Article

Sacrificial Pseudoreplication in LEED Cross-Certification Strategy Assessment: Sampling Structures

Svetlana Pushkar

Department of Civil Engineering, Ariel University, Ariel 40700, Israel; svetlanap@ariel.ac.il; Tel.: +972-3-9066-410

check for updates

Abstract: The study aims to suggest sampling structures to avoid sacrificial pseudoreplication in the evaluation of Leadership in Energy and Environmental Design (LEED)-certified projects. The sampling includes two structures that exclude sacrificial pseudoreplication and one structure that leads to sacrificial pseudoreplication: (i) The state is the sampling frame in which LEED projects are treated as primary sampling units; (ii) The US is the sampling frame, the state is the primary sampling unit in which LEED projects are treated as evaluation units; and (iii) The US is the sampling frame in which LEED projects are pooled from different states and treated as primary sampling units. The three sampling structures are applied to the evaluation of the Silver-to-Gold cross-certification performances of LEEDv3 for new construction and LEEDv3 for existing buildings. The same cross-certification strategy was revealed if either structure (i) or structure (ii) was applied, while it was poorly estimated and misinterpreted if structure (iii) was applied, i.e., sacrificial pseudoreplication had occurred.

Keywords: LEED cross-certification strategies; sacrificial pseudoreplication; sampling primary units; evaluation units; sampling frame

1. Introduction

We first start with statistical terminology necessary for explaining the study (Table 1).

Table 1. Statistical terminology.

Terminology	Reference
The sampling frame is a collection of primary sampling units "accessible for sampling in the population of interest"	[1]
The primary sampling unit is statistically independent of other primary sampling units within the same sampling frame	[1]
Evaluation units are nested in the primary sampling unit and are statistically dependent of those primary sampling unit	[1]
Simple pseudoreplication occurs when multiple measurements are performed on individual primary sampling units and treated statistically as if each represents a separate primary sampling unit	[2]
Temporal pseudoreplication occurs when performing on individual primary sampling units multiple measurements, which are performed over time and treated statistically as if each represents a different primary sampling unit	[2]
Sacrificial pseudoreplication occurs when the evaluation units from different primary sampling units are pooled and each evaluation unit is treated as a primary sampling unit	[2]

"Pseudoreplication is a serious type of statistical error that is unfortunately common in all the sciences. It was originally defined in the context of manipulative experiments but can also occur in observational studies" [3]. Leadership in Energy and Environmental Design (LEED)-certified projects data analysis is an observational study [4]. Therefore, LEED project data analysis should be considered regarding the problem of pseudoreplication.

The term pseudoreplication was coined by Hurlbert [5] and defined as " ... a particular combination of experimental design (or sampling) and statistical analysis which is inappropriate for testing the hypothesis of interest". Pseudoreplication typically occurs when the number of observations or the number of data points are treated inappropriately as independent replicates [5–7]. Pseudoreplication contains several types: simple, temporal, and sacrificial. One of the most common types is sacrificial pseudoreplication [6]. This study focuses on the sacrificial pseudoreplication that arises when the evaluation units from different primary sampling units are pooled to obtain "[but not universal] ... exaggeration of both the strength of the evidence for a true difference between treatments and of the precision with which any difference that does exist has been estimated" [2]. Recently, when Wu et al. [8] and Wu et al. [9] analyzed the LEED project data, they utilized pooling procedures that achieved extremely low P values. An evaluation of LEED project data and different sampling structures (without and with pooling) is discussed below.

A collection of Certified, Silver, Gold, and Platinum LEED-certified projects in the US Green Building Council (USGBC) website [10] is a unique database that has been used by researchers to perform retrospective analyses of the LEED-certified building certification, category, and cross-certification performances [4,8,9,11,12]. Such analyses produce feedback that can be a useful information for experts who oversee further improvements in the seven LEED categories such as sustainable sites (SS), water efficiency (WE), energy and atmosphere (EA), material and resources (MR), indoor environmental quality (EQ), innovation in design (ID), and regional priority (RP).

In three recent cases [4,8,9], a single-unit design structure of the analysis according to the term originally suggested by Hurlberd [7] (p. 652) was used. In these studies, all LEED projects were treated as statistically independent units using a non-parametric Wilcoxon–Mann–Whitney (WMW) test [13]. The WMW test was chosen, because LEED data are related to an ordinal measurement scale.

Wu et al. [8] pooled all LEED projects within the US to determine cross-certification performances of LEED-NCv2.2-certified projects in 2007–2015. Sample sizes for both Silver and Gold LEED projects were $n_1 = 1798$ and $n_2 = 2469$, respectively [8] (p. 174, Table 8). Based on the absolute effect size (mean credit increase, [MCI]) and statistical difference (*p*-values) of the pooled data, the authors concluded that at least three categories, EA (MSI = 2.72 and $p = 0.000$), EQ (MCI = 1.50 and $p = 0.000$), and SS (MCI = 1.12 and $p = 0.000$), were propelling categories for moving the projects from Silver to Gold certification [8] (p. 173, Table 6).

Wu et al. [9] recently carried out pooling of all LEED projects within the US to determine cross-certification performances of LEED-NCv3 projects in 2007–2015. Sample sizes for both Silver and Gold LEED-certified projects were $n_1 = 1310$ and $n_2 = 1201$, respectively [9] (p. 375, Table 9). Based on MCI and *p*-values of the pooled data, the authors concluded that at least three categories, EA (MCI = 5.46 and $p = 0.000$), SS (MCI = 2.57 and $p = 0.000$), and WE (MCI = 0.84 and $p = 0.000$), were propelling categories in order to move from Silver to Gold certification [9] (p. 375, Table 7).

In both studies [8,9], large sample sizes were used to determine the statistical difference via *p*-values. This can lead to statistical error, i.e., fallacy of large sample size (Hurlbert and Lombardi 2009). Hurlbert and Lombardi [14] allegorically noted: "it lurks quietly in the darkness, waiting for researchers to pass by who are too focused on obtaining adequate sample sizes. If sample sizes are too large, one may be 'in danger' of getting very low *p*-values and establishing the sign and magnitude of even small effects with too much confidence."

In this respect, Pushkar and Verbitsky [4] more recently performed pooling of all LEED projects within individual states to determine cross-certification performances of Silver-Gold LEED-NCv3 certificated projects in 10 US states (CA, $n_1 = 42$, $n_2 = 58$; FL, $n_1 = 31$, $n_2 = 11$; GA, $n_1 = 11$, $n_2 = 9$; IL,

n_1 = 13, n_2 = 19; MA, n_1 = 19, n_2 = 14; NY, n_1 = 14, n_2 = 11; OH, n_1 = 23, n_2 = 8; TX, n_1 = 27, n_2 = 9; VA, n_1 = 38, n_2 = 11; WA, n_1 = 10, n_2 = 11) in 2016. It should be noted that because LEED data is ordinal data, the median credit increase (MdCI) in the Pushkar and Verbitsky [4] was used instead of the mean credit increase (MCI) that was suggested in the Wu et al. [9] study. In particular, the EA category was propelled in six of 10 states (CA [MdCI = 10.0 and p = 0.01); FL [MdCI = 5.0 and p = 0.016]; IL [MdCI = 7.0 and p = 0.004]; MA [MdCI = 2.5 and p = 0.001]; VA [MdCI = 5.0 and p = 0.001]; and WA [MdCI = 8 and p = 0.001]; SS was propelling category in two of 10 states (WA [MdCI = 4.5 and p = 0.001]; OH [MdCI = 4.5 and p = 0.001]), and the WE was a propelling category in one of 10 states (CA [MdCI = 2.0 and p = 0.013) [4]. Thus, in the Pushkar and Verbitsky [4] study, results of EA, SS, and WE were somewhat contradictory to the results revealed in the Wu et al. [9] study.

One possible reason for such contradictory results is that the different US states have adopted different green and energy codes along with their different versions [15,16]. For example, one of the main energy regulations, ASHRAE 90.1 Standard (Energy Standard for Buildings Except Low-Rise Residential Buildings), is not under national regulation. Hence, states are willing to adopt any version of the ASHRAE 90.1; in this way, different ASHRAE 90.1 versions are adopted in different US states [15]. However, these different ASHRAE 90.1 versions that are accepted in different US states can have different influences on one of the highly weighted LEED categories, EA [4]. In addition, ASHRAE 90.1 [17] version has usually moved dynamically through the years in the same state. For example, in CA: 2011—ASHRAE 90.1-2007, 2014—ASHRAE 90.1-2010, and 2017—ASHRAE 90.1-2013; and in OH: 2011—ASHRAE 90.1-2004, 2014—ASHRAE 90.1-2007, and 2017—ASHRAE 90.1-2010 [17]. Thus, an adopted ASHRAE 90.1 version can be changed one time in three years in the same state, while typical duration of the building construction is approximately two years. For this reason, to minimize the evaluation of LEED projects in the same state with different ASHRAE 90.1 versions, we propose that the time collection of the LEED projects must not exceed one year. Therefore, due to both the spatial and temporal influences of ASHRAE 90.1 energy code on EA category, LEED projects certified in different years cannot be pooled in one US group as suggested by Wu et al. [8] and Wu et al. [9].

In the context of statistical terminology, a single-unit design allows three methods of analysis: (i) If a sampling frame contains primary sampling units, and the primary sampling unit does not contain evaluation units, then primary sampling units are treated as statistically independent units; (ii) If a sampling frame contains primary sampling units, and a primary sampling unit contains evaluation units, then evaluation units must be averaged within each primary sampling unit, and then primary sampling units are treated as statistically independent units. However, (iii) If a sampling frame contains primary sampling units that contain evaluation units, and evaluation units are pooled from two or more primary sampling units and treated as primary sampling units, it will lead to the problem of sacrificial pseudoreplication [5]. Ignoring sampling structure of the collected data leads to sacrificial pseudoreplication and can lead to "artificially inflated degrees of freedom, giving the illusion of having a more powerful test than the data support" [1].

In our case, pooling LEED projects from different states of the US and treating them as primary sampling units can lead to erroneous conclusions about cross-certification strategies [4]. Thus, based on the Kozlov and Hurlbert [18] study, to avoid erroneous conclusions about cross-certification strategies, the following logical structure should be accepted: any two LEED projects in the same state share more similar "green building policy" conditions compared to any two LEED projects from the different states. It is suggested that states in the US are "natural" groups or clusters. If a single-design structure in the present study is used, then the primary sampling units are only used in the inferential statistical analysis.

The goal of the present study is to correct the statistical malaise over evaluating and interpreting the LEED certification process. Correcting statistical analysis included two possible evaluation structures: (i) individual state sampling frame analysis and (ii) the US sampling frame analysis. These two structures were contrasted to a third evaluation structure: (iii) sacrificial pseudoreplication analysis.

Application of the three evaluation levels was demonstrated on a case study of LEED-NCv3 and LEED-EBv3 projects certified in four US states in 2016. For both certification schemes, only the most popular Silver and Gold certifications [4] were considered. As a result, the cross-certification analyses were performed so as to reveal which LEED categories are Silver-to-Gold propelling. This will serve as valuable information for understanding the strategies employed by green building practitioners in moving projects toward higher certification level.

2. Materials and Methods

2.1. Design of the Study

Only Silver and Gold projects of LEED-NCv3 and LEED-EBv3 in the US in 2016 were considered for the analysis, which was performed in the following three consequent steps:

1. In this analysis, each of the three structures was evaluated using either approximate or exact WMW tests: (i) individual state sampling frame analysis used an approximate WMW test, (ii) the US sampling frame analysis used an exact WMW test, and (iii) sacrificial pseudoreplication analysis used an approximate WMW test. To perform an approximate WMW test [11], the balanced sample sizes for Silver (n_1) and for Gold (n_2) projects were ($n_1 = n_2 \geq 9$). For states with suitable sample sizes for LEED projects, an approximate WMW test was used. Then, LEED projects from each state were randomly selected. To perform an exact WMW test, minimum sample size was ($n_1 = n_2 = 4$) [19] (p. 19). To obtain primary sampling units, the median Silver and Gold project in each state was computed and collected.
2. USGBC scorecards of the randomly selected projects were retrieved from the USGBC website new construction building directory [20] and existing building directory [21]. Then, information on the retrieved projects regarding the awarded points in the six main categories (SS, WE, EA, MR, EQ, and ID) was accumulated. Eventually the RP points were also accumulated and redistributed among the five relevant main categories.
3. For each of the three structures, (i) individual state sampling frame analysis (the state is sampling frame in which LEED projects are treated as primary sampling units), (ii) the US sampling frame analysis (the US is the sampling frame, the state is the primary sampling unit in which LEED projects were treated as evaluation units), and (iii) sacrificial pseudoreplication analysis (the US is the sampling frame in which LEED projects were pooled from different states and treated as primary sampling units), robust statistical analysis for comparison between the randomly selected Silver- and Gold-certified projects in ordinal scale was performed.

2.2. Statistical Analysis

LEED data are presented in ordinal scales. Based on that fact, the following nonparametric tests, Cliff's δ [22] and WMW test [13], were used to compare the two unpaired groups. The data are presented as the median ± interquartile range (IQR, 25th–75th percentile).

Cliff's δ was used to measure the substantive significance (effect size) between two unpaired groups. Cliff's δ [22] (p. 495) is expressed as

$$\delta = \#(x_1 > x_2) - \#(x_1 < x_2)/(n_1 n_2)$$

in which x_1 and x_2 are scores within group 1 and group 2, respectively; n_1 and n_2 are the sizes of the sample groups, group 1 and group 2; and # indicates the number of times.

Cliff's δ ranges between −1 and + 1; positive (+) values indicate that group 1 is larger than group 2, 0 indicates equality or overlap, and negative (−) values indicate that group 2 is larger than group 1 [22].

The effect size is considered (i) negligible if $|\delta| < 0.147$, (ii) small if $0.147 \leq |\delta| < 0.33$, (iii) medium if $0.33 \leq |\delta| < 0.474$, or (iv) large if $|\delta| \geq 0.474$ [23]. According to Cohen [24] (p. 156), "a medium effect is visible to the naked eye of a careful observer. A small effect is noticeably smaller than medium

but not so small as to be trivial. A large effect is the same distance above the medium as small is below it." It should be noted that the effect size is not "iron-clad criteria" [25] but is only a general rule of thumb that might be followed in the absence of knowledge of the area [26].

WMW test. The WMW test was used to determine statistical difference (*p*-value) between two unpaired groups. It should be noted that WMW tests could be applied in two forms: approximate form or extract form [27]. If sample sizes were $n_1 = n_2 \geq 9$, then an approximate WMW test was used [28] (p. 56). Mann and Whitney [13] (p. 50) noted that when sample sizes achieve $n_1 = n_2 = 8$, then "at this point the distribution is almost normal". If sample sizes were $n_1 = n_2 = 4$, an exact WMW test was used [19] (p. 19). In both tests, a two-tailed *p*-value was applied. Only balanced sample size cases were studied.

Neo-Fisherian significance assessments. In the current study, for standard types of significance assessment, the hybrid of the Paleo–Fisherian and Neyman–Pearsonian paradigms [i.e., null hypothesis significance tests (NHST)] are replaced by a neo-Fisherian assessment, as recommended by Hurlbert and Lombardi [14,29]. The neo-Fisherian paradigm (1) does not fix α, (2) does not describe *p*-values as 'significant' or 'nonsignificant', (3) does not accept null hypotheses based on high *p*-values but only suspends judgment, (4) interprets significance tests according to "three-valued logic", and (5) presents effect size information in conjunction with significance tests. Analyses conducted under this paradigm are termed neo-Fisherian significance assessments (NFSAs) [14]. NFSAs are used to interpret the signs and magnitudes of the statistical effects [14]. Based on NFSAs, precise *P*-values were evaluated and shown according to a three-valued logic as follows: "it seems to be positive" (i.e., there seems to be a difference between group 1 and group 2), "it seems to be negative" (i.e., there does not seem to be a difference between group 1 and group 2), and "judgment is suspended" regarding the difference between group 1 and group 2 [14,29].

3. Results

The number of the selected states with the total number of LEED projects and number of randomly selected LEED projects is presented in Table 2. Ordinal font size numbers show the total number of LEED's projects. Numbers in parentheses display randomly selected LEED's projects to perform the balanced non-parametric statistical analysis. The resulting state number with Silver- and Gold-certified projects was $n_{1states} = n_{2states} = 4$ and included the states of California (CA), Washington (WA), Virginia (VA), and Texas (TX) (Table 2). It was sufficient to perform an exact WMW test in the (ii) US sampling frame analysis [19]. The number of Silver- and Gold-selected projects $n_{1projects} = n_{2projects}$ was different for each of the four states, but it was sufficient to perform an approximate WMW test in (i) individual state sampling frame analysis and (iii) sacrificial pseudoreplication analysis.

Table 2. The distribution of LEED-NCv3 and LEED-EBv3 Silver- and Gold-certified projects within four US states in 2016.

State	LEED-NCv3: Sample Size (n)		LEED-EBv3: Sample Size (n)	
	Silver	Gold	Silver	Gold
California (CA)	42 (12)	58 (12)	24 (12)	81 (12)
Virginia (VA)	38 (11)	11 (11)	12 (12)	12 (12)
Washington (WA)	10 (10)	11 (10)	11 (9)	9 (9)
Texas (TX)	27 (9)	9 (9)	12 (10)	10 (10)

Cross-certification performances, namely, from Silver to Gold certification level in both the LEED-NCv3 and LEED-EBv3 in the four states, CA, VA, WA, and TX, were statistically evaluated under (i) individual state sampling frame analysis (an approximate WMW test), (ii) the US sampling frame analysis (an exact WMW test), and (iii) sacrificial pseudoreplication (an approximate WMW test).

3.1. LEED-NC 2009 Certified Projects

The individual state sampling frame analysis. Strategies of each of the four states toward moving from Silver to Gold for LEED-NCv3 are presented in Table 3. According to these strategies, the EA is the most accepted propelling category, which was involved in all four states. These results confirm the popularity of the EA propelling category in cross-certification performance of LEED-NCv3-certified projects, which was revealed by other studies [4,9]. In particular, Pushkar and Verbitsky [4], who analyzed LEED-NCv3 certified projects in 2016, revealed that EA was the most propelling category in Florida (FL), Illinois (IL), and Massachusetts (MA), while EQ was the most propelling category in New York (NY).

Table 3. The individual state-sampling frame analysis: median $\pm$ interquartile range (IQR, 25th–75th percentile) of six categories of LEED-NCv3 in four US states.

LEED Category	State	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
			Silver	Gold			
SS	CA	12	18.5 ± 5.5	22.5 ± 5.5	-0.347	0.157	suspended
	VA	11	18.0 ± 4.8	21.0 ± 6.8	-0.273	0.293	negative
	WA	10	18.5 ± 2.0	23.0 ± 3.0	-0.850	0.001	positive
	TX	9	15.0 ± 7.7	17.0 ± 7.3	-0.198	0.507	negative
WE	CA	12	5.0 ± 1.0	5.0 ± 2.0	-0.188	0.451	negative
	VA	11	6.0 ± 3.5	7.0 ± 3.5	-0.050	0.869	negative
	WA	10	5.0 ± 2.0	4.5 ± 1.0	0.060	0.850	negative
	TX	9	6.0 ± 2.7	6.0 ± 3.5	-0.247	0.401	negative
EA	CA	12	9.0 ± 4.5	17.5 ± 10.0	-0.757	0.002	positive
	VA	11	10.0 ± 3.8	15.0 ± 6.5	-0.785	0.002	positive
	WA	10	8.5 ± 5.0	16.0 ± 2.0	-0.930	0.0005	positive
	TX	9	12.0 ± 10.3	14.0 ± 5.0	-0.296	0.310	negative
MR	CA	12	5.0 ± 1.0	5.5 ± 2.0	-0.097	0.707	negative
	VA	11	7.0 ± 1.8	7.0 ± 2.0	0.025	0.948	negative
	WA	10	5.0 ± 1.0	5.0 ± 2.0	0.180	0.519	negative
	TX	9	6.0 ± 3.0	7.0 ± 1.5	-0.074	0.825	negative
EQ	CA	12	9.5 ± 3.5	10.5 ± 3.5	-0.014	0.977	negative
	VA	11	9.0 ± 3.8	10.0 ± 3.0	-0.124	0.646	negative
	WA	10	9.5 ± 2.0	10.0 ± 2.0	-0.430	0.112	suspended
	TX	9	9.0 ± 1.0	12.0 ± 1.8	-0.654	0.022	positive
ID	CA	12	3.5 ± 3.0	5.0 ± 1.0	-0.396	0.105	suspended
	VA	11	4.0 ± 3.0	5.0 ± 1.8	-0.355	0.167	suspended
	WA	10	5.0 ± 2.0	4.5 ± 2.0	0.080	0.791	negative
	TX	9	4.0 ± 2.3	5.0 ± 1.3	-0.321	0.268	negative

SS, EQ, and ID are moderately accepted propelling categories (Table 3); these were involved in two of the four states. Again, the SS and EQ popularity for Silver–Gold moving was revealed early by Wu et al. [9], who outlined SS as one of the main propelling categories and by Pushkar and Verbitsky [4], who named both SS and EQ as the moderate propelling categories. ID was also reported as a somewhat moderate propelling category for Silver–Gold cross-certification performance [9].

According to the present research, WE and MR were not involved in any of the four states in Silver–Gold performance (Table 3). These two categories were noted as the worst-performing categories in LEED-NCv3 cross-certification performance [4,9]. As was explained by Wu et al. [9], there is a difficulty in accessing MR points due to a low possibility of reducing construction material consumption.

The US sampling frame analysis. Total tendency in the strategy employed by the four states applying the US sampling frame analysis toward moving from Silver to Gold for LEED-NCv3 is presented in Table 4. According to the revealed strategies, the EA and EQ were the most accepted

propelling categories, SS and ID were the moderately accepted propelling categories, and WE and MR were not accepted propelling categories at all. In general, the results of the US sampling frame analysis (Table 4) match the results of the individual state sampling frame analysis (Table 3). The only exception is the EQ category. In the individual state sampling frame analysis (Table 3), EQ was grouped with the moderately accepted propelling categories (SS, EQ, and ID), and in the US sampling frame analysis (Table 4), it was grouped with the most accepted propelling categories (EA and EQ).

Table 4. The US sampling frame analysis: median $\pm$ interquartile range (IQR, 25th–75th percentile) of six categories of LEED-NCv3 in four US states.

LEED Category	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
		Silver	Gold			
SS	4	18.3 ± 2.0	21.8 ± 3.8	-0.625	0.171	suspended
WE	4	5.5 ± 1.0	5.5 ± 1.8	0.000	1.000	negative
EA	4	9.5 ± 2.3	15.5 ± 2.3	-1.000	0.029	positive
MR	4	5.5 ± 1.5	6.3 ± 1.8	-0.250	0.742	negative
EQ	4	9.3 ± 0.5	10.3 ± 1.3	-1.000	0.029	positive
ID	4	4.0 ± 0.8	5.0 ± 0.3	-0.688	0.142	suspended

The sacrificial pseudoreplication analysis. Total tendency in strategy employed by the four states on the basis of the analysis when projects from different states were pooled in one US sampling frame for LEED-NCv3 is presented in Table 5. According to these revealed strategies, the SS, EA, EQ, and ID were the most accepted propelling categories, while WE and MR were not accepted propelling categories at all. In general, the results revealed that when projects from different states were pooled in one US sampling frame (Table 5), the results evaluated in the individual state sampling frame analysis were different (Table 3). The differently evaluated results were for SS, EQ, and ID categories: in the individual state sampling frame analysis (Table 3), these three categories were grouped under the moderately accepted propelling categories (SS, EQ, and ID), and in the analysis where different states were pooled in one US sampling frame (Table 5), these categories were grouped under the most accepted propelling categories (SS, EA, EQ, and ID).

Table 5. The sacrificial pseudoreplication analysis: median $\pm$ interquartile range (IQR, 25th–75th percentile) of six categories of LEED-NCv3 in four US states.

LEED Category	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
		Silver	Gold			
SS	42	18.0 ± 5.0	21.5 ± 7.0	-0.375	0.003	positive
WE	42	5.0 ± 3.0	5.0 ± 2.0	-0.103	0.420	negative
EA	42	9.5 ± 5.0	16.0 ± 5.0	-0.709	0.00000002	positive
MR	42	6.0 ± 2.0	6.0 ± 2.0	0.035	0.788	negative
EQ	42	9.0 ± 3.0	10.5 ± 3.0	-0.274	0.031	positive
ID	42	4.0 ± 2.0	5.0 ± 2.0	-0.296	0.020	positive

3.2. LEED-EB 2009 Certified Projects

The individual state sampling frame analysis. Strategies of each of the four states toward moving from Silver to Gold for LEED-EB 2009 are presented in Table 6. According to the results, all six categories were propelling: SS and EQ were the most accepted, and WE, EA, MR, and IO were moderately accepted propelling categories. Thus, in contrast to unemployed WE and MR in a case of new projects certified under LEED-NCv3 (Table 3), these categories were active in a case of renovated buildings certified under LEED-EBv3 (Table 6). It can be suggested that site and building restrictions are weaker in new building design and construction than in the renovation of existing buildings. Therefore, in selecting preferred strategy design, a team has more flexibility in new buildings than

in renovated buildings. As a result, in new building certification, difficult points such as MR points were not considered as a preferred Silver-Gold cross-certification strategy (Table 3), while in renovated buildings, all possible strategies were employed (Table 6).

Table 6. The individual state-sampling frame analysis: median ± interquartile range (IQR, 25th–75th percentile) of six categories of LEED-EBv3 in four US states.

LEED Category	State	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
			Silver	Gold			
SS	CA	12	12.0 ± 6.5	15.0 ± 5.5	−0.424	0.083	suspended
	VA	12	7.5 ± 6.0	14.0 ± 6.0	−0.660	0.007	positive
	WA	9	11.0 ± 3.8	18.0 ± 5.0	−0.988	0.0005	positive
	TX	10	11.0 ± 2.0	16.0 ± 5.0	−0.590	0.028	positive
WE	CA	12	6.0 ± 4.5	9.0 ± 3.0	−0.368	0.133	suspended
	VA	12	9.0 ± 7.0	8.5 ± 2.0	0.299	0.225	negative
	WA	9	7.0 ± 2.5	8.0 ± 1.5	−0.370	0.199	suspended
	TX	10	6.0 ± 2.0	8.0 ± 3.0	−0.690	0.010	positive
EA	CA	12	18.5 ± 7.5	22.5 ± 4.0	−0.604	0.013	positive
	VA	12	19.0 ± 5.0	22.0 ± 5.5	−0.479	0.049	positive
	WA	9	19.0 ± 9.8	18.0 ± 5.0	0.173	0.566	negative
	TX	10	19.0 ± 6.0	21.0 ± 7.0	−0.320	0.241	negative
MR	CA	12	3.5 ± 1.5	4.5 ± 1.0	−0.347	0.156	suspended
	VA	12	3.0 ± 2.0	4.0 ± 1.0	−0.444	0.068	suspended
	WA	9	5.0 ± 1.3	6.0 ± 2.0	−0.309	0.288	negative
	TX	10	3.0 ± 2.0	3.5 ± 1.0	−0.240	0.383	negative
EQ	CA	12	7.5 ± 2.0	8.0 ± 1.5	0.056	0.840	negative
	VA	12	6.5 ± 3.0	8.5 ± 2.0	−0.590	0.015	positive
	WA	9	6.0 ± 1.3	8.0 ± 3.0	−0.420	0.145	suspended
	TX	10	7.0 ± 2.0	8.5 ± 2.0	−0.590	0.028	positive
IO	CA	12	6.0 ± 0.5	6.0 ± 0.5	−0.021	0.953	negative
	VA	12	5.0 ± 1.5	6.0 ± 0.5	−0.583	0.016	positive
	WA	9	4.0 ± 1.3	6.0 ± 2.0	−0.321	0.267	negative
	TX	10	6.0 ± 1.0	6.0 ± 1.0	−0.030	0.939	negative

The US sampling frame analysis. Total tendency in strategy employed by the four states applying the US sampling frame analysis toward moving from Silver to Gold for LEED-EBv3 is presented in Table 7. According to the revealed strategies, the SS and EQ were the most accepted propelling categories; WE, EA, MR, and ID were the moderately accepted propelling categories. The US sampling frame analysis results (Table 7) were the same as the results evaluated with the individual state sampling frame analysis (Table 6).

Table 7. The US sampling frame analysis: median ± interquartile range (IQR, 25th–75th percentile) of six categories of LEED-EBv3 in four US states.

LEED Category	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
		Silver	Gold			
SS	4	11.0 ± 2.3	15.5 ± 2.5	−1.000	0.029	positive
WE	4	6.5 ± 2.0	8.3 ± 0.8	−0.563	0.286	suspended
EA	4	19.0 ± 0.3	21.5 ± 2.8	−0.500	0.314	suspended
MR	4	3.3 ± 1.3	4.3 ± 1.5	−0.563	0.257	suspended
EQ	4	6.8 ± 1.0	8.3 ± 0.5	−1.000	0.029	positive
IO	4	5.5 ± 1.5	6.0 ± 0.0	−0.500	0.429	suspended

It is interesting that EQ was one of the most propelling categories in both new buildings (certified under LEED-NCv3) and renovated buildings (certified under LEED-EBv3). However, an additional most propelling category was different in these schemes: EA—in LEED-NCv3, and SS—in LEED-EBv3. It should be noted that SS and EA have the same number of achievable points, 26 pt and 35 pt,

respectively, in both LEED-NCv3 and LEED-EBv3 [30,31]. However, there are differences in some inherent credits between LEED-NCv3 and LEED-EBv3 [30,31]. For example, the SS category in LEED-NC 2009 [30] includes four separate rigid sub-credits for SSc4 Alternative Transportation (12 pt) (SSc4.1 Public Transportation Access, SSc4.2 Bicycle Storage & Changing Rooms, SSc4.3 Low-Emitting & Fuel-Efficient Vehicles, and SSc4.4 Parking Capacity). However, the SS category in LEED-EB 2009 [31] includes the one flexible SSc4 Alternative Commuting Transportation credit (15 pt), which covers the issue of reduction in regular conventional commuting trips (when individual conventional automobiles are replaced with any other human-powered conveyances, for example, mass transit, walking, or bicycles). SSc4 credit is more flexible and more point-accounted in LEED-EB 2009 [31], but it is more rigid and less point-accounted in LEED-NC 2009 [30]. This is possibly the reason why SS was one of the most propelling categories in LEED-EB 2009 [31] but not in LEED-NC 2009 [30].

The sacrificial pseudoreplication analysis. Total tendency in strategy employed by the four states on basis of the analysis when projects from different states were pooled in one US sampling frame for LEED-EBv3 is presented in Table 8. According to the revealed strategies, the SS, EA, MR, EQ, and IO were the most accepted propelling categories, while WE was the moderately propelling category. The results of different pooled states in one US sampling frame (Table 4c) were not reflected in the results evaluated with the individual state sampling frame analysis, in which only SS and EQ were the most propelling categories (Table 6).

Table 8. The sacrificial pseudoreplication analysis: median $\pm$ interquartile range (IQR, 25th–75th percentile) of six categories of LEED-NCv3 in four US states.

LEED Category	Sample Size; $n_1 = n_2$	Certified Projects		Cliff's, δ	*p*-Value	NFSAs
		Silver	Gold			
SS	43	11.0 ± 5.8	16.0 ± 5.8	-0.639	0.0000003	positive
WE	43	7.0 ± 4.0	8.0 ± 2.0	-0.233	0.063	suspended
EA	43	19.0 ± 6.0	22.0 ± 5.8	-0.302	0.016	positive
MR	43	3.0 ± 1.8	4.0 ± 1.8	-0.299	0.017	positive
EQ	43	7.0 ± 2.0	8.0 ± 2.0	-0.381	0.002	positive
IO	43	5.0 ± 2.0	5.0 ± 2.0	-0.246	0.050	positive

It is interesting that the most propelling categories in both new buildings (the SS, EA, EQ, and ID) and renovated buildings (the SS, EA, MR, EQ, and IO) were almost the same. There was only exception: MR category. MR was not revealed as a propelled category in LEED-NCv3 certified projects (Table 5), while it was revealed as a propelled category in LEED-EBv3 certified projects (Table 8). However, both of these results should be considered as false, due to deliberately applied wrong analysis that could lead to sacrificial pseudoreplication. Perhaps, this is why these results that are closer to the results reported by Wu et al. [8], who concluded that SS, EA, and EQ were revealed as propelled categories from Silver to Gold for LEED-NCv2.2 certified projects. It can be suggested that such similarity in results was revealed because the same method for statistical evaluation was used here, in the sacrificial pseudoreplication analysis sub-section, and in the study of Wu et al. [8].

4. Conclusions

Silver to Gold cross-certification in both the LEED-NCv3 and LEED-EBv3 projects in the four states (CA, VA, WA, and TX) was statistically evaluated applying (i) the individual state sampling frame analysis, (ii) the US sampling frame analysis, and (iii) the sacrificial pseudoreplication analysis. The following was concluded:

- The individual state sampling frame analysis. Different cross-certification strategies were revealed in LEED-NCv3 and LEED-EBv3-certified projects. For newly constructed projects certified under LEED-NCv3, four of the six categories were employed, in which EA was the most popular propelling category; SS, EQ, and ID were the intermediately popular propelling categories; and

WE and MR were completely unpopular categories. For renovated projects certified under LEED-EBv3, all categories were employed, in which SS and EQ were the most popular propelling categories; WE and EA were the intermediately popular propelling categories; and MR and IO were the least popular propelling categories. Thus, in new versions of LEED-NC and LEED-EB, experts should encourage building practitioners to also focus on currently less popular categories toward more equal high achievements in all of the five main environmental categories.

- The US sampling frame analysis. For both LEED-NCv3- and LEED-EBv3-certified projects, the cross-certification strategy revealed in (ii) the US sampling frame analysis was the same as the strategy revealed in (i) the individual state sampling frame analysis. This means that the US sampling frame (a median LEED project in each individual state) analysis can be recommended as an appropriate test for revealing total tendency in the LEED strategy of both new and renewed projects.

- The sacrificial pseudoreplication analysis. For both the LEED-NCv3- and LEED-EBv3-certified projects, the cross-certification strategy revealed that when projects from different states were pooled in one US sampling frame (the sacrificial pseudoreplication analysis), the strategy revealed on the individual state sampling frame analysis was different. This means that the sacrificial pseudoreplication analysis cannot be recommended as an appropriate test for revealing total tendency in LEED category strategy of both new and renewed projects.

5. Implications of This Study

This study outlines the importance of correct applied statistical methods based on the individual state sampling frame analysis as opposed to the incorrect one based on the US sampling frame through revealed building practice in certification under LEED-NCv3 and LEED-EBv3. The results of this individual state-prevailing trend can help LEED researchers to further correct "a serious type of statistical error" [2] in current evaluations of LEED certified projects. As a result, more realistic feedback from LEED certification can help LEED experts in further versions of rating schemes to move towards more sustainable building.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Picquelle, S.J.; Mier, K.L. A practical guide to statistical methods for comparing means from two-stage sampling. *Fish. Res.* **2011**, *107*, 1–13. [CrossRef]
2. Hurlbert, S.H. The ancient black art and transdisciplinary extent of pseudoreplication. *J. Comp. Psychol.* **2009**, *123*, 434–443. [CrossRef] [PubMed]
3. Hurlbert, S.H.; Lombardi, C.M. Research methodology: Experimental design sampling design, statistical analysis. In *Encylopedia of Animal Behavior*; Bekoff, M.M., Ed.; Greenwood Press: London, UK, 2004; Volume 2, pp. 755–762.
4. Pushkar, S.; Verbitsky, O. LEED-NCv3 silver and gold certified projects in the US: An observational study. *J. Green Build.* **2018**, *13*. in press.
5. Hurlbert, S.H. Pseudoreplication and the Design of Ecological Field Experiments. *Ecol. Monogr.* **1984**, *54*, 187–211. [CrossRef]
6. Hurlbert, S.H. On misinterpretations of pseudoreplication and related matters: A reply to Oksanen. *Oikos* **2004**, *104*, 591–597. [CrossRef]
7. Hurlbert, S.H. Pseudofactorialism, response structures and collective responsibility. *Austral Ecol.* **2013**, *38*, 646–663. [CrossRef]
8. Wu, P.; Mao, C.; Wang, J.; Song, Y.Z.; Wang, X.Y. A decade review of the credits obtained by LEED v2.2 certified green building projects. *Build. Environ.* **2016**, *102*, 167–178. [CrossRef]

9. Wu, P.; Song, Y.; Shou, W.; Chi, H.; Chong, H.Y.; Sutrisna, M. A comprehensive analysis of the credits obtained by LEED 2009 certified green buildings. *Renew. Sustain. Energy Rev.* **2017**, *68 (Pt 1)*, 370–379. [CrossRef]

10. USGBC Website. Available online: http://www.usgbc.org/projects/ (accessed on 21 October 2017).

11. Fuerst, F. Building momentum: An analysis of investment trends in LEED and Energy Star-certified properties. *J. Retail Leis. Prop.* **2009**, *8*, 285–297. [CrossRef]

12. Ma, J.; Cheng, J.C.P. Data-driven study on the achievement of LEED credits using percentage of average score and association rule analysis. *Build. Environ.* **2016**, *98*, 121–132. [CrossRef]

13. Mann, H.B.; Whitney, D.R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [CrossRef]

14. Hurlbert, S.H.; Lombardi, C.M. Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Ann. Zool. Fenn.* **2009**, *46*, 311–349. [CrossRef]

15. Sun, X.; Brown, M.A.; Cox, M.; Jackson, R. Mandating better buildings: A global review of building codes and prospects for improvement in the United States. *Wiley Interdiscip. Rev. Energy Environ.* **2016**, *5*, 188–215. [CrossRef]

16. Hsieh, L.-H.C.; Noonan, D. Strategic Behavior in Certifying Green Buildings: An Inquiry of the Non-building Performance Value. *Environ. Manag.* **2017**, *60*, 231–242. [CrossRef] [PubMed]

17. ASHRAE 90.1. Commercial Adoption, States that Have Adopted a Statewide Commercial Energy Code that Meets or Exceeds an ASHRAE 90.1 Standard, Commercial Code Status. Available online: http://bcapcodes.org/code-status/commercial/ (accessed on 15 September 2017).

18. Kozlov, M.; Hurlbert, S.H. Pseudoreplication, chatter, and the international nature of science: A response to D.V.Tatarnikov. *J. Fundam. Biol. (Moscow)* **2006**, *67*, 145–152.

19. Fay, M.P.; Proschan, M.A. Wilcoxon-Mann-Whitney or *t*-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat. Surv.* **2010**, *4*, 1–39. [CrossRef] [PubMed]

20. USGBC. LEED for New Construction Projects Directory. Available online: http://www.usgbc.org/projects/new-construction (accessed on 10 October 2017).

21. USGBC. LEED for Existing Building Projects Directory. Available online: https://www.usgbc.org/projects/existing-buildings (accessed on 10 October 2017).

22. Cliff, N. Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychol. Bull.* **1993**, *114*, 494–509. [CrossRef]

23. Zhang, F.; Mockus, A.; Keivanloo, I.; Zou, Y. Towards building a universal defect prediction model with rank transformed predictors. *Empir. Softw. Eng.* **2016**, *21*, 2107–2145. [CrossRef]

24. Cohen, J. A power primer. *Psychol. Bull.* **1992**, *112*, 155–159. [CrossRef] [PubMed]

25. Durlak, J.A. How to select, calculate, and interpret effect sizes. *J. Pediatr. Psychol.* **2009**, *34*, 917–928. [CrossRef] [PubMed]

26. Volker, M.A. Reporting effect size estimates in school psychology research. *Psychol. Schools* **2006**, *43*, 653–672. [CrossRef]

27. Fagerland, M.W.; Sandvik, L. The Wilcoxon–Mann–Whitney test under scrutiny. *Stat. Med.* **2009**, *28*, 1487–1497. [CrossRef] [PubMed]

28. Marx, A.; Backes, C.; Meese, E.; Lenhof, H.P.; Keller, A. EDISON-WMW: Exact dynamic programing solution of the Wilcoxon–Mann–Whitney test. *GPB* **2016**, *14*, 55–61. [CrossRef] [PubMed]

29. Hurlbert, S.H.; Lombardi, C.M. Lopsided reasoning on lopsided tests and multiple comparisons. *Aust. N. Z. J. Stat.* **2012**, *54*, 23–42. [CrossRef]

30. LEED-NC 2009. *LEED for New Construction and Major Renovations*; U.S. Green Building Council: Washington, DC, USA, 2009.

31. LEED-EB 2009. *LEED for Existing Buildings: Operations and Maintenance*; U.S. Green Building Council: Washington, DC, USA, 2009.