Journal of Cultural Economics
https://doi.org/10.1007/s10824-023-09493-8

ORIGINAL ARTICLE

Evaluating the quality of UNESCO World Heritage List: a comparison with the Baedeker's guidebooks

Martina Dattilo[1] · **Fabio Padovano**[2,3]

Received: 28 March 2023 / Accepted: 20 October 2023

Abstract
This study verifies whether the number of criteria of Outstanding Universal Value (OUV) satisfied by a site in the UNESCO World Heritage List (WHL) can be considered as an ordinal measure of its quality against the alternative hypotheses that: a) quality can be measured just dichotomously, by inclusion in the WHL); b) the multiplicity of existing OUV is just meant to capture alternative aesthetic criteria expressed by different cultures. This issue is important for both scientific and policy reasons. To avoid problems of endogeneity and reverse causality, we examine the correlation between the number of satisfied criteria and the evaluation of the site's quality made by an authoritative travel guidebook that pre-existed UNESCO, the Baedeker's guide of the early twentieth century. Exploiting a newly assembled dataset on 234 UNESCO World Heritage Sites (WHS) in 10 European countries from 11 Baedeker's guidebooks, from 1899 to 1911, we proxy the Baedeker's evaluations of quality by four measures: (1) total number of citations of the site; (2) weighted number of citations; (3) average length of the paragraphs with at least one citation; and (4) sentiment expressed in the text. All these measures appear positively and significantly correlated with the number of UNESCO criteria that the site satisfies, using a variety of strategies and robustness checks, confirming that they are an informative ordinal proxy for the quality of UNESCO WHS. Moreover, this analysis brings evidence to bear on the debate about the formation and persistence of UNESCO experts' evaluations over time.

Keywords UNESCO World Heritage · Measurement of quality · Baedeker's travel guidebooks · Text analysis

JEL Classification D11 · D13 · D91 · H87 · Z11 · Z18

Paper presented at the European Workshop on Applied Cultural Economics 2022 in Turin and at the PhD seminar of the Lancaster University Management School. We thank Enrico Bertacchini, David Masclet, Marianne Lumeau, François Moreau, Yvon Rocaboy, Giovanna Segre and two anonymous referees for helpful comments on previous versions of the paper. The usual *caveat* applies.

Extended author information available on the last page of the article

Published online: 09 December 2023

1 Introduction

The UNESCO World Heritage List (WHL), which counts 1157 heritage sites in 2023, includes all the sites that UNESCO has declared to be part of the world's cultural and natural heritage, to be preserved and transmitted to the future generations of mankind (UNESCO, 1972). To be included in the WHL, a site should meet at least one out of ten between cultural and natural criteria of Outstanding Universal Value (OUV), set by UNESCO themselves. This reduction of a multiplicity of criteria to a dichotomous decision, based on the satisfaction of just one criterion, generates several unsolved issues, all of which are relevant for scholarly, policy and economic motives. The goal of this paper is to make advances towards the solution of these issues.

The first and foremost of these problems is whether the number of OUV that each site satisfies can be considered as a reasonable ordinal proxy of the quality of the site itself. For example, can the fact that the historic centre of Rome is included in the WHL with 5 out of 6 cultural criteria satisfied, while the site of Ivrea meets only one of them, be reckoned as an informative ordinal measure of the relative quality of the two sites? Likewise, can we state the same about the Grand Canyon, which satisfies all four natural OUV criteria, compared to the Coast of Devonshire, which has been admitted because it satisfies only one?

Acceptance of this interpretation would enable the literature on the UNESCO WHL to make significant progress, since the multivariate evaluations of the sites made by the UNESCO experts would allow us to go beyond the now prevailing still ordinal, but purely dichotomous evaluation of quality, based simply on inclusion. As a matter of fact, being included in the WHL implicitly means "higher/sufficient quality", while being excluded denotes "lower/insufficient quality". If quality is evaluated in a purely dichotomous way, there is no possibility to compare the sites that are included in the WHL. They are all equally as outstanding.

The main justification for the dichotomous evaluation of quality is that the multiplicity of the UNESCO criteria is allegedly meant to encompass alternative aesthetic criteria expressed by different cultures, all of which are supposed to be represented in a list that has a worldwide dimension (Jokilehto, 2006). It is allegedly for this reason that UNESCO accepts into the list sites that satisfy just one criterion. In such a case, exploiting the multiplicity of OUV criteria to proxy the quality of the sites would be inappropriate and no ordinal evaluation of the accepted sites would be possible.

Yet, the wording with which each criterion of OUV is defined is neither culture-specific, nor it refers to any particular aesthetic standard. On the contrary, they are conceived in a quite general and encompassing manner, so that any site, expression of any culture, can potentially be evaluated as satisfying any of them.[1] It is precisely in this sense that they are considered to be "universal"(UNESCO,

[1] The wording of the OUV criteria is as follows. Cultural criterion 1 (aesthetic value): "[The site] represents a masterpiece of human creative genius". Cultural criterion 2 (aesthetic, historical, technical values): "[the site] exhibits an important interchange of human values, over a span of time or within a cultural area of the world, on developments in architecture or technology, monumental arts, town-planning or landscape design". Natural criterion 1: "[the site] contains superlative natural phenomena or areas of exceptional natural beauty and aesthetic importance", and so on. Table 9 in the Appendix reports the definitions of the ten UNESCO criteria of OUV.

2013).[2] Furthermore, over time UNESCO has introduced some corrections to the original formulation of the criteria, to make sure that they adequately reflect the variety of cultures of the world (UNESCO, 2017). Given the wording of the criteria, the risk of a cultural or aesthetic bias may then emerge at the later stage of the judgements of whether the site fits into any of these criteria made by the UNESCO experts and country representatives. Yet, the generality of the definitions of the criteria should ensure that the ordinal evaluation of the quality of the already accepted sites according to the number of criteria satisfied is a methodology per se free from cultural biases. If we accept this argument, proxying quality by the number of OUV criteria satisfied becomes again legitimate.

It is worth bearing in mind that the evaluation of the quality of the UNESCO sites is an important issue for several reasons. A first and obvious one is related to tourism policy and politics. Since the inclusion in the WHL is generally considered a signal of outstanding quality of a site (Adie, 2017), obtaining the UNESCO WHS label can be a crucial determinant of consumers' choices; such a recognition is especially valuable in a market characterized by imperfect information on the side of consumers, like the touristic one (Keane, 1997). Even though the empirical analyses on this point yield rather mixed results, still the idea that the inclusion of a site in the UNESCO WHL increases the touristic revenues is widely held among practitioners and politicians (Adie et al., 2018; Bertacchini et al., 2023; Buckley, 2004; Patuelli et al., 2016; Rakic, 2007; Van Blarcom & Kayahan, 2011; Yang et al., 2010).

A second set of reasons is scholarly. The concept of quality is an important, yet difficult to characterize, dimension of consumers' choices in microeconomics theory; it obviously affects the demand for a commodity but it seldom appears in models. For culture-related goods, assessing the quality dimension is particularly important, since in this domain product differentiation is often so extreme to make many works of art and cultural experiences essentially unique (Ginsburgh & Weyers, 1999; Waldfogel, 2012). In this case, it is the evaluation of quality to drive consumers' preferences and choices.[3]

Yet, contrary to quantity, which is observable, quality needs to be approximated; hence, recognizing the number of satisfied criteria as a proxy for the quality of the site is an important matter. Moreover, papers that have adopted this approach (Dattilo et al., 2022) have shown that it may upset some established results in the literature about the UNESCO WHL that instead adopt the dichotomous definition of quality based on inclusion. For instance, when quality is so defined, lobbying is found to play an important role; this obviously questions the credibility of the list itself as a standard for the patrimony of mankind (Bertacchini & Saccone, 2012; Frey & Steiner, 2011). When, instead, quality is measured in a multivariate way by the number of criteria satisfied, the estimates suggest that countries tend to "lobby

[2] We acknowledge that, especially outside the economic literature, the debate on this matter is considered to be still open. Indeed, some authors still maintain that the Convention apparently favours an Eurocentric point of view (Labadi, 2013). A recent empirical analysis, however, (Dattilo et al., 2020) fails to find any evidence that the criteria are actually biased in favour of European heritage.

[3] In the present analysis, to pre-empt all semantic disquisitions, the concept of "quality" is to be interpreted in a broad manner, i.e. as the dimension over which individuals express their preferences and make their choices when the quantity of the commodity is essentially single.

at the margin", i.e. they use their political influence only to push through sites of not universal renown (e.g. the coast of Devonshire), that can satisfy only a limited number of criteria, if any. For those that instead receive universal recognition (e.g. the Grand Canyon), the political weight of the country seems not to play a role. If lobbying is used only at the margin, the whole UNESCO WHL site appears less distorted and more reliable.

While proxying the sites' quality as the sum of the satisfied criteria may seem straightforward, in fact it is not. To assess the legitimacy and correctness of this criterion, we need an independent counterfactual, against which comparing the evaluations made by the UNESCO experts. Finding such a counterfactual is plagued by problems of endogeneity and reverse causality. If, for instance, we selected the number of tourists visiting a site, this might affect UNESCO's decision to include the site on the list, but also, in reverse order, inclusion in the WHL might influence the number of visitors. The same problems arise when considering the literature, e.g. tourist guidebooks, websites, cultural publications and the like, as an independent benchmark to evaluate the quality of the UNESCO sites. Again, a contemporary guidebook may devote more pages to a tourist attraction because it is a WHS; likewise, a site may enter the list because it has received positive evaluations from important guidebooks.

To overcome these problems, as an independent standard of quality we elect the *Baedeker's Travel Guidebooks*, the most authoritative and comprehensive guidebook of the beginning of the twentieth century; in a way, the Baedeker's represent the origin and the model of contemporary travel guidebooks and websites. Two are the motivations underlying our choice. First, since UNESCO did not exist at the time of the publication of the Baedeker's guidebooks, their selection as a standard for quality avoids problems of endogeneity and reverse causality that contemporary guides suffer from. Second, Baedeker's guidebooks existed for most countries within Europe, which, compared to the rest of the world, represent a geographical area characterized by a high degree of cultural homogeneity. Such homogeneity provides a "cultural *ceteris paribus* condition" which is crucial for the analysis, since it allows to consider the number of OUV criteria satisfied as a proxy for quality rather than as a plurality of aesthetic standards referred to different cultural contexts. Then, if the correlation holds within Europe, it can be potentially applied (upon verification) also in other contexts.

The results of the analysis show that the number of criteria satisfied by the UNESCO WHS is, indeed, correlated with four alternative measures of relevance expressed in the Baedeker's travel guidebooks, namely, the (weighted and unweighted) number of citations, the length of the description of the site and the sentiment expressed in the text. This result, corroborated by a battery of robustness tests, confirms that the number of satisfied criteria is a reasonable ordinal approximation of the quality of the WHS.

The rest of the paper is organized as follows. The following section describes the Baedeker's travel guidebooks and motivates their selection as an independent counterfactual to appraise the quality of the sites. Section 3 illustrates the data and the metrics for the evaluation of quality that we have derived from the texts of the guidebooks. Section 4 discusses the empirical strategy and the econometric issues

involved in the analysis. Its results and robustness tests are reported in Sect. 5. Section 6 concludes.

2 The Baedeker guidebooks

As Koshar (1998) recalls, the English middle-class professionals and intellectuals of the nineteenth century played a fundamental role in shaping the tastes that define, still today, what can be considered as cultural heritage. When these individuals started to travel across Europe, they needed to organize their tours more efficiently than the earlier generations of European aristocrats, who devoted a full year, if not more, of their education to the so-called Grand Tour of Europe. The English middle-class tourists, instead, had substantial but not unlimited economic resources and, most of all, had tighter time constraints. For this reason, guidebooks started to focus on "what ought to be seen" rather than "what could be seen".

Probably the first to understand this fundamental change in perspective was Karl Baedeker (and four generations of Baedekers after him), who created a tri-lingual set of travel guides, covering around forty countries that existed at those times. This series quickly became the accepted international paradigm for guidebooks; their authors were considered "the arbiters of artistic tastes" (Bruce, 2010). Many are the reasons for their success. First, the Baedeker's guidebooks eliminated all the nationalistic comments, creating an almost completely objective and reliable guide (Bruce, 2010; Koshar, 1998).[4] Second, they set uniform standards of evaluation, which made the Baedeker's guidebooks a valid substitute for the absence of a single travel guidebook covering multiple countries. This allowed the Baedeker's to be much more informative and detailed than their competitors. Moreover, the Baedeker's guidebooks were the first to make a selection and a ranking of historical heritages, thus initiating a common practice in contemporary times. Finally, they popularized the practice of using asterisks to denote sites of extraordinary quality.[5] Historical buildings and monuments were among the prized sites in tourists' itineraries, as they represent almost 30% of the points of interest. The similarities with the UNESCO WHL are evident, considering that around 35% of the WHS belong to this category (Labadi, 2007).[6]

[4] In the Murray guidebooks, for example, the pride in the English sense of freedom and comfort is evident in the description of the travel preparations and accommodations.

[5] Mariana Starke was an important precursor. Indeed, her two-volume guidebook contains exclamation marks to rank important sights.

[6] These figures refer to the year 2005.

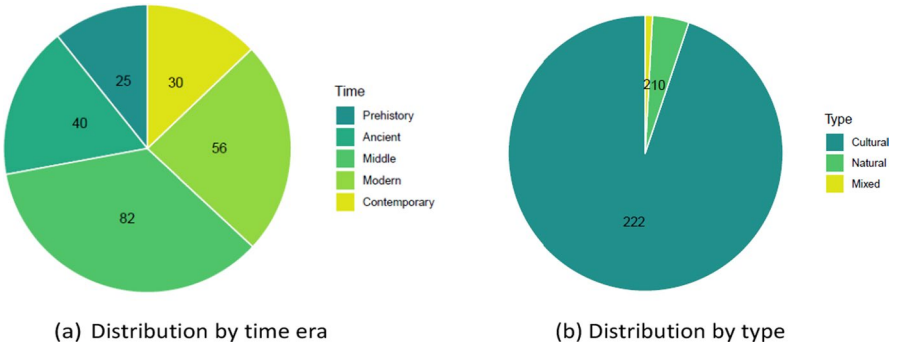

Fig. 1 Sample distribution of WHS by time era and type

3 Data and measurements

3.1 Distribution of sites and characteristics of the sample

To verify the evaluation of (future) UNESCO WHS by (pre-existing) Baedeker's guidebooks, we select 11 volumes of the series.[7] They were published in a restricted time span (from 1889 to 1911[8]) and described areas of similar extensions, ensuring a good comparability for our analysis. This procedure restricts our sample to the heritage located in 10 European countries.[9] In addition, we have removed sites whose name is a common English name, such as the city of Bath, since its value could be overstated by the references to the common word "bath". Furthermore, sites placed outside the continental Europe, e.g. such as the "French Austral Lands and Seas" for France, are excluded from the sample of WH sites considered. Indeed, they are likely not to be present in the Baedeker guidebooks and would rather bias our analysis. As a result, our dataset contains 234 UNESCO WHS, whose list is presented in Appendix 2. Figure 1 plots the distribution of the sites in our sample with respect to the time of construction and the type of the site, i.e. cultural, natural or mixed. The distribution emerging from our dataset is coherent with the finding of the previous literature on the UNESCO WHL: Middle Age cultural sites represent the large majority of the European WHS. This evidence represents a first insight that the sample under

[7] Specifically, "Austria-Hungary: with excursions to Cetinje, Belgrade and Bucharest: handbook for travellers" 1911; "Belgium and Holland including the Grand-Duchy of Luxembourg", 1910; "Northern France; Handbook For Travellers", 1899; "Southern France including Corsica", 1902; "Great Britain: Handbook For Travellers", 1906; "Italy, from the Alps to Naples", 1904; "Southern Italy and Sicily, with Excursions into the Liparia Islands, Malta, Sardinia, Tunis, and Corfu", 1903; "Spain and Portugal", 1908; "Northern Germany as far as the Bavarian and Austrian Frontier", 1910; "Southern Germany", 1902; "The Rhine from Rotterdam to Constance", 1906.

[8] We limit the analysis to the 1911 edition because the outbreak of World War One reduced the possibility to travel and revamped nationalistic sentiments also in the Baedeker's guidebook, making the comparison of heritage sites from different countries less balanced.

[9] The countries included in the study are: Austria, Hungary, France, Spain, Portugal, the Netherlands, Belgium, Italy, Germany, the United Kingdom.

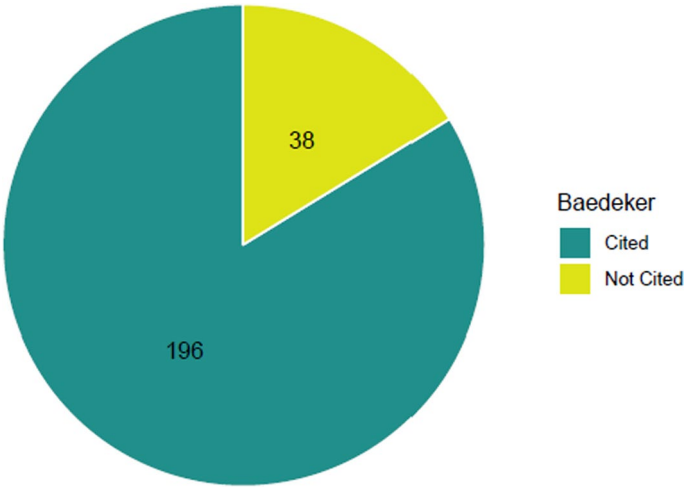

Fig. 2 WH sites representation in the Baedeker's guidebooks

scrutiny is indeed representative of the entire population of European WHS. Additionally, more than 80% of the total of the European UNESCO WHS (precisely 196) are cited at least once in the Baedeker's guidebooks (Fig. 2), which reinforces the hypothesis that the guidebooks constitute a relevant source of information about outstanding heritage.

Despite the fact that the large majority of sites are mentioned in the Baedeker's guidebooks, a still relevant portion, i.e. 38 sites, is not included (Fig. 2). Therefore, we should exclude that these observations are systematically different from the included ones, or our analysis could be affected by a potential bias. That does not seem to be the case, however: Fig. 3 shows the distribution of the sites, both included and excluded, with respect to the time of construction (Fig. 3a) and the type (Fig. 3b). Figure 3a shows that the distribution of the sites with respect to the time of construction is quite similar in the two subsamples: the proportion of sites built in prehistoric or contemporary era is higher for non-cited sites, but the two subsamples are equally representative of all time periods, in similar proportions. The same evidence emerges from Fig. 3b: the proportion of mixed and natural sites slightly increases when passing from the group of included sites to that of excluded ones, but this slight variation does not affect the global distribution.

These simple descriptive statistics point out three main issues. First, the sample considered can be deemed to be representative of the population of the European WHS, at least as far as the type of the site and the time of its construction are concerned. Second, the selection made by the Baedeker's when writing the books seems to be coherent with the UNESCO WHL. Third, in our sample, this consistency seems to be independent from the time of construction or type of heritage. Yet, in order to verify the second hypothesis, a more in-depth analysis is needed: one should consider not only the mere reference to the site in the guidebooks, but also the author's perception of the sites; this in turn requires adequate metrics. The

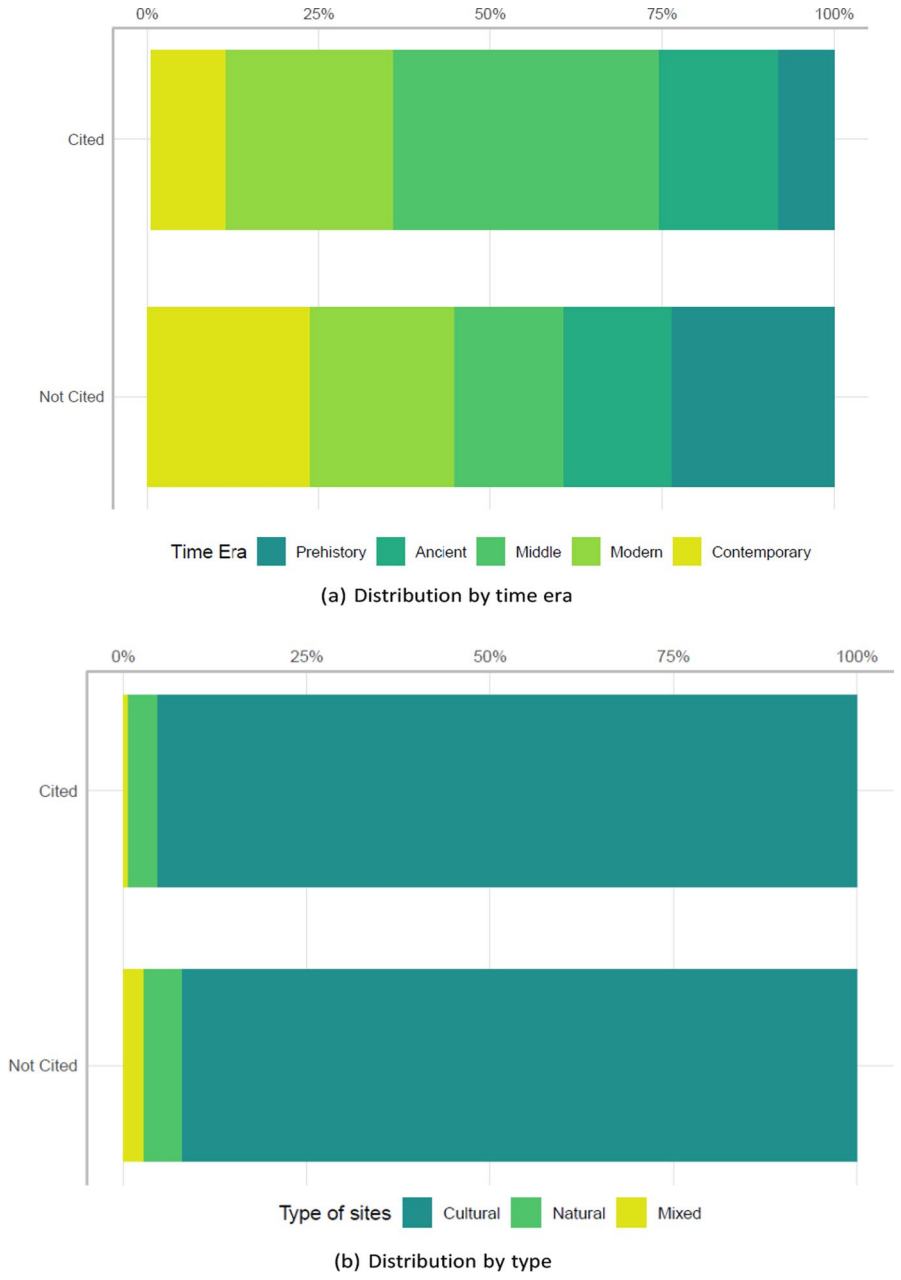

Fig. 3 Representation in the Baedeker's guidebooks

following section describes the intuition behind and the process of extraction of these indicators.

3.2 Measuring the sites' quality in the Baedeker guidebooks

The WHS are analysed considering different aspects of the tastes of the authors of the Baedeker's guidebooks. An original algorithm processes the PDF version of the books and extracts the data. This approach has several advantages: first, it avoids measurement errors, which are likely to appear when filling a long dataset by hand. Second, it reduces the subjectivity of the analysis: only the definition of the variables must be specified a priori, which ensures that each observation is evaluated according to the same standards. Third, the study could be reproduced and verified at any moment, in compliance with the scientific method. The only issue that the algorithm fails to overcome is digitalisation errors, i.e. incorrect transpositions of the text from the paper to the PDF format. We make sure, however, to perform the text analysis minimizing our dependency on this type of error.

In order to produce a reliable analysis, two main issues must be overcome. For some countries, like the United Kingdom, the Baedeker's features a single guidebook, while for others, such as Italy, France and Germany, the algorithm must merge the texts of more than one book. Considering Italy as a matter of example, the data extracted from two volumes, one covering the North and the Centre of the country, another the South, are aggregated into a single set of measures in the final data frame.[10] Secondly, the official names of the UNESCO WHS are often long and elaborate. A clear example is the official name of the Last Supper by Leonardo da Vinci: "Church and Dominican Convent of Santa Maria delle Grazie with The Last Supper by Leonardo da Vinci". Dealing with such titles could be complex when implemented in an R code. Therefore, as a first step, we simplify the official names, in order to have a title with just one or two words in lowercase. In the example, the name is reduced to cenacolo".[11] When multiple cities or places are involved, the title is subsetted into multiple names. For example, "Portovenere, Cinque Terre, and the Islands (Palmaria, Tino and Tinetto)" is subsetted into "portovenere", "cinque terre","palmaria","tinetto".

The measures of quality that we have extracted through this approach are described in the following sections.

[10] There are no significant overlaps of sites between multiple guidebooks of a single country, which ensures that the large majority of sites is cited only in one guidebook. In the few cases where a site is cited in more than one guidebook of a given country, that should be reflected in a higher value of quality of cultural heritage of that site (e.g. the city of Palermo is extensively described in the guide of Southern Italy but it receives an occasional citation for its historical significance in the guidebook of Northern Italy too). Consequently, the two (or three) measures obtained from different guidebooks are summed up to compose the final score. We have also provided some robustness check of alternative methods of aggregation. In Appendix 3 we report the regression results of Eq. 5, where the independent variable is the maximum value (instead of the sum) of the indices obtained in the different guidebooks. The results do not change qualitatively.

[11] To be implemented in our algorithm, the names of the WHS must be reported in lowercases.

3.3 Number of citations

Our first variable of interest, $Citation_i$, is the number of citations the WH site i has received in the Baedeker's guidebooks. We assume that the higher the number of citations, the greater the importance the guidebook's author attributed to the site. If the site i is not cited in the guidebooks, then the value of ln*Citation* is equal to 0. To minimize a possible bias in favour of large cities, whose names are more likely to appear in the title of a chapter, we have removed the title of each page from the scanned text. The distribution of $Citation_i$, shown in the first row of Fig. 4, suggests that a logarithmic scale would be more appropriate to capture its variation. Thus, ln$Citation_i$, the logarithmic transformation of *Citation*, computed as in Eq. 1, will also be used. From the second line of Fig. 4 it is evident that this variable better represents the variability in our sample.

$$\ln Citation_i = \begin{cases} 0 & if\ Citation_i = 0 \\ log_{10}\left(Citation_i\right) & if\ Citation_i > 0 \end{cases} \tag{1}$$

Despite the similarity in size, the guidebooks may differ with respect to the total number of sites cited. For example, the index of the North of France guidebook contains 3366 items, while that for the South of France lists 6590 heritage sites. To overcome this disparity, we have computed ln$Icitation_i$ as described in Eq. 2, where *Tweight* is the number of index entries in each book. The distribution of this variable, plotted in the third row of Fig. 4, appears similar to the one of ln$Citation_i$, confirming once more that the Baedeker's guidebooks are easily comparable between them.

$$\ln Icitation_i = \begin{cases} 0 & if\ Citation_i = 0 \\ \log_{10} \frac{100000}{Tweight} \times Citation_i & if\ Citation_i > 0 \end{cases} \tag{2}$$

3.4 Length of the paragraph

Because the author of the Baedeker's guidebooks could express his appreciation for a (then-to-be) WHS by allocating more space to its description, we test the hypothesis that the longer the paragraph describing a site, i.e. the greater the number of details and information that the author judges interesting for the reader, the higher should be the quality of the site. We focus on paragraph length instead of sentence length for two main reasons. First, in the context of travel guidebooks, a sentence is usually insufficient to fully describe a heritage site, while the length of the paragraph better reflects the type of information we want to collect. Secondly, since paragraphs are separated by a clear blank space while the separation of sentences could be less clear cut, focusing on paragraphs ensures that our analysis is reliable and independent from any digitalization error. Hence, from the whole text we extract all the paragraphs j that contain at least one citation of the site i considered.

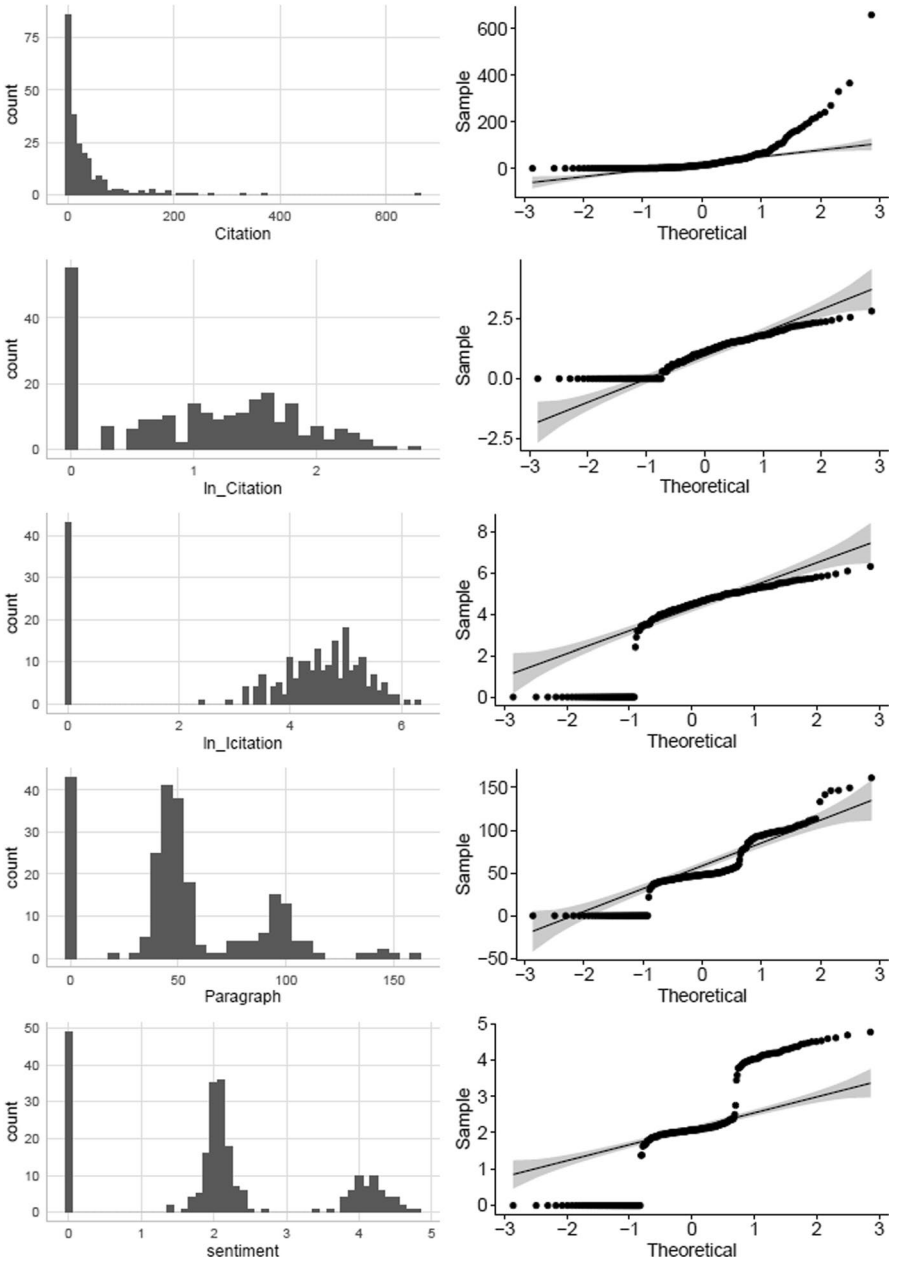

Fig. 4 Distributions of the variables. Note: the left-hand column of the graph plots the frequency of each variable considered; the right-hand column shows the comparison between the actual distribution of the variable considered (vertical axis) and the theoretical normal distribution (horizontal axis). *Source*: computation by the authors

We compute $Paragraph_i$ as the average number of words w included in these paragraphs. The variable is computed in the same way as in Eq. (3), where T is the total number of paragraphs for which $Citation_i > 0$. To simplify the interpretation, the variable is divided by 10, so that if *Paragraph* increases by 1 point, it corresponds to a variation of 10 words.

The distribution of this variable, displayed in the fourth row of Fig. 4, appears to be concentrated around four values: when a site is not cited in the text, *Paragraph* is equal to zero; a short paragraph is likely to contain around 500 words; a medium-size paragraph is around 1000 words; while a long paragraph contains around 1500 words on average. As this distribution is probably due to the author's style, we should not expect a normal distribution.

$$Paragraph_i = \frac{1}{10 \times T} \sum_{j=1}^{T} w_j \tag{3}$$

3.5 The sentiment of the text

Finally, we capture the author's appreciation for a WHS by looking at the vocabulary he adopts in the description. The straightforward idea is that the more positive are the expressions used in the description of the site, the higher should be its quality. The literature focuses on two methods to perform the "sentiment analysis" of a text: the lexicon-based method and the machine learning method. The former approach relies on lexicons, i.e. pre-defined lists of words, with each word assigned a score reflecting the emotion of interest. This allows measuring the emotional content of a text based on the prevalence of negative or positive words. The main drawback of the lexicon-based technique is that it ignores all the contextual characteristics specific to the text in which the word appears (Shapiro et al., 2020).[12] An alternative and frequently adopted algorithm is *Vader*, developed by Hutto and Gilbert (2014). It accounts for the word's context within the sentence; yet, since it is developed at the sentence level and is specific for social media language, it does not seem appropriate for the Baedeker content and format. To avoid this problem, we resort to the lexicon approach, using the *AFFIN* lexicon of the *tidytext* R package (Silge & Robinson, 2016; Nielsen, 2011b), a largely used algorithm (Sharma et al., 2018; Benchimol et al., 2022). As the guidebooks are already a selection of "the best sites", we focus only on laudatory expressions. Additional tests considering also the negative, non-laudatory words are presented in the robustness section. In detail, to compute $Sentiment_i$, we extract from the guidebooks all the paragraphs containing at least one citation of the site considered and then we analyse the vocabulary adopted. As a result, we obtain the variable $Sentiment_i$ computed as the Eq. (4), where N_j is the number of positive words in paragraph j; f is the frequency of the word w in the paragraph j and v is a "laudatory weight" with values ranging from 1 (slightly

[12] Applications of this method can be found in political economy (Shapiro and Wilson, 2022) and in financial economics (Correa et al., 2021; Fraiberger, 2016; Heston and Sinha, 2017; Nyman et al., 2021).

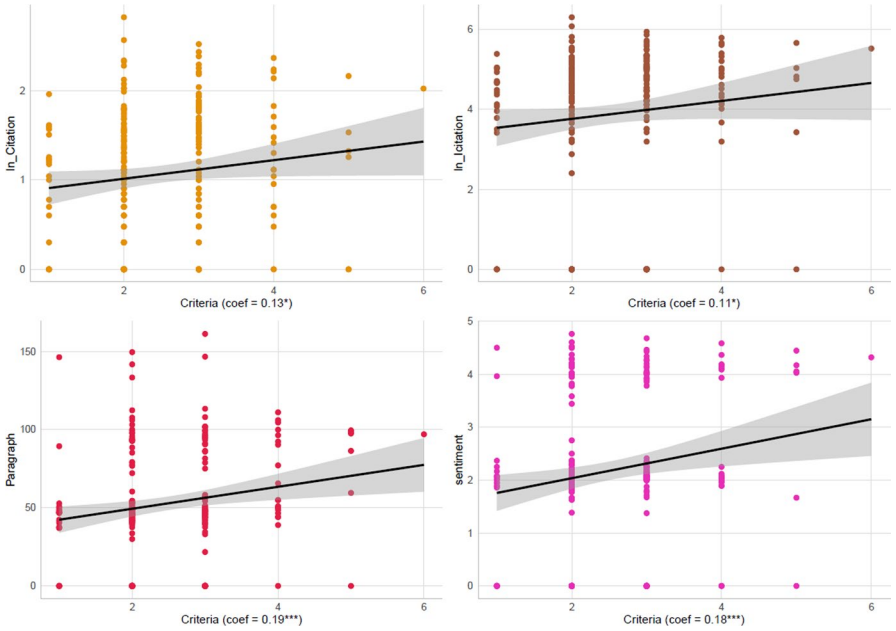

Fig. 5 Correlation between different measures of quality. Note: the plot shows the correlation between Criteria and the measure of quality extracted from the Baedeker guidebooks. The value of the correlation coefficient is reported below each table, following the label of the x-axis (Criteria). The significance level is expressed as follows: $*p<0.1$; $**p<0.05$;.$***p<0.01$

positive) to 5 (really positive), which are assigned by the *AFFIN* dictionary[13] to the word w. Again, T is the total number of paragraphs for which $Citation_i > 0$.

$$Sentiment_i = \frac{1}{T}\sum_{j=1}^{T}\frac{1}{N_j}\sum_{w=1}^{N_j} f_{wj}v_{wj} \tag{4}$$

The distribution of *Sentiment* is shown in the last row of Fig. 4. Similarly to *Paragraph*, the frequency distribution is concentrated around 2 values. Again, we explain this behaviour as a consequence of the author's style. As Fig. 5 shows, all four variables of Baedeker's tastes are positively and significantly correlated with the number of criteria satisfied by the WHS under consideration. This result gives the first insight that these variables can be valid proxies of the quality of UNESCO sites. However, we complement this preliminary finding with a more systematic and insightful empirical strategy.

[13] AFFIN is a list of English words rated for valence with an integer between minus five (negative) and plus five (positive) (Nielsen, 2011a). From 2009 to 2011 Nielsen, the author of the dictionary, has aggregated and expanded pre-existing dictionaries such as Original Balanced Affective Word List (http://www.sci.sdsu.edu/CAL/wordlist/origwordlist.html), the Urban Dictionary (http://www.urbandictionary.com) and The Compass DeRose Guide to Emotion Words (http://www.derose.net/steve/resources/emotionwords/ewords.html).

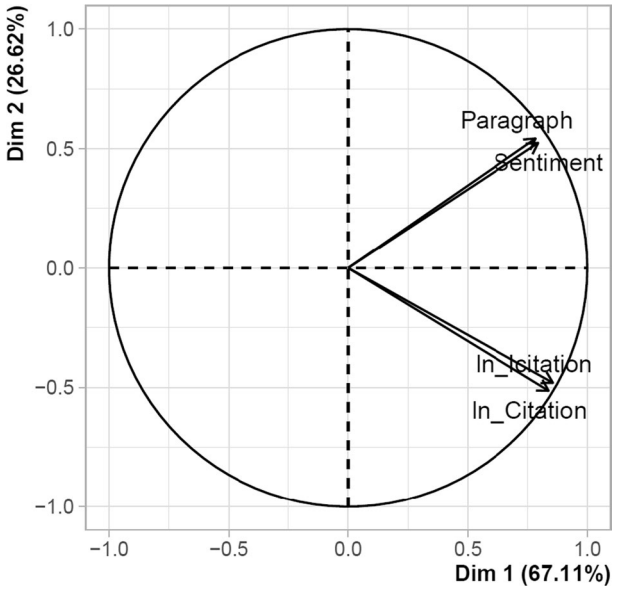

Fig. 6 PCA analysis of the four Baedeker's indexes

3.6 Comparison between the indexes

It is important to verify that these indexes do not evaluate quality in the same way; in other words, they are not perfect substitutes one for the other. To verify that this is not the case, we have performed a Principal Component Analysis of the four indexes specified above. The results, reported in Figs. 6 and 7, that not all the four variables considered are capturing similar aspects of the quality of the cultural heritage, as the plot of Fig. 6 shows. In particular, ln*Citation* and ln*Icitation* are close substitutes, yet they are orthogonal to *Paragraph* and *Sentiment*, which represent different dimensions of the quality of cultural heritage. Moreover, the groups of sites emerging from the two main principal components reported in Fig. 8 reflect the non-normal distribution of the variables *Paragraph* and *Sentiment*.

4 Empirical strategy

The goal of the analysis is testing the hypothesis that the evaluation of the heritage sites in the Baedeker's travel guidebooks is consistent with the evaluation of their quality based on the number of UNESCO OUV criteria satisfied. To this end, we have performed a straightforward test: we have regressed $Criteria_i$, i.e. the number of criteria satisfied by site *i*, on the four independent variables expressing the Baedeker's evaluations, i.e. ln*Citation*, ln*Icitation*, *Paragraph* and *Sentiment*. The explanatory variables are included sequentially, resulting in four different regressions. We

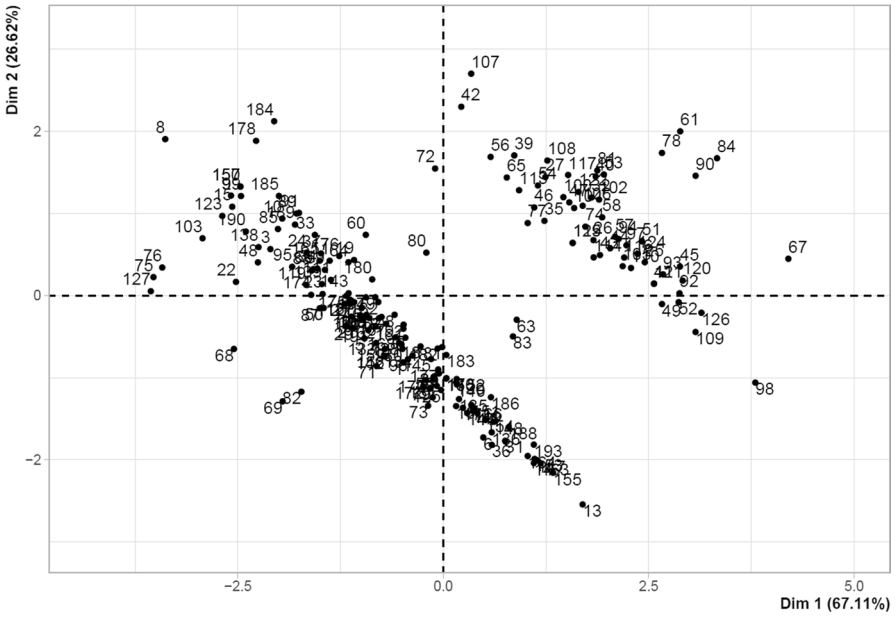

Fig. 7 Distribution of sites along the two PCA dimensions

first estimate Eq. 5, representing the empirical model, using a simple OLS model, where Γ represents the country fixed effect.[14]

$$Criteria_{ij} = \beta_0 + \beta_1 Baedeker_{ij} + \beta_3 \Gamma_j + u_{ij} \tag{5}$$

Criteria is a discrete and ordered variable: $Criteria_{ij} = 1$ means that the site i in country j has the lowest quality, while if $Criteria_{ij} = 6$, the site i in country j has the highest quality in our sample.[15] In this case, following Wooldridge (2010), the Ordered Logit Model should be more precise in estimating the regression coefficients. Actually, in the OLS model we assume that Q_{ij}, i.e., the quality of the site i in country j, is a continuous variable, but we can only observe whether Q attains certain thresholds δ_k:

[14] We have estimated both the OLS and the Ordered Logit models also without the country fixed effects. The results, available upon request, are qualitatively similar.

[15] The site with the highest quality in our sample is *Venice and its Lagoon,* a cultural site. Our sample contains only two mixed sites, i.e. *Ibiza, Biodiversity and Culture* and *St Kilda*, with both Criteria = 5. Theoretically the maximum score that a natural site can attain in terms of quality is $Criteria_i = 4$. In our sample, however, this value is never reached as the natural sites with the highest number of criteria satisfied are the *Gulf of Porto* and *Doñana National Park*, where Criteria = 3.

$$Criteria = \begin{cases} 1 \ if \ Q \leq \delta_1 \\ 2 \ if \ \delta_1 \leq Q \leq \delta_2 \\ 3 \ if \ \delta_2 \leq Q \leq \delta_3 \\ 4 \ if \ \delta_3 \leq Q \leq \delta_4 \\ 5 \ if \ \delta_4 \leq Q \leq \delta_5 \\ 6 \ if \ \delta_5 \leq Q \end{cases} \tag{6}$$

5 Results

5.1 Main results

Tables 1 and 2 show the results of the estimates of Eq. 5. The coefficients of the linear model are all positive and significant at least at the 10% level. Since *Criteria* is an integer number, however, the magnitude of the coefficients of the linear models is hard to interpret. To overcome this problem, the coefficients of Table 2, presenting the estimates of the Ordered Logit model, are expressed as odds ratios (OR) of the regressions. They are calculated as $OR = exp^{\beta}$, where β is the original regression coefficient. The results show that, when the raw number of citations increases by 1%, the likelihood of being classified into the superior criterion increases by 0.40 percentage points (ordered logit model 1). This percentage decreases to 0.14% when we consider ln*Icitation*, but it is still positive and significant (logit model 2). The site's odds of receiving one more criterion is 1% higher when the average length of the paragraph describing the site increases by 10 words (logit model 3). On the other way, the likelihood of being evaluated on the superior criterion increases by 26% when the site receives, on average, one additional *Sentiment* point. In addition, based on the R^2 and the AIC value, the models containing *Paragraph* fit the data better than the other ones.

Overall, the estimates confirm that the evaluations made in Baedeker's guidebooks are consistent with those that the UNESCO experts have rendered more than 60 years later, thus legitimizing the interpretation of the number of criteria that a WHS satisfies as an ordinal measure of its quality.[16]

5.2 Robustness tests

In addition to the simple specifications exposed in the empirical strategy, we complement our analysis by a series of robustness checks. First, we verify if the positive correlation between $Criteria_i$ and the other measures of quality is somehow linked to the presence in our sample of both natural and cultural sites. As we have seen

[16] We have performed an additional robustness check of our main specification (equation 5), including clustered standard errors at the country level. While the results present higher p values of the coefficients of interest, they are coherent with our main results.

Table 1 Determinant of criteria: linear model

	Dependent variable			
	Criteria			
	1	2	3	4
ln_Citation	0.175**			
	(0.086)			
ln_Icitation		0.062*		
		(0.033)		
Paragraph			0.006***	
			(0.002)	
Sentiment				0.125**
				(0.049)
Constant	1.736***	1.701***	1.662***	1.663***
	(0.156)	(0.157)	(0.145)	(0.150)
Country FE	YES	YES	YES	YES
Observations	234	234	234	234
R2	0.123	0.119	0.143	0.134
Adjusted R2	0.084	0.079	0.104	0.095

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

natural sites may satisfy a smaller number of criteria, 4 against the 6 for cultural sites. On the other hand, the Baedeker's guides are mostly focused on cities and monuments, while parks and landscapes receive a lower attention. For this reason, we test our model on a subsample of UNESCO WHL containing only cultural sites.

Tables 3 and 4 show the results of the estimates of this additional test. The sample is reduced from 234 to 222 observations and the estimates are still coherent with our main results. The coefficients of both the linear and the logit models are all positive and significant. The magnitudes of the coefficients of Table 4 are extremely similar to the ones of Table 2. The likelihood of receiving one more criterion when the raw number of citations increases by 1% rises from 0.41 to 0.47 (model 1 of Table 4). The same small increase in magnitude can be observed for the weighted number of citations, which pass from 0.14 to 0.18. It is worth noticing that restricting our sample to only cultural sites reduces the p-value of ln*Icitation*, which is significant at the 5% level in model 2 of Table 4. The odds of receiving one more criterion when the average length of the paragraph increases by ten words is constant at 1%. Similarly, the likelihood of receiving one additional criterion is confirmed at 26% when the sentiment of site *i* increases by 1 point.

Next, in Tables 5 and 6 we control for the time when the site had been built. Some sites of the UNESCO WHL could not be mentioned by the Baedeker's guides because either they did not exist at the time of the publication, like the case of some contemporary sites (e.g. the industrial archaeological site of Ivrea), or because they had not been not discovered yet, as in the case of some prehistoric sites (like the Caves of Lascaux, discovered in 1940). The multivariate qualitative variable *Time* is therefore constructed only for cultural sites; it takes values from 1, if the site

Table 2 Determinant of criteria: ordered logit model

	Dependent variable			
	Criteria			
	1	2	3	4
ln_Citation	0.40**			
	(0.18)			
ln_Icitation		0.14*		
		(0.07)		
Paragraph			0.01***	
			(0.00)	
Sentiment				0.26**
				(0.10)
Country FE	YES	YES	YES	YES
AIC	597.73	599.20	593.91	596.28
Log Likelihood	−283.86	−284.60	−281.96	−283.14
Num. obs	234	234	234	234
McFadden's R2	0.05	0.05	0.06	0.05

*$p<0.1$; **$p<0.05$; ***$p<0.01$

dates back to the Prehistoric era, to 5, if the site has been built in the contemporary period, according to the classification of Fig. 1. The squared value of this variable is also considered, to check whether nonlinear effects are present. The OLS estimates, reported in Table 5, indeed seem to support a nonlinear relationship between *Time* and our dependent variable $Criteria_i$, while this is only marginally the case for the logit model (Table 6). The main results, however, remain coherent with our main analysis, since all the indicators of quality extracted from the Baedeker's guides remain positive and significant.

Finally, we have verified whether the result for the variable $Sentiment_i$ depends on the way the index has been constructed. Following the same lexicon approach, we have tested three alternative specifications of the variable described in paragraph 3.5: $Sentiment\_all_i$, $Sentiment\_positive_i$, and $Sentiment\_paragraph_i$. Specifically, comparing the results of $Sentiment\_all_i$ and $Sentiment\_positive_i$ allows to verify whether negative evaluations (i.e., negative expressions of sentiment) which are excluded from the $Sentiment\_positive_i$ index, may bias our estimates. Although fairly rare, negative evaluations of cultural sites can be found in the Baedeker's travel guidebooks. Furthermore, with the index $Sentiment\_paragraph_i$ we test whether the frequency with which an evaluation is repeated in any paragraph, which may be affected by the writing style of the author, may be another source of bias.

The indexes are specified as follows:

$$sentiment\_all = \frac{1}{n}\sum_{i=1}^{n} f_i V_i \tag{7}$$

where n is the number of words in all the paragraphs containing the site's name; V is a value from −5 (really negative) to +5 (really positive) assigned by the *AFFIN* dictionary to the word i, f is the frequency of the word i.

Table 3 Robustness checks: cultural sites—linear model

	Dependent variable			
	Criteria			
	1	2	3	4
ln_Citation	0.206**			
	(0.081)			
ln_Icitation		0.077**		
		(0.031)		
Paragraph			0.005***	
			(0.002)	
Sentiment				0.116**
				(0.047)
Constant	1.719***	1.667***	1.669***	1.675***
	(0.156)	(0.152)	(0.145)	(0.151)
Country FE	YES	YES	YES	YES
Observations	222	222	222	222
R2	0.163	0.159	0.174	0.163
Adjusted R2	0.124	0.119	0.135	0.123

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 4 Robustness checks: cultural sites—logit model

	Dependent variable:			
	Criteria			
	1	2	3	4
Ln_Citation	0.47**			
	(0.19)			
Ln_Icitation		0.18**		
		(0.08)		
Paragraph			0.01***	
			(0.00)	
Sentiment				0.26**
				(0.11)
Country FE	YES	YES	YES	YES
AIC	542.62	543.27	539.78	543.22
Log Likelihood	−256.31	−256.63	−254.89	−256.61
Num. obs	222	222	222	222
McFadden's R2	0.07	0.06	0.07	0.06

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

$$sentiment\_positive = \frac{1}{N}\sum_{i=1}^{N} f_i v_i \tag{8}$$

Table 5 Robustness checks: time – linear model

	Dependent variable:			
	Criteria			
	1	2	3	4
Ln_Citation	0.155*			
	(0.084)			
Ln_Icitation		0.056*		
		(0.029)		
Paragraph			0.005***	
			(0.002)	
Sentiment				0.093**
				(0.047)
Time	0.439**	0.448**	0.429**	0.452**
	(0.222)	(0.226)	(0.214)	(0.215)
Time2	−0.079**	−0.081**	−0.079**	−0.082**
	(0.035)	(0.036)	(0.033)	(0.033)
Constant	1.201***	1.159***	1.174***	1.152***
	(0.350)	(0.345)	(0.341)	(0.344)
Country FE	YES	YES	YES	YES
Observations	222	222	222	222
R2	0.180	0.178	0.193	0.183
Adjusted R2	0.133	0.130	0.147	0.136

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

where N is the number of only positive words in all the paragraphs containing the site's name; v is a value from 1 (slightly positive) to 5 (really positive) assigned by the *AFFIN* dictionary to the word i, f is the frequency of the word i.

$$sentiment\_paragraph = \frac{1}{T}\sum_{j=1}^{T}\frac{1}{n_j}\sum_{w=1}^{N_j} f_{wj}V_{wj} \tag{9}$$

where n_j is the number of words in the paragraph j (containing the site's name); Vw_j is a value from −5 (really negative) to +5 (really positive) assigned by the *AFFIN* dictionary to the word w in the paragraph j, f is the frequency of the word w in the paragraph j and T is the number of paragraphs containing the site's name. This specification is the closest to the one used in the main regression analysis, which however included only positive evaluations.

The results, reported in Tables 7 and 8, confirm the same signs found in the main regressions: all the coefficients are positive and significant at least at the 10% level. Yet, the estimates are fully comparable with those obtained in the main regressions also in term of magnitude: the likelihood of receiving one additional criterion lays in a really small range (between 26 to 30%) regardless the formula used to compute *Sentiment*.

Table 6 Robustness checks: time—logit model

	Dependent variable:			
	Criteria			
	1	2	3	4
Ln_Citation	0.35*			
	(0.20)			
Ln_Icitation		0.13*		
		(0.08)		
Paragraph			0.01***	
			(0.00)	
Sentiment				0.21*
				(0.11)
Time	0.99	0.98	0.97	1.03*
	(0.62)	(0.62)	(0.61)	(0.61)
Time2	−0.18*	−0.18*	−0.18*	−0.19*
	(0.10)	(0.10)	(0.10)	(0.10)
Country FE	YES	YES	YES	YES
AIC	542.42	542.90	538.75	542.05
Log Likelihood	−254.21	−254.45	−252.38	−254.03
Num. obs	222	222	222	222
Iterations	6	6	6	6
McFadden's R2	0.07	0.07	0.08	0.07

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 7 Robustness checks: sentiment metrics—linear model

	Dependent variable		
	Criteria		
	1	2	3
Sentiment all	0.150*		
	(0.080)		
Sentiment positive		0.124***	
		(0.045)	
Sentiment paragraph			0.136*
			(0.076)
Constant	1.741***	1.668***	1.745***
	(0.153)	(0.147)	(0.155)
Country FE	YES	YES	YES
Observations	234	234	234
R2	0.120	0.138	0.118
Adjusted R2	0.080	0.099	0.079

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 8 Robustness checks: sentiment metrics—logit model

	Dependent variable:		
	Criteria		
	1	2	3
Sentiment all	0.30*		
	(0.17)		
Sentiment positive		0.26***	
		(0.10)	
Sentiment paragraph			0.28*
			(0.16)
Country FE	YES	YES	YES
AIC	599.83	595.33	600.02
Log Likelihood	−284.92	−282.67	−285.01
Num. obs	234	234	234
Iterations	6	6	6
McFadden's R2	0.05	0.05	0.05

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

6 Conclusions

In this paper we show that the number of criteria satisfied by the UNESCO WHS is correlated with four different measures of appreciation expressed in the Baedeker's travel guidebooks. These evaluations have not only the advantage of being authoritative, but they also avoid problems of endogeneity and reverse causation because the Baedeker's guidebooks were published more than half a century before the creation of the UNESCO WHL. Furthermore, they refer to a culturally homogeneous area, Europe, which legitimizes the interpretation of the plurality of UNESCO criteria as indicators of quality rather than of cultural diversity.

To verify the hypothesis under test, we have constructed an original data set based on the Baedeker's guidebooks published at the beginning of the twentieth century. These results empirically prove that the number of OUV criteria of UNESCO WHS is an informative ordinal proxy of the quality of the site. This confirmation is important because, given the single quantity of each UNESCO site, the preferences of tourists, of policymakers' and of the UNESCO experts about the sites can be captured only by another, more informative metric, that we broadly define as "quality". The broadness of the content of this concept of quality is defined by the specific criteria (aesthetic, historical, representativeness) that each site satisfies. The authority of the Baedeker's guidebooks as standard of evaluation reinforces the credibility of the UNESCO WHL, which has often been questioned due to the lobbying activities that may play a role in determining its composition.

Proxying a site's quality by the number of OUV criteria satisfied opens several possibilities in terms of re-examination of several research topics discussed in the UNESCO literature—and possibly in others as well. To make just a few examples, the analysis of the impact of the inclusion of a site in the UNESCO WHL on tourism flows, property values and local growth has been extensively examined, with mixed results, but none of these studies has ever considered the quality of the accepted sites. A possibility to be explored is that sites which satisfy a smaller number of criteria, and are thus relatively less well known to the general public, receive a larger boost from the inclusion on the list than sites of greater quality, which are already well known. This may explain the sometimes contradictory findings of the literature on the economic consequences of the UNESCO WHL. Moreover, the results of our study could also be interpreted as a forecaster of future inclusions in the WHL, exploiting the ability of the Baedeker's metrics to predict the potential of a site to meet the UNESCO standards. The list of possible applications of our proxy is certainly much longer.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10824-023-09493-8.

Funding Open access funding provided by Università degli Studi Roma Tre within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors hereby declare that neither of them has a conflict of interests of whatsoever nature about the research conducted in this paper.

References

Adie, B. A. (2017). Franchising our heritage: The UNESCO World Heritage Brand. *Tourism Management Perspectives, 24*, 48–53.

Adie, B. A., Hall, C. M., & Prayag, G. (2018). World heritage as a placebo brand: A comparative analysis of three sites and marketing implications. *Journal of Sustainable Tourism, 26*, 399–415.

Benchimol, J., Kazinnik, S., & Saadon, Y. (2022). Text mining methodologies with R: an application to Central Bank texts. *Machine Learning with Applications,* 8, 100286.

Bertacchini, E. E., Revelli, F. & Zotti, R. (2023). Lord, how I want to be in that number! On the blessing of UNESCO World Heritage listing. *CESifo Working Paper* n. 10293

Bertacchini, E. E., & Saccone, D. (2012). Toward a political economy of world heritage. *Journal of Cultural Economics, 36*, 327–352.

Bruce, D. M. (2010). Baedeker: The perceived "inventor'of the formal guidebook-A bible. *Giants of Tourism,* 93.

Buckley, R. (2004). The effects of World Heritage listing on tourism to Australian national parks. *Journal of Sustainable Tourism, 12*, 70–84.
Correa, R., Garud, K., Londono, J. M., & Mislang, N. (2021). Sentiment in Central Banks' financial stability reports. *Review of Finance, 25*, 85–120.
Dattilo, M., Padovano, F., & Rocaboy, Y. (2022). More is worse: The evolution of quality of the UNESCO World Heritage List and its determinants. *Journal of Cultural Economics,* 47(1), 71–96.
Dattilo, M., Padovano, F. & Rocaboy, Y. (2020). Is beauty defined by victors? An analysis of colonial sites of the UNESCO WHL, Economics Working Paper, Condorcet Center for Political Economy, CREM-CNRS 2020-04-ccr.
Fraiberger, S. (2016). News Sentiment and Cross-Country Fluctuations. In *Proceedings of the First Workshop on NLP and Computational Social Science* (pp. 125-131).
Frey, B. S., & Steiner, L. (2011). World Heritage List: Does it make sense? *International Journal of Cultural Policy, 17*, 555–573.
Ginsburgh, V., & Weyers, S. (1999). On the perceived quality of movies. *Journal of Cultural Economics, 23*, 269–283.
Heston, S. L., & Sinha, N. R. (2017). News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal, 73*, 67–83.
Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media, 8*, 216–225.
Jokilehto, J. (2006). World heritage: Defining the outstanding universal value. *City and Time, 2*, 1–10.
Keane, M. J. (1997). Quality and pricing in tourism destinations. *Annals of Tourism Research, 24*, 117–130.
Koshar, R. (1998). 'What ought to be seen': Tourists' guidebooks and national identities in modern Germany and Europe. *Journal of Contemporary History, 33*, 323–340.
Labadi, S. (2007). *World Heritage: Challenges for the millennium*. UNESCO.
Labadi, S. (2013). *UNESCO, cultural heritage, and outstanding universal value: Value-based analyses of the world heritage and intangible cultural heritage conventions*. AltaMira Press.
Nielsen, F. A. (2011a). AFINN. Informatics and mathematical modelling, Technical University of Denmark.
Nielsen, F. A. (2011b). A new anew: Evaluation of a word list for sentiment analysis in microblogs. https://doi.org/10.48550/arXiv.1103.2903
Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control, 127*, 104–119.
Patuelli, R., Mussoni, M., & Candela, G. (2016). The effects of world heritage sites on domestic tourism: A spatial interaction model for Italy. *Journal of Geographical Systems, 15*, 369–402.
Rakic, T. (2007). World heritage: Issues and debates. *Tourism: An International Interdisciplinary Journal, 55*, 209–219.
Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2020). Measuring news sentiment. *Journal of Econometrics, 228*, 221–243.
Shapiro, A. H., & Wilson, D. J. (2022). Taking the FED at its word: A new approach to estimating Central Bank objectives using text analysis. *The Review of Economic Studies, 89*, 2768–2805.
Sharma, N. & Privaryat. (2018). Sentiment analysis using tidytext package in R. In 2018 *First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp 577–580.
Silge, J., & Robinson, D. (2016). Tidytext: Text mining and analysis using tidy data principles in R. *Journal of Open Source Software, 1*, 37.
UNESCO. (1972). *Convention concerning the protection of the World cultural and natural Heritage*. UNESCO World Heritage Center.
UNESCO (2013). *Operational guidelines for the implementation of the World Heritage Convention*. UNESCO World Heritage Center.
UNESCO. (2017). *Operational guidelines for the implementation of the World Heritage Convention*. UNESCO World Heritage Center.
Van Blarcom, B. L., & Kayahan, C. (2011). Assessing the economic impact of a UNESCO World Heritage designation. *Journal of Heritage Tourism, 6*, 143–164.
Waldfogel, J. (2012). Copyright protection, technological change, and the quality of new products: Evidence from recorded music since Napster. *Journal of Law and Economics, 55*, 715–740.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Yang, C.-H., Lin, H.-L., & Han, C. C. (2010). Analysis of international tourist arrivals in China: The role of World Heritage Sites. *Tourism Management, 31*, 827–837.

Authors and Affiliations

Martina Dattilo[1] · Fabio Padovano[2]

✉ Fabio Padovano
fabio.padovano@univ-rennes1.fr

[1] Department of Economics and Statistics "Cognetti de Martiis", Università degli Studi di Torino, Turin, Italy

[2] DSP, Università Roma Tre, Rome, Italy

[3] Condorcet Center for Political Economy, CREM CNRS UMR 6211, University of Rennes, Rennes, France