^o\right) = O\left(a^{-\frac{km}{2}}\right). \tag{26}$$

If $z > (a^{m+1}/t^o)$ then the probability in (26) is zero. If $z \leq (a^{m+1}/t^o)$ then $(t^o)^{-\frac{k}{4}} z^{-\frac{k}{4}} \geq a^{-\frac{k}{4}(m+1)}$. Hence

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{a^m<n\leq a^{m+1}} nt_n^{-1} > z, t_n \geq t^o\right) = z^{-\frac{k}{4}} O\left(a^{-\frac{k}{4}(m-1)}\right).$$

Summing m over the positive integers yields

$$\sup_{0<\beta<\beta_a^*} \mathbb{P}\left(\sup_{n>a} nt_n^{-1} > z, t_n \geq t^o\right) = z^{-\frac{k}{4}} O\left(\frac{1}{\log(a)}\right)$$

and (25) follows. At stopping

$$\sum_{i=1}^{\tau} x_i y_i \geq a t_\tau^{\frac{3}{4}} \sqrt{\hat{\sigma}_\tau} \geq a t_\tau^{\frac{5}{8}}.$$

Then

$$\left|\beta\left(\frac{1}{\hat{\beta}_\tau} - \frac{1}{\beta}\right)\right| = \left|\frac{\beta t_\tau - \sum_{i=1}^{\tau} x_i y_i}{\sum_{i=1}^{\tau} x_i y_i}\right| \leq \left|\frac{\sum_{i=1}^{\tau} x_i e_i}{a t_\tau^{\frac{5}{8}}}\right|.$$

Since $p' < \frac{4kp}{5p+4k}$ then $k > \frac{5p'p}{4(p-p')} = \frac{5p'}{4} + \frac{5(p')^2}{4(p-p')}$. Choose p'' such that $p' < p'' < p$ and $k > \frac{5p'}{4} + \frac{5(p')^2}{4(p''-p')} = \frac{5p'p''}{4(p''-p')}$. By Holder's Inequality

$$\begin{aligned} &\sup_{0<\beta<\beta_a^*} \mathbb{E}\left|\beta\left(\frac{1}{\hat{\beta}_\tau} - \frac{1}{\beta}\right)\right|^{2p'} \\ &= \sup_{0<\beta<\beta_a^*} \mathbb{E}\left(\left|\frac{\sum_{i=1}^{\tau} x_i e_i}{a t_\tau^{\frac{5}{8}}}\right|^{2p'}; \tau \leq a\right) + \sup_{0<\beta<\beta_a^*} \mathbb{E}\left(\left|\frac{\sum_{i=1}^{\tau} x_i e_i}{a t_\tau^{\frac{5}{8}}}\right|^{2p'}; \tau > a\right) \end{aligned}$$

$$
\begin{aligned}
&\le \sup_{0<\beta<\beta_a^*} a^{-2p'}(t_2^o)^{-\frac{5}{2}p'}\mathbb{E}\left(\sup_{2\le n\le a}\left|\sum_{i=1}^n x_ie_i\right|^{2p'};\tau\le a\right)\\
&\quad+\sup_{0<\beta<\beta_a^*} a^{-2p'}\mathbb{E}\left(\sup_{n\ge a}\left|n^{-\frac{5}{8}}\sum_{i=1}^n x_ie_i\right|^{2p'}\left(\frac{\tau}{t_\tau}\right)^{\frac{5}{4}p'};\tau>a\right)\\
&\le \sup_{0<\beta<\beta_a^*} a^{-2p'}(t_2^o)^{-\frac{5}{2}p'}\left[\mathbb{E}\sup_{2\le n\le a}\left(\sum_{i=1}^n x_ie_i\right)^{2p}\right]^{\frac{p'}{p}}\left[\mathbb{P}(\tau\le a)\right]^{\frac{p-p'}{p}}\\
&\quad+\sup_{0<\beta<\beta_a^*} a^{-2p'}\left[\mathbb{E}\sup_{n>a}\left|n^{-\frac{5}{8}}\sum_{i=1}^n x_ie_i\right|^{2p''}\right]^{\frac{p'}{p''}}\left[\mathbb{E}\sup_{n>a}\left(\frac{\tau}{t_\tau}\right)^{\frac{5p'p''}{4(p''-p')}};\tau>a\right]^{\frac{p''-p'}{p''}}\\
&\le a^{-2p'}\left[O\left(a^p\right)\right]^{\frac{p'}{p}}\left[O\left(a^{-p}\right)\right]^{\frac{p-p'}{p}}+a^{-2p'}o(1)O(1)\\
&= O\left(a^{-\min\{p,2p'\}}\right).
\end{aligned}
$$

The bounds on the first two expectations follow from Lemma 3, the rate on the probability is calculated in Lemma 8 and the last expectation is bounded by (25). □

5 Assumption (P3')

Lemma 13 *Suppose the predictors satisfy (P1), (P2) and (P3') then (P3) holds.*

Proof. Assumption (P3') implies

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}(x_i^2|W_{i-1})\ge m_x.$$

By (P2) and Jensen's inequality for conditional random variables

$$\sup_{0<\beta<\beta_a^*}\sum_{i=1}^a \mathbb{E}|x_i^2-\mathbb{E}(x_i^2|W_{i-1})|^k=O(a).$$

By Lemma 1 and for $z_o>1/m_x$,

$$
\begin{aligned}
\sup_{0<\beta<\beta_a^*}\mathbb{P}\left(\sup_{n>a} nt_n^{-1}>z_o\right) &\le \sup_{0<\beta<\beta_a^*}\mathbb{P}\left(\sup_{n>a}-\frac{1}{n}\sum_{i=1}^n x_i^2>-\frac{1}{z_o}\right)\\
&\le \sup_{0<\beta<\beta_a^*}\mathbb{P}\left(\sup_{n>a}\frac{1}{n}\left|\sum_{i=1}^n x_i^2-\mathbb{E}(x_i^2|W_{i-1})\right|>m_x-\frac{1}{z_o}\right)\\
&= O\left(a^{-\frac{k}{2}}\right).
\end{aligned}
$$

□

Acknowledgments: This work is the first chapter of my Ph.D. dissertation written at The University of Michigan under the supervision of Robert W. Keener. The problem and the