Econometric Theory, **0**, 2026, 1–33.
doi:10.1017/S0266466626100681

JACKKNIFE EMPIRICAL LIKELIHOOD FOR INFINITE-ORDER *U*-STATISTICS

QING WANG
Wellesley College

YICHUAN ZHAO
Georgia State University

TING ZHANG
University of Georgia

Infinite-order *U*-statistics play a vital role in modern statistics. However, given their degree k_n depends on the sample size n, quantifying their variability is quite challenging, as fixed-degree theories cannot be directly adopted. In this article, we propose a jackknife empirical likelihood (EL) framework that can be used to construct confidence intervals for an infinite-order *U*-statistic. We establish sufficient conditions for the limiting chi-squared distribution of the JEL ratio statistic and introduce an efficient partition-sampling realization for practical implementation. To evaluate the performance of the proposed methodology, we conduct simulation studies in the context of subsampling-based ensemble methods and the maximum test, respectively, comparing its performance with four competing methods. The simulation studies yield satisfactory results across the board. Finally, we demonstrate the practical application of our method using a United States Forest Service timber auctions dataset to predict the highest bid at an auction.

1. INTRODUCTION

Empirical likelihood (EL) was first introduced by Owen (1988), aiming at tackling the problem of confidence interval (CI) construction for the mean. Later, Owen (1990, 1991) extended its applications to confidence region constructions and linear models. Since then, EL methods have been seen in a wide spectrum of statistical applications, such as quantile estimation (Chen and Hall, 1993; Zhou and Jing, 2003a), estimation for the difference of quantiles (Zhou and Jing, 2003b; Yang, Yau, and Zhao, 2014), mean residual life with biased and censored data (Alemdjrodo and Zhao, 2020), and receiver operating characteristic curve (ROC; Claeskens et al., 2003; Qin and Zhou, 2006; Liu and Zhao, 2012; Yang and Zhao, 2012), just to name a few. For a comprehensive review of recent developments in EL, its applications, and its related computational issues, see Chen and Keilegom

We would like to thank the Joint Editor-in-Chief, Professor Patrik Guggenberger, and the two anonymous reviewers for their constructive feedback and helpful comments, which have substantially improved the quality and presentation of the article. Address correspondence to Qing Wang, Department of Mathematics and Statistics, Wellesley College, USA, e-mail: qwang@wellesley.edu.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (https://creativecommons.org/licenses/by-nc-nd/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use and/or adaptation of the article.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

(2009), Xue and Zhu (2012), Nordman and Lahiri (2014), Lazar (2021), and Liu and Zhao (2023).

Several variants of the traditional EL method have been developed, including Bayesian EL (Lazar, 2003), bias-corrected EL (Zhu and Xue, 2006), adjusted EL (Chen, Variyath, and Abraham, 2008), and jackknife EL (JEL) (JEL; Jing, Yuan, and Zhou, 2009), among others. In particular, to overcome the computational challenge under nonlinear constraints, Jing et al. (2009) proposed JEL, a hybrid methodology that adopts jackknife pseudo-values into the construction of the empirical log-likelihood ratio. Under certain regularity conditions, they derived an asymptotic chi-squared distribution for the empirical log-likelihood ratio statistic. Furthermore, they illustrated the performance of JEL in the context of fixed-degree, one-sample or two-sample U-statistics.

Building upon the seminal work of Jing et al. (2009), we extend the scope of JEL to a novel statistical setting. Specifically, this article investigates JEL in the context of infinite-order U-statistics (Frees, 1989). Unlike traditional Wald-type methods, the JEL framework bypasses the need for a plug-in variance estimator, which is particularly beneficial for infinite-order U-statistics where explicit variance estimation can be computationally intensive or analytically infeasible. This allows for a more robust and computationally efficient inference framework in this setting. Additionally, JEL preserves the data-driven shape of confidence regions, which is especially relevant given the potentially non-symmetric asymptotic behavior associated with infinite-order kernels.

Infinite-order U-statistics have numerous applications in marketing (Gómez and Maravall, 1998), finance (Nowotarski and Weron, 2015), and econometrics (Bonhomme and Jolivet, 2009; Arellano, Blundell, and Bonhomme, 2017). These applications leverage techniques, such as subsampling-based ensemble methods, cross-validation, and subsampling distributions, among others. For example, Bonhomme and Jolivet (2009) used ensemble methods to account for unobserved heterogeneity in credit risk assessment. Nowotarski and Weron (2015) proposed a quantile regression averaging (QRA) model, which aggregates forecasts from multiple models to construct robust prediction intervals for electricity spot prices. However, the generalization from fixed-degree U-statistics to infinite-order U-statistics is not trivial, as in the latter case, the degree k_n of the U-statistic depends on n, and therefore it grows as n goes to infinity. Consequently, existing theoretical results and proofs for fixed-degree U-statistics cannot be extended to the infinite-order setting in a straightforward fashion. DiCiccio and Romano (2022) stated the sufficient conditions for the validity of a limiting normal distribution of an infinite-order U-statistic. Subsequently, Peng, Coleman, and Mentch (2022) proved the central limit theorem (CLT) under a so-called generalized U-statistic framework, relaxing one of the core conditions specified in DiCiccio and Romano (2022). More recently, Mayala et al. (2026) pointed out that a specific Lindeberg condition is needed to establish the limiting normal distribution for subsampling infinite random forests. Additionally, Mayala et al. (2025) also extended the work of Peng et al. (2022) to handle imbalanced classification tasks.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

In this article, we establish the sufficient conditions for the fundamental theory of JEL in the context of infinite-order U-statistics. We confirm that such conditions can be verified explicitly in practice, such as in the case of the maximum test. In addition, our simulation studies suggest that the proposed methodology is robust to the theoretical conditions—it yields satisfactory results in the context of subsampling-based ensemble methods, where the kernel function of the infinite-order U-statistic cannot be expressed in an explicit form. As a further endeavor to alleviate computational cost, we discuss an efficient computational strategy that can be utilized to calculate the jackknife pseudo-values (Efron and Stein, 1981) based on incomplete, infinite-order U-statistics without extra computational effort. Furthermore, we evaluate the performance of the proposed methodology through simulation studies in a number of settings and demonstrate its practical applications using a United States Forest Service (USFS) dataset that predicts the highest bid in dollars at a timber auction. Our main contributions to existing literature are three-fold: first, we develop a general methodology for CI construction of an infinite-order U-statistic that is computationally efficient to realize; second, we push the boundaries of existing U-statistic literature and establish the theoretical framework of JEL for infinite-order U-statistics; and third, we exhibit the applications of our proposed methodology in several interesting scenarios with satisfactory results.

The rest of the article proceeds as follows: we begin in Section 2 by presenting the proposed JEL method for infinite-order U-statistics, including its theoretical foundation and application to incomplete U-statistics. Section 3 then evaluates the method's effectiveness in constructing CIs for ensemble methods, utilizing subbagging and sub-random forest algorithms (Bühlmann and Yu, 2002; Zaman and Hirose, 2009) with linear and multivariate adaptive regression spline (MARS) models. To illustrate the method's real-world applicability, Section 4 presents a case study predicting the highest bid at a timber auction using a real dataset from the USFS. Section 5 concludes with a summary of our contributions and findings. Additionally, a simulation study on the maximum test, where the main theorem's conditions are explicitly verified, is included in Appendix C. The detailed technical proofs of the lemmas supporting the main theorem's derivation are deferred to the Supplementary Material. Additional details on verifying the conditions for the maximum test are also provided there.

2. JACKKNIFE EMPIRICAL LIKELIHOOD FOR INFINITE-ORDER *U*-STATISTICS

2.1. General Framework

Consider an independent, identically distributed (i.i.d.) sample, $\mathcal{X}_n = (X_1, \ldots, X_n)$, possibly vector valued, drawn from some underlying distribution F. Suppose the parameter of interest θ_{k_n} can be written as a statistical functional of a kernel function ϕ_n in the following form:

$$\theta_{k_n} = E[\phi_n(X_{i_1}, \ldots, X_{i_{k_n}})],$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

where ϕ_n is permutation invariant in its k_n components, and its degree k_n is dependent on the sample size n. That is, $k_n \to \infty$ as $n \to \infty$. Different from the classical fixed-degree U-statistics (Hoeffding, 1948; Lee, 1990), the parameter of interest in this setting depends on k_n and n, so we denote it by θ_{k_n} for clarity. Given a target parameter θ_{k_n} and an i.i.d. sample $\mathcal{X}_n$, one can construct an unbiased estimator for θ_{k_n} using a complete, infinite-order U-statistic (Frees, 1989), denoted by U_n:

$$U_n = \binom{n}{k_n}^{-1} \sum_{1 \le i_1 < \cdots < i_{k_n} \le n} \phi_n(X_{i_1}, \ldots, X_{i_{k_n}}), \tag{1}$$

where $(X_{i_1}, \ldots, X_{i_{k_n}})$ represents a subsample of size k_n, and the summation is taken over all possible data subsets of size k_n. In the context of growing dimensions, θ_{k_n} does not necessarily converge. However, existing literature has shown the convergence of $U_n - \theta_{k_n}$ under certain regularity conditions, see, for instance, Mentch and Hooker (2016), DiCiccio and Romano (2022), Peng et al. (2022), and Mayala et al. (2025, 2026).

Applications of infinite-order U-statistics span diverse areas, including maximum tests (Lehmann and Romano, 2005), Hodges–Lehmann estimators (Hodges and Lehmann, 1963), model comparison techniques via cross-validation (Wang and Lindsay, 2014, 2017; Wang and Cai, 2021), subsampling distribution (Politis, Romano, and Wolf, 1999), and subsampling-based ensemble methods (Bühlmann and Yu, 2002; Mentch and Hooker, 2016; Wang and Wei, 2022). In particular, in the context of subsampling-based ensemble methods, one first fits individual base learners based on data subsets of size k_n, drawn independently from a learning dataset of size n, then obtains predictions for a fixed test instance x_0 using each of the fitted learners, and in the end returns the aggregated prediction. For inferential purposes, it is crucial to evaluate the variability of an infinite-order U-statistic estimator, such as to quantify the uncertainty of predictions obtained from a subsampling-based ensemble estimator.

Some recent work on variance estimation of subsampling-based ensemble methods includes Mentch and Hooker (2016, 2017), Wager, Hastie, and Efron (2014), Wager and Athey (2018), Peng, Mentch, and Stefanski (2021), Zhou, Mentch, and Hooker (2021), Peng et al. (2022), and Wang and Wei (2022). However, due to the dependence of k_n on n, most theories and results of the conventional, fixed-degree U-statistics are no longer valid. As a result, it is practically challenging to estimate the variance of an infinite-order U-statistic. Existing methods on this topic do not seem to provide promising numerical results, unless k_n is of a much lower order of n. In this article, we aim to bridge this gap by extending the work of Jing et al. (2009) and employ JEL method to infinite-order U-statistics. The numerical results in Section 3 indicate that our developed methodology yields favorable outcomes across a wide range of applications, exhibiting superior performance compared to competing methods in the existing literature. In addition, its performance is quite robust to the order of k_n relative to n. It is worth of attention that the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

extension of Jing et al.'s work (Jing et al., 2009) is non-trivial, as the theoretical mounting of JEL in this context requires careful probabilistic treatments to handle the dependence of k_n on n, which will be discussed and presented in later sections and the Supplementary Material.

Given a complete infinite-order U-statistic as defined in equation (1), denote the jackknife pseudo-values as

$$\hat{V}_{n,i} = nU_n - (n-1)U_{n-1}^{(-i)}, \text{ for } 1 \le i \le n,$$

where $U_{n-1}^{(-i)}$ is a delete-one U-statistic. Specifically, we define $U_{n-1}^{(-i)}$ as a U-statistic of degree k_n using the same kernel function ϕ_n, constructed from a delete-one dataset excluding the ith observation ($i \in \{1, \ldots, n\}$):

$$U_{n-1}^{(-i)} = \binom{n-1}{k_n}^{-1} \sum_{\substack{1 \le i_1 < \cdots < i_{k_n} \le n \\ i \notin \{i_1, \ldots, i_{k_n}\}}} \phi_n(X_{i_1}, \ldots, X_{i_{k_n}}).$$

It is easy to show that $(1/n)\sum_{i=1}^n U_{n-1}^{(-i)} = U_n$.

As noted by Shi (1984), $\hat{V}_{n,i}$'s are dependent random variables but asymptotically independent under some mild regularity conditions. Following the notations in Jing et al. (2009), let $\mathbf{p} = (p_1, \ldots, p_n)$ be a vector of probabilities for jackknife pseudo-values $\hat{V}_{n,i}$'s ($1 \le i \le n$) such that $\sum_{i=1}^n p_i = 1$ and $0 \le p_i \le 1$ for $1 \le i \le n$. It is easy to check that $U_n = (1/n)\sum_{i=1}^n \hat{V}_{n,i}$, and therefore $E(\hat{V}_{n,i}) = \theta_{k_n}$ and $\theta_{\mathbf{p}} = \theta_{k_n}$. Then, the EL, evaluated at θ_{k_n}, can be written as

$$L(\theta_{k_n}) = \max\left\{\prod_{i=1}^n p_i : \sum_{i=1}^n p_i = 1 \text{ and } \sum_{i=1}^n p_i \hat{V}_{n,i} = \theta_{k_n}\right\}.$$

And, the JEL ratio, evaluated at θ_{k_n}, is given by

$$R(\theta_{k_n}) = \frac{L(\theta_{k_n})}{n^{-n}} = \max\left\{\prod_{i=1}^n (np_i) : \sum_{i=1}^n p_i = 1 \text{ and } \sum_{i=1}^n p_i \hat{V}_{n,i} = \theta_{k_n}\right\}.$$

By applying the Lagrange multiplier, under the condition that $\min_{1\le i\le n} \hat{V}_{n,i} \le \theta_{k_n} \le \max_{1\le i\le n} \hat{V}_{n,i}$, we have

$$p_i = \frac{1}{n} \times \frac{1}{1 + \lambda(\hat{V}_{n,i} - \theta_{k_n})}, \tag{2}$$

where λ satisfies

$$f(\lambda) := \frac{1}{n}\sum_{i=1}^n (\hat{V}_{n,i} - \theta_{k_n}) / \left[1 + \lambda(\hat{V}_{n,i} - \theta_{k_n})\right] = 0. \tag{3}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

As a result, the jackknife empirical log-likelihood ratio statistic can be expressed as

$$\log R(\theta_{k_n}) = -\sum_{i=1}^{n} \log\Big[1+\lambda(\hat{V}_{n,i}-\theta_{k_n})\Big]. \tag{4}$$

A key contribution of this article is the development of a theoretical framework for JEL applied to infinite-order U-statistics. More specifically, under some regularity conditions, $-2\log R(\theta_{k_n})$ has a limiting χ_1^2 distribution, which is stated in Theorem 1 and proved in Appendix B. The Supplementary Material presents the proofs for Lemmas 1–6, which establish the theoretical foundation for the proof of Theorem 1.

Before formally stating the main theorem, we introduce some necessary notation.

Let $h^{(1)}(x_1) = \int \phi_n(x_1,\ldots,x_{k_n})dF(x_2)\ldots dF(x_{k_n}) - \theta_{k_n}$ and recursively define

$$h^{(c)}(x_1,\ldots,x_c) = \int \phi_n(x_1,\ldots,x_{k_n})dF(x_{c+1})\ldots dF(x_{k_n}) - \sum_{j=1}^{c-1}\sum_{(i_1,\ldots,i_j)\in\mathcal{S}^{(j)}_{1,\ldots,c}} h^{(j)}(x_{i_1},\ldots,x_{i_j}),$$

where $\mathcal{S}^{(j)}_{1,\ldots,c}$ is the set that contains all index combinations of size j from $\{1,\ldots,c\}$. Also, we write

$$\eta_{1,2k_n-1} = \operatorname{cov}(\Phi_n^{(1)}(X_1,X_2,\ldots,X_{2k_n-1}),\Phi_n^{(1)}(X_1,X_{2k_n},\ldots,X_{4k_n-3})), \text{ where}$$

$$\Phi_n^{(1)}(x_1,\ldots,x_{2k_n-1}) = \frac{1}{(2k_n-1)!}\sum_{|\{x_{i_1},\ldots,x_{i_{k_n}}\}\cap\{x_{i_{k_n}},\ldots,x_{i_{2k_n-1}}\}|=1} \phi_n(x_{i_1},\ldots,x_{i_{k_n}})\phi_n(x_{i_{k_n}},\ldots,x_{i_{2k_n-1}}).$$

The summation in the second equation is taken over all $(2k_n-1)!$ permutations of $(x_1,\ldots,x_{2k_n-1})$.

We make the following assumptions that are needed in Lemmas 1–6 in the Supplementary Material for the proof of our main result.

A1. The U-statistic U_n is non-degenerate with $E\big[\phi_n^2(X_1,\ldots,X_{k_n})\big] < \infty$.

A2. The degree k_n satisfies $k_n/n^2 \to 0$ as $n\to\infty$.

A3. Let $\sigma_c^2 = \operatorname{var}\big[h^{(c)}(X_1,\ldots,X_c)\big]$, $1\le c\le k_n$, and we assume that $\varsigma_{k_n} = \max_{1\le c\le k_n}\sigma_c = o(n^{1/2}k_n^{-2})$.

A4. Under Assumption A1, further assume that $k_n^3 n^{-3/2}\eta_{1,2k_n-1}^{1/2} = o(1)$ for $4k_n-3\le n$ with a sufficiently large n.

A5. Suppose that there exist a sequence $\iota_n > 2k_n$ and a $\delta\in(0,1)$ such that

$$\limsup_{n\to\infty} n^{-(\delta\iota_n/2-k_n-1)}E\left(|\phi_n(X_1,\ldots,X_{k_n})|^{\iota_n}\right) < \infty.$$

Remark 1. Let $\phi_n^{(d)} = E\big[\phi_n(X_1,\ldots,X_{k_n})|X_1,\ldots,X_d\big]$ for $1\le d\le k_n$. From Lee (1990) (see equation (9) on page 32), we have

$$\sigma_c^2 = \sum_{d=1}^{c}(-1)^{c-d}\binom{c}{d}\operatorname{var}(\phi_n^{(d)}) \text{ for } 1\le c\le k_n.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

Therefore, a sufficient condition for A3 is to require $\max_{1\le d\le k_n}\mathrm{var}(\phi_n^{(d)}) = o\left(\frac{n}{k_n^4\cdot 2^{k_n}}\right)$, which is satisfied in the example of the maximum test (see Appendix C). Assumption A4 is imposed to bound the remainder term of the jackknife pseudo-values, leveraging its equivalent expression connected to a set of U-statistics of degree $2k_n - c$ for $1 \le c \le k_n$ (see Proposition 1 and Lemma 5 in the Supplementary Material for details).

Furthermore, the statement and proof of the main theorem are based on the following two additional conditions:

B1. Without loss of generality, assume that $E\left[\phi_n(X_1,\ldots,X_{k_n})\right] = 0$.

B2. Assume U_n admits an asymptotic normal distribution, that is,

$$\frac{\sqrt{n}(U_n - \theta_{k_n})}{\sqrt{k_n^2\zeta_{1,k_n}}} \xrightarrow{D} N(0,1),$$

where $\zeta_{1,k_n} = \mathrm{cov}(\phi_n(X_1,\ldots,X_{k_n}), \phi_n(X_1, X_{k_n+1},\ldots,X_{2k_n-1}))$.

Condition B2 is essential for proving the limiting chi-squared distribution in the main theorem (Theorem 1 below). In existing literature (e.g., DiCiccio and Romano, 2022; Peng et al., 2022; Mayala et al., 2026), various sets of regularity conditions on the convergence rate of k_n have been established. For example, Peng et al. (2022) stated that, if $\frac{k_n}{n}\frac{\zeta_{k_n,k_n}}{k_n\zeta_{1,k_n}} \to 0$, where $\zeta_{k_n,k_n} = \mathrm{var}(\phi_n)$, the CLT of U_n holds. In the most recent work, Mayala et al. (2026) highlighted the necessity of a Lindeberg condition to establish the CLT for subsampling infinite random forests. In Appendix C, we verify that the Lindeberg condition specified in Mayala et al. (2026) holds for the maximum test. Therefore, within the technical framework of Mayala et al. (2026), we can still formally demonstrate the validity of condition B2 for the maximum test.

THEOREM 1. *Let $X_1, X_2, \ldots$ be i.i.d. random variables, possibly vector valued, that are drawn from a common distribution function F. Assume Assumptions A1–A5 and Conditions B1 and B2. In addition, assume that*

$$\zeta_{1,k_n} = cov(\phi_n(X_1, X_2, \ldots, X_{k_n}), \phi_n(X_1, X_{k_n+1}, \ldots, X_{2k_n-1})) < \infty.$$

If, furthermore, ζ_{1,k_n} is bounded from below by a positive constant, then

$$-2\log R(\theta_{k_n}) \xrightarrow{D} \chi_1^2.$$

Following Theorem 1, to construct a $100(1-\alpha)\%$ asymptotic CI for θ_{k_n}, one needs to solve the following two equations simultaneously:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{V}_{n,i} - \theta_{k_n}}{1 + \lambda(\hat{V}_{n,i} - \theta_{k_n})} = 0, \tag{5}$$

$$-2\log R(\theta_{k_n}) = \chi_1^2(1-\alpha), \tag{6}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

where $\chi_1^2(1-\alpha)$ denotes the $100(1-\alpha)$th percentile of χ_1^2. It can be shown that the JEL ratio statistic, that is, $-2\log R(\theta_{k_n})$ given in equation (4), is monotonously decreasing for $\theta_{k_n} \in (\min\{\hat{V}_i\}, \bar{\hat{V}}) = (\min\{\hat{V}_i\}, U_n)$ and monotonously increasing for $\theta_{k_n} \in (\bar{\hat{V}}, \max\{\hat{V}_{n,i}\}) = (U_n, \max\{\hat{V}_{n,i}\})$, where $\bar{\hat{V}} = (1/n)\sum_{i=1}^n \hat{V}_{n,i} = U_n$ (see Lemma 4 in the Supplementary Material for the proof). Hence, one can apply the bi-section method (Wu, 2005) using these ranges as initial search intervals to identify the two roots of equations (5) and (6) simultaneously.

Remark 2. A milder condition on ϕ_n (Assumption A5) is to only require that there exists a sequence $\iota_n > 2k_n$ with $\lim_{n\to\infty} n^{-(\iota_n/2-k_n)} E\left(|\phi_n(X_1,\ldots,X_{k_n})|^{\iota_n}\right) = 0$. More discussion on this can be found in Lemma 6 in the Supplementary Material. The condition on ζ_{1,k_n} in the main theorem is essential for bounding the remainder term of the JEL log-likelihood ratio statistic $-2\log R(\theta_{k_n})$. Although we acknowledge that this condition can be challenging to verify in certain complex practical settings, our numerical results consistently demonstrate that the developed inferential framework is remarkably robust to deviations from this constraint. In addition, all conditions for Theorem 1 are explicitly verifiable. We demonstrate their applicability in Appendix C, where we show that these requirements are strictly satisfied in the case of the maximum test statistic.

Remark 3. In the recent work of Chiang, Matsushita, and Otsu (2025), the authors also extended the JEL framework to infinite-order U-statistics with applications to honest random forests, including a valuable discussion on the "sparse-subsampling" scenario and a modified JEL procedure under sparse-subsampling. However, our study offers several distinct contributions. First, we introduce a partition-sampling scheme that enhances the computational efficiency of jackknife pseudo-values. Second, Chiang et al. (2025) focused on honest random forests and discusses sparse-subsampling asymptotics, while our work provides a broader scope of applications to infinite-order U-statistics, including examples on bagging and the maximum test. Finally, we provide explicit verification of the regularity conditions for the maximum test and demonstrate the practical utility of our method through a real-world application using USFS timber auction data (see Section 4), while Chiang et al. (2025) is based on simulation studies within honest random forests.

2.2. Application to Incomplete, Infinite-Order *U*-Statistics

Since U_n's degree k_n may grow with n, computing U_n can be quite expensive or even infeasible in practice. As a result, one often considers its incomplete counterpart in real-life applications. For example, in the case of subsampling-based ensemble methods, such as subbagging and sub-random forest (Bühlmann, 2003; Zaman and Hirose, 2009), one often fits a simple regression or classification tree (Breiman et al., 1984) on B independently drawn subsamples, where B may be much smaller than $\binom{n}{k_n}$. This results in an incomplete, infinite-order U-statistic. We denote it by

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

$$U_{n,B} = \frac{1}{B}\sum_{b=1}^{B}\phi_n(\mathcal{X}_{S_b}),$$

where ϕ_n represents the tree fitting algorithm in the given context, and $\mathcal{X}_{S_1}, \ldots, \mathcal{X}_{S_B}$ are data subsets of size k_n, where $\mathcal{X}_{S_b} = \{X_i : i \in S_b,\ S_b \subset \{1, \ldots, n\},\ \text{and } |S_b| = k_n\}$, drawn independently from a given learning dataset $\mathcal{X}_n$. $U_{n,B}$ serves as a Monte Carlo approximation of the complete, infinite-order U-statistic U_n. The theoretical foundations of our developed method, such as the asymptotic normal distribution of U_n, remain valid for $U_{n,B}$ provided that $B/n \to \infty$. A formal statement of the main theorem in the case of incomplete, infinite-order U-statistics can be found in Corollary 1 below. We also discuss these asymptotic requirements in the context of the numerical studies presented later in the article. In what follows, we discuss the application of JEL to incomplete, infinite-order U-statistics through a computationally efficient partition-sampling scheme (Wang and Lindsay, 2014).

Suppose we want to aggregate predictions over B individual learners fitted on random subsamples (e.g., decision trees). Without loss of generality, assume that n is divisible by k_n. The partition-sampling strategy proceeds as follows: we start with randomly partitioning the learning dataset of size n into n/k_n blocks of data subsets, each of size k_n, and then repeat the random partitioning for $B/(n/k_n) = Bk_n/n$ times to obtain a total number of B subsamples. For each block, we fit a simple tree algorithm and make predictions at a fixed test instance, denoted as x_0. In the end, we average over the B predictions to obtain $U_{n,B}$, that is, an incomplete, infinite-order U-statistic.

To calculate the delete-one U-statistics based on the generated B subsamples, we first recognize that for a fixed index i ($1 \le i \le n$), there is one and only one block of data subset per partition that contains observation i. Therefore, building on the same set of partitioned subsamples, we can realize each of its delete-1 variants as an average over $(n/k_n - 1) \times (Bk_n/n) = B - Bk_n/n$ subsamples. In particular, we do not need to refit these individual trees. Therefore, calculating n additional delete-1 U-statistics, that is, $U^{(-i)}_{n-1,B}$'s, essentially does not incur any additional computational cost. (Note that although we reserve the same subscript B when denoting the constructed delete-1 incomplete U-statistics, each of them is calculated using less than B subsamples.) We show in our later simulation studies that this strategy makes the proposed methodology quite computationally efficient to realize.

Following the aforementioned definitions of $U_{n,B}$ and $U^{(-i)}_{n,B}$, we define the jackknife pseudo-values based on incomplete U-statistics in an analogous form as below:

$$\hat{V}_{B,i} = nU_{n,B} - (n-1)U^{(-i)}_{n-1,B}.$$

Without loss of generality, we assume that the kernel functions for $U_{n,B}$ and $U^{(-i)}_{n-1,B}$ are identical. This is a reasonable assumption in practice: for example, if k_n is a fraction of n, rounding k_n to its ceiling will always result in the same value for k_n

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

and k_{n-1}, which leads to the same kernel function for both ϕ_n and ϕ_{n-1}. Then, the following results hold.

PROPOSITION 1. *If the kernel functions for $U_{n,B}$ and $U_{n-1,B}$ are identical, we have*

$$(1/n)\sum_{i=1}^{n} U_{n-1,B}^{(-i)} = U_{n,B}, \text{ and } (1/n)\sum_{i=1}^{n} \hat{V}_{B,i} = U_{n,B}.$$

The proof for Proposition 1 can be found in Appendix A.

COROLLARY 1. *Assume the same assumptions and conditions as in Theorem 1. If the subsample size B satisfies $B/n \to \infty$, the log-likelihood ratio statistic computed through incomplete U-statistics still admits the following limiting chi-squared distribution:*

$$-2\log R_B(\theta_{k_n}) \xrightarrow{D} \chi_1^2,$$

where $\log R_B(\theta_{k_n}) = -\sum_{i=1}^{n} \log\left[1 + \lambda(\hat{V}_{B,i} - \theta_{k_n})\right]$.

Theorem 3.1 in DiCiccio and Romano (2022) states that the asymptotic normal distribution still holds for $U_{n,B}$, provided that $B/n \to \infty$. Based on this result, the limiting chi-squared distribution stated in Theorem 1 can be generalized to incomplete, infinite-order U-statistics. Note that conditions for B may be formulated differently in specific contexts, such as subsampling infinite random forests (Mayala et al., 2026) and honest random forests (Chiang et al., 2025). The statement in Corollary 1 is essentially analogous to Theorem 1 in Chiang et al. (2025). Consequently, one can construct a CI for θ_{k_n} based on incomplete, infinite-order U-statistics, following the same procedure as described in Section 2.1.

Remark 4. We acknowledge that the modified JEL method in Chiang et al. (2025) yields an improved finite-sample performance in the context of honest random forests when the number of subsamples B is not sufficiently large (i.e., their so-called "sparse-subsampling" scenario), as demonstrated in their simulation study. However, as our simulations reveal, if B is sufficiently large, the standard JEL-based CI eventually approaches the target coverage probability.

3. SIMULATION STUDIES

In this section, we evaluate the performance of the proposed method to construct asymptotic CIs of an infinite-order U-statistic in three different settings: (1) subsampling-based ensemble methods with data generated from a simple linear regression (SLR) model, (2) subsampling-based ensemble methods with data generated from a MARS model, and (3) the maximum test. We present the first two simulations in Section 3 and defer the last example to Appendix C. For the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

maximum test, one can express the kernel function of its associated infinite-order U-statistic in a closed form, and therefore it is possible to verify the conditions of Theorem 1 explicitly. More elaborations on this can be found in Appendix C.

3.1. Simple Linear Regression Model

In the first simulation study, we consider an SLR model, given by

$$Y_i = 2X_i + \epsilon_i,$$

where $X_1, \ldots, X_n$ are i.i.d. Uniform($[0,2]$) random variables, and $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. $N(0,1)$. We set the learning sample size n to be 500 and the subsample size k_n to be either $0.05n$, $0.10n$, $0.20n$, or $0.50n$. In addition, we used $B = 100$, $1,000, 5,000$, or $10,000$ random subsamples of size k_n to build each ensemble estimator, namely, $U_{n,B}$. Individual trees in each ensemble were realized using the R function `randomForest` by setting `ntree` $= 1$. This function is part of the `randomForest` package (Liaw and Wiener, 2002) in R (R Core Team, 2024) that employs Breiman's random forest algorithm. Since there is only one feature, sub-random forest is identical to subbagging in this case.

For each simulated dataset, we implemented the proposed JEL method to construct 95% CIs for $\theta_{k_n}(x_0) = E[U_{n,B}(x_0)] = E(\phi_n)$ at test instance $x_0 = 1$, where ϕ_n represents the tree-fitting algorithm realized based on a training sample of size k_n. Because in practice computing the complete infinite-order U-statistic U_n is unfeasible, we approximated the true parameter value $\theta_{k_n}(x_0)$ by the empirical mean of $U_{n,B}$ with $B = 10^4$ out of 500 randomly generated datasets. Table 1 summarizes the approximate true parameter values, the average lower bound (LB) and upper bound (UB), standard deviations of LB and UB, and the average width of the proposed CIs realized by JEL. All results were based on $R = 500$ independently simulated datasets, and we adopted the bi-section algorithm in Wu (2005) to identify each CI.

To compare our method with existing approaches, we also constructed 95% CIs using an asymptotic normal distribution:

$$(U_{n,B} - 1.96 \times \mathrm{SD}(U_{n,B}), \quad U_{n,B} + 1.96 \times \mathrm{SD}(U_{n,B})), \tag{7}$$

where the standard deviation $\mathrm{SD}(U_{n,B})$ was estimated by one of the following four methods: the unbiased variance estimator (denoted by "Unbiased") and its negative-corrected variant (denoted by "Unbiased +") (Wang and Wei, 2022), the infinitesimal jackknife (IJ; Jaeckel, 1972; Zhou et al., 2021), and the internal variance estimation method (IV; Mentch and Hooker, 2016). These four methods were also studied in Wang and Wei (2022), where the unbiased variance estimator and its variant were originally proposed. Specifically, the Unbiased + estimator replaces any negative variance estimates by a positive-definite variance estimate, as recommended in Wang and Wei (2022). The IV method is essentially a double-layer Monte-Carlo method. Therefore, to ensure comparability, we set the total

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 1. Simulation results of the proposed method under the SLR model

		Number of trees = 100					Number of trees = 1,000				
		Mean		Interval	SD		Mean		Interval	SD	
k_n	$\theta_{k_n}(x_0)$	LB	UB	Width	LB	UB	LB	UB	Width	LB	UB
$0.05n$	1.995	1.065	2.945	1.880	0.173	0.176	1.626	2.363	0.737	0.126	0.122
$0.10n$	1.991	0.648	3.343	2.695	0.257	0.258	1.462	2.521	1.059	0.181	0.182
$0.20n$	1.986	−0.014	3.995	4.009	0.430	0.442	1.197	2.771	1.574	0.289	0.292
$0.50n$	1.984	−1.883	5.852	7.735	1.241	1.301	0.519	3.447	2.928	0.624	0.665
		Number of trees = 5,000					Number of trees = 10^4				
		Mean		Interval	SD		Mean		Interval	SD	
k_n	$\theta_{k_n}(x_0)$	LB	UB	Width	LB	UB	LB	UB	Width	LB	UB
$0.05n$	1.995	1.728	2.261	0.533	0.122	0.119	1.744	2.246	0.502	0.123	0.119
$0.10n$	1.991	1.605	2.378	0.773	0.182	0.180	1.626	2.356	0.730	0.183	0.182
$0.20n$	1.986	1.413	2.560	1.147	0.291	0.300	1.444	2.530	1.086	0.293	0.303
$0.50n$	1.984	0.957	3.013	2.056	0.610	0.659	1.026	2.943	1.917	0.617	0.662

number of iterations (i.e., the inner-loop iterations multiplied by the outer-loop iterations) equal to the specified value of B in the simulation.

The results based on the proposed JEL method are summarized in Table 1. In addition, we evaluated the coverage probability of different methods, calculated as the proportion of the 500 independent simulations where the constructed CI contains the true parameter value $\theta_{k_n}(x_0)$. These results are summarized in Table 2. Moreover, we record the average computational time (in seconds) for each method across the 500 simulations in Table 3. The simulations were performed on a single personal computing machine without parallel programming.

Tables 1 and 2 reveal that the proposed method performs well for constructing CIs of the ensemble method in the SLR setting. For a given subsample size k_n, as the number of trees B increases, the overall performance of the proposed method improves, that is, it attains a narrower CI and smaller standard deviations for both the LB and UB. In addition, the coverage probability approaches the nominal level of 0.95 with a greater B. When k_n is relatively large, for example, $k_n = 0.2n$ or above, one needs to increase the number of trees in order to ensure a reliable CI. The standard deviations of the UB and LB decrease significantly when B increases from 100 to 1,000. However, increasing B further does not appear to lead to significant additional reduction in the standard deviations of the LB and UB. In general, using $B = 5,000$ trees for each ensemble seems to provide satisfactory results for the coverage probability across different subsample sizes under the SLR model.

Compared to the other competing methods shown in Table 2, our proposed method achieves coverage probabilities closer to the target 95% level in all cases

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 2. Comparison of coverage probabilities of different methods under the SLR model

			Competing methods			
$n = 500$	No. of trees	Proposal	Unbiased	Unbiased +	IJ	IV
$k_n = 0.05n$	100	1.000	0.862	0.870	1.000	1.000
	1,000	0.996	0.950	0.952	0.992	0.994
	5,000	0.968	0.952	0.952	0.966	0.980
	10,000	0.962	0.970	0.970	0.958	0.978
$k_n = 0.10n$	100	1.000	0.896	0.900	1.000	1.000
	1,000	0.998	0.970	0.970	0.996	1.000
	5,000	0.964	0.970	0.970	0.940	0.986
	10,000	0.950	0.974	0.974	0.932	0.996
$k_n = 0.20n$	100	1.000	0.976	0.976	1.000	1.000
	1,000	1.000	0.990	0.990	0.994	0.998
	5,000	0.966	0.990	0.990	0.910	0.998
	10,000	0.940	0.992	0.992	0.896	1.000
$k_n = 0.50n$	100	1.000	1.000	1.000	0.984	1.000
	1,000	0.996	1.000	1.000	0.614	1.000
	5,000	0.948	1.000	1.000	0.438	1.000
	10,000	0.920	1.000	1.000	0.422	1.000

except one occasion, particularly as the number of trees or the degree of the U-statistic increases. Notably, the performance of competing methods deteriorates significantly with a larger U-statistic degree, primarily due to violation of the assumption on k_n for the asymptotic normal distribution of U_n. Furthermore, the coverage probabilities of the unbiased estimation method and its variant tend to increase with the number of trees, but they do not appear to approach the true level as B becomes larger. It is also worth mentioning that the unbiased variance estimator can produce negative variance estimates, as noted in (Wang and Lindsay (2014) and Wang and Wei (2022). In particular, the chance of getting a negative variance estimate can be as high as 8.4%, as shown in Table 2 of Wang and Wei (2022). In comparison, the IJ method corresponds to coverage probabilities that are significantly below the target level when B or k_n increase. The IV method, on the other hand, has a large positive bias for estimating the true variance of the U-statistic, resulting in overly liberal CIs, especially when k_n is not small.

As detailed in Table 3, computational cost is primarily driven by the ensemble size (i.e., the number of trees per ensemble) and less affected by the U-statistic degree k_n. Our proposed method exhibits computational efficiency comparable to Unbiased and Unbiased +, while the IV method appears to be slightly less expensive to realize. Conversely, the cost of the IJ method scales with the learning sample size, rendering it 3–4 times as expensive in this setting. These computational trends

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 3. Comparison of average computational time (in seconds) across different methods under the SLR model

			Competing methods			
$n = 500$	No. of trees	Proposal	Unbiased	Unbiased +	IJ	IV
$k_n = 0.05n$	100	0.6	0.5	0.6	2.7	0.2
	1,000	5.5	5.2	5.8	21.8	1.9
	5,000	30.7	29.2	29.2	89.6	12.2
	10,000	63.2	58.5	58.5	182.4	23.2
$k_n = 0.10n$	100	0.6	0.6	0.7	3.0	0.2
	1,000	6.1	5.5	5.5	23.5	1.9
	5,000	36.7	29.8	29.8	100.8	12.7
	10,000	72.8	59.5	59.5	207.6	23.7
$k_n = 0.20n$	100	0.7	0.6	0.6	3.2	0.2
	1,000	6.6	5.8	5.8	24.3	2.0
	5,000	39.6	31.5	31.5	104.2	13.3
	10,000	76.9	63.2	63.2	216.5	24.7
$k_n = 0.50n$	100	0.9	0.8	0.8	3.4	0.3
	1,000	8.5	7.9	7.9	27.7	3.1
	5,000	50.1	40.2	40.2	122.0	15.1
	10,000	97.5	85.1	85.1	268.8	28.0

were consistent across both the SLR model and the MARS model. Therefore, we omit the results for computational comparison in the subsequent section.

3.2. MARS Model

In the second simulation study, we consider the MARS model, also known as Friedman #1, which was proposed by Friedman (1991) and also studied in Bühlmann (2003), Mentch and Hooker (2016), and Wang and Wei (2022):

$$Y_i = f(\mathbf{X}_i) + \epsilon_i \qquad (i = 1, \ldots, n),$$

where $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are i.i.d. Uniform$([0,1]^5)$, and $\epsilon_1, \ldots, \epsilon_n$ are i.i.d. $N(0,1)$. The regression function is given by $f(\mathbf{x}) = 10\sin(\pi x_1 x_2) + 20(x_3 - 0.05)^2 + 10x_4 + 5x_5$.

We set the learning sample size to $n = 500$ or $2,000$ and the subsample size $k_n = 0.05n$, $0.10n$, $0.20n$, and $0.50n$. To assess the effect of ensemble size, we varied the number of trees (B) from 100 to 10^5. Given each simulated dataset, we calculated the proposed 95% CIs for the expected prediction at the test point $\mathbf{x}_0$ where all five features were set to 0.5. Individual trees were built using the `randomForest` function with `ntree` $= 1$ in R . For subbagging, all five features were used for splitting at each internal node, while sub-random forests used $\lfloor 5/3 \rfloor = 1$ random feature for each binary splitting, as suggested by Breiman (2001).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

To establish a baseline for evaluating CIs, we approximated the true value of the parameter $\theta_{k_n}(\mathbf{x}_0)$ by the empirical mean of $U_{n,B}$ with $B = 10^4$ trees over 500 random datasets. All these approximate true parameter values are reported in Tables 4–7.

The complexity of the MARS data-generation process significantly increases the difficulty in evaluating the variability of the ensemble estimator, a phenomenon also observed by Wang and Wei (2022). Consistent with our earlier observations in Section 3.1, the overall performance of the proposed method improves as the number of trees B increases, given a fixed subsample size k_n. In Tables 4 and 6, when the learning sample size is $n = 500$, with $B = 10^5$ trees, the coverage probability of the constructed CI is fairly close to the nominal level of 0.95 in both the subbagging and sub-random forest scenarios. In comparison, results in Tables 5 and 7 reveal that when the learning sample size is $n = 2,000$, the coverage probability of the constructed CI all remains at 100% at $B = 100$ to 10^4. However, as the number of trees B increases further, the coverage probability improves. In particular, when $B = 10^5$, the coverage probability is around 97%–98% for most of the settings under the MARS model. Although not included in the simulation results in this article, we did notice in our numerical investigations that this figure further approaches 95% when B is increased to 10^6. It seems that for a more complex model, one needs to set B to be large enough to ensure satisfactory performance of the CI construction. This conclusion can be explained by the theoretical foundation of the developed methodology, that is, the asymptotic chi-squared distribution of $-2\log R(\theta_{k_n})$ relies on the asymptotic normality of the infinite-order U-statistic. And, as DiCiccio and Romano (2022) noted, the limiting normal distribution holds for an incomplete, infinite-order U-statistic under the condition that $n/B \to 0$. In other words, as the sample size n grows, one needs to increase B to better satisfy the condition for the asymptotic normality of an incomplete, infinite-order U-statistic.

We further evaluated the coverage probabilities of our proposed method against the four previously discussed competitors under the MARS model using subbagging and sub-random forests estimators when $n = 500$ (see Tables 8 and 9). Consistent with the SLR results, our method outperforms all competing methods except in a few instances. In particular, the unbiased method and its variant suffer from overcoverage at larger degrees k_n. Furthermore, in this setting, the chance that the unbiased variance estimator yields negative values is as high as 20%, as reported in Wang and Wei (2022). In addition, the IV method performs poorly by producing a coverage probability that is at or close to 100% even with a relatively small degree k_n. Conversely, the coverage probability of the IJ method declines sharply as B increases, eventually falling way below the target 95% level. Similar conclusions can be made for $n = 2,000$. From a different aspect, conclusions on their computational costs as discussed in Section 3.1 still hold in the MARS model: IJ is significantly more expensive to calculate than the other methods; its computational cost also increases with a larger sample size or a higher-dimensional dataset.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 4. Simulation results of the proposed method for subbagging under the MARS model ($n = 500$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.535	12.417	22.614	10.197	0.763	0.794	1.000	0.7
$0.10n$	17.480	10.634	24.272	13.638	0.965	0.966	1.000	0.8
$0.20n$	17.727	8.065	27.373	19.308	1.358	1.394	1.000	0.8
$0.50n$	18.373	0.218	36.546	36.328	2.980	2.932	1.000	1.1
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.535	15.761	19.322	3.561	0.431	0.454	1.000	6.8
$0.10n$	17.480	15.114	19.858	4.744	0.530	0.578	1.000	12.9
$0.20n$	17.727	14.429	21.049	6.620	0.692	0.753	1.000	14.9
$0.50n$	18.373	12.341	24.431	12.090	1.180	1.234	1.000	11.5
		Number of trees = 5,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.535	16.465	18.622	2.157	0.406	0.424	0.998	66.9
$0.10n$	17. 480	16.098	18.895	2.797	0.509	0.550	0.992	73.5
$0.20n$	17.727	15.881	19.630	3.749	0.649	0.697	0.994	82.0
$0.50n$	18.373	15.222	21.618	6.396	1.002	1.069	0.998	98.1
		Number of trees = 10^4						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.535	16.591	18.501	1.910	0.407	0.424	0.978	134.7
$0.10n$	17. 480	16.278	18.726	2.448	0.506	0.543	0.984	147.7
$0.20n$	17.727	16.155	19.375	3.220	0.648	0.698	0.986	157.2
$0.50n$	18.373	15.791	21.091	5.300	0.989	1.067	0.992	197.2
		Number of trees = 10^5						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.535	16.723	18.379	1.656	0.401	0.420	0.962	1,303.2
$0.10n$	17.480	16.464	18.550	2.086	0.499	0.540	0.960	1,343.2
$0.20n$	17.727	16.051	19.292	3.241	0.870	0.874	0.954	1,357.9
$0.50n$	18.373	16.446	20.521	4.075	1.005	1.078	0.950	1,610.8

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 5. Simulation results of the proposed method for subbagging under the MARS model ($n = 2,000$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.643	8.770	26.483	17.713	0.943	0.870	1.000	2.4
$0.10n$	18.091	5.736	30.414	24.678	1.319	1.277	1.000	2.9
$0.20n$	18.616	0.753	36.432	35.679	2.106	2.039	1.000	3.6
$0.50n$	19.202	−14.301	52.701	67.002	5.067	4.978	1.000	6.3
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.643	14.774	20.517	5.743	0.366	0.383	1.000	25.4
$0.10n$	18.091	14.087	22.068	7.981	0.494	0.470	1.000	29.5
$0.20n$	18.616	12.883	24.333	11.450	0.669	0.658	1.000	39.4
$0.50n$	19.202	8.508	29.911	21.403	1.363	1.421	1.000	71.9
		Number of trees = 5,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.643	16.250	19.040	2.790	0.326	0.332	1.000	129.1
$0.10n$	18.091	16.179	20.001	3.822	0.422	0.419	1.000	150.3
$0.20n$	18.616	15.913	21.342	5.429	0.544	0.552	1.000	196.0
$0.50n$	19.202	14.210	24.212	10.002	0.911	0.949	1.000	343.7
		Number of trees = 10^4						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.643	16.573	18.725	2.152	0.323	0.327	1.000	249.2
$0.10n$	18.091	16.640	19.561	2.921	0.412	0.414	1.000	294.5
$0.20n$	18.616	16.580	20.680	4.100	0.537	0.538	1.000	371.9
$0.50n$	19.202	15.514	22.926	7.412	0.838	0.882	1.000	615.8
		Number of trees = 10^5						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	17.643	16.984	18.327	1.343	0.317	0.322	0.964	4,296.2
$0.10n$	18.091	17.243	18.990	1.747	0.409	0.411	0.964	4,690.6
$0.20n$	18.616	17.480	19.834	2.354	0.529	0.543	0.974	5,568.1
$0.50n$	19.202	17.347	21.200	3.853	0.820	0.896	0.984	8,195.3

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 6. Simulation results of the proposed method for sub-random forest under the MARS model ($n = 500$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.309	12.623	24.012	11.389	0.784	0.913	1.000	0.7
$0.10n$	18.279	10.239	26.392	16.153	0.994	1.143	1.000	0.7
$0.20n$	18.301	6.743	29.816	23.073	1.410	1.522	1.000	0.8
$0.50n$	18.429	−3.157	39.980	43.137	2.901	3.093	1.000	1.0
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.309	16.384	20.240	3.856	0.416	0.434	1.000	0.7
$0.10n$	18.279	15.613	20.952	5.339	0.491	0.524	1.000	7.6
$0.20n$	18.301	14.483	22.102	7.619	0.621	0.667	1.000	8.5
$0.50n$	18.429	11.397	25.508	14.111	0.962	1.065	1.000	10.7
		Number of trees = 5,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.309	17.223	19.403	2.180	0.389	0.397	0.992	33.6
$0.10n$	18.279	16.847	19.720	2.873	0.459	0.472	0.996	38.5
$0.20n$	18.301	16.330	20.274	3.944	0.572	0.600	0.998	41.5
$0.50n$	18.429	15.007	21.870	6.863	0.804	0.866	1.000	50.4
		Number of trees = 10^4						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.309	17.378	19.246	1.868	0.379	0.387	0.988	68.2
$0.10n$	18.279	17.088	19.480	2.392	0.463	0.472	0.992	71.8
$0.20n$	18.301	16.707	19.911	3.204	0.571	0.598	0.992	77.6
$0.50n$	18.429	15.778	21.099	5.321	0.801	0.835	0.996	93.2
		Number of trees = 10^5						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.309	17.551	19.082	1.531	0.378	0.384	0.976	1,246.4
$0.10n$	18.279	17.302	19.269	1.967	0.436	0.429	0.974	1,298.5
$0.20n$	18.301	17.137	19.496	2.359	0.573	0.600	0.974	1,414.9
$0.50n$	18.429	16.756	20.156	3.400	0.819	0.862	0.974	1,571.8

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 7. Simulation results of the proposed method for sub-random forest under the MARS model ($n = 2,000$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.265	7.629	28.908	21.279	1.073	1.196	1.000	2.4
$0.10n$	18.338	3.632	33.103	29.471	1.416	1.513	1.000	2.8
$0.20n$	18.459	−2.604	39.520	42.124	2.158	2.376	1.000	3.6
$0.50n$	18.639	−21.183	58.440	76.623	5.384	5.363	1.000	6.1
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.265	14.860	21.673	6.813	0.337	0.380	1.000	26.1
$0.10n$	18.338	13.598	23.068	9.470	0.422	0.452	1.000	30.3
$0.20n$	18.459	11.695	25.229	13.534	0.565	0.590	1.000	40.8
$0.50n$	18.639	5.922	31.355	25.433	1.100	1.131	1.000	72.0
		Number of trees = 5,000						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.265	16.663	19.861	3.198	0.288	0.301	1.000	134.3
$0.10n$	18.338	16.138	20.539	4.401	0.346	0.363	1.000	157.6
$0.20n$	18.495	15.342	21.571	6.229	0.443	0.453	1.000	204.3
$0.50n$	18.639	12.871	24.416	11.545	0.690	0.691	1.000	351.7
		Number of trees = 10^4						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.265	17.702	19.459	1.757	0.281	0.289	0.998	267.3
$0.10n$	18.338	16.714	19.965	3.251	0.344	0.356	1.000	309.3
$0.20n$	18.495	16.183	20.741	4.558	0.430	0.441	1.000	392.2
$0.50n$	18.639	14.484	22.800	8.316	0.625	0.644	1.000	644.3
		Number of trees = 10^5						
		Mean		Interval	SD		Coverage	Time
k_n	$\theta_{k_n}(\mathbf{x}_0)$	LB	UB	Width	LB	UB		
$0.05n$	18.265	17.630	18.901	1.271	0.277	0.285	0.976	3,267.4
$0.10n$	18.338	17.527	19.157	1.630	0.338	0.346	0.984	4,749.5
$0.20n$	18.459	17.398	19.539	2.141	0.422	0.433	0.986	5,435.9
$0.50n$	18.639	16.939	20.383	3.444	0.589	0.621	0.996	9,650.1

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 8. Comparison of coverage probabilities of different methods under the MARS model, using the subbagging estimator

			Competing methods			
$n = 500$	No. of trees	Proposal	Unbiased	Unbiased +	IJ	IV
$k_n = 0.05n$	100	1.000	0.788	0.820	1.000	1.000
	1,000	1.000	0.862	0.906	0.998	0.998
	5,000	0.998	0.954	0.958	0.994	0.998
	10^4	0.978	0.988	0.988	0.970	0.998
	10^5	0.962	1.000	1.000	0.946	0.976
$k_n = 0.10n$	100	1.000	0.808	0.852	1.000	1.000
	1,000	1.000	0.978	0.978	0.998	1.000
	5,000	0.992	0.994	0.994	0.988	1.000
	10^4	0.984	0.998	0.998	0.970	1.000
	10^5	0.960	1.000	1.000	0.932	0.988
$k_n = 0.20n$	100	1.000	0.956	0.958	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000
	5,000	0.994	1.000	1.000	0.984	1.000
	10^4	0.986	1.000	1.000	0.964	1.000
	10^5	0.954	1.000	1.000	0.900	0.996
$k_n = 0.50n$	100	1.000	1.000	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	0.998	1.000
	5,000	0.998	1.000	1.000	0.910	1.000
	10^4	0.992	1.000	1.000	0.830	1.000
	10^5	0.950	1.000	1.000	0.608	1.000

Remark 5. In real-life applications, the ensemble size, B, can be chosen by monitoring the stability of the CI's bounds. Our simulations reveal that the JEL-based CI becomes stable for a sufficiently large B. Hence, to balance accuracy with computational cost, we recommend selecting the smallest B where further increases no longer significantly change the CI. Furthermore, we note that a more complex underlying data-generating process may require a higher value of B to ensure stability.

4. REAL DATA EXAMPLE

In this section, we present a real data example to demonstrate the practical applications of the proposed methodology. We consider a dataset on USFS timber auctions, where publicly owned forests were sold to private sector organizations through auctions. Previous research that concerns timeber auctions include Baldwin, Marshall, and Richard (1997), Haile (2001), and Haile and Tamer (2003). For this particular illustration, we focus on sealed-bid auctions in Forest Service

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 9. Comparison of coverage probabilities of different methods under the MARS model, using the sub-random forest estimator

			Competing methods			
$n = 500$	No. of trees	Proposal	Unbiased	Unbiased +	IJ	IV
$k_n = 0.05n$	100	1.000	0.764	0.818	1.000	1.000
	1,000	1.000	0.820	0.880	1.000	1.000
	5,000	0.992	0.956	0.966	1.000	1.000
	10^4	0.988	0.984	0.984	0.984	1.000
	10^5	0.972	1.000	1.000	0.972	0.986
$k_n = 0.10n$	100	1.000	0.788	0.822	1.000	1.000
	1,000	1.000	0.970	0.980	1.000	1.000
	5,000	0.996	0.998	0.998	0.998	1.000
	10^4	0.992	1.000	1.000	0.986	1.000
	10^5	0.974	1.000	1.000	0.952	0.994
$k_n = 0.20n$	100	1.000	0.960	0.968	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000
	5,000	0.998	1.000	1.000	1.000	1.000
	10^4	0.992	1.000	1.000	0.986	1.000
	10^5	0.974	1.000	1.000	0.904	0.994
$k_n = 0.50n$	100	1.000	1.000	1.000	1.000	1.000
	1,000	1.000	1.000	1.000	1.000	1.000
	5,000	1.000	1.000	1.000	0.982	1.000
	10^4	0.996	1.000	1.000	0.920	1.000
	10^5	0.974	1.000	1.000	0.723	1.000

Regions 1 and 5 between 1982 and 1990. The same dataset was also analyzed in Liu and Luo (2007).

We used the same trimmed dataset as in Liu and Luo (2007), where salvage sales and sales for small businesses were excluded. Additionally, we also discarded observations that have only one bidder for the given auction. The final trimmed dataset has 1,714 observations. For ease of implementation of the partition-sampling scheme, we randomly sampled 1,710 auctions as our learning dataset and selected one of the remaining observations as the test sample (denoted by $\mathbf{x}_0$).

The primary response of interest is the log-transformed highest bid in dollars (Y). To account for various auction-specific characteristics, we included predictors such as the log-transformed size of the tract in acres ($X_1 :=$ log_acres), log-transformed estimated timber volume in MBF ($X_2 :=$ log_vol_sum), log-transformed per-MBF values for appraisal ($X_3 :=$ AppValue_avg), log-transformed estimated selling price ($X4 :=$ log_SellValue_avg), estimated harvesting costs ($X_5 :=$ log_cost_avg), log-transformed estimated manufacturing costs ($X_6 :=$ log_MtgCost_avg), the Herfindahl–Hirschman Index ($X_7 :=$ HHI),

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

and the Forest Service region ($X_8 :=$ D5 which is a binary indicator with level 1 representing Region 5 and 0 representing Region 1). We utilized subbagging and sub-random forest estimators to predict the log-transformed highest bid in dollars of the test instance. In this context, the subbagging estimator can be written in the form of an incomplete, infinite-order U-statistics as follows:

$$U_{n,B}^{\text{bag}} = \frac{1}{B}\sum_{b=1}^{B}\phi_n(\mathcal{X}_{S_b}),$$

where ϕ_n represents the regression tree algorithm where all eight predictors are considered at each binary splitting, and $\mathcal{X}_{S_b}$ is a random subsample of size k_n taken from the given learning dataset. Similarly, the sub-random forests estimator can be written as

$$U_{n,B}^{\text{rf}} = \frac{1}{B}\sum_{b=1}^{B}\phi_n^*(\mathcal{X}_{S_b}),$$

where ϕ_n^* denotes the tree fitting algorithm where only a random subset of three features, out of eight, is considered at each internal splitting.

We considered four choices of the subsample size, namely, $k_n = 0.05n, 0.10n$, $0.20n$, and $0.50n$. In addition, we chose three different values of B: 100, 1,000, and 5,000. Our numerical explorations showed that an ensemble size of $B = 5,000$ is sufficient for the given dataset. Therefore, we did not explore larger values for B. For each choice of k_n and B, we calculated the U-statistic, that is, $U_{n,B}^{\text{bag}}$ or $U_{n,B}^{\text{rf}}$, which is an unbiased estimator for the expected value of log bid for the test instance, that is, $\theta_{k_n}(\mathbf{x}_0)$. Then, we constructed the proposed 95% JEL-based CI. All results are summarized in Table 10. The results reveal that the width of the constructed 95% CI decreases as the number of trees increases; this is consistent with our observations in the simulation studies in Section 3. In particular, the change in interval width is dramatic when B increases from 100 to 1,000, which is even more visible when fitting a sub-random forests estimator.

5. SUMMARY

This article introduces a general framework for applying JEL to infinite-order U-statistics. The central contribution is proving that the jackknife empirical log-likelihood ratio statistic, built upon infinite-order U-statistics, asymptotically follows a chi-squared distribution under specific regularity conditions. The article validates these conditions for the maximum test and proposes an efficient computational method for constructing jackknife pseudo-values using incomplete, infinite-order U-statistics. We present simulation studies in the context of subsampling-based ensemble and the maximum test to construct CIs for ensemble methods. The simulation results suggest that the proposed methodology exhibits prominent performance across different simulation settings. In addition, it shows robust performance even when the asymptotic conditions cannot be explicitly

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE 10. Constructed 95% confidence interval for the test instance at different values of k_n and B, using either the subbagging or the sub-random forests estimators

Subbagging	No. of trees	$U_{n,B}^{\text{bag}}$	95% CI	Interval width
$k_n = 0.05n$	100	10.175	(8.623, 11.713)	3.090
	1,000	10.209	(9.743, 10.674)	0.931
	5,000	10.199	(9.962, 10.436)	0.474
$k_n = 0.10n$	100	10.315	(8.311, 12.305)	3.994
	1,000	10.348	(9.729, 10.966)	1.237
	5,000	10.297	(10.011, 10.585)	0.574
$k_n = 0.20n$	100	10.394	(8.159, 12.623)	4.464
	1,000	10.374	(9.622, 11.125)	1.503
	5,000	10.373	(10.043, 10.706)	0.663
$k_n = 0.50n$	100	10.423	(7.579, 13.258)	5.679
	1,000	10.432	(9.284, 11.582)	2.298
	5,000	10.437	(9.911, 10.963)	1.052
Sub-RF	No. of trees	$U_{n,B}^{\text{rf}}$	95% CI	Interval width
$k_n = 0.05n$	100	10.075	(8.211, 11.939)	3.728
	1,000	10.023	(9.436, 10.611)	1.175
	5,000	9.985	(9.726, 10.245)	0.519
$k_n = 0.10n$	100	10.019	(7.556, 12.486)	4.930
	1,000	10.082	(9.312, 10.850)	1.538
	5,000	10.063	(9.702, 10.424)	0.722
$k_n = 0.20n$	100	10.120	(6.705, 13.543)	6.838
	1,000	10.104	(9.037, 11.170)	2.133
	5,000	10.117	(9.598, 10.636)	1.038
$k_n = 0.50n$	100	10.163	(3.648, 16.681)	13.033
	1,000	10.188	(8.174, 12.197)	4.023
	5,000	10.219	(9.329, 11.111)	1.782

verified, or when the degree of the U-statistic k_n grows at a faster rate than what is required in the main theorem. In addition, we further showcase our method by a real data example, using a USFS timber auction dataset.

A. PROOF OF PROPOSITION 1

Proof of Proposition 1. To show $(1/n)\sum_{i=1}^{n} U_{n-1,B}^{(-i)} = U_{n,B}$, we need to first note that

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

$$\frac{1}{n}\sum_{i=1}^{n} U_{n-1,B}^{(-i)} = \frac{1}{n(B-Bk_n/n)}\sum_{i=1}^{n}\sum_{\substack{1\le b\le B\\ X_i\notin\mathcal{S}_b}} \phi_n(\mathcal{S}_b).$$

where ϕ_n denotes the kernel function for both $U_{n,B}$ and $U_{n-1,B}$.

Each block of data subset contains k_n observations. And, for the delete-i calculation of $U_{n-1,B}^{(-i)}$, Bk_n/n blocks of such data subsets are removed. Therefore, when summing over i where $X_i \notin \mathcal{S}_b$, each data subset is counted repeatedly for $n-k_n$ times. That is,

$$\sum_{i=1}^{n}\sum_{\substack{1\le b\le B\\ X_i\notin\mathcal{S}_b}} \phi_n(\mathcal{S}_b) = (n-k_n)\sum_{b=1}^{B}\phi_n(\mathcal{S}_b),$$

where $\mathcal{S}_1,\ldots,\mathcal{S}_B$ are the B data subsets obtained through the previous partition sampling procedure. Hence,

$$\frac{1}{n(B-Bk_n/n)}\sum_{i=1}^{n}\sum_{\substack{1\le b\le B\\ X_i\notin\mathcal{S}_b}} \phi_n(\mathcal{S}_b) = \frac{1}{B}\sum_{b=1}^{B}\phi_n(\mathcal{S}_b) = U_{n,B}.$$

Furthermore, since $(1/n)\sum_{i=1}^{n} U_{n-1,B}^{(-i)} = U_{n,B}$, it is easy to see that

$$\frac{1}{n}\sum_{i=1}^{n}\hat{V}_{n,i} = nU_{n,B} - \frac{n-1}{n}\sum_{i=1}^{n}U_{n-1,B}^{(-i)} = nU_{n,B} - (n-1)U_{n,B} = U_{n,B}. \qquad \square$$

B. PROOF OF THEOREM 1

Building upon Lemmas 1–6 in the Supplementary Material, we detail the proof of Theorem 1 below.

COROLLARY B.1. *Let* $W_n = \max_{1\le i\le n}|\hat{V}_{n,i}-\theta_{k_n}|$. *If* $E\{\phi_n^2(X_1,\ldots,X_{k_n})\}<\infty$, *then with probability one,* $W_n = o_p(k_n n^{1/2})$ *and* $n^{-1}\sum_{i=1}^{n}|\hat{V}_{n,i}-\theta_{k_n}|^3 = o_p(k_n^3 n^{1/2}\zeta_{1,kn}+k_n n^{1/2})$.

Proof. First, we can re-express U_n as follows:

$$U_n = \frac{n-k_n}{n}U_{n-1}^{(-i)} + \binom{n}{k_n}^{-1}\sum_{(j_1,\ldots,j_{k_n-1})\in\mathcal{S}_{-i}^{(k_n-1)}}\phi_n(X_i,X_{j_1},\ldots,X_{j_{k_n-1}})$$

for any fixed $i\in\{1,\ldots,n\}$. As a result,

$$\begin{aligned}\hat{V}_{n,i} &= nU_n-(n-1)U_{n-1}^{(-i)}\\ &= (n-k_n)U_{n-1}^{(-i)} + n\binom{n}{k_n}^{-1}\sum_{(j_1,\ldots,j_{k_n-1})\in\mathcal{S}_{-i}^{(k_n-1)}}\phi_n(X_i,X_{j_1},\ldots,X_{j_{k_n-1}}) - (n-1)U_{n-1}^{(-i)}\\ &= k_n\binom{n-1}{k_n-1}^{-1}\sum_{(j_1,\ldots,j_{k_n-1})\in\mathcal{S}_{-i}^{(k_n-1)}}\phi_n(X_i,X_{j_1},\ldots,X_{j_{k_n-1}}) - (k_n-1)U_{n-1}^{(-i)}.\end{aligned}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

Thus, $|\hat{V}_{n,i}-\theta_{k_n}| \le (2k_n-1)\max|\phi_n(X_1,\ldots,X_{k_n})|+|\theta_{k_n}| = (2k_n-1)H_n+|\theta_{k_n}|$. From Lemma 6, $W_n = o_p(k_n n^{1/2})$. Together with the result in Lemma 5, we have

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}|\hat{V}_{n,i}-\theta_{k_n}|^3 &\le W_n\frac{1}{n}\sum_{i=1}^{n}|\hat{V}_{n,i}-\theta_{k_n}|^2\\ &= o_p(k_n n^{1/2})\left\{k_n^2\zeta_{1,k_n}+o(k_n^2\zeta_{1,k_n})+o_p(1)\right\}\\ &= o_p(k_n^3 n^{1/2}\zeta_{1,k_n}+k_n n^{1/2}).\end{aligned}$$

□

Proof of the Main Theorem. In the following, we complete the proof of the main theorem. □

Proof. From Lemma 4, with probability tending to one, the true parameter θ_{k_n} satisfies $\min_{1\le i\le n}\hat{V}_{n,i} < \theta_{k_n} < \max_{1\le i\le n}\hat{V}_{n,i}$. Thus, the solution to equations (B.1) and (B.2) exists and is unique:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{V}_{n,i}-\theta_{k_n}}{1+\lambda(\hat{V}_{n,i}-\theta_{k_n})}=0, \tag{B.1}$$

$$-2\log R(\theta_{k_n}) = \chi_1^2(1-\alpha). \tag{B.2}$$

In what follows, we first show the root of equation (B.1). From equation (B.1), we have

$$\begin{aligned}0 &= |f(\lambda_n)|\\ &= \frac{1}{n}\left|\sum_{i=1}^{n}(\hat{V}_{n,i}-\theta_{k_n})-\lambda_n\sum_{i=1}^{n}\frac{(\hat{V}_{n,i}-\theta_{k_n})^2}{1+\lambda_n(\hat{V}_{n,i}-\theta_{k_n})}\right|\\ &\ge \frac{|\lambda_n|}{n}\sum_{i=1}^{n}\frac{(\hat{V}_{n,i}-\theta_{k_n})^2}{1+\lambda_n(\hat{V}_{n,i}-\theta_{k_n})}-\frac{1}{n}\left|\sum_{i=1}^{n}(\hat{V}_{n,i}-\theta_{k_n})\right|\\ &\ge \frac{|\lambda_n|S_n}{1+|\lambda_n|W_n}-\left|\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_{n,i}-\theta_{k_n})\right|.\end{aligned}$$

By condition B2 and the fact that $(1/n)\sum_{i=1}^{n}\hat{V}_{n,i}=U_n$, the second term is $O_p(\sqrt{k_n^2\zeta_{1,k_n}/n})$. From Lemma 5, $S_n = k_n^2\zeta_{1,k_n}+o(k_n^2\zeta_{1,k_n})+o_p(1)$. It follows that $|\lambda_n|/(1+|\lambda_n|W_n) = O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})$. Hence, by Corollary B.1,

$$\begin{aligned}|\lambda_n|(1-O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})W_n) &= O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})\\ |\lambda_n|(1-O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})o(k_n n^{1/2})) &= O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})\\ |\lambda_n|(1-o_p(\zeta_{1,k_n}^{-1/2})) &= O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2}).\end{aligned}$$

Under the assumption that $\zeta_{1,k_n}\ge c_0$ for some positive real constant c_0, we have

$$|\lambda_n| = O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2}). \tag{B.3}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

Let $\gamma_i = \lambda_n(\hat{V}_{n,i} - \theta_{k_n})$. From Corollary B.1 and equation (B.3),

$$\max_{1\le i\le n} |\gamma_i| = O_p(k_n^{-1} n^{-1/2} \zeta_{1,k_n}^{-1/2}) o(k_n n^{1/2}) = o_p(1). \tag{B.4}$$

When expanding equation (B.1), we have

$$\begin{aligned} 0 = f(\lambda_n) &= \frac{1}{n}\sum_{i=1}^{n} (\hat{V}_{n,i} - \theta_{k_n})(1 - \gamma_i + \gamma_i^2/(1+\gamma_i)) \\ &= \frac{1}{n}\sum_{i=1}^{n} \hat{V}_{n,i} - \theta_{k_n} - S_n\lambda_n + \frac{1}{n}\sum_{i=1}^{n} (\hat{V}_{n,i} - \theta_{k_n})^3 \lambda_n^2/(1+\gamma_i), \end{aligned}$$

where the last term is bounded from above by

$$\begin{aligned} &(1/n)\sum_{i=1}^{n} \left\{ |\hat{V}_{n,i} - \theta_{k_n}|^3 \lambda_n^2 |1+\gamma_i|^{-1} \right\} \\ &\quad \le \left\{ \frac{1}{n}\sum_{i=1}^{n} |\hat{V}_{n,i} - \theta_{k_n}|^3 \right\} \times \lambda_n^2 \times \frac{1}{1 + \max_i |\gamma_i|} \\ &\quad = o_p(k_n^3 n^{1/2} \zeta_{1,k_n} + k_n n^{1/2}) \times O_p(k_n^{-2} n^{-1} \zeta_{1,k_n}^{-1}) \times \frac{1}{1 + o_p(1)} \\ &\quad = o_p(k_n n^{-1/2} + k_n^{-1} n^{-1/2} \zeta_{1,k_n}^{-1}). \end{aligned}$$

Under the assumption that $\zeta_{1,k_n} \ge c_0 > 0$ and $k_n = o(n^{1/2})$, the last line above becomes $o_p(1)$.

Therefore, we may write

$$\lambda_n = S_n^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \hat{V}_{n,i} - \theta_{k_n} \right) + \beta = S_n^{-1}(U_n - \theta_{k_n}) + \beta, \tag{B.5}$$

where $|\beta| = o_p(k_n^{-2} \zeta_{1,k_n}^{-1})$. By equation (B.4), we may expand $\log(1+\gamma_i) = \gamma_i - \gamma_i^2/2 + \eta_i$, where, for some finite $B > 0$, we have $P(|\eta_i| \le B|\gamma_i|^3; 1 \le i \le n) \to 1$ as $n \to +\infty$.

Recall that, in the main text, we define

$$p_i = \frac{1}{n} \times \frac{1}{1 + \lambda_n(\hat{V}_{n,i} - \theta_{k_n})}, \tag{B.6}$$

where λ satisfies equation (B.1), that is, $f(\lambda) := \frac{1}{n}\sum_{i=1}^{n} (\hat{V}_{n,i} - \theta_{k_n})/\{1 + \lambda(\hat{V}_{n,i} - \theta_{k_n})\} = 0$.

By plugging equations (B.5) and (B.6) into $-2\log R(\theta_{k_n})$, where $\log R(\theta_{k_n}) = \sum_{i=1}^{n} \log\left\{1 + \lambda(\hat{V}_{n,i} - \theta_{k_n})\right\}$, we have

$$\begin{aligned} -2\log R(\theta_{k_n}) &= -2\sum_{i=1}^{n} \log(np_i) = 2\sum_{i=1}^{n} \log(1+\gamma_i) \\ &= 2\sum_{i=1}^{n} \gamma_i - \sum_{i=1}^{n} \gamma_i^2 + 2\sum_{i=1}^{n} \eta_i \end{aligned}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

$$= 2n\lambda_n(U_n - \theta_{k_n}) - nS_n\lambda_n^2 + 2\sum_{i=1}^{n}\eta_i$$
$$= \frac{n(U_n - \theta_{k_n})^2}{S_n} - nS_n\beta^2 + 2\sum_{i=1}^{n}\eta_i.$$

For the first term, we have $n(U_n - \theta_{k_n})^2/S_n \xrightarrow{D} \chi_1^2$ by condition B2 and Lemma 5. For the second term, it follows from Lemma 5 and equation (B.5) that

$$|-nS_n\beta^2| = n\left\{k_n^2\zeta_{1,k_n} + o(k_n^2\zeta_{1,k_n}) + o_p(1)\right\} o_p(k_n^{-4}\zeta_{1,k_n}^{-2}) = o_p(nk_n^{-2}\zeta_{1,k_n}^{-1}).$$

Under the assumption that $\zeta_{1,k_n} \geq c_0$ and $k_n = o(n^{1/2})$, we have $|-nS_n\beta^2| = o_p(1)$.

For the last term, by Corollary B.1 and equation (B.5), we have that with probability one,

$$\begin{aligned}
|\sum_{i=1}^{n}\eta_i| &\leq \sum_{i=1}^{n}|\eta_i| \leq B|\lambda_n|^3\sum_{i=1}^{n}|\hat{V}_{n,i} - \theta_{k_n}|^3 \\
&= \left(O_p(k_n^{-1}n^{-1/2}\zeta_{1,k_n}^{-1/2})\right)^3 o(k_n^3 n^{3/2}\zeta_{1,kn} + k_n n^{3/2}) \\
&= O_p(k_n^{-3}n^{-3/2}\zeta_{1,k_n}^{-3/2})o(k_n^3 n^{3/2}\zeta_{1,kn} + k_n n^{3/2}) \\
&= o_p(\zeta_{1,k_n}^{-1/2} + k_n^{-2}\zeta_{1,k_n}^{-3/2}) \\
&= o_p(1).
\end{aligned}$$

In the end, by combining all the above results, it follows that $-2\log R(\theta_{k_n}) \xrightarrow{D} \chi_1^2$. □

C. SIMULATION STUDY ON THE MAXIMUM TEST

We present an example on the maximum test that is also considered in DiCiccio and Romano (2022) (see their Examples 2.1 and 4.1). Different from the applications to subbagging and sub-random forest as discussed in Section 3 of the article, in this example, the associated kernel function of the infinite-order U-statistic (i.e., the maximum test statistic) can be expressed in a closed form. As a result, we can verify the conditions of Theorem 1 explicitly.

Assume $X_1, \ldots, X_n$ are independent normal random variables with mean μ_i and variance 1. Suppose one is interested in testing whether $H_0 : \mu_1 = \cdots = \mu_n = 0$ against H_a : some of the μ_i's are nonzero. As noted in DiCiccio and Romano (2022), if the index set i's for $\mu_i \neq 0$ is sparse, a better approach than the uniformly most powerful invariant (UMPI) level- α test is to perform a maximum test. The maximum test has found its applications in the econometrics of structural breaks (Andrews, 1993), the detection of financial bubbles (Phillips, Wu, and Yu, 2011), and the detection of high-dimensional signal (Cai, Liu, and Xia, 2014). In these applications, the maximum test is preferred due to the sparsity of signals.

For ease of discussion, the problem can be formulated as follows: for some $\epsilon > 0$ (ϵ can be either fixed or dependent on n. In the latter case, $\epsilon = \epsilon_n \to 0$ as $n \to \infty$.) and $k = k_n$, we want to identify the maximum level- α test such that at least k_n of the μ_i's satisfy $\mu_i > \epsilon$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

Then, the associated test statistic can be written as an infinite-order U-statistic with a kernel function given by

$$\phi_n = \exp\left(\epsilon \sum_{i=1}^{k_n} X_i\right)$$

for some small $\epsilon > 0$. One rejects the null hypothesis if the U-statistic takes a large value. Under the null hypothesis, $\theta_{k_n} = E(U_n) = \exp(\epsilon^2 k_n/2)$. In addition, it can be derived that

$$\zeta_{1,k_n} = e^{\epsilon^2 k_n}\left(e^{\epsilon^2} - 1\right),$$
$$\zeta_{k_n,k_n} = e^{2\epsilon^2 k_n} - e^{\epsilon^2 k_n}.$$

For any fixed $\epsilon > 0$, if $(\log n)^{-1} k_n \to 0$ as $n \to \infty$, then ζ_{1,k_n} is bounded away from zero for all large n, and

$$\frac{k_n}{n}\left(\frac{\zeta_{k_n,k_n}}{k_n \zeta_{1,k_n}} - 1\right) = \frac{k_n}{n}\left[\frac{e^{\epsilon^2 k_n} - 1}{k_n(e^{\epsilon^2} - 1)} - 1\right] \to 0.$$

The latter result recovers the limiting normal distribution of U_n based on the central limit theorem derived in Peng et al. (2022). The detailed proof for this scenario can be found in Lemma 7 in the Supplementary Material. On the other hand, if $\epsilon_n \to 0$ as $n \to 0$, Lemma 8 in the Supplementary Material states the conditions to ensure that ζ_{1,k_n} is bounded away from zero and a limiting normal distribution of U_n can be established in the given context. In this case, because ζ_{1,k_n} is bounded below by a positive constant, the Lindeberg condition (see equation (H1) in Mayala et al., 2026), namely, $n\zeta_{1,k_n} \to \infty$ as $n \to \infty$, holds. Therefore, under the theoretical framework of Mayala et al. (2026), the asymptotic normality of U_n is still valid.

Furthermore, it can be confirmed that all the other conditions in Theorem 1 are met. For example, the technical conditions for Lemma 6 in the Supplementary Material can be verified as presented below. First, we can show that

$$E\left(|\phi_n(X_1, \ldots, X_{k_n})|^{\iota_n}\right) = \prod_{i=1}^{k_n} E\left[\exp(\iota_n \epsilon X_i)\right] = \exp(k_n \epsilon^2 \iota_n^2/2),$$

and thus

$$\frac{E\left(|\phi_n(X_1, \ldots, X_{k_n})|^{\iota_n}\right)}{n^{\delta\iota_n/2 - k_n - 1}} = n^{k_n+1}\left(\frac{e^{k_n\epsilon^2\iota_n}}{n^{\delta}}\right)^{\iota_n/2}.$$

If we choose $\iota_n \le \kappa \log n$ for some $\kappa > 0$, then it can be shown that

$$\frac{E\left(|\phi_n(X_1, \ldots, X_{k_n})|^{\iota_n}\right)}{n^{\delta\iota_n/2 - k_n - 1}} \le n^{(\kappa k_n \epsilon^2 - \delta)\iota_n/2 + k_n + 1},$$

and thus the choice of $\iota_n = \frac{2(k_n+1)}{\delta - \kappa k_n \epsilon^2}$ leads to condition (4) in Lemma 7 in the Supplementary Material. To guarantee the existence and compatibility of such a choice of ι_n, it suffices to require that

$$\frac{2(k_n+1)}{\delta - \kappa k_n \epsilon^2} \le \kappa \log n.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE C1. Simulation results on maximum test statistic ($\epsilon = 0.01$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		(seconds)
$0.05n$	1.000	0.987	1.014	0.027	0.005	0.005	0.990	5.284
$0.10n$	1.001	0.973	1.029	0.056	0.010	0.011	0.990	5.465
$0.20n$	1.001	0.944	1.059	0.115	0.020	0.022	0.996	6.172
$0.50n$	1.003	0.838	1.169	0.331	0.050	0.062	0.998	7.474
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		(seconds)
$0.05n$	1.000	0.990	1.011	0.021	0.005	0.005	0.962	50.418
$0.10n$	1.001	0.980	1.021	0.041	0.01	0.01	0.962	52.419
$0.20n$	1.001	0.960	1.043	0.083	0.019	0.021	0.960	58.587
$0.50n$	1.003	0.897	1.111	0.214	0.045	0.056	0.966	54.983
		Number of trees = 2,000						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		(seconds)
$0.05n$	1.000	0.990	1.010	0.020	0.005	0.005	0.962	94.175
$0.10n$	1.001	0.981	1.021	0.040	0.010	0.010	0.954	101.786
$0.20n$	1.001	0.961	1.042	0.081	0.019	0.021	0.960	110.261
$0.50n$	1.003	0.901	1.106	0.205	0.045	0.055	0.956	113.298

Since $k_n\epsilon^2 = O(1)$, there exists a $c_2 > 0$ such that $k_n\epsilon^2 \le c_2$ for all large n, and, as a result,

$$\frac{2(k_n+1)}{\delta - \kappa k_n \epsilon^2} \le \frac{2(k_n+1)}{\delta - \kappa c_2}$$

holds for all large n. Therefore, it suffices to require that

$$\frac{2(k_n+1)}{\delta - \kappa c_2} \le \kappa \log n.$$

Let $\kappa = \delta/(2c_2)$, which minimizes the restriction on k_n, then equation (4) in the Supplementary Material holds if

$$k_n \le \frac{\delta^2}{8c_2}\log n - 1.$$

Since $\delta \in (0,1)$ can be chosen arbitrarily close to one, it suffices to require that

$$k_n < \frac{1}{8c_2}\log n - 1.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

TABLE C2. Simulation results on maximum test statistic ($\epsilon = 0.1$)

		Number of trees = 100						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		(seconds)
$0.05n$	1.025	0.892	1.167	0.275	0.045	0.063	0.992	5.253
$0.10n$	1.051	0.765	1.355	0.590	0.079	0.146	0.994	5.428
$0.20n$	1.105	0.472	1.769	1.297	0.126	0.368	0.984	5.708
$0.50n$	1.284	−0.826	3.401	4.227	0.577	1.770	0.958	6.969
		Number of trees = 1,000						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		
$0.05n$	1.025	0.924	1.133	0.209	0.045	0.057	0.956	47.468
$0.10n$	1.051	0.841	1.273	0.432	0.083	0.127	0.960	51.003
$0.20n$	1.105	0.655	1.574	0.919	0.133	0.315	0.954	57.053
$0.50n$	1.284	−0.085	2.664	2.749	0.161	1.372	0.886	71.064
		Number of trees = 2,000						
		Mean		Interval	SD		Coverage	Time
k_n	θ_{k_n}	LB	UB	Width	LB	UB		
$0.05n$	1.025	0.926	1.131	0.205	0.045	0.057	0.960	88.829
$0.10n$	1.051	0.845	1.268	0.423	0.083	0.127	0.956	108.075
$0.20n$	1.105	0.667	1.562	0.895	0.134	0.312	0.956	110.372
$0.50n$	1.284	−0.030	2.609	2.639	0.130	1.338	0.882	135.783

Hence, the conditions for Lemma 7 in the Supplementary Material hold, and so does Theorem 1, provided that k_n is of a desired order of n. Further details on verifying the conditions of Theorem 1 can be found in Section 2 of the Supplementary Material. In addition, the simulation results in Section 3 also suggest that the performance of the developed methodology is quite robust, even if the condition on k_n's growth rate is not strictly satisfied.

We generated $R = 500$ independent samples of size $n = 100$ under H_0, that is, $X_1, \ldots, X_n$ are i.i.d. $N(0, 1)$. For each of the simulated dataset, we approximated the complete infinite-order U-statistic with B random subsets of size k_n, where B was taken to be either 100, 1,000, or 2,000, and $k_n = 0.05n, 0.10n, 0.20n$, or $0.50n$. We set $\epsilon = 0.01$ or 0.1 for illustration. Then, we constructed 95% CIs for θ_{k_n}, under the jackknife empirical likelihood framework, based on each dataset. Tables C1 and C2 summarize the simulated mean and standard deviation of the CIs, as well as the coverage probability out of 500 iterations. In addition, we also reported the computational time in seconds for constructing a single CI on average.

When the theoretical conditions are satisfied, the results demonstrate that the coverage probability of the JEL-based CI approaches the nominal level 0.95 as B increases. In addition, the performance of the proposed method is pretty robust to the deviation from the constraint on k_n. For instance, when k_n increases to $0.2n$, a level that does not fully

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

satisfy the required order of k_n in Theorem 1, the coverage probability of the JEL-based CI is still fairly close to 0.95 when $B = 1,000$ or $2,000$. Only when $\epsilon = 0.1$ and $k_n = 0.5n$, the coverage probability of the CI is noticeably below the target level.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at https://doi.org/10.1017/S0266466626100681.

FUNDING STATEMENT

The work of Y.Z. is partially supported by the Simons Foundation (MPS-TSM-00638679).

COMPETING INTEREST STATEMENT

The authors declare that no competing interests exist.

REFERENCES

Alemdjrodo, K., & Zhao, Y. (2020). New empirical likelihood inference for the mean residual life with length-biased and right-censored data. *Journal of Nonparametric Statistics*, 32, 1029–1046.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856.

Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, 85(3), 693–734.

Baldwin, L. H., Marshall, R. C., & Richard, J. F. (1997). Bidder collusion at forest service timber sales. *Journal of Political Economy*, 105, 657–699.

Bonhomme, S., & Jolivet, G. (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics*, 24(5), 763–795.

Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. (1984). *Classification and regression trees*. (1st ed.) Wadsworth.

Bühlmann, P. (2003). *Bagging, subagging and bragging for improving some prediction algorithms*. Eidgenössische Technische Hochschule (ETH).

Bühlmann, P., & Yu, B. (2002). Analyzing bagging. *Annals of Statistics*, 30(4), 927–961.

Cai, T. T., Liu, W., & Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(2), 349–372.

Chen, J., Variyath, A. M., & Abraham, B. (2008). Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics*, 17(2), 426–443.

Chen, S., & Hall, P. (1993). Smoothed empirical likelihood confidence intervals for quantiles. *Annals of Statistics*, 21, 1166–1181.

Chen, S. X., & Keilegom, I. V. (2009). A review on empirical likelihood methods for regression. *TEST*, 18, 415–447.

Chiang, H. D., Matsushita, Y., & Otsu, T. (2025). Empirical likelihood for random forests and ensembles. Preprint, arXiv:2511.13934.

Claeskens, G., Jing, B. Y., Peng, L., & Zhou, W. (2003). Empirical likelihood confidence regions for comparison distributions and roc curves. *The Canadian Journal of Statistics*, 31, 173–190.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

DiCiccio, C., & Romano, J. (2022). CLT for *U*-statistics with growing dimension. *Statistica Sinica*, 32, 323–344. https://doi.org/10.5705/ss.202020.0048

Efron, B., & Stein, C. (1981). The jackknife estimate of variance. *Annals of Statistics*, 9(3), 586–596.

Frees, E. W. (1989). Infinite order *U*-statistics. *Scandinavian Journal of Statistics*, 16(1), 29–45.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *Annals of Statistics*, 19(1), 1–67.

Gómez, V., & Maravall, A. (1998). Seasonal adjustment and signal extraction in economic time series. Technical report, Banco de España.

Haile, P. A. (2001). Auctions with resale markets: An application to us forest service timber sales. *American Economic Review*, 91, 399–427.

Haile, P. A., & Tamer, E. (2003). Inference with an incomplete model of English auctions. *Journal of Political Economy*, 111, 1–51.

Hodges, J. L., & Lehmann, E. L. (1963). Estimation of location based on ranks. *Annals of Mathematical Statistics*, 34(2), 598–611.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19(3), 293–325.

Jaeckel, L. A. (1972). The infinitesimal jackknife. Technical report, Bell Telephone Laboratories.

Jing, B., Yuan, J., & Zhou, W. (2009). Jackknife empirical likelihood. *Journal of American Statistical Association*, 104(487), 1224–1232.

Lazar, N. A. (2003). Bayesian empirical likelihood. *Biometrika*, 90, 319–326.

Lazar, N. A. (2021). A review of empirical likelihood. *Annual Review of Statistics and Its Application*, 8, 329–344.

Lee, A. J. (1990). *U-statistics: Theory and practice*. Marcel Dekker.

Lehmann, E. L., & Romano, J. P. (2005). *Testing statistical hypotheses*. (3rd ed.) Springer.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.

Liu, N., & Luo, Y. (2007). A nonparametric test for comparing valuation distributions in first-price auctions. *International Economic Review*, 58(3), 857–888.

Liu, P., & Zhao, Y. (2023). A review of recent advances in empirical likelihood. *Wiley International Reviews Computational Statistics*, 15(3). https://doi.org/10.1002/wics.1599

Liu, X., & Zhao, Y. (2012). Semi-empirical likelihood inference for the roc curve with missing data. *Journal of Statistical Planning and Inference*, 142, 3123–3133.

Mayala, M., Scornet, E., Tillier, C., & Wintenberger, O. (2025). Asymptotic normality of infinite centered random forests-application to imbalanced classification. Preprint, arXiv:2506.08548.

Mayala, M., Wintenberger, O., Tillier, C., & Dombry, C. (2026). Infinite random forests for imbalanced classification tasks. *Electronic Journal of Statistics*, 20(1), 2314–2365.

Mentch, L., & Hooker, G. (2016). Quantifying uncertainty in random forest via confidence intervals and hypothesis tests. *Journal of Machine Learning Research*, 17, 1–41.

Mentch, L., & Hooker, G. (2017). Formal hypothesis tests for additive structure in random forest. *Journal of Computational and Graphical Statistics*, 26(3), 589–597.

Nordman, D. J., & Lahiri, S. N. (2014). A review of empirical likelihood methods for time series. *Journal of Statistical Planning and Inference*, 155, 1–18.

Nowotarski, J., & Weron, R. (2015). Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Computational Statistics*, 30, 791–803.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75, 237–249.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions. *Annals of Statistics*, 18, 90–120.

Owen, A. B. (1991). Empirical likelihood for linear models. *Annals of Statistics*, 19, 1725–1747.

Peng, W., Mentch, L., & Stefanski, L. (2021). Bias, consistency, and alternative perspectives of the infinitesimal jackknife. *Statistica Sinica*. https://doi.org/10.5705/ss.202022.0131

Peng, W., Coleman, T., & Mentch, L. (2022). Rates of convergence for random forests via generalized *U*-statistics. *Electronic Journal of Statistics*, 16(1), 232–292.

Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s NASDAQ: When did periodic exuberance constitute bubbles? *International Economic Review*, 52(1), 201–226.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681

Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. Springer.
Qin, G., & Zhou, X. H. (2006). Empirical likelihood inference for the area under the ROC curve. *Biometrics*, 62, 613–622.
R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
Shi, X. (1984). The approximate independence of jackknife pseudo-values and the bootstrap methods. *Journal of Wuhan Institute of Hydraulic and Electric Engineering*, 2, 83–90.
Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of American Statistical Association*, 113(523), 1228–1242.
Wager, S., Hastie, T., & Efron, B. (2014). Confidence intervals for random forests: The jackknife and the infinitesimal jackknife. *The Journal of Machine Learning Research*, 15(1), 1625–1651.
Wang, Q., & Cai, X. (2021). An efficient variance estimator for cross-validation under partition-sampling. *Statistics*, 55(3), 660–681. https://doi.org/10.1080/02331888.2021.1943393
Wang, Q., & Lindsay, B. (2017). Pseudo-kernel method in U-statistic variance estimation with large kernel size. *Statistica Sinica*, 27(3), 1155–1174.
Wang, Q., & Lindsay, B. G. (2014). Variance estimation of a general U-statistic with application to cross-validation. *Statistica Sinica*, 24(3), 1117–1141.
Wang, Q., & Wei, Y. (2022). Quantifying uncertainty of subsampling-based ensemble methods under a U-statistic framework. *Journal of Statistical Computation and Simulation*, 92(17), 3706–3726.
Wu, C. (2005). Algorithms and r codes for the pseudo empirical likelihood method in survey sampling. *Statistica Sinica*, 31, 239–243.
Xue, L., & Zhu, L. (2012). Empirical likelihood in some nonparametric and semiparametric models. *Statistics and Its Interface*, 5, 367–378.
Yang, H., & Zhao, Y. (2012). Smoothed empirical likelihood for roc curves with censored data. *Journal of Multivariate Analysis*, 109, 254–263.
Yang, H., Yau, C., & Zhao, Y. (2014). Smoothed empirical likelihood inference for the difference of two quantiles with right censoring. *Journal of Statistical Planning and Inference*, 146, 95–101.
Zaman, F., & Hirose, H. (2009). Effect of subsampling rate on subbagging and related ensembles of stable classifiers. In Chaudhury, S., Mitra, S., Murthy, C. A., Sastry, P. S., Pal, S. K. (eds) *Pattern recognition and machine intelligence*. PReMI 2009. (pp. 44–49). Lecture Notes in Computer Science, vol 5909. Springer, Berlin, Heidelberg.
Zhou, W., & Jing, B. Y. (2003a). Adjusted empirical likelihood method for quantiles. *Annals of the Institute of Statistical Mathematics*, 55, 689–703.
Zhou, W., & Jing, B. Y. (2003b). Smoothed empirical likelihood confidence intervals for the difference of quantiles. *Statistica Sinica*, 13, 83–95.
Zhou, Z., Mentch, L., & Hooker, G. (2021). V-statistics and variance estimation. *Journal of Machine Learning Research*, 22(287), 1–48.
Zhu, L., & Xue, L. (2006). Empirical likelihood confidence regions in a partially linear single-index model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68, 549–570.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 23:39:53, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100681