



HAL
open science

Measuring quality of cultural heritage: consequences for the literature on the UNESCO World Heritage List

Martina Dattilo

► **To cite this version:**

Martina Dattilo. Measuring quality of cultural heritage: consequences for the literature on the UNESCO World Heritage List. Economics and Finance. Université de Rennes, 2022. English. NNT: 2022REN1G004. tel-04212155

HAL Id: tel-04212155

<https://theses.hal.science/tel-04212155>

Submitted on 20 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse de doctorat de

L'UNIVERSITE DE RENNES 1

Ecole Doctorale n°597

Sciences Economiques et sciences De Gestion

Spécialité : Sciences Economiques

Par

Martina DATTILO

**Measuring quality of cultural heritage: consequences for
the literature on the UNESCO World Heritage List**

Thèse présentée et soutenue à RENNES, le 13 Decembre 2022

Unité de recherche : CNRS UMR-6211 CREM

Rapporteurs avant soutenance :

Enrico BERTACCHINI Associate Professor, Università degli Studi di Torino

Francois MOREAU Professeur, Université Paris 13

Composition du Jury :

Président : David MASCLET Professeur, Université de Rennes 1

Examineurs : Enrico BERTACCHINI Associate Professor, Università di Torino

David MASCLET Professeur, Université de Rennes 1

Marianne LUMEAU Maître de Conférence, Université de Rennes 1

Co-directeur de thèse : Fabio PADOVANO Professeur, Université de Rennes 1

DSP, Università Roma Tre

Co-directeur de thèse : Yvon ROCABOY Professeur, Université de Rennes 1

The University of Rennes 1 does not intend to endorse nor deprecate any of the views expressed in this Ph.D thesis. These views are to be regarded only as those of their author.

L'Université de Rennes 1 n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Acknowledgments

First of all, I want to thank the first persons that believed in me and made my vague ideas a concrete reality, my two supervisors, Proff. Fabio Padovano and Yvon Rocaboy. From the beginning, they trusted my capability to do research and develop original ideas; they respected me like a peer, made me responsible for my decisions and always asked for my opinions. I should acknowledge that dealing with a headstrong person like me is not always easy, but they were always extremely patient and careful. I hope this is just the start of a long and successful collaboration!

Then, I would like to thank the members of the evaluation committee of this thesis and especially the ones who have accepted to be the rapporteurs: Francois Moreau and Enrico Bertacchini, other than Marianne Lumeau and David Masclet.

I also want to thank the University of Rennes 1 and the past and present directors of the CREM (Centre de Recherche en Économie et Management) laboratory, Franck Moraux, David Masclet, Fabien Moizeau and Laurent Denant-Boëmont, that provides me with all the means to efficiently conduct my research and were always willing to give me suggestions on how I could improve my work. Many thanks go to the administration of the CREM, in the persons of Anne L'Azou and Naila Louise-Rose, and a special one goes to Cécile Madoulet that was always there for me, not only to solve practical issues but also to listen and assist. I also want to thank the *école doctorale EDGE*, in the persons of Héléne Jean, Thierry Pénard and Eric Malin, for the friendly support at each step. Together with the *Region Bretagne* and the CREM laboratory, it financially sponsored my experience as a visiting researcher at the University of Birmingham, an extremely enriching opportunity, fundamental for my development as a researcher.

A number of other teachers and professors have had an important impact on

my research and teaching capabilities. I want to thank Prof. Eric Darmon who has always trusted me and the teaching material I have provided. His passion for delivering lectures is a true inspiration. Then, I want to thank Dr James Rockey who welcomed me to the University of Birmingham. He was always available to advise me, read my research works and discuss any issue. Then, I also want to thank Prof Davide Vannoni, director of the Master in Economics at the University of Turin, who, before anyone else and against the opinion of the administrative office, supports my application to the double degree program between the University of Turin and the University of Rennes 1, making my experience at Rennes possible. Moreover, to all the other professors of the faculty of economics of the Rennes 1 who mentor me or collaborate with me, I want to express my profound gratitude.

And now my colleagues of the PhD program! Many thanks to Romain Gate who was there at the really beginning and the really end of this journey and to Jimmy Merlet who fills our days with his willingness to laugh and joke. Thanks to Jacques (a super office mate!), Thibaut Arpinon and Sébastien Irigoyen for helping me during my year as an ATER: I would have never made it without you! (I will repay my debt in form of candies). I am full of gratitude to Etienne Dagorn et Matthieu Pourieux for all the pertinent comments that substantially improved my work. I learn a lot working with you: you push me to do my best and, in short, to be a better researcher. Thanks to Alejandra Giraldo for being, not only a colleague, but also a true friend. Thanks to Roberto Brunetti; this page is too short for detailing all the memories we share, but they have a special place in my heart (and good luck for the Nobel prize!). And like them, many thanks to Thibault, Vincent, Nathalie, Enora, Maeva, Louise, Lucile, Francesco, Ahmet, Shervin, Youssoufa and all the others that were, at the same time, colleagues, friends, coauthors and advisors. You make this journey a shared experience and, even in the worst moments, I never feel alone.

Many thanks to all my friends that shared my joys and anxieties. Thanks to my family who helped me colouring the Winnie Po calendar in the second grade, finishing the homework of Science, Art and Geography in middle school, moving in and out of every apartment and accepting me for who I am. Thanks to my brother for so many things I am not able to list, but basically for being there every time I needed help and cheering me up every time I felt that I was not enough. We are

an unbeatable team! Last but not least, thanks to my boyfriend, friend, colleague and mentor, Andrea, who not only has supported me emotionally, has made a real scientific contribution to all my ideas, always sees the best in me, but mostly, for doing it without any limitations or conditions.

One last special thank goes, again, to my mother. When I was a 13 years old little girl, for the first time I was confronted with a research project, or what could be called a research project in middle school, free to choose the topic and the methodology. As I did not know what topic to choose, I asked my mother for suggestions. She told me that the occasions when we can freely speak and people are there with the only scope to listen to us are rare in life. Therefore, the best is to say something you really want to say. I can see today as this talk had a profound impact on my life and my decisions as a researcher. I have developed the profound belief that each topic we are researching is somehow personal. Since we are spending so much time and effort working on our research, the fact that it is really relevant to society and to us, in the end, is the only thing that matter. So, when I look back at the past 10 years of my education, I can see the coherence of my path and how the issue of aesthetic beauty and its impact on society has been walking next to me all this time. For this reason, for the incredible luck I had ending up in an incredible city like Rennes and meeting all the amazing persons cited above, these years spent doing the PhD were the best of my life and the positive and joyful moments largely outweighed the hard times. Thanks!

Résumé

Les politiques visant à préserver et à promouvoir le patrimoine culturel impliquent souvent la création de listes et d'inventaires, qui peuvent atteindre des dimensions considérables (Klamer et al., 2013). En Italie, 45 025 propriétés sont considérées comme des biens architecturaux protégés (Ministero dei beni e delle attivit'a culturali, 2015) ; tandis qu'en France, la liste des monuments historiques compte 45 588 sites (Ministère de la Culture, 2021). Au niveau international, l'UNESCO est la principale autorité de définition du patrimoine culturel. En effet, depuis 1972, la préservation du patrimoine de l'humanité est sa mission principale. À cette fin, la Liste du Patrimoine Mondial (LPM) ne doit contenir que du patrimoine de qualité exceptionnelle, une sorte de "best of" du patrimoine culturel et naturel de l'Humanité. En 2022, 897 sites du patrimoine culturel, 218 sites naturels et 39 sites mixtes ont reçu le label de l'UNESCO, pour un total de 1154 sites inscrits à la LPM de l'UNESCO. La désignation est généralement associée à des avantages économiques directs ou indirects (Benhamou, 2013, Klamer et al., 2013), ce qui rend l'inscription attrayante pour de nombreux acteurs et potentiellement sujette à des activités de lobbying. Malgré l'abondante littérature sur les différents aspects de la LPM de l'UNESCO (Aa, 2005 ; Bertacchini et Saccone, 2012 ; Bertacchini et al. 2015, 2016 ; Frey et al., 2010 ; Frey et Steiner, 2011 ; Stainer et Frey, 2011 ; Parenti et De Simone, 2015), la question de l'évolution de la qualité des sites a rarement été examinée jusqu'à présent. Il est donc intéressant de vérifier si et comment les conclusions de cette littérature changent lorsqu'une certaine conceptualisation de la qualité est prise en considération.

Cette thèse contribue à la littérature en fournissant un aperçu de la manière dont la qualité du patrimoine culturel peut être évaluée. En effet, nous considérons comme indicateur de qualité le nombre de critères de Valeur Universelle Exceptionnelle (VUE) qu'un site remplit au moment de son inscription à la LMP. Cet

indicateur constitue une mesure simple et directe de la qualité des sites du patrimoine mondial (PM) de l'UNESCO, qui réduit la subjectivité. Grâce à cet indicateur, nous examinons de manière critique certains points de vue établis dans la littérature à la LPM de l'UNESCO. Par exemple, nous montrons que l'augmentation du nombre de sites inscrits sur la LPM réduit la qualité marginale des sites nouvellement inscrits (chapitre 1). En exploitant un ensemble de données original, nous testons la robustesse de cet indicateur et nous montrons qu'il est corrélé avec d'autres mesures possibles de la qualité (chapitre 2). Ainsi, l'indicateur que nous proposons s'avère être proche de la qualité des sites, telle qu'exprimée dans les guides de voyage Baedeker, source faisant autorité au début du XX^{ème} siècle, c'est-à-dire à une période très éloignée de la création de la LPM de l'UNESCO. En outre, cette thèse clarifie le rôle des experts dans l'établissement de ces normes et suggère comment ces mesures pourraient être utilisées pour l'évaluation des politiques (chapitre 3). En prenant en considération les concepts d'intégrité et d'authenticité, en plus des simples critères de VUE, nous sommes en mesure de construire un indice encore plus complet de la qualité des sites proposés pour inscription. Il nous permet d'évaluer le rôle des experts de l'UNESCO dans la définition et l'application des critères d'inscription. En contrôlant la qualité de cette manière, on constate que les experts de l'UNESCO sont impartiaux et qu'ils attribuent le même nombre de critères au patrimoine indigène et colonial, ce qui exclut les craintes d'un parti pris culturel pro-européen dans la sélection des sites du PM.

Bien qu'elle soit largement adoptée, la réglementation sur la préservation du patrimoine culturel par le biais de politiques d'inscription doit être examinée avec soin. En particulier dans les pays développés, les bâtiments et les monuments n'ont pas été créés dans le but premier d'élargir le capital culturel du pays (Bonet, 2013), mais ils en font partie à travers le processus d'inscription (Peacock, 1995). Pour cette raison, Towse (2019) caractérise la demande de patrimoine culturel comme étant induite par les fournisseurs. Intuitivement, cet aspect du marché du patrimoine culturel a des conséquences importantes, notamment lorsque le statut des sites du patrimoine culturel nécessite des investissements publics. En effet, comme les ressources financières sont limitées et que l'ensemble du capital culturel ne peut être préservé, la sélection est généralement basée sur la "qualité" ou la "valeur" des sites. Plus généralement, la qualité est habituellement la principale raison de consommer un bien culturel et la différenciation des produits est si extrême que

les produits culturels sont souvent uniques. Le concept de qualité, cependant, est difficile à caractériser et il est surtout subjectif. Au total, cette ambiguïté renforce le pouvoir discrétionnaire des experts (Klamer et al., 2013), et ouvre la porte aux spéculations et aux activités de lobbying. Dans le cas de l'UNESCO, la sélection efficace des sites est une préoccupation majeure de l'organisation, car elle lui permet de préserver sa réputation internationale.

Pour répondre à ces questions, les chercheurs et les décideurs politiques ont besoin de données évaluant la qualité des biens culturels et, plus particulièrement, des sites du patrimoine culturel. Ces mesures sont difficiles à concevoir pour deux raisons principales. Premièrement, il n'y a pas d'accord sur les éléments qui devraient être contenus dans un tel indicateur. Deuxièmement, la plupart des statistiques culturelles se concentrent principalement sur les aspects économiques des industries créatives, laissant le secteur du patrimoine indûment abordé. En effet, bien qu'elles soient fondamentales pour de nombreuses raisons, dont l'évaluation des politiques, les données sur cette importante sous-section du secteur culturel sont loin d'être satisfaisantes (Towse, 2019).

Chapitre 1 : Plus c'est pire : Diminution de la qualité marginale de la liste du patrimoine mondial de l'UNESCO

Le chapitre 1 analyse empiriquement l'évolution de la qualité des sites inscrits sur la LPM de l'UNESCO de 1972 à 2016. Nous exploitons le fait que, pour figurer sur la liste, chaque site doit satisfaire à au moins un des dix critères de valeur universelle exceptionnelle, sur lesquels l'UNESCO fonde son évaluation. Étant donné que les critères prennent en compte les différentes raisons pour lesquelles un site peut mériter d'être inclus dans la LPM (comme sa pertinence historique, artistique, représentative, etc.), nous utilisons le nombre de critères satisfaits au moment de l'inscription comme indicateur de la qualité du site. Les résultats de ce chapitre montrent que les pays proposent des sites de qualité décroissante avec le nombre de sites inscrits. Selon notre interprétation, cela est lié à l'existence d'une quantité fixe de capital culturel dont la qualité marginale est décroissante : si un pays augmente constamment le nombre de sites dans la LPM alors que le stock de capital culturel reste constant, la qualité marginale des sites diminuera. Ainsi, nous observons l'effet le plus fort pour les pays ayant plus de 10 sites inscrits sur la LPM, un

sous-échantillon qui représente 12 % des pays et 51 % du total des sites. Contrairement aux études précédentes qui se concentraient uniquement sur le nombre de sites inscrits sur la liste, la prise en compte de la qualité montre que la puissance de lobbying du pays n'a pas d'importance pour l'inscription sur la LPM, alors que la qualité de son administration en a une. Pourtant, les sites de qualité élevée et incontestée ne devraient pas avoir besoin d'une forte pression politique pour être inclus dans la LPM. Ces résultats sont robustes aux tests de stabilité des critères d'évaluation de l'UNESCO dans le temps et aux changements d'estimateurs économétriques.

Chapitre 2 : Évaluation de la qualité de la liste du patrimoine mondial de l'UNESCO : Une comparaison avec les guides Baedeker

Bien qu'intuitive et simple, l'idée de compter le nombre de critères satisfaits comme indicateur de la qualité d'un site ne fait pas l'unanimité. On pourrait faire valoir, par exemple, que ces critères sont censés rendre compte de la diversité culturelle des sites, plutôt que de la qualité individuelle d'un site. Une évaluation plus approfondie de la question est donc justifiée. Le chapitre 2 vérifie si le nombre de critères de VUE que les sites du patrimoine mondial de l'UNESCO satisfont est un indicateur raisonnable de leur qualité. Pour éviter les problèmes d'endogénéité et de causalité inverse, nous examinons la corrélation entre le nombre de critères satisfaits et l'évaluation faite par un guide de voyage faisant autorité, à savoir le guide Baedeker du début du XX^{ème} siècle. Puisqu'à l'époque de la publication des guides Baedeker l'UNESCO n'existait pas, leur sélection comme norme de qualité évite les problèmes d'endogénéité et de causalité inverse dont peut souffrir tout autre indicateur d'appréciation des consommateurs ou des experts. À cette fin, nous construisons un ensemble de données original concernant 10 pays européens à partir de 11 guides Baedeker's, allant de 1899 à 1911. Nous évaluons l'appréciation du patrimoine culturel par Baedeker à travers quatre mesures différentes : le nombre total de citations, le nombre pondéré de citations, la longueur moyenne des paragraphes qui contiennent au moins une citation et le sentiment exprimé dans le texte. Nous montrons que ces quatre mesures sont positivement et significativement corrélées avec le nombre de critères auxquels le site répond, ce qui confirme qu'elles constituent un bon indicateur de la qualité des sites du patrimoine mondial de l'UNESCO. Ce résultat est robuste à différents modèles économétriques et à l'introduction des effets fixes par pays.

Chapitre 3 : Y a-t-il un biais pro-européen dans la sélection des sites de l'UNESCO ? Un nouveau test basé sur les sites coloniaux

Outre l'argument de la supériorité du pouvoir de lobbying, la capacité des pays européens à imposer leurs normes esthétiques est un autre argument fréquent en faveur de la surreprésentation de l'Europe dans la LPM de l'UNESCO, puisque ces normes déterminent en fin de compte le choix des sites candidats à la LPM. Dans le chapitre 3, nous abordons cette question en examinant la qualité perçue des sites dans des pays où ceux-ci peuvent être soit indigènes, soit de type européen, c'est-à-dire construits par des colonisateurs européens. La qualité des sites est calculée à la fois comme la somme des critères de la VUE, le même indicateur que celui utilisé dans les chapitres 2 et 3, et par une analyse textuelle des rapports du Conseil international des monuments et des sites (ICOMOS), permettant ainsi de considérer également l'intégrité et l'authenticité comme une composante de la qualité. Ces évaluations sont principalement motivées par des normes esthétiques car elles sont réalisées à un stade précédant le processus de lobbying pour l'inclusion des sites dans la LPM. Après avoir contrôlé de nombreux facteurs potentiellement conditionnants, les estimations font ressortir que ni les différences de qualité ou de probabilité d'inscription n'existent entre les sites coloniaux et indigènes, ce qui suggère que les experts de l'ICOMOS semblent être impartiaux.

Contents

General Introduction	1
1 - More Is Worse: Decreasing Marginal Quality Of The UNESCO WHL	13
1 Introduction	14
2 Literature review	17
2.1 Evaluation of quality	17
2.2 Studies about the politics of UNESCO	19
3 The decision-making process behind the UNESCO WHL	20
3.1 The UNESCO selection procedure	20
3.2 Criteria of Outstanding Universal Value	22
4 Empirical strategy	26
4.1 Dependent variable and estimation issues	26
4.2 Baseline model and explanatory variables	27
5 Estimation of the baseline model	31
6 Robustness checks	34
6.1 Alternative specifications for N_{it}	34
6.2 Disaggregating countries by the number of sites	36
6.3 Stability of criteria	38
6.4 Specification of the model	40
7 Conclusions	43
References	44
2 - Evaluating the Quality of UNESCO World Heritage List	47
1 Introduction	48
2 The Baedeker guidebooks	50
3 Data and measurements	51
3.1 Number of citations	53

3.2	Length of the paragraph	54
3.3	The sentiment of the text	55
4	Empirical Strategy	59
5	Results	60
6	Conclusions	64
	References	65
 Appendix A - Evaluating the Quality of UNESCO World Heritage		
	List	69
 3 - Is there a pro-European bias in the selection of UNESCO sites?		83
1	Introduction	84
2	Literature Review	86
3	Nomination and evaluation process	88
3.1	Nomination process	88
3.2	ICOMOS evaluation process	89
4	Empirical Strategy	90
4.1	Dependent variable	90
4.2	Independent variable: Colonization	92
5	Results of the baseline estimates	95
6	Robustness checks	98
6.1	Type of sites	98
6.2	Heterogeneity across continents	104
6.3	Experts' personal bias	108
6.4	ICOMOS recommendation	114
7	Conclusion	119
	References	121
 Appendix B - Is there a pro-European bias in the selection of		
	UNESCO sites?	123
B.1	List of Countries	124
B.2	Colonization determinants	125
B.3	List of variables	126
 General Conclusion		127
	References	129

General Introduction

”Et si je connais, moi, une fleur unique au monde, qui n’existe nulle part, sauf dans ma planète, et qu’un petit mouton peut anéantir d’un seul coup, ce n’est pas important ça ?”

Le Petit Prince, Antoine De Saint-Exupéry

Policies aiming to preserve and promote cultural heritage often involve the creation of lists and inventories, which can reach considerable dimensions (Klamer et al., 2013). The UNESCO organisation is the primary cultural heritage designation authority, with 1154 World Heritage (WH) sites in 2022. In addition to the internationally recognised sites, many countries define their own list of protected heritage. In Italy, 45 025 properties are considered protected architectural assets (Ministero dei beni e delle attività culturali, 2015); while in France, the list of historical monuments counts 45 588 sites (Ministère de la Culture, 2021). The designation is usually associated with direct or indirect economic benefits (Benhamou, 2013, Klamer et al., 2013), which makes the inclusion appealing for many actors and potentially subject to lobbying activities.

This thesis contributes to the cultural economics literature by providing insights on how the quality of cultural heritage can be proxied. Indeed, we consider the number of criteria of Outstanding Universal Value (OUV) that a site met at the moment of the inscription as an indicator of quality. It constitutes a simple and straightforward measure of the quality of UNESCO WH sites, which reduces subjectivity. Thanks to this index, we critically investigate some established views of the literature on the UNESCO World Heritage List (WHL). For example, we show that enlarging the number of sites inscribed in the WHL reduces the marginal qual-

ity of newly inscribed sites (Chapter 1). Exploiting an original dataset, we test the robustness of this indicator and we show that it is correlated with other possible measures of quality (Chapter 2). In detail, our index is found to be in line with the quality of the sites, as expressed in the Baedeker travel guidebooks, authoritative source of the beginning of the XXth century, i.e. in a period very distant from the creation of the UNESCO WHL. In addition, this thesis clarifies the role of experts in setting those standards and suggests how these measures could be used for policy evaluations (Chapter 3). Taking into consideration the concepts of *uniqueness*, *integrity* and *authenticity*, in addition to the mere criteria of OUV, we are able to construct an even more comprehensive index of the quality of the nominated sites. It allows us to assess the role of UNESCO experts in setting and applying the criteria for inscription. Controlling for quality in such a way, UNESCO experts are found to be impartial and assign the same number of criteria to native and colonial heritage, thus excluding concerns of a pro-European cultural bias in the selection of WH sites.

Despite being so widely adopted, regulating the preservation of cultural heritage through listing policies should be carefully examined. Especially in developed countries, buildings and monuments were not created with the primary scope of enlarging the country's cultural capital (Bonet, 2013), but they become part of it through the process of inscription (Peacock, 1995). For this reason, Towse (2019) characterises the demand for cultural heritage as supplier-induced. Intuitively, this aspect of the cultural heritage market has important consequences, especially when the status of cultural heritage sites requires public investments. Indeed, since financial resources are limited and not all cultural capital can be preserved, the selection is usually based on the “quality” or “value” of the sites. More generally, quality is usually the main reason to consume a cultural good and product differentiation is so extreme that cultural products are often unique. The concept of quality, however, is difficult to characterise and it is mostly subjective. Eventually, this ambiguity strengthens the discretion of those who actually make the decision concerning the inscription, i.e. the experts (Klamer et al., 2013), and opens the door to speculations and lobbying activities. In the UNESCO case, efficiently selecting the sites is a principal concern for the organisation, since it allows it to preserve its well-established international reputation.

In order to assess these questions, researchers and policy-makers need data assessing the quality of cultural goods and, specifically, cultural heritage sites. These measurements are hard to conceive for two main reasons. First, there is no agreement on which elements should be contained in such an index. Secondly, most cultural statistics focus mainly on the economic aspects of creative industries, leaving the heritage sector improperly addressed (Deloumeaux, 2013). Indeed, despite being fundamental for many reasons including the policy evaluation, data on this important subsection of the cultural sector are far from satisfactory (Towse, 2019).

The Economics of Cultural Heritage

In the absence of appropriate data on the quality and availability of cultural heritage sites, the economics of cultural heritage literature has largely focused on the demand side, and, more specifically on cultural heritage tourism ¹. This phenomenon has its roots in the tradition of the Grand Tour, the educational travel across landmarks of Western culture followed by the upper European classes (Bonet, 2013, Brodsky-Porges, 1981). The educational purposes, together with an increasing interest in the fine arts, pushed students and their tutors toward cities and learning centres (Brodsky-Porges, 1981), motivated mainly by the destination's cultural environment (Towner, 1985). In the XIXth century, this practice evolved into contemporary mass tourism, but, in this context, the location's cultural value had little or no importance. In recent times, however, the more exigent demand and the rising competition with new destinations have stimulated the search for added-value products and cultural attractions have gained importance and consideration (Bonet, 2013).

More generally speaking, the tourism industry is an increasingly significant source of revenue for public and private actors. In fact, the revenues from international tourism account for almost 3% of the European Union's GDP (World Tourism Organization, 2021). Despite the dramatic fall in revenues due to the SARS-Covid pandemic, international tourism saw a strong rebound in the first five months of 2022, recovering by almost half (46%) of pre-pandemic 2019 levels (World Tourism Organization, 2022). Europe, in particular, welcomed more than four times as many

¹We should mention that also cultural tourism data, similarly to data on the touristic sector in general, can be rare, insufficient and subject to different definitions. A huge effort has been made in recent years to improve the objectivity and granularity of this data

international arrivals as in the first five months of 2021 (+350%), boosted by strong intra-regional demand (World Tourism Organization, 2022).

For some European countries such as Italy, Germany and France, however, the cultural tourism sector alone was estimated to be greater than 3% of countries' GDP in 2008, reaching a peak of almost 7.5% for Spain (World Tourism Organization, 2018). In particular, monuments and heritage sites attract most of the cultural tourism demand, well above the tourism of festivals, live performances and contemporary art (Bonet, 2013). In fact, 43.4 % of the European adult population visited at least one cultural site in 2015; while the percentage of Europeans attending live performances was estimated to be 42.8% in the same year (Eurostat, 2016). For a significant portion of tourists, the cultural environment represents a primary reason for travelling (30% on average) and the market size is even larger if we consider visitors undertaking cultural activities as a secondary motivation (World Tourism Organization, 2018). Although relevant in terms of visitor numbers, the cultural heritage sector is even more important in terms of expenditure: cultural tourists tend to be relatively high spending, generating income and employment (World Tourism Organization, 2018). Evidence has shown that this market has grown in Europe in recent years (excluding the pandemic period), as European countries reported an average increase of cultural tourists' arrival of 20% from 2014 to 2018 (World Tourism Organization, 2018). All these figures make the cultural heritage sector among the first in the cultural tourism industry in terms of importance.

These performances justify the scholars' and policy-makers' growing interest in the economics of cultural heritage. Policy-makers' attention is especially devoted to the public good argument and the consequent need for government intervention. The issue, however, does not encounter universal agreement among cultural economists. If outdoor built and natural heritage clearly possesses public-good characteristics, museum visitors can be easily excluded from consumption (Towse, 2019, Benhamou, 2013) and therefore the need for subsidies is harder to justify. In addition, in well-known touristic destinations, congestion causes rivalry in consumption. In the extreme case of over-tourism, a too high number of visitors is detrimental to the site preservation and the quality of the visit. Indeed, in the presence of asymmetric information and increasing touristic pressure on historical cities, the producer of cultural services may act as a monopolist and reduce the quality of the

product supplied, with important repercussions on the reputation of the destination (Caserta and Russo, 2002). Indeed, since cultural tourism demand is primarily generated by a high-quality cultural environment, if the value-generation capacity of a certain destination decreases, less money can be reinvested in conservation and, as a consequence, the destination loses its attractiveness. Between the different policies proposed, introducing a pricing regulation to limit the tourists flow appears to be appropriate (Towse, 2019). An astonishing example is the entry tax implemented by the municipality of Venice which should reduce over-tourism and enhance the city's preservation. Experts even consider restricting access to important heritage sites, such as St Mark's Square (Bertocchi et al., 2020).

The issue of quality in Economics and in Cultural Economics

The issue of quality, and its implication for market competition, has been a concern for different disciplines in the economics literature. The idea that products can differ with respect to some characteristics can be found for the first time in Hotelling's "Stability in Competition" in 1929. Indeed, he analysed the competition between firms in a market in which buyers were dispersed geographically and in which firms competed by varying locations in addition to setting prices. This seminal paper gave birth to a flourishing literature analysing the causes and consequences of product differentiation. A few years later, in 1933, Edwin Chamberlin produced another relevant piece of economics theory by publishing his classic "Theory of Monopolistic Competition". The essay provided a theory of product variety where firms differentiated their products in an optimal way.

Another significant step forward in the literature's definition of quality and product differentiation was made in 1966 when Lancaster introduced a seminal idea: what differentiates products is the characteristics that they possess. In short, his theory assumes each product to be a bundle of different characteristics. Consumers have preferences over each element of the bundle. Therefore, each characteristic returns a particular utility. In equilibrium, the consumer selects the bundle with the marginal utility equal to the implicit price of the characteristics.

Even if this approach has regained attention in recent empirical analysis (Bajari and Benkard, 2005), the current dominant literature understands the problem of product differentiation via two main classes of models. As in Belleflamme and Peitz (2015), the first one, called the *representative consumer approach*, assumes that all consumers are identical. One example is the linear demand model, which assumes that any single consumer is representative of all other consumers. Indeed, the resulting aggregate demand is simply the rescaled individual demand. For the *discrete choice approach*, on the contrary, the main source of product differentiation is the heterogeneity among consumers: they decide to buy or not a particular product (making a discrete choice), based on their preferences over product characteristics. Between the class of discrete choice models, we can distinguish between vertical and horizontal product differentiation.

Vertical product differentiation and the issue of quality are closely related. To understand this concept we must assume, first, that a characteristics space corresponding to the set of goods on offer exists. Secondly, we should assume that all the points of this space lie on the same ray vector through the origin. Under this assumption, a good that is further out along this ray is better, i.e. it has a higher quality (Belleflamme and Peitz, 2015). A sample of goods satisfying these assumptions is said to be *vertically differentiated*. As a consequence, we can rank them in terms of some quality index. At the equilibrium, if all the goods have the same price, all consumers will buy the product with the highest quality.

A critical assumption of the model of vertical differentiation is that consumers can identify and rank the quality of the goods. Indeed, how the information available influences equilibrium and consumers' choice in relation to the quality of the products became a primary concern during the '70 and '80. Among the first to raise this concern, Nelson (1970) introduced the concept of *experience goods*, for which products' quality can only be acknowledged by consumption. Indeed, when consumers do not observe the quality of the product but producers do, asymmetric information can arise (Akerlof, 1970). Not in all cases, however, low-quality product crowds out high-quality ones: consumers' expectations about quality, or firm reputation, are determinant of market power and can be used as an incentive to keep high-quality standards (Shapiro, 1982, 1983, Rogerson, 1983).

Most cultural goods and works of art are experience goods. Indeed, many cultural economics scholars have investigated the role of expert opinion in mitigating this lack of information. The role of experts in determining what is worthwhile art is sometimes controversial. If some scholars sustain that some experts' judgment, even if subjective, is necessary (Ginsburgh and Weyers, 1999) others prefer to let each individual define what "Art" is (Frey, 1994). Nevertheless, many papers have found a significant effect of experts' judgment on consumer demand in the domain of cinema (Nelson et al., 2001, Deuchert et al., 2005, Reinstein and Snyder, 2005) and books (Ponzo and Scoppa, 2015). These evaluations, however, are not always accurate and could be improved (Ashenfelter and Jones, 2013). In conclusion, the discussion concerning the evaluation of the quality of cultural and not-cultural goods is far from being settled.

Outline

Since 1972, preserving the heritage of mankind is the principal mission of the UNESCO organisation. To this aim, the WHL should only contain heritage of outstanding quality, a selection of the "very best". As of 2022, 897 cultural heritage sites, 218 natural sites and 39 mixed sites have been awarded with the UNESCO label, for a total of 1154 sites inscribed in the UNESCO WHL. As emphasised at the beginning, this thesis contributes to the cultural economics literature by highlighting how shifting the emphasis from quantity to quality can modify our knowledge of the cultural heritage sector as well as our interpretation of policies and regulations. Indeed, despite the extensive literature on various aspects of the UNESCO WHL (van der Aa, 2005, Bertacchini and Saccone, 2012, Bertacchini et al., 2017, 2016, Steiner and Frey, 2012, Frey et al., 2013), the issue of the evolution of the quality of the WH sites has rarely been examined so far. It is therefore interesting to verify if and how the conclusions of this literature change when our conceptualization of quality is taken into consideration. To this aim, we produce original data and statistics, which provide the basis for a critical analysis of UNESCO WHL and its policies.

Chapter 1 empirically analyzes the evolution of the quality of the sites included in the UNESCO WHL from 1972 till 2016. We exploit the fact that, to enter the list, each site must satisfy at least one of ten non-alternative criteria of OUV, upon which UNESCO base their evaluation. Since the criteria capture different reasons why a site might deserve to be included in the WHL (such as its historical, artistic, representative relevance and the like), we use the number of criteria satisfied at the moment of the inscription as a proxy for the quality of the site. The results of this chapter show that countries propose sites of decreasing quality over time. In our interpretation, this is linked to the existence of fixed stock of cultural capital: if a country constantly increases the number of sites in the WHL while the stock of cultural capital remains constant, the sites' marginal quality will decrease. Indeed, we observe the strongest effect for countries with more than 10 sites inscribed in the WHL². Contrary to previous studies focusing just on the number of sites included in the list, considering quality shows that the country's lobbying power does not matter for inclusion in the WHL, while the quality of its administration does. Yet sites of high and undisputed quality should not need much political push to be included in the WHL. These results are robust to tests of the stability of the UNESCO evaluation criteria over time and to changes in econometric estimators.

Yet, although intuitive and straightforward, the idea of counting the number of criteria satisfied as a proxy for the site's quality does not meet universal acceptance. It could be argued, for instance, that these criteria are meant to capture cultural diversity among sites, rather than the individual quality of a site (Jokilehto, 2006). Hence a more in-depth evaluation of the issue is needed. Chapter 2 verifies if the number of criteria of OUV that UNESCO WH sites satisfy is a reasonable proxy for their quality. To avoid problems of endogeneity and reverse causality, we examine the correlation between the number of criteria satisfied and the evaluation made by an authoritative travel guidebook, i.e., the Baedeker's guidebook of the early XXth century. Since at the time of the publication of the Baedeker guidebooks the UNESCO did not exist, their selection as a standard for quality avoids problems of endogeneity and reverse causality that any other indicator of consumers' or experts' appreciation may suffer from. To this end, we construct an original dataset considering 10 European countries from 11 Baedeker's guidebooks, ranging from 1899

²This subsample represents 12% of the countries and 51% of the total sites.

to 1911. We evaluate Baedeker's tastes for cultural heritage through four different measures: the total number of citations, the weighted number of citations, the average length of the paragraphs which contain at least one citation and the sentiment expressed in the text. We show that all these four measures are positively and significantly correlated with the number of criteria that the site satisfies, confirming that they are a good proxy for the quality of UNESCO WH sites. This result is robust to different econometrics models and to the inclusion of country fixed effects.

Together with the superior lobbying power argument, the European countries' ability to impose their aesthetic standard is another frequent argument in favour of Europe's over-representation in the UNESCO WHL; since aesthetic standards ultimately determines the choice of the candidate sites for the WHL. In Chapter 3, we address this issue by examining the perceived quality of sites in countries where these can be either indigenous or European-like, i.e. built by European colonizers. The quality of sites is computed both as the sum of criteria of OUV, the same index used in Chapters 2 and 3, and by a textual analysis of the reports of the International Council on Monuments and Sites (ICOMOS), allowing to consider also the integrity and the authenticity as a component of quality. These evaluations are mainly driven by aesthetic standards because they are made at a stage preceding the lobbying process for the inclusion of the sites in the WHL. After controlling for numerous potentially conditioning factors, the estimates point out that neither differences in quality nor in the probability of inscription exist between colonial and native sites, suggesting that ICOMOS experts appear to be impartial.

This dissertation is structured as follows: for each chapter the body of the chapter is first presented, followed by a dedicated appendix when available. Then, the dissertation proceeds to the next chapter until the general conclusion. Each chapter contains a reference section exposing the related bibliography.

References

- Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.
- Ashenfelter, O. and Jones, G. V. (2013). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics*, 8(3):285–293.
- Bajari, P. and Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of political economy*, 113(6):1239–1276.
- Belleflamme, P. and Peitz, M. (2015). *Industrial organization: markets and strategies*. Cambridge University Press.
- Benhamou, F. (2013). Public intervention for cultural heritage: normative issues and tools. In *Handbook on the economics of cultural heritage*, pages 3–16. Edward Elgar Publishing.
- Bertacchini, E., Liuzza, C., and Meskell, L. (2017). Shifting the balance of power in the unesco world heritage committee: an empirical assessment. *International journal of cultural policy*, 23(3):331–351.
- Bertacchini, E., Liuzza, C., Meskell, L., and Donatella, S. (2016). The politicization of unesco world heritage decision making. *Public Choice*, 167.
- Bertacchini, E. E. and Saccone, D. (2012). Toward a political economy of world heritage. *Journal of cultural economics*, 36(4):327–352.
- Bertocchi, D., Camatti, N., Giove, S., and van der Borg, J. (2020). Venice and overtourism: simulating sustainable development scenarios through a tourism carrying capacity model. *Sustainability*, 12(2):512.
- Bonet, L. (2013). Heritage tourism. In *Handbook on the economics of cultural heritage*. Edward Elgar Publishing.
- Brodsky-Porges, E. (1981). The grand tour travel as an educational device 1600–1800. *Annals of tourism research*, 8(2):171–186.

- Caserta, S. and Russo, A. P. (2002). More means worse: Asymmetric information, spatial displacement and sustainable heritage tourism. *Journal of Cultural Economics*, 26(4):245–260.
- Deloumeaux, L. (2013). Current challenges in cultural statistics: a focus on heritage. *Handbook on the Economics of Cultural Heritage*.
- Deuchert, E., Adjamah, K., and Pauly, F. (2005). For oscar glory or oscar money? *Journal of Cultural Economics*, 29(3):159–176.
- Eurostat (2016). Cultural statistics.
- Frey, B. (1994). Art: The economic point of view. In *Cultural economics and cultural policies*, pages 3–16. Springer.
- Frey, B. S., Pamini, P., and Steiner, L. (2013). Explaining the world heritage list: an empirical study. *International Review of Economics*, 60(1):1–19.
- Ginsburgh, V. and Weyers, S. (1999). On the perceived quality of movies. *Journal of Cultural Economics*, 23(4):269–283.
- Jokilehto, J. (2006). World heritage: Defining the outstanding universal value. *City & time*, 2(2):1–10.
- Klamer, A., Mignosa, A., and Lyudmila, L. (2013). Cultural heritage policies: a comparative perspective. In *Handbook on the economics of cultural heritage*. Edward Elgar Publishing.
- Ministero dei beni e delle attività culturali (2015). Minicifre della cultura 2014.
- Ministère de la Culture (2021). Patrimostat 2021.
- Nelson, P. (1970). Information and consumer behavior. *Journal of political economy*, 78(2):311–329.
- Nelson, R. A., Donihue, M. R., Waldman, D. M., and Wheaton, C. (2001). What’s an oscar worth? *Economic inquiry*, 39(1):1–6.
- Peacock, A. (1995). A future for the past: The political economy of heritage. In *Proceedings of the British Academy, Volume 87: 1994 Lectures and Memoirs*, pages 189–243.

GENERAL INTRODUCTION

- Ponzo, M. and Scoppa, V. (2015). Experts' awards and economic success: evidence from an italian literary prize. *Journal of Cultural Economics*, 39(4):341–367.
- Reinstein, D. A. and Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The journal of industrial economics*, 53(1):27–51.
- Rogerson, W. P. (1983). Reputation and product quality. *The Bell Journal of Economics*, pages 508–516.
- Shapiro, C. (1982). Consumer information, product quality, and seller reputation. *The Bell Journal of Economics*, pages 20–35.
- Shapiro, C. (1983). Premiums for high quality products as returns to reputations. *The quarterly journal of economics*, 98(4):659–679.
- Steiner, L. and Frey, B. S. (2012). Correcting the imbalance of the world heritage list: Did the unesco strategy work? *Journal of International Organizations Studies*, 3(1):25–40.
- Towner, J. (1985). The grand tour: A key phase in the history of tourism. *Annals of tourism research*, 12(3):297–333.
- Towse, R. (2019). *A textbook of cultural economics*. Cambridge University Press.
- van der Aa, B. J. (2005). Preserving the heritage of humanity? obtaining world heritage status and the impacts of listing. Ph. D. Dissertation, Faculty of Spatial Science, University of Groningen, The Netherlands.
- World Tourism Organization (2018). Tourism and culture synergies.
- World Tourism Organization (2021). Yearbook of tourism statistics.
- World Tourism Organization (2022). World tourism barometer.

Chapter 1

More Is Worse: Decreasing Marginal Quality Of The UNESCO World Heritage List¹

Contents

1	Introduction	14
2	Literature review	17
3	The decision-making process behind the UNESCO WHL	20
4	Empirical strategy	26
5	Estimation of the baseline model	31
6	Robustness checks	34
7	Conclusions	43
	References	44

¹This chapter is an adapted version of Dattilo et al. (2022), a paper coauthored with Proff. Fabio Padovano and Yvon Rocaboy

” Quality means individuality, is therefore difficult. But unless we go deeper now, quantity at expense of quality will be our national tragedy—the rise of mediocrity into higher places.”

A Testament(1957), Frank Lloyd Wright

1 Introduction

Since 1972, UNESCO recognizes sites that constitute “... *parts of the cultural and natural heritage (that) are of outstanding interest and therefore need to be preserved as a part of the world heritage of mankind as a whole*” (UNESCO, 1972). As of 2020, 1121 sites, 869 of the cultural and 213 of the natural type, have been included in the WHL. The sheer size reached by this list calls into question the evolution of its quality over time. To appreciate the problem, in 1979 the Grand Canyon was among the first natural sites to receive the UNESCO recognition; 22 years later, in 2001, it was the turn of the East Coast of Devonshire. Even if one might agree that both sites deserve to be preserved, many would recognize that the unique and astonishing scenery of the first outshines the second (van der Aa, 2005). Likewise, the centre of Rome, included in 1980, can hardly be compared with the industrial archaeology of Ivrea, listed in 2018. It is therefore worth verifying whether the marginal “quality” of the sites decreases as the stock of the world cultural capital still available, i.e., the sites potentially eligible but not yet included in the WHL, shrinks.

Analyzing the evolution of the quality of the WHL is important for at least two reasons. The first is related to the reputation of the UNESCO WHL itself. Keeping a high average quality of the sites inscribed in the WHL is crucial for the credibility of the organization and for the successful preservation of the world’s cultural heritage. It is no accident that UNESCO foresees the possibility that a site is removed from the list if it fails to maintain certain standards. The second motive is scholarly. Despite the extensive political economy literature about the rent-seeking activities characterizing UNESCO’s decision-making process (Bertacchini and Saccone, 2012, Bertacchini et al., 2017, 2016, Parenti and De Simone, 2015), the issue of the evolution of the quality of the WH sites has never been examined so far. These studies have examined the distribution of the sites using quantitative indicators, mainly the number of sites per country. Indeed, without controlling for the quality of the

sites, associating the number of a country's sites to lobbying variables, such as the country's presence in the UNESCO committee, might overstate the importance of rent-seeking in the selection process. It is therefore interesting to verify if and how the conclusions of this literature change when our conceptualization of quality is taken into consideration. For instance, some scholars claim that the geographic distribution of sites allegedly suffers from a pro-European numerical bias (Frey et al., 2013, Steiner and Frey, 2012). This argument may lose relevance once this distribution is weighted by the quality of the sites. Finally, the evaluation of quality has always represented an important and still problematic analytical challenge in the economics of arts and culture, as well as in mainstream economics (Ginsburgh and Weyers, 1999, Ginsburgh, 2003, Throsby, 1990). This chapter aims to contribute to this analytical endeavour.

We define the "quality" of the sites included in the UNESCO WHL in a way that is both straightforward and that minimizes the impact of subjective evaluations of quality. We exploit the fact that, to enter the list, each site must satisfy at least one of ten non-alternative "criteria of outstanding universal value", upon which UNESCO base their evaluation. These criteria capture different dimensions of "quality", i.e., different reasons why a site might deserve to be included in the WHL. We hold that the greater the number of criteria that each site satisfied when it was accepted in the WHL, the greater its quality. Coming back to the previous example, the center of Rome satisfies 5 criteria out of 6 for cultural sites, while the industrial city of Ivrea only 1. Among the natural sites, the Grand Canyon satisfies 4 criteria out of 4, while the Coastline of Devonshire (UK) only one.

This specific (and by no means unique) definition of quality has several advantages. First, it is based on the original evaluation of the site made by UNESCO itself. The appraisal of quality cannot, therefore, be attributed to the preferences of the analyst (i.e., the authors of this chapter) or of any specific expert involved in the review of the site. The eventual inclusion in the WHL is the outcome of a quite complicated process; such complexity minimizes the importance of each individual's subjective assessment, and of the associated biases. Second, the criteria adopted by UNESCO have remained rather constant over time. Third, contrary to most alternative evaluation methods, based on the individuals' willingness to pay (e.g., the number of tourists attracted or contingent evaluations) the one we propose is less

exposed to endogeneity bias. We maintain that this method to assess quality marks an improvement with respect to the existing literature, which either ignores the issue or proxies it via a dichotomous variable, which merely says that sites included in the WHL are considered of higher interest than the excluded ones.

To anticipate the results, our estimates lend support to the hypothesis that, as the number of UNESCO sites of a country increases, their marginal quality decreases. Since every new site that enters the WHL reduces the stock of the country's cultural capital still available for future recognition, countries are eventually compelled to propose new sites of lower quality. This effect is more evident for countries with more than 10 sites in the WHL, representing 12% of the countries and 51% of the total sites. Furthermore, we find that the efficiency of the country's bureaucracy, rather than its lobbying power, plays an important role in the inclusion of low-quality sites in the WHL. High-quality sites, instead, do not need either an efficient state administration or political pressure to be enlisted. This result is at variance with the already cited public choice literature on the UNESCO WHL that, looking at the evaluation of the sites' quality in a dichotomous way (i.e. inclusion in the WHL or not), usually found that lobbying affects the selection of WH sites.

The rest of the chapter is organized as follows. Section 2 reviews the literature on the evaluation of quality in cultural economics and about the "political economy" of the UNESCO WHL. Section 3 illustrates the process through which UNESCO selects the sites to be included in the WHL and the criteria that each site must satisfy to be recognized. Section 4 discusses the empirical strategy, the variables included in the specification of the empirical model and the econometric issues associated with the estimates. In Section 5 the estimates' baseline results are presented, while section 6 illustrates the robustness checks. Finally, Section 7 summarizes the main conclusions of the analysis.

2 Literature review

2.1 Evaluation of quality

Although standard microeconomic models of consumers' choices are expressed in terms of quantities, quality is indeed an important, yet difficult to characterize, dimension of such choices. It crucially affects the demand of a commodity, especially for culture-related goods, where product differentiation is often so extreme as to make many works of art and cultural experiences essentially unique (Ginsburgh and Weyers, 1999, Waldfogel, 2012). In empirical analyses, however, quality is difficult to characterize because it is a subjective and not directly observable concept; as such it needs to be approximated. These difficulties are one of the reasons why the consideration of quality has been so far neglected in the empirical literature.

Two alternative approaches exist for the empirical assessment of the quality of cultural and artistic goods (Ginsburgh, 2003)². The first one decomposes the evaluation of quality into several dimensions, then establishes criteria to rate each dimension, and finally aggregates the scores. The criteria for evaluation, as Throsby (1990) stresses, should be “generally agreed,” and provide the foundations for the subsequent application of aesthetic judgments; yet the identification of “generally agreed,” i.e., non (excessively) subjective criteria, is quite hard. On the one hand, this approach has the important advantage for empirical analysis of expressing the characteristics of cultural and artistic goods along some metric; yet the researcher's value judgments in the identification of the characteristics that determine quality and in their cardinal evaluation make the resulting metric highly subjective and arbitrary.

The second approach envisages the evaluation of quality as a two-step procedure. The first step consists in resorting to experts' evaluations of quality; the second verifies the ability of these evaluations to endure the test of time, to minimize the role of fashion and of short-lived opinions in the evaluation of quality; furthermore,

²These approaches have ancient historical roots in the philosophy and aesthetics. One of the first expressions of the first approach can be found in De Piles (1708) *Cours de Peinture par Principes*. Hume's *Four Dissertations* (1757) provide a clear description of the second approach.

the test of time is a way to compare the original experts' opinions with consumers' (or the general public's) preferences. Being less subjective and more amenable to empirical analysis than the first, this approach has been more often used, especially in the domain of music (Ginsburgh and Noury, 2008, Ginsburgh and Van Ours, 2003), cinema (Nelson et al., 2001, Deuchert et al., 2005, Reinstein and Snyder, 2005) and even literature (Ponzo and Scoppa, 2015). These studies usually exploit awards and ratings assigned during competitions as proxies for experts' opinions and look at commercial success as a proxy for consumers' preferences. In cultural sectors where no competition is taking place, such as the wine industry and the performing arts, researchers conduct experts' opinion surveys (Ashenfelter and Jones, 2013, Tobias, 2004).

The demand for cultural heritage is often evaluated using stated preferences (Bedate et al., 2004, Alberini and Longo, 2006, Ruijgrok, 2006). This approach, however, presents huge limitations, as it drastically depends on the survey's structure and on the response rate. In addition, marginal changes in cultural goods are difficult to conceive and often evoke opposed responses, depending on the individuals' preferences (Noonan, 2003).

Our study actually adopts a mix of these two methodologies for the evaluation of quality. On the one hand, it includes experts' opinions, as in the second approach, since UNESCO resorts to committees and panels of experts to evaluate whether a site satisfies the eligibility criteria. On the other hand, these criteria are expressed in a binary scale, reflecting a multiplicity of characteristics that the sites must possess and are eventually aggregated; all these are quantitative features typical of the first methodology. Furthermore, compared to other settings examined in the literature, in the case of the UNESCO WHL the influence of fashions and/or the reactions to current events has little effect on experts' opinions, since cultural heritage is per se recognized after a long period of time. In addition, as we shall see, the final ruling by UNESCO is the outcome of a complex decision-making process fragmented between many different veto players, upon which each individual subjective evaluation has little bearing. All these features contribute to minimizing subjectivity in the evaluation of quality. Finally, when compared with other methods adopted in the literature, our idea of summing the number of criteria presents the advantage of being straightforward and transparent.

2.2 Studies about the politics of UNESCO

The procedure through which UNESCO selects the sites to be included in the WHL has been extensively studied in both cultural economics and public choice studies. Many studies of both strands concur that there is a problem of “inequality” in the composition of the WHL, i.e. the alleged over-representation of European sites in the WHL, especially in the case of cultural sites (Bertacchini and Saccone, 2012, Bertacchini et al., 2017, 2016, Frey et al., 2013, Steiner and Frey, 2012). Steiner and Frey (2012) in particular claim that this inequality has increased from 1978 to 2007, reflecting the UNESCO’s inability to raise the share of sites from non-European countries, notwithstanding the implementation of their “Global Strategy for a Representative, Balanced and Credible World Heritage List” since 1994. This strategy introduced a series of measures aimed at re-balancing the geographic representativeness of WHL, with quotas of sites imposed on European countries.

Other studies explain the unequal geographic distribution of sites in the WHL arguing that European countries either care more about the WHL and therefore propose many more sites than non-European ones; or that they enjoy more political power in the UNESCO selection committee. Such influence would subjugate an independent evaluation of the sites’ quality to the political logic of rent-seeking, thus generating a pro-European bias in selection (Bertacchini et al., 2009, Frey et al., 2013). Against this conclusion, van der Aa (2005) observes that, up to the year 2000, Europe had 46% of the sites included in the WHL, but also 45% of the sites rejected, which is hardly evidence of a bias. He also argues that any argument in favour of “greater equality” or of rebalancing the geographic distribution of the WH sites starts from the undemonstrated premise that cultural capital is actually homogeneously distributed around the world; the lack of a benchmark for a “balanced” distribution of sites makes the notion of a pro-European bias unwarranted.

To some extent, these debates exist because most papers in the literature fail to properly and explicitly consider the quality of the sites in their analyses. Virtually, all empirical studies in this literature use dummy variables that consider whether a site has been included in the WHL or not. So far, a positive correlation between a

country’s number of sites and its presence in the UNESCO selection committee is usually considered as evidence of rent-seeking (Bertacchini et al., 2016). But this conclusion may be spurious without controlling for the quality of the sites approved. As already said, sites of outstanding value do not need any political pressure to be included in the WHL (e.g., Paris), whereas others of lesser renown might do. This information cannot be conveyed by a dichotomous variable. Likewise, any evaluation of how world heritage sites are distributed across the world must consider the assessment of their quality made by the UNESCO itself, not just the end result of the decision-making process; it must also somehow control for the distribution of cultural and natural capital stock across the world, to provide some benchmark against which evaluating whether a bias in fact exists.

3 The decision-making process behind the UNESCO WHL

3.1 The UNESCO selection procedure

The UNESCO Convention of 1972 regulates the process through which UNESCO attributes the label “World Heritage” to a site. Two branches within UNESCO are in charge of the WHL: the General Assembly, which includes all member countries of the UNESCO³, and the World Heritage Committee, the executive body composed of 21 representatives that remain in charge for six years. Representatives’ tenures in the Committee are staggered and rotating; every two years some countries enter into the Committee in place of the existing ones⁴. The distribution of seats is based on geographic location, with the aim of “*ensuring an equitable representation of the different regions and cultures of the world*”⁵. Conversely, to enter into the

³Membership in the UNESCO does not necessarily coincide with membership in the UN; the United States, for instance, quitted the UNESCO once in 1984 and then in 2018, while always remaining a member of the UN.

⁴This number is actually variable, because countries may voluntarily decide to reduce the length of their mandate to maximize turnover

⁵Seats are allocated as follows: 2 for Western European and North America, 2 for Eastern Europe, 2 for Latin America, 3 for Asia and Pacific, 4 for Africa and 2 for the Arab States. We also note that the regions as presented here are defined by UNESCO for its activities, and do not necessarily reflect the actual geographical location of countries.”

General Assembly a country must sign the *Convention concerning the Protection of the World Cultural and Natural Heritage*. This treaty requires the member countries to provide a “compulsory contribution” to the World Heritage Fund, computed as a fixed yearly percentage of its total contributions to the UN, which cannot exceed 1%. A country may however decide to push its contributions beyond such a limit⁶ and make *voluntary contributions*.

Upon joining the UNESCO, a member country is encouraged to submit a tentative list of natural and cultural sites located within its borders. This list anticipates the sites that the country may propose for inclusion in the WHL in the next five to ten years. Two independent advisory bodies (actually, two NGOs), formally external to UNESCO, evaluate the proposed sites: the International Council on Monuments and Sites (ICOMOS), for the cultural sites; and the International Union for Conservation of Nature (IUCN), for the natural ones. These bodies may provide four alternative recommendations: “inscription”, “referral”, “deferral” or “not to inscribe”. A recommendation of “not to inscribe” implies that the country cannot present that site ever again. The “referral” and “deferral” evaluations encourage the country to provide minor changes (in the case of “referral”) or substantial revisions (in the case of “deferral”) and resubmit the candidature at a later session. Upon consideration of the recommendations of the advisory bodies, the Committee takes the final decision; a site is inscribed if it obtains a majority of 2/3 of the present members, who cast their vote through a secret ballot. It is especially at this stage that rent-seeking activities take place.

At the time of the promulgation of the Convention, no specific limits were imposed on the number of nominations, neither per country, nor per year⁷. In 1994, however, the UNESCO Committee approved the “Global Strategy for a Representative, Balanced and Credible World Heritage List” and since 2000 they introduced a series of measures aimed at re-balancing the geographic representativeness of WHL. These consisted in an overall limit of 30 nominations examined per year and one nomination proposed per country. In 2004, these limits were relaxed to two nominations per country, provided that at least one concerned a natural site, and to 45

⁶UNESCO (1972), art. 16 n. 2.

⁷For instance, in 1997 Italy scored a record of ten new sites included in the WHL.

nominations examined per year. The limits have remained stable from 2004 to the present day.

3.2 Criteria of Outstanding Universal Value

According to the Convention, in order to be included in the WHL, one of the parties involved in the decision-making process (either the country, the Advisory Board, or the Committee itself) must prove that the site is of Outstanding Universal Value from the point of view of history, art, science or nature. These aspects justify the preservation of sites as part of the UNESCO world heritage. The efforts made by UNESCO experts to render this definition as objective as possible are evident. The yearly Operational Guideline (UNESCO, 2017) and Jokilehto (2006) point out that *Outstanding* means that, in comparison with the generally documented cultural heritage, they belong to the very best or are “representative of the best”; while *Universal* means that these outstanding values can be acknowledged as such in general and worldwide and be of common importance for present and future generations. As these definitions are still too generic to drive the evaluation of new proposals, They are further spelt in ten criteria, six for cultural sites and four for natural ones, which express as many *values* that the UNESCO recognizes (Jokilehto, 2006). Table 1.1 illustrates these criteria. Among all, the idea of historical value emerges as the main one, as it is an integral part of the majority of the criteria, from 1 to 8. It emerges through expressions such as “over a span of time” (2), “human history” (4) or “stages of earth’s history” (8). The aesthetic /artistic value also plays an important role. It could be found in expressions such as “creative genius” (1), “monumental arts” (2) or “artistic and literary work”. The idea of representativeness is also recurrent, especially in criteria from (3) to (6) delineated as civilizations, buildings, human settlement and environment relations, and traditions. Similarly, the scientific, ethnological and anthropological values appear in different criteria.

Two points clearly emerge from this table. First, it is reasonable to maintain that not all sites have the same quality, as they do not satisfy the same number of criteria. Second, all criteria are binary, i.e., each of them can be either fulfilled or not, with no possibility of a “partial satisfaction”. This greatly simplifies our quantitative evaluation of the quality of the sites. Yet, to be able to compare the quality of sites

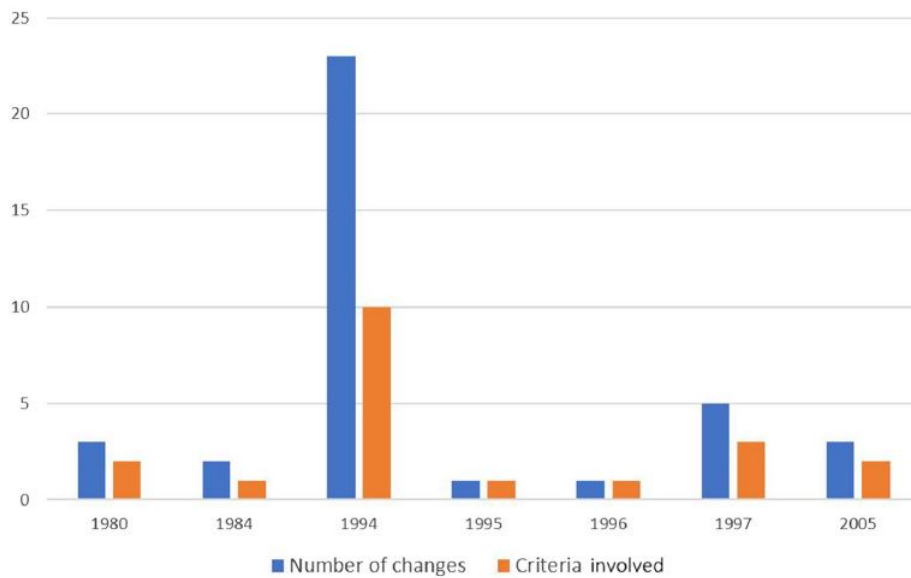
Table 1.1: Criteria for Cultural and Natural sites in 2018

N.	Cultural Criteria	Value involved
1	Represents a masterpiece of human creative genius	Aesthetic
2	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	Aesthetic, Historical, Technical
3	Bears a unique or at least exceptional testimony to a cultural tradition or to a civilization which is living, or which has disappeared	Historical, Representative
4	Is an outstanding example of a type of building, architectural or technological ensemble or landscape which illustrates (a) significant stage(s) in human history;	Historical, Representative, Technical
5	Is an outstanding example of a traditional human settlement, land-use, or sea-use which is representative of a culture (or cultures), or human interaction with the environment especially when it has become vulnerable under the impact of irreversible change	Historical, Scientific
6	Is directly or tangibly associated with events or living traditions, with ideas, or with beliefs, with artistic and literary works of outstanding universal significance. (The Committee considers that this criterion should preferably be used in conjunction with other criteria);	Representative
N.	Natural Criteria	Value involved
7	Contains superlative natural phenomena or areas of exceptional natural beauty and aesthetic importance	Aesthetical
8	Offers outstanding examples representing major stages of Earth's history, including the record of life, significant on-going geological processes in the development of landforms, or significant geomorphic or physiographic features	Historical, Scientific
9	Offers outstanding examples representing significant on-going ecological and biological processes in the evolution and development of terrestrial, fresh water, coastal and marine ecosystems and communities of plants and animals	Representative, Scientific
10	Contains the most important and significant natural habitats for in-situ conservation of biological diversity, including those containing threatened species of Outstanding Universal Value from the point of view of science or conservation.	Scientific

over time, as it is the aim of this study, the criteria adopted by UNESCO must have also remained stable through the sample period. The definitions of the UNESCO criteria have in fact somewhat evolved over time in different stages, as Figure 1.1 illustrates. The issue is to assess to what extent these changes are purely semantic or have in fact produced consequences. On this point the literature seems to lean towards the semantic view. Labadi (2013), for example, judges that the evolution of the criteria was “non linear, but rather complex and circular, having been at various points the results of contradictory recommendations and decisions” and can therefore be altogether neglected. Steiner and Frey (2012) have not found changes in the distribution of sites following changes in criteria, including the apparently major one of the “Global Strategy” of 1994. Be that as it may, we prefer not to have any *a priori* in our analysis and investigate the issue empirically in Section 6.

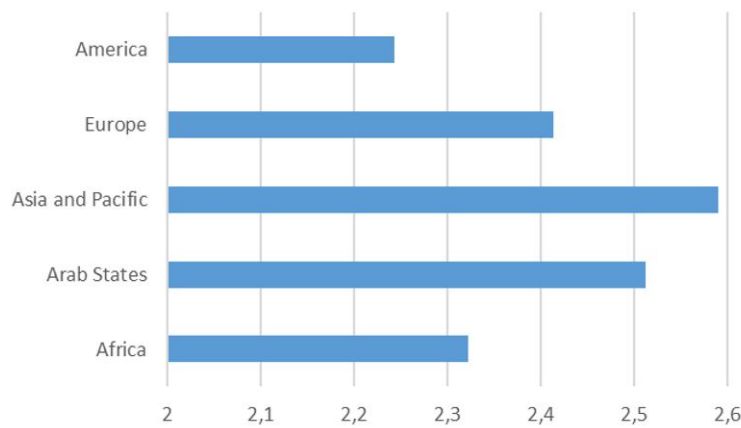
Figure 1.2 illustrates the distribution of the mean values of the number of criteria that each site satisfies across the UNESCO geographical areas. Although Europe holds the highest number of sites, it is not the area with the highest average quality: Asia and Arabia reach higher average scores. This is indeed *prima facie* evidence that the marginal quality of the WHL is decreasing. We can illustrate this negative relationship by means of a scatter plot between the number of sites of each country and the correlation coefficient between the quantity and quality of its sites. Figure 1.3 shows, on the vertical axis, the value of the correlation coefficient between the number of sites already inscribed and the quality of the marginal site; the horizontal axis instead reports the number of sites. Beyond 14 sites (if we consider Brazil as an outlier), the correlation coefficient becomes negative, i.e. an additional site lowers average quality. The diagram confirms that it is worth analyzing this negative relationship by means of regression analysis in the context of a more complex model, to obtain a more precise assessment of the evolution of the marginal quality of the UNESCO WHL over time.

Figure 1.1: Evolution of criteria over time



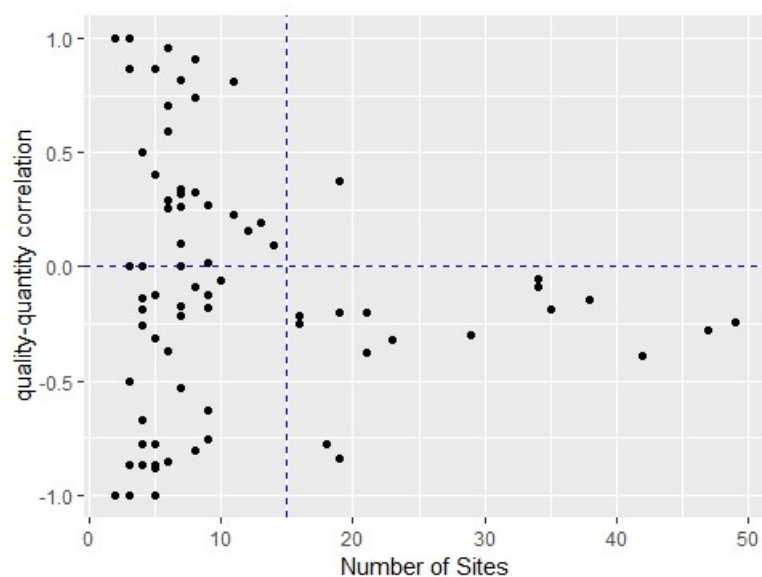
Note: the graph presents in blue the total number of semantic changes by year and in orange the number of criteria involved.

Figure 1.2: Mean quality of sites by UNESCO geographical area



Note: the graph plots in blue the average quality of sites proxied by the number of criteria satisfied at the inscription.

Figure 1.3: Correlation between quantity and quality of UNESCO sites



Note: each dot represents one country's correlation coefficient between the quality and quantity of its WH sites. The blue lines divide the plot into four sections: (from top-left) low quantity and increasing quality, high quantity and increasing quality, high quantity and decreasing quality, low quantity and decreasing quality.

4 Empirical strategy

4.1 Dependent variable and estimation issues

The hypothesis under test is that, as the number of sites that a country has in the WHL increases, the quality of the marginal site decreases. This amounts to estimating the derivative $(\partial Q_{it})/(\partial N_{it})$, where N_{it} is the total number of sites that country i has in WHL at year t , and Q_{it} , the endogenous variable, is the corresponding average quality of the sites. A decreasing marginal quality implies a negative sign of the derivative. The sample includes 180 countries between 1978 and 2016. To calculate Q_{it} , we exploit the binary nature of criteria, assigning a value of 1 if criterion y is satisfied and 0 otherwise; we have first summed the y , thus obtaining a measure of quality for each site; then, since country i may have more than one site approved per year, we have divided the sum of the scores by the

number of sites enlisted in each year, thus obtaining an average quality of the sites enlisted by the country in that year.

Although this specification has the advantage of simplicity, it also creates a number of econometric issues. First, modelling Q_{it} as yearly averages prevents us from considering it as a count variable, which excludes the possibility of estimating negative binomial and/or zero-inflated models. Moreover, our raw data have a panel structure where almost 90% of the observations are zeros, because quite often no new sites are recorded for a country/year combination. The frequency of zero values generates problems of estimation and interpretation. First, it makes the probability of observing a strictly positive quality highly related to the probability of having a site enlisted, as in years when one or more sites are included, both Q_{it} and N_{it} increase. To solve this problem, we proxy N_{it} by the lagged number of the total number of sites within the WHL that country i has at year t (variable $Sites_{it-1}$). Moreover, to avoid the concern that the results be driven by a single specification of the main independent variable, we have proxied N_{it} also by the number of years that a country i has been a member of the UNESCO at time t (variable $Tenure$). The idea is that a longer membership should result in a greater number of sites. Second, in any year t the zeroes may reflect either the fact that the country did not propose any site, or that they were rejected. Defining a proper instrument that is able to distinguish these two events and is also independent from the sites' quality is difficult. We have therefore estimated the model including only the strictly positive observations, which yields an unbalanced panel.

4.2 Baseline model and explanatory variables

Although Figure 1.3 shows a negative correlation between the number of sites and their marginal quality, other conditioning factors may affect the dependent variable and must be controlled for. The first is the country's cultural capital, which is not homogeneously distributed across geographical areas, because of the different histories of civilization of each country. Processes of internal competition between touristic sites may lead countries with larger and more heterogenous stocks of cultural capital to propose sites of lower quality than countries with smaller and more homogeneous stocks. A second conditioning factor is the country's lobbying power

at UNESCO, which, as we have argued before, is likely to be exerted more for sites of relatively low quality. Thirdly, also the efficiency of the country’s public administration may affect the number of sites included in the WHL, because the preparation of the proposal and the explanation of how the site satisfies the UNESCO criteria are all bureaucratic tasks. Our baseline model is therefore specified as follows:

$$Q_{it} = \beta_0 + \beta_1 N_{it} + \beta_2 CulturalK_{it} + \beta_3 Lobby_{it} + \beta_4 Bureaucracy_{it} + u_{it} \quad (1.1)$$

For each of the control variables of equation 1.1 appropriate proxies must be found. As a measure of cultural capital (*CulturalK*), the literature (e.g., Stainer and Frey, 2013) generally use the area and population of the country. The idea is that the country’s population reflects its potential to produce cultural goods, while in larger countries it should be more likely to find sites worth including in the list; furthermore, both measures are easily available for a large number of countries. Yet, as cultural heritage is a good originated in the past, historical proxies are more appropriate. We therefore also consider the historical population in the year 1500 and GDP per capita in the year 1820, from Maddison Historical Statistics (2020 release). The motivation is that the larger was the country’s population in the past, the greater should be its historical human capital and therefore the cultural capital still available today. Likewise, the higher was GDP per capita in the past, the more resources a country could devote to the production of cultural capital. The drawback of historical proxies is that in the Maddison Historical Statistics these variables are not available for all countries. We must therefore distinguish between two types of countries: “high cultural capital countries” (HCK), for which historical data are available and “low cultural capital countries” (LCK), for which they are not. The idea behind this classification is that only more developed civilizations with a high level of human capital were able to generate information about their historical GDP and population. To the remaining countries we assign a value of 0. To avoid any possibility of misrepresentation of reality, we test three specifications: one with *POP_1500* and *GDP_1820*, which encompasses the entire sample and treats the lack of information as a sign of low level of cultural capital, attributing a 0 value to them; another with *POP_1500hk* and *GDP_1820hk*, which includes only HCK countries and considers the missing information as data not available; and a last one

on the subsample of LCK countries, that includes the current *Population*, since it is the only control variable always available for those countries.

To capture the effects of lobbying, we consider two types of variables: the country's membership in the selection Committee, and the money flows from each country to UNESCO. In particular, *Committee* is a cumulative variable equal to the total number of mandates the country had fulfilled till time t when it is a member of the selection Committee and 0 otherwise. This specification allows not only to capture the effects of the inclusion of the country in the Committee, but also those of its permanence in the selection process in terms of experience accumulated and connections established. The other two proxies for lobbying are based on monetary flows: the first, *Expect_contr* is the sum of compulsory and voluntary contributions, i.e., the country's total contributions to UNESCO. As these contributions should have an effect only after they are budgeted, the variable is lagged one period. Second, to capture the entire contributive history of a country, we have computed the variable *Unpaid_contr*. When a member country has paid all compulsory contributions, this variable is 0; otherwise, it is equal to the difference between the contributions due and those actually paid. Just like *Expect_contr*, *Unpaid_contr* is lagged one period.

Finally, we adopt the Government Effectiveness Index (variable *Gov_eff* from the Worldwide Governance Indicators of the World Bank) to proxy the efficiency of the country's bureaucracy. *Gov_eff* captures the (perceptions of the) efficiency of the country's public and civil services. These scores are aggregated into a single index, in units of a standard normal distribution, ranging from 0 to 1. As such, *Gov_eff* is the best proxy for the country's ability to prepare a proposal for inclusion of a site into the WHL. Table 1.2 reports the correlation matrix between selected variables, Table 1.3 the descriptive statistics and Table 1.4 summarizes the expected signs of the explanatory variables.

All panel models are estimated using random effects. Intuitively, we cannot estimate a fixed effect model because of the presence of dummy variables or variables constant over time in the model. In any event, the Lagrange multiplier test for the choice of the econometric model presented in Table 10 supports the application of a random effect model.

Table 1.2: Correlation Matrix

	<i>Tenure</i>	<i>Committee</i>	<i>Quality</i>	<i>Pop</i>	<i>POP_1500</i>	<i>Area</i>	<i>Exp_contr</i>	<i>Unpaid_contr</i>	<i>Gov_eff</i>	<i>Sites_{t-1}</i>	<i>GDP_1820</i>
<i>Tenure</i>	1	0.17	0.037	0.13	0.085	0.18	0.11	0.047	0	0.32	0.18
<i>Committee</i>		1	0.19	0.23	0.22	0.23	0.2	0.034	0.17	0.34	0.25
<i>Quality</i>			1	0.29	0.3	0.2	0.23	0.062	0.12	0.39	0.26
<i>Pop</i>				1	0.96	0.44	0.22	0.11	0	0.52	0.18
<i>POP_1500</i>					1	0.31	0.17	0.02	0.05	0.52	0.19
<i>Area</i>						1	0.32	0.32	0.04	0.42	0.14
<i>Exp_contr</i>							1	0.58	0.33	0.5	0.47
<i>Unpaid_contr</i>								1	0.09	0.15	0.17
<i>Gov_eff</i>									1	0.31	0.55
<i>Sites_{t-1}</i>										1	0.59
<i>GDP_1820</i>											1

Table 1.3: Descriptive Statistics

	<i>Mean</i>	<i>Median</i>	<i>St.dev</i>	<i>Min</i>	<i>Max</i>	<i>N of obs</i>
<i>Quality</i>	2.37	2.00	0.77	1.00	6.00	701
<i>Tenure</i>	11.08	8.00	11.16	0	42.00	7020
<i>Sites_{t-1}</i>	2.942	1.000	5.68	0	49.000	7020
<i>Area</i>	7.24	1.18	19.85	0	170.98	6992
<i>Population</i>	321.56	61.97	1212.6	0.12	13786.65	6988
<i>POP_1500hk</i>	7103	1250	20642.80	100	110000	1951
<i>GDP_1820hk</i>	752.41	642.02	341.72	83.33	1837.98	1872
<i>POP_1500</i>	1974	0	11336.33	0	110000	7020
<i>GDP_1820</i>	200.6	0	376.63	0	1838.0	7020
<i>Committe</i>	0.2248	0	0.67	0	5.0000	7020
<i>Exp_contr</i>	14211	294	63551.60	0	927085	6374
<i>Unpaid_contr</i>	5980	0	57280.13	-104741	1420606	4291
<i>Gov_eff</i>	0.49	0.45	0.21	0	1	2982

Table 1.4: Expected signs

<i>Variable</i>	<i>Expected sign</i>
<i>Sites_{t-1}</i>	Negative
<i>Tenure</i>	Negative
<i>Area</i>	Positive
<i>Population</i>	Positive
<i>POP_1500hk</i>	Positive
<i>GDP_1820hk</i>	Positive
<i>POP_1500</i>	Positive
<i>GDP_1820</i>	Positive
<i>Committee</i>	Negative
<i>Exp_contr</i>	Negative
<i>Unpaid_contr</i>	Positive
<i>Gov_eff</i>	Negative
<i>Change94</i>	Not significant
<i>Change0103</i>	Positive
<i>Change05</i>	Negative

5 Estimation of the baseline model

Table 1.5 illustrates the results of our econometric analysis, where the number of sites N_{it} is proxied by $Sites_{it-1}$. Only the observations where the dependent variable has nonzero values are reported; this reduces the sample to 580 observations⁸. Evidently, the most important result is that the coefficient on $Sites_{it-1}$ is negative and statistically significant in all models; in other words, the estimated sign of the derivative $(\partial Q_{it})/(\partial N_{it})$ is negative for the entire sample (model 1) and the selected subsamples (models from 2 to 4). This lends support to our fundamental hypothesis that countries that have more sites are experiencing a diminishing marginal quality of the newly accepted sites.

As for the controls, the proxies for lobbying (*Committee*, *Exp_contr* and *Unpaid_contr*), although with the expected signs, are generally not statistically signifi-

⁸Countries generally inscribed more than one site each year, reducing substantially the number of observations. In addition, transnational sites, i.e. sites shared between more than one country, are excluded from the sample.

cant: only in model 4 *Committee* is marginally significant. This suggests that when the quality of the sites is considered instead of the simple inclusion in the WHL, lobbying loses its explanatory potential, probably because the inclusion of sites that are universally recognized as world heritage does not require much political pressure. Lobbying may instead be relevant only for the marginal sites, i.e., those whose admittance to the WHL thanks only to their quality is uncertain. Also, the efficiency of the country's public administration shows the expected negative sign (model 3), confirming that more efficient bureaucracies are better able to have relatively low quality sites approved into the WHL. When the sample is restricted to countries with a small stock of cultural capital (for which historical population and GDP are not available), the coefficients on the efficiency of the bureaucracy lose significance (model 4), probably because these countries have very few sites to propose, so that there are not enough observations to detect a negative correlation on *Gov_eff*.

The proxies for the stock of cultural capital based on historical data reveal that, when the entire sample of countries for which such data is available is considered, population in the year 1500 seems to exert a positive impact on the quality of sites (model 1 and 2). When instead the sample is restricted only to high cultural capital countries, GDP per capita plays a more relevant, and still positive, role (model 3). Finally, when the lack of historical data forces us to use current values of country area and population, as it is usually done in the literature, these variables never turn out statistically significant (model 4).

Table 1.5: Regression results. Number of sites proxied by $Sites_{t-1}$

	<i>Dependent variable:</i>			
	<i>Quality</i>			
	Model 1 Complete sample	Model 2 HCK countries	Model 3 HCK countries	Model 4 LCK countries
$Sites_{t-1}$	-0.01793*** (0.005151)	-0.01926*** (0.005883)	-0.007341** (0.00388)	-0.08722** (0.0431)
<i>Committee</i>	0.00756 (0.03897)	-0.006512 (0.04829)	0.02592 (0.05428)	0.1786* (0.1035)
<i>Expect_contr</i>	0.0000007 (0.0000006)	0.0000008 (0.0000007)		0.000003 (0.000008)
<i>Unpaid_contr</i>	-0.0000004 (0.0000004)	-0.0000006 (0.0000008)	-0.0000002 (0.0000007)	-0.000001 (0.0000007)
<i>Gov_eff</i>			-2.1535*** (0.4974)	-0.2045 (0.5386)
<i>Pop_1500</i>	0.0000088*** (0.000003)			
<i>GDP_1820</i>	0.0001895 (0.000117)			
<i>Pop_1500hk</i>		0.000008** (0.0000034)	0.000004 (0.0000028)	
<i>GDP_1820hk</i>		-0.000009 (0.00026)	0.0006181** (0.000295)	
<i>Area</i>	-0.001938 (0.001724)	-0.001512 (0.00278)	0.0002128 (0.002577)	0.004783 (0.005163)
<i>Population</i>				0.001493 (0.005163)
<i>Intercept</i>	2.3080*** (0.06717)	2.5784*** (0.2512)	3.3499*** (0.2978)	2.3601*** (0.2619)
Adj. R ²	0.072239	0.034685	0.11125	0.14991
F-statistic	6.88437***	1.94965*	3.98641***	4.21297**
Observations	580	308	168	129

Standard errors in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01. 33

6 Robustness checks

The econometric issues discussed above require estimating a series of variants of the baseline model, to minimize the risk of spurious correlations or misspecifications of the model and/or inappropriate estimation techniques.

6.1 Alternative specifications for N_{it}

First, we verify whether the estimated results remain fundamentally the same when the explanatory variable of interest, N_{it} , is proxied by another variable. To this end, we use the number of years that country i has been a member of the UNESCO at time t (variable *Tenure*). The idea behind *Tenure* is that a longer membership should result in a greater number of sites. Proxying N_{it} by *Tenure* has the further advantage of avoiding risks of multicollinearity with the other covariates, all of which have a positive effect on the number of sites; *Tenure* instead is positively correlated with the number of sites but not with the other variables, as the correlation matrix of Table 1.2 shows.

Table 1.6 reports the results; they are quite similar to those already obtained in the baseline model. Once more, the estimated coefficients on *Tenure* are always negative and statistically significant in models from 5 to 7. As time goes by, countries that have ratified the UNESCO Convention earlier (and that are therefore likely to have more sites) include sites of lower quality in the list. The effect is stronger when only high cultural capital countries are considered. The estimated coefficients on the variables measuring lobbying, the quality of the public administration and the stock of cultural capital confirm the results already obtained with N_{it} , proxied by $Sites_{it-1}$.

Table 1.6: Robustness checks. Number of sites proxied by *Tenure*

	<i>Dependent variable:</i>			
	<i>Quality</i>			
	Model 5 Complete sample	Model 6 HCK countries	Model 7 HCK countries	Model 8 LCK countries
<i>Tenure</i>	-0.008144** (0.003632)	-0.01892*** (0.005025)	-0.02446*** (0.007711)	0.00358 (0.01)
<i>Committee</i>	0.003171 (0.03922)	-0.00267 (0.04811)	0.03854 (0.05259)	0.09598 (0.1007)
<i>Expect_contr</i>	0.0000003 (0.0000005)	0.0000005 (0.0000007)		-0.0000026 (0.000007)
<i>Unpaid_contr</i>	-0.0000003 (0.0000004)	-0.0000004 (0.0000008)	-0.0000003 (0.0000007)	-0.00000035 (0.0000006)
<i>Gov_eff</i>			-2.5247*** (0.4992)	-0.3293 (0.5373)
<i>Pop_1500</i>	0.000006** (0.000003)			
<i>GDP_1820</i>	0.0001167 (0.000114)			
<i>Pop_1500hk</i>		0.0000054* (0.0000032)	0.0000012 (0.0000025)	
<i>GDP_1820hk</i>		-0.0001259 (0.0002591)	0.0006283** (0.0002714)	
<i>Area</i>	-0.00668 (0.00171)	-0.001261 (0.00281)	0.002122 (0.002589)	-0.000315 (0.004487)
<i>Population</i>				-0.00008158 (0.0004)
<i>Intercept</i>	2.3848*** (0.081)	2.8178*** (0.2663)	4.0507*** (0.3726)	2.3012*** (0.30835)
Adj. R ²	0.055078	0.041979	0.15788	0.11389
F-statistic	5.4931***	2.5171**	5.4728***	3.35011***
Observations	580	308	168	129

Standard errors in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01.

6.2 Disaggregating countries by the number of sites

Second, we check whether the sign of the derivative $(\partial Q_{it})/(\partial N_{it})$ remains the same for countries with a large and a small number of sites in the WHL. The idea is to verify whether the process of diminishing marginal quality is stronger for countries with a large number of sites, controlling for the stock of cultural capital.

To this end, we disaggregate the sample by the number of sites that a country has in the WHL. We set the threshold number of sites at 10 in 2016, to obtain a subsample that represents the top 10% of the distribution of sites by country and almost 50% of the sites included in the WHL. Table 1.7 presents the estimates with N_{it} proxied by $Sites_{it-1}$; they reveal that the negative and statistically significant coefficients found in Tables 1.5 and 1.6 are mainly driven by the countries with more than 10 sites. Models 9 and 11 show that the correlation is always negative and statistically significant when countries have more than 10 sites, while models 10 and 12 instead reveal that this effect disappears for countries below that threshold.

Likewise, greater government efficiency has a negative effect on quality only for countries with more than 10 sites (model 11), confirming that more efficient bureaucracies can have more sites of lower quality approved. The remaining results do not significantly change; Population seems to have a positive scale effect on quality (models 9 and 10), and so do the historical proxies for the stock of cultural capital (models 11 and 12). The lobbying variables are once again not significant.

Table 1.7: Robustness checks. Sample divided by the number of sites per country

	<i>Dependent variable:</i>			
	<i>Quality</i>			
	Model 9 Country sites>10	Model 10 Country sites<10	Model 11 Country sites>10	Model 12 Country sites<10
<i>Sites_{t-1}</i>	-0.01943*** (0.00576)	-0.01784 (0.0264)	-0.01591* (0.0083)	-0.04298 (0.0362)
<i>Committee</i>	0.00483 (0.047)	-0.02551 (0.08044)	0.02026 (0.0515)	0.08092 (0.0974)
<i>Expect_contr</i>	0.00000055 (0.0000006)	-0.0000029 (0.0000052)		
<i>Unpaid_contr</i>	-0.0000005 (0.0000004)	0.0000003 (0.000006)	-0.0000006 (0.0000005)	-0.0000027 (0.0000068)
<i>Gov_eff</i>			-1.31*** (0.3999)	-0.5024 (0.4058)
<i>POP_1500</i>			0.0000059** (0.0000023)	0.0001467** (0.0000562)
<i>GDP_1820</i>			0.000835*** (0.000311)	0.0000086 (0.00022)
<i>Area</i>	-0.003 (0.00205)	0.001197 (0.00937)	-0.001342 (0.00184)	0.0115 (0.0114)
<i>Population</i>	0.00007** (0.0000279)	0.0004** (0.000188)		
<i>Intercept</i>	2.6696*** (0.144)	2.2288*** (0.0842)	2.7134*** (0.285)	2.4615*** (0.2098)
Adj. R ²	0.029433	0.046893	0.11572	0.063214
F-statistic	1.9404*	3.3601***	3.65462***	2.54477**
Observations	273	291	143	162

Standard errors in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01.

6.3 Stability of criteria

The actual invariance over time of the criteria to include a site in the WHL needs to be verified, to ensure intertemporal consistency in the evaluations of the quality of the sites. On the basis of the evidence illustrated in Figure 1.1 and the procedures explained in Section 3, we select three breakpoints: the year 1994, when there was a peak in the change of the wording of the definitions of criteria; the years 2001-03, when a new admittance procedure restricted the number of new sites to one per country and 30 in total every year; the year 2005, when it was introduced the possibility of mixed sites (partly natural and cultural) and the number of sites was re-expanded to two per country and 45 in total ⁹.

Table 1.8 presents the results of the control of the stability of criteria over time. We organize the analysis in two steps; in models 13-15 we test whether any of the three beak points has a direct effect on the quality of the sites; in models 16-18 we verify whether including the proxies for the number of sites modifies this result. In model 13 the dummy for the changes introduced in 1994 turns out negative and significant, which suggests that these changes apparently reduced the quality of the sites subsequently included in the list. Yet, once we control for the number of sites (model 16), the change of criteria of 1994 does not seem to be relevant, since the negative quantity-quality relationship subject of our study holds. In other words, we find evidence that the changes in the definitions of the criteria approved in 1994 did not refrain countries with more sites to have new ones of lower marginal quality from being approved into the WHL. Interestingly, we observe an increase in the quality of sites between 2001 and 2003, as a consequence of the restrictions imposed on the number of nominations per country (models 14 and 17). Probably this restriction created an incentive to submit sites of higher quality, to minimize the possibility of receiving a rejection among the proposed sites. In a complementary way, following the relaxation in the UNESCO policy for site nominations in 2005, we observe a widespread reduction of the average quality after that year, regardless of the number of sites that a country had (models 15 and 18). The negative quantity-quality relationship is thus corroborated because imposing a limitation on the number of

⁹Alternatively, we perform an “unrestricted” test of the stability of the criteria by introducing a set of dummy variables that capture a series of five years intervals. The estimates do not change in a qualitative way.

sites that could be nominated seems to increase the quality of sites included in the list and vice versa. In other words, our hypothesis is valid in both directions.

Table 1.8: Robustness checks. Stability in the definition of criteria, breakpoints

	<i>Dependent variable:</i>					
	Model 13	Model 14	Model 15	<i>Quality</i>		Model 18
				Model 16	Model 17	
<i>Sites_{t-1}</i>				-0.01334** (0.00564)	-0.01821*** (0.00505)	-0.01197** (0.00559)
<i>Committee</i>	0.001039 (0.0388)	-0.0021 (0.0383)	0.006724 (0.0388)	0.01189 (0.0389)	0.01482 (0.0383)	0.0151 (0.039)
<i>Exp_contr</i>	0.0000005 (0.0000005)	0.0000003 (0.0000005)	0.0000004 (0.0000005)	0.0000009* (0.0000005)	0.000001* (0.0000005)	0.0000009 (0.0000005)
<i>Unpaid_contr</i>	-0.0000003 (0.0000004)	-0.0000003 (0.0000004)	-0.0000004 (0.0000004)	-0.0000004 (0.0000004)	-0.0000004 (0.0000004)	-0.0000005 (0.0000004)
<i>Area</i>	-0.001708 (0.00178)	-0.001373 (0.00168)	-0.001595 (0.00181)	-0.001423 (0.00179)	-0.001233 (0.00174)	-0.001432 (0.00181)
<i>Population</i>	0.0000569** (0.0000237)	0.0000515* (0.0000223)	0.0000583** (0.0000244)	0.0000789*** (0.0000256)	0.000084*** (0.0000251)	0.0000785*** (0.00002615)
<i>Change94</i>				-0.11 (0.092)		
<i>Change0103</i>		0.5418*** (0.129)			0.5824*** (0.0128)	
<i>Change05</i>			-0.2591*** (0.0777)			-0.1875** (0.0844)
<i>Intercept</i>	2.4542*** (0.08)	2.2769*** (0.056)	2.4023*** (0.0626)	2.4296*** (0.081)	2.3223*** (0.0582)	2.4113*** (0.0626)
Adj. R ²	0.052281	0.059478	0.068338	0.066103	0.096811	0.078238
F-statistic	5.7238***	6.47493***	7.5729***	5.94148***	8.88314***	7.24583***
Observations	580	580	580	580	580	580

Standard errors in parentheses. Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01.

6.4 Specification of the model

As a final robustness check, we have estimated a cross-section model to exclude the possibility that our results depend on the model specification. Given the different structure of the dataset, we are obliged to estimate a second equation:

$$Q_i = \beta_0 + \beta_1 N_i + \beta_2 CulturalK_i + \beta_3 Lobby_i + \beta_4 AVquality_i + u_i \quad (1.2)$$

The dependent variable Q_i is the quality of a single site i whose value, just like in the analysis conducted so far, equals the number of criteria that the site satisfies. Having removed from the sample the sites excluded from the WHL, whose values would have been zero, Q_i is a positive integer with a lower bound equal to 1.

Like in the estimates of equation 1.1, we proxy the number of sites N_i , by $Sites_{it-1}$ and $Tenure_i$; yet the cross section specification allows us also to include $Year$, i.e., the year in which a site is included in the WHL. If the sign of $(\partial Q_{it})/(\partial N_{it})$ is negative, sites enlisted in more recent years should be of lower quality; the expected sign on $Year$ is negative. As measures of cultural capital, we select $Area$ and $Population$ to keep the number of observations as large as possible; we include the same lobbying variables of the baseline model. Moreover, the focus on single sites of the cross-section specification eliminates the possibility to explain the evolution of the quality of the sites on a country basis. To limit such a drawback, we add the country's average quality, in order to evaluate the marginal evolution of the quality of the sites.

The cross-sectional specification of equation 1.2 has several additional advantages: first, it allows keeping more than one observation for every year, without the need to compute country averages of the quality of the sites. Second, it rules out the problem of missing values. Third, as the sample contains only positive integers, we can test our hypothesis using a count data model. Since Q is not over-dispersed and its mean and variance show quite similar values, we assume a Poisson distribution. The descriptive statistics related to the cross-section dataset are shown in Table 1.10.

Table 1.11 shows that changing the specification of the model from a panel

to a cross-section, where single sites rather than country averages are considered, does not qualitatively change the results. In all the estimates $Sites_{t-1}$ is negatively correlated with the quality of the sites and so is $Tenure$. In the cross-sectional model also the proxy $Year$ has a negative and statistically significant coefficient. The results on the other explanatory variables remain basically unchanged.

Table 1.9: Tests for random effect

<i>Lagrange Multiplier Test - Honda</i>	
Model 1 - With zeros normal = 14.242, p-value < 2.2e-16 alternative hypothesis: significant effects	Model 2 - Without zeros normal = 5.1121, p-value = 1.593e-07 alternative hypothesis: significant effects
<i>Lagrange Multiplier Test - Breusch-Pagan</i>	
Model 1 - With zeros chisq = 202.84, df = 1, p-value < 2.2e-16 alternative hypothesis: significant effects	Model 2 - Without zeros chisq = 26.134, df = 1, p-value = 3.185e-07 alternative hypothesis: significant effects

Table 1.10: Descriptive statistics cross-sectional dataset

	<i>Mean</i>	<i>Median</i>	<i>St.dev</i>	<i>Min</i>	<i>Max</i>
<i>Quality</i>	2.424	2.000	1.01197	1.000	7.000
<i>Site_{t-1}</i>	8.31	4.00	10.3850	0	47.00
<i>Year</i>	1996	1997	10.6909	1978	2016
<i>Committee</i>	0.7818	0	1.0885	0	5.0000
<i>Exp_contr</i>	61048	14216	111459.7927	0	804756
<i>Unpaid_contr</i>	14217	0	97222.9591	66769	1420606
<i>Area</i>	21.0748	4.4740	37.1761	0.0006	170.9825
<i>Population</i>	1377.440	384.695	3097.984	0.208	13786.650
<i>AVquality</i>	2.400	2.467	0.4832	1.000	5.000
<i>Tenure</i>	13.14	10.00	10.59	0	41.00

Table 1.11: Robustness checks. Cross-sectional estimations

	<i>Dependent variable:</i>		
	Model 19 Poisson	<i>Quality</i> Model 20 Poisson	Model 21 Poisson
<i>Sites_{t-1}</i>	-0.0063** (0.0026)		
<i>Tenure</i>		-0.0051** (0.0022)	
<i>Year</i>			-0.0051** (0.0021)
<i>Committee</i>	0.004 (0.0228)	0.0008 (0.0226)	-0.0048 (0.0225)
<i>Expect_contr</i>	0.0000002 (0.0000002)	0.0000001 (0.0000002)	0.00000004 (0.0000002)
<i>Unpaid_contr</i>	-0.0000001 (0.0000003)	-0.0000001 (0.0000002)	-0.0000001 (0.0000002)
<i>Area</i>	0.0004 (0.0008)	0.0006 (0.0008)	0.0004 (0.0008)
<i>Population</i>	0.000006 (0.000009)	0.000002 (0.000008)	-0.0000006 (0.0000008)
<i>AVquality</i>	0.4255*** (0.0469)	0.4275*** (0.0476)	0.4284*** (0.0478)
<i>Intercept</i>	-0.1281 (0.1184)	-0.0963 (0.1198)	10.0818** (4.1051)
AIC	2505.9	2506.1	2505.6
Observations	812	812	812

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7 Conclusions

Our study uses a simple and straightforward definition of quality of the sites of the UNESCO WHL. The estimates based on this proxy lend empirical support to our main research question: as the number of sites that a country has in the WHL increases, their marginal quality decreases. In other words, since the stock of cultural and natural capital is fixed, new entries into the WHL appear to be of lower quality than earlier ones. This negative quantity-quality relationship is particularly evident for countries with more than 10 sites. Quite importantly, this result seems robust after controlling for the stock of cultural capital, the lobbying power of the UNESCO member countries and the (rather semantic) changes in the criteria for the evaluation of quality that UNESCO has adopted during the 1972-2018 time interval. Finally, this relationship shows up also in the opposite sense, as countries reacted by raising the average quality of their newly proposed sites in years when limitations on the number of sites that could be proposed became more stringent. The results are quite robust to changes in the estimating techniques and in the specification of the data.

This research, however, raises several new questions and scenarios for future research, as it deals with an issue, the assessment of quality that is at the same time important and difficult to handle both in cultural and in mainstream economics. A first topic that will have to be revisited in the literature on the UNESCO WHL in the light of our research is the role that rent-seeking plays in the assignment of the new sites. Lobbying seems decisive in sites whose quality is barely sufficient to enter the list, contrary to the current consensus in the literature, that countries always resort to rent-seeking. Another open question is determining the precise number of sites beyond which the average quality of the whole WHL starts to decrease. A reduction of the average quality of the WHL would call into question the credibility and usefulness of the UNESCO policy to add more sites to the list.

References

- Alberini, A. and Longo, A. (2006). Combining the travel cost and contingent behavior methods to value cultural heritage sites: Evidence from armenia. *Journal of Cultural Economics*, 30(4):287–304.
- Ashenfelter, O. and Jones, G. V. (2013). The demand for expert opinion: Bordeaux wine. *Journal of Wine Economics*, 8(3):285–293.
- Bedate, A., Herrero, L. C., and Sanz, J. Á. (2004). Economic valuation of the cultural heritage: application to four case studies in spain. *Journal of cultural heritage*, 5(1):101–111.
- Bertacchini, E., Donatella, S., and Santagata, W. (2009). Enhancing the valorisation of unesco world heritage sites : a pigouvian approach. Technical report, Università di Torino, department of economics ”S. Cogneetti de Martiis”.
- Bertacchini, E., Liuzza, C., and Meskell, L. (2017). Shifting the balance of power in the unesco world heritage committee: an empirical assessment. *International journal of cultural policy*, 23(3):331–351.
- Bertacchini, E., Liuzza, C., Meskell, L., and Donatella, S. (2016). The politicization of unesco world heritage decision making. *Public Choice*, 167.
- Bertacchini, E. E. and Saccone, D. (2012). Toward a political economy of world heritage. *Journal of cultural economics*, 36(4):327–352.
- Dattilo, M., Padovano, F., and Rocaboy, Y. (2022). More is worse: the evolution of quality of the unesco world heritage list and its determinants. *Journal of Cultural Economics*, pages 1–26.
- De Piles, R. (1708). *Cours de peinture par principes composé par M. de Piles*. chez Jacques Estienne.
- Deuchert, E., Adjamah, K., and Pauly, F. (2005). For oscar glory or oscar money? *Journal of Cultural Economics*, 29(3):159–176.
- Frey, B. S., Pamini, P., and Steiner, L. (2013). Explaining the world heritage list: an empirical study. *International Review of Economics*, 60(1):1–19.

- Ginsburgh, V. (2003). Awards, success and aesthetic quality in the arts. *Journal of Economic perspectives*, 17(2):99–111.
- Ginsburgh, V. and Noury, A. G. (2008). The eurovision song contest. is voting political or cultural? *European Journal of Political Economy*, 24(1):41–52.
- Ginsburgh, V. and Weyers, S. (1999). On the perceived quality of movies. *Journal of Cultural Economics*, 23(4):269–283.
- Ginsburgh, V. A. and Van Ours, J. C. (2003). Expert opinion and compensation: Evidence from a musical competition. *American Economic Review*, 93(1):289–296.
- Jokilehto, J. (2006). World heritage: Defining the outstanding universal value. *City & time*, 2(2):1–10.
- Labadi, S. (2013). Unesco, cultural heritage, and outstanding universal value: Value-based analyses of the world heritage and intangible cultural heritage conventions.
- Nelson, R. A., Donihue, M. R., Waldman, D. M., and Wheaton, C. (2001). What’s an oscar worth? *Economic inquiry*, 39(1):1–6.
- Noonan, D. S. (2003). Contingent valuation and cultural resources: A meta-analytic review of the literature. *Journal of cultural economics*, 27(3):159–176.
- Parenti, B. and De Simone, E. (2015). Explaining determinants of national unesco tentative lists: an empirical study. *Applied Economics Letters*, 22(15):1193–1198.
- Ponzo, M. and Scoppa, V. (2015). Experts’ awards and economic success: evidence from an italian literary prize. *Journal of Cultural Economics*, 39(4):341–367.
- Reinstein, D. A. and Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The journal of industrial economics*, 53(1):27–51.
- Ruijgrok, E. C. (2006). The three economic values of cultural heritage: a case study in the netherlands. *Journal of cultural heritage*, 7(3):206–213.
- Steiner, L. and Frey, B. S. (2012). Correcting the imbalance of the world heritage list: Did the unesco strategy work? *Journal of International Organizations Studies*, 3(1):25–40.

- Throsby, C. D. (1990). Perception of quality in demand for the theatre. *Journal of cultural economics*, 14(1):65–82.
- Tobias, S. (2004). Quality in the performing arts: Aggregating and rationalizing expert opinion. *Journal of Cultural Economics*, 28(2):109–124.
- UNESCO (1972). Convention concerning the protection of the world cultural and natural heritage.
- UNESCO (2017). Operational guidelines for the implementation of the world heritage convention.
- van der Aa, B. J. (2005). Preserving the heritage of humanity? obtaining world heritage status and the impacts of listing. Ph. D. Dissertation, Faculty of Spatial Science, University of Groningen, The Netherlands.
- Waldfogel, J. (2012). Copyright protection, technological change, and the quality of new products: Evidence from recorded music since napster. *The journal of law and economics*, 55(4):715–740.

Chapter 2

Evaluating the Quality of UNESCO World Heritage List: a comparison with the Baedeker's guidebooks¹

Contents

1	Introduction	48
2	The Baedeker guidebooks	50
3	Data and measurements	51
4	Empirical Strategy	59
5	Results	60
6	Conclusions	64
	References	65

¹This chapter is an adapted version of a working paper coauthored with Proff Fabio Padovano and Yvon Rocaboy

*”Go, little book, God send thee good passage,
And specially let this be thy prayere
Unto them all that thee will read or hear,
Where thou art wrong, after their help to call,
Thee to correct in any part or all.”*

Go, little book, Geoffrey Chaucer ²

1 Introduction

As extensively discussed in the first chapter, to be included in the UNESCO WHL a site should meet at least one out of ten cultural and natural criteria of Outstanding Universal Value set by the UNESCO itself. Albeit with a few dissenting opinions (Frey and Steiner, 2011, Yang et al., 2019, Cellini, 2011), the literature and policy-makers generally concur that being a UNESCO World Heritage site generates significant benefits in terms of touristic revenues and public attention (Adie et al., 2018, VanBlarcom and Kayahan, 2011, Patuelli et al., 2016, Yang et al., 2010, Buckley, 2004, Rakic, 2007). Indeed, the inclusion in the WHL is generally considered a signal of sites’ outstanding quality (Adie, 2017) which could be a crucial determinant of consumers’ choices in a market, such as the touristic one, characterized by imperfect information (Keane, 1997).

The concept of quality is, indeed, an important, yet difficult to characterize, dimension of consumers’ choices, as the introduction of this thesis suggests. Contrary to quantity, quality needs to be approximated and these evaluations do not always encounter universal acceptance. The definition of the criteria for inclusion in the WHL is, therefore an important matter since they must assure that the selected sites are, indeed, of Outstanding Universal Value.

Scholars and national delegates to UNESCO have questioned the criteria formulated by the organization on two dimensions. The first is the validity and appropriateness of such criteria (Musitelli, 2002, Bertacchini and Saccone, 2012) which led

²This citation introduces each English version of the Baedeker series. It encourages the reader to look for any error in the guidebook, and, more subtly, to keep acquiring knowledge about each destination.

over time to the creation of new categories of sites and to a revision of the selection policies. The second type of critique, while accepting the existing criteria for evaluation of the proposed sites, points out that the mere inclusion is not a very informative signal of the quality of a site. As we highlight in the previous chapter of this thesis, the multiplicity of the criteria could convey more information about the quality of the site. Therefore, whether this indicator is robust and we can consider the number of satisfied criteria, more than the mere inscription, as an indicator of the “quality” or “value” of a site is a legitimate question.

Yet, although intuitive and straightforward, the idea of counting the number of criteria satisfied as a proxy for the site’s quality does not meet universal acceptance. It could be argued, for instance, that these criteria are meant to capture cultural diversity among sites, rather than the individual quality of a site (Jokilehto, 2006). Hence a more in-depth evaluation of the issue is warranted.

In this chapter, we test the hypothesis that the higher the number of criteria satisfied by the site, the higher its quality, as evaluated by UNESCO experts. We tackle this issue by examining whether the number of criteria satisfied by the WH sites coincides with an independent standard of quality. To this end, we select the Baedeker travel guidebooks, the most authoritative and comprehensive guidebook of the beginning of the XXth century, as that standard; in a way, they are the origin and model of contemporary travel guidebooks. The motivations underlying our choice are essentially two. First, since at the time of the publication of the Baedeker guidebooks the UNESCO did not exist, their selection as a standard for quality avoids problems of endogeneity and reverse causality that contemporary guides may suffer from. In other words, a contemporary guidebook may devote more pages to a tourist attraction because it is a WH site; or a site may enter the list because it has received positive evaluations from important guidebooks. Second, Baedeker guidebooks existed for most European countries. This continent is a geographical area characterized by a high degree of cultural homogeneity, compared to the rest of the world. Such homogeneity allows us to consider the number of OUV criteria satisfied as a proxy for quality rather than as a plurality of standards applicable to a variety of heterogeneous cultural contexts.

The analysis carried out in this chapter shows that the number of criteria satisfied by the UNESCO WH sites is, indeed, correlated with alternative measures of relevance expressed in the Baedeker travel guidebooks, like the number of citations, the length of the description of the site and the sentiment expressed in the text. This result confirms that the quality of the WH sites is of different degrees and that these degrees can be sensibly approximated by the number of satisfied criteria. Moreover, there seems to be little variation between the evaluation of cultural heritage expressed by the UNESCO WHL and the tastes of the Baedeker experts at the beginning of the XXth century.

Since the scope and definition of the criteria of OUV have been accurately discussed in Chapter 1; the present chapter will continue with the description of the Baedeker guidebooks and their relevance for the cultural tourism literature (Section 2). Section 3 presents the methodology applied for the extraction of the data and the definition of the metrics capturing the quality of the sites. The empirical strategy is detailed in Section 4; while Section 5 discusses the results of the analysis. A summary of the results obtained concludes the chapter (Section 6).

2 The Baedeker guidebooks

As Koshar (1998) highlights, the English middle-class professionals and intellectuals of the XIXth century have played a fundamental role in shaping the tastes that define, still today, what can be considered as cultural heritage. When these groups started to travel across Europe, they needed to organize their tours more efficiently than 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 time. 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 spanning 180 years and covering around forty countries existing

at those times. This series quickly became the accepted international paradigm for guidebooks and their author was considered “the arbiter of artistic tastes” (Bruce, 2010). First, the Baedekers eliminated all the nationalist comments, creating an almost completely objective and reliable guide (Koshar, 1998, Bruce, 2010)³. Second, they set quite uniform standards of evaluation, which made the Baedeker guidebooks a valid substitute for the absence of a single travel guidebook covering multiple countries. This allowed the Baedeker to be much more informative and detailed than their competitors. Moreover, the Baedeker guidebook was the first to select and list historical heritage, a fundamental practice in the market for cultural heritage still today. Indeed, he used asterisks to indicate sites of extraordinary quality⁴. Historical buildings and monuments were among the prized sites in the tourist’s itinerary and represent around 30% of the points of interest. The similarities with the UNESCO WHL are evident considering that around 35% of the sites belong to this category (Labadi, 2007)⁵.

To verify the evaluation of UNESCO WH sites in Baedeker’s guidebooks, we consider 11 volumes of the series⁶, edited from 1889 to 1911. We limit the edition to 1911 because the First World War reduced the possibility to travel and revamped nationalistic sentiments also in the Baedeker guide, making the comparison of heritage from different countries less balanced.

3 Data and measurements

Our final data frame contains 247 UNESCO WH sites, for 10 European coun-

³In the Murray guidebooks, for example, the pride in the English sense of freedom and comfort is evident in the description of the travel preparation and accommodations.

⁴Mariana Starke was an important precursor. Indeed, her two-volume guidebook contains exclamation marks to rank important sights.

⁵Figures refer to the year 2005.

⁶“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.

tries⁷, which are analysed considering different aspects of the tastes of the authors of the Baedeker guidebooks. An original algorithm (developed by the author) extracts the data processing the PDF version of the books⁸. The advantages of this approach are evident: 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: the definition of the variables should be specified *a priori*, ensuring that each observation is evaluated following 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, eventually, digitalisation errors, i.e. incorrect transposition of the text from the paper to the PDF format. We make sure, however, to perform the text analysis minimizing our dependency on this kind of error.

In order to produce a reliable analysis, two main issues had to be overcome. In some cases, such as the United Kingdom a single book is available; while for other countries, such as Italy, France and Germany, the algorithm must merge more than one text. Considering Italy as a matter of example, the data extracted from two volumes, one covering the north and the centre of the country and another the south, are aggregated in a single measure in the final data frame. Secondly, the official names of the UNESCO WH sites are often long and elaborate. A clear example is the official name of the *Cenacolo*: “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. The former example is then simply reduced to “cenacolo”⁹. 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”¹⁰.

⁷Countries included in the study: Austria, Hungary, France, Spain, Portugal, the Netherlands, Belgium, Italy, Germany, United Kingdom

⁸The texts are available at <https://archive.org/>. The digitalisation is sponsored by the Kahle/Austin Foundation

⁹To be implemented in our algorithm, WH sites’ names must be lowercase.

¹⁰Same remark as above: titles must be lowercase.

The measures of quality we extracted with this approach are detailed in the following sections.

3.1 Number of citations

Our first variable of interest $Citation_i$, is the number of citations the WH site i has received in Baedeker’s guidebooks. We assume that the higher the number of citations, the greater the importance that the author has attributed to the site. 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 text. 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”. Also, the sites placed outside the geographical Europe, e.g. the former colonies, are excluded from the sample. The distribution of $Citation_i$, shown in the first row of Figure 2.1, 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 equation 2.1, will be used instead. From the second line of Figure 2.1 is evident that this latter variable better captures the variability in our sample.

$$ln_Citation_i = \begin{cases} 0 & \text{if } Citation_i = 0 \\ \log_{10}(Citation_i) & \text{if } Citation_i > 0 \end{cases} \quad (2.1)$$

Despite the similarity in size, the guidebooks could differ with respect to the total number of sites cited. For example, the index of the North of France guidebook contains 3366 items; while the South of France one lists 6590 heritage sites. To overcome this disparity we have computed $ln_Icitation_i$ as described in equation 2.2: where $Tweight$ is the number of index entries in each book. The distribution of this variable, plotted in the third row of Figure 2.1, appears to be close to the distribution of $ln_Citation_i$, showing again that the Baedeker’s books considered are easily comparable between them.

$$ln_Icitation_i = \begin{cases} 0 & \text{if } Citation_i = 0 \\ \log_{10} \frac{100000}{Tweight} * Citation_i & \text{if } Citation_i > 0 \end{cases} \quad (2.2)$$

3.2 Length of the paragraph

As the author of the Baedeker’s guidebooks could express his appreciation for a WH site 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 that the author deems interesting for the reader, the higher the quality of the site should be. 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 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 sentence separation could be less evident, focusing on paragraphs makes sure our analysis is reliable and independent from any digitalization errors. We, then, extract from the whole text all the paragraphs containing at least one citation of the site considered. We compute $Paragraph_i$ as the average number of words within these paragraphs. The variable is computed as in equation 2.3, where T is the total number of paragraphs for which $Citation_{ij} > 0$. The distribution of this variable, displayed in the fourth row of Figure 2.1, 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 500 words; a medium-size paragraph is around 1000 words; while a long paragraph contains around 1500 words on average. This distribution is not surprising since a paragraph with just 1 or 2 words is extremely unlikely, as it is one 2000 words long. This distribution is probably due to the author’s style and we should not, therefore, expect a normal distribution. As a consequence, the categorical variable $Paragraph_c$ could be more efficient in capturing this kind of distribution. We define it as equal to 0 when $Paragraph_i$ is equal to 0; equal to 1 when $Paragraph_i$ is between 1 and 750; equal to 2 if $Paragraph_i$ is between 751 and 1250, and equal to 3 if $Paragraph_i$ is greater than 1250.

$$Paragraph_i = \frac{1}{T} \sum_{j=1}^T W_j \quad (2.3)$$

3.3 The sentiment of the text

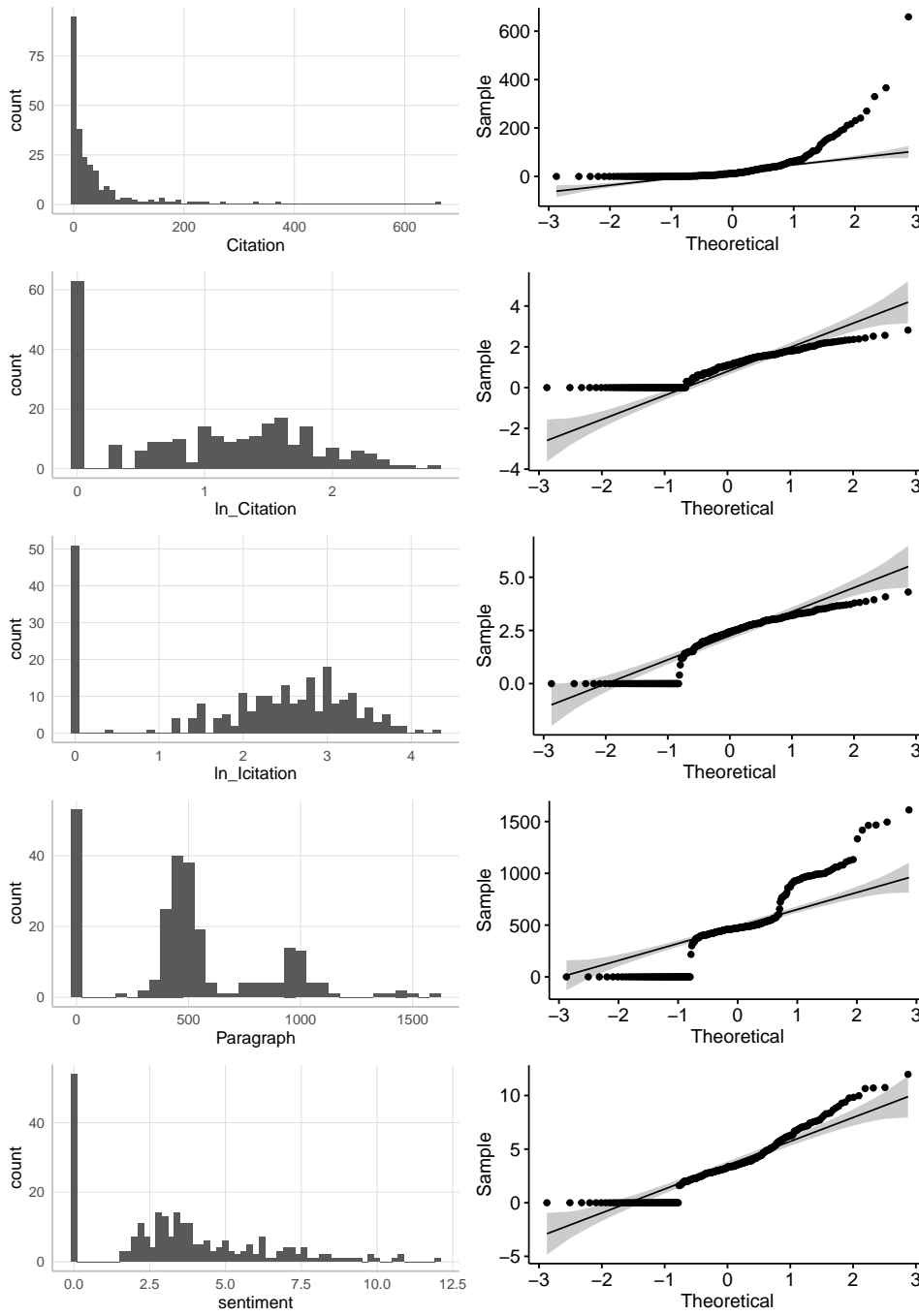
Finally, we capture the author’s tastes for a WH site by looking at the vocabulary he adopts in its description. The idea is that the more positive the expressions used in the description of the site, the higher should be its quality. The literature focuses on two methods to perform the “sentiment analysis” of a text: the lexicon-based and the machine learning one (Hutto and Gilbert, 2014). 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 all the contextual characteristics specific to the text in which the word appears are ignored (Shapiro et al., 2020). Application of this method could be found in Political Economy (Shapiro and Wilson, 2022) and Financial Economics (Heston and Sinha, 2017, Fraiberger, 2016, Correa et al., 2021, Nyman et al., 2021). An example of the machine learning approach is the algorithm developed by Hutto and Gilbert (2014), called Vader, which accounts for the word’s context within the sentence. Since Vader is developed at the sentence level and specific for social media language, it is not appropriate for the Baedeker content and format. Instead, we exploit the *affin* lexicon of the *tidytext* R package (Silge and Robinson, 2016, Nielsen, 2011b), a largely used and valid alternative (Sharma et al., 2018, Benchimol et al., 2022). Since the guidebooks are already a selection of “the best sites” we focus only on laudatory expression. 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 analyzed the vocabulary adopted. As a result, we obtain the variable *sentiment* computed as the equation 2.4; where N_j is the number of positive words in the paragraph j ; f is the frequency of the word w in the paragraph j and v is a value from 1 (slightly positive) to 5 (really positive) assigned by the *affin* dictionary¹¹ to the word w in the paragraph j . The distribution of *sentiment* is shown in the last row of Figure 2.1.

¹¹*AFINN* 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, Finn Årup Nielsen, the author of the dictionary, has aggregated and expanded preexisting 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.derosene.net/steve/resources/emotionwords/ewords.html>)

$$sentiment_i = \frac{1}{T} \sum_{j=1}^T \frac{1}{N_j} \sum_{w=1}^{N_j} f_{wj} v_{wj} \quad (2.4)$$

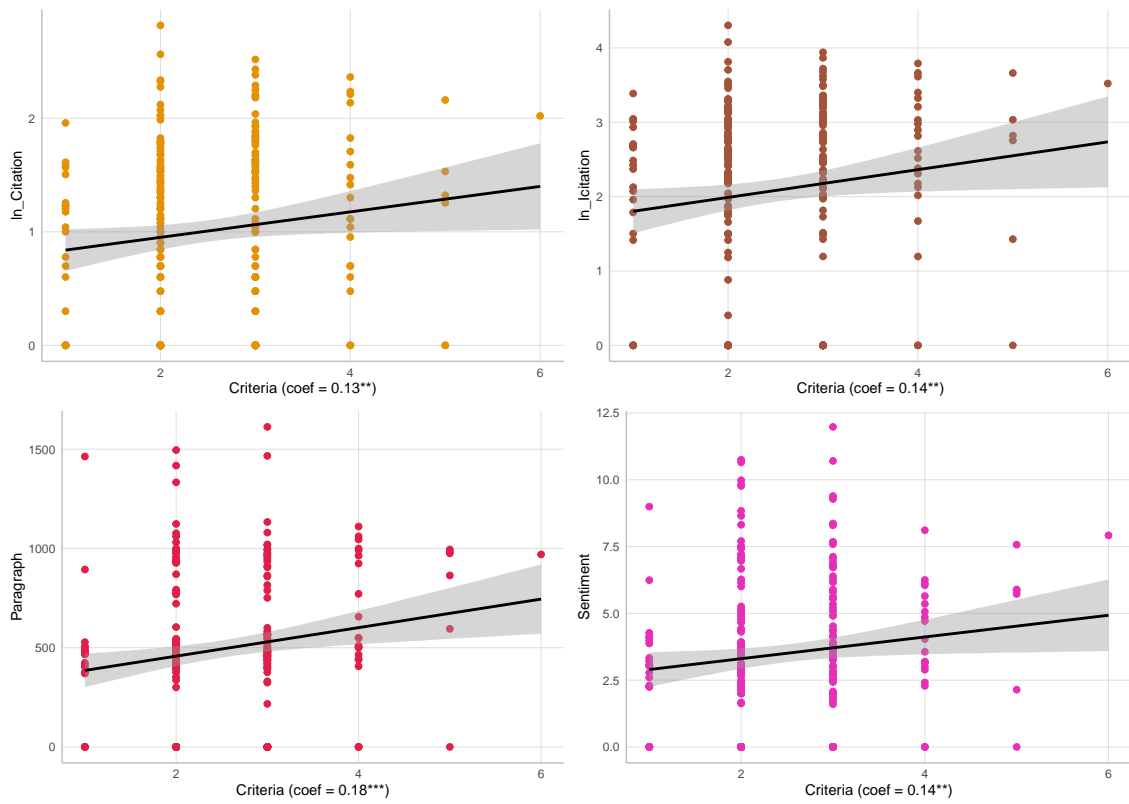
As Figure 2.2 shows, all four variables of Baedeker's tastes are positively and significantly correlated with the number of criteria satisfied by the WH sites considered. This result gives the first insight that the quality of UNESCO sites can indeed be validly proxied by these variables. However, we complement this first finding with a more systematic and robust empirical strategy.

Figure 2.1: Distribution of quality (Baedeker 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).

Figure 2.2: 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: * , $p < 0.1$; ** , $p < 0.05$; *** , $p < 0.01$

4 Empirical Strategy

The goal of the analysis is to test the hypothesis that the evaluation of the heritage sites in the Baedeker travel guidebooks is consistent with the evaluation of their quality via the number of UNESCO OUV criteria satisfied¹². To this end, we have performed two tests. The first is straightforward: we regress $Criteria_i$, i.e. the number of criteria satisfied by site i , on the four independent variables expressing Baedeker's tastes, i.e. $ln_Citation$, $ln_Icitation$, $Paragraph$ and $sentiment$. The explanatory variables are included sequentially, resulting in four different regressions. We estimate equation 2.5, representing the empirical model, first using a simple OLS model.

$$Criteria_i = \beta_0 + \beta_1 * Baedeker_i + u_i \quad (2.5)$$

$Criteria$, however, is intended as a discrete and ordered variable: $Criteria_i = 1$ means that the site i has the lowest quality, while if $Criteria_i = 6$, site i has the highest quality in our sample. In this case, following (Wooldridge, 2010) the Ordered Logit Model should be more precise in estimating the regression coefficients. Indeed, we assume that Q_i , i.e. the quality of the site i , is a continuous variable, but we can only observe whether Q attains certain thresholds δ_k :

$$\begin{aligned} Criteria &= 1, & \text{if } Q &\leq \delta_1 \\ Criteria &= 2, & \text{if } \delta_1 &\leq Q \leq \delta_2 \\ Criteria &= 3, & \text{if } \delta_2 &\leq Q \leq \delta_3 \\ Criteria &= 4, & \text{if } \delta_3 &\leq Q \leq \delta_4 \\ Criteria &= 5, & \text{if } \delta_4 &\leq Q \leq \delta_5 \\ Criteria &= 6, & \text{if } \delta_5 &\leq Q \end{aligned} \quad (2.6)$$

Both the OLS and the ordered Logit models are also estimated including the country fixed effects.

¹²As the Baedeker guidebooks predate the UNESCO by almost half a century, our analysis also sheds light on whether these authors have influenced the discipline of UNESCO criteria. We prefer, however, not to make any strong (e.g. causal) statement on the subject.

The second test verifies if, once the evaluations of the Baedeker guidebooks are introduced in models explaining the quality of the UNESCO WH sites, the results still go through. So far, the only empirical analysis to have examined the quality of UNESCO WH sites is the one presented in the previous chapter. We, therefore, estimate again equation 1.2, replacing the dependent variables used in the original model with the four measures of Baedeker’s tastes, i.e. *ln_Citation*, *ln_Icitation*, *Paragraph* and *sentiment*. The main explanatory variables remain unchanged¹³; the main variables of interest are proxies for the time when a site is included, while the controls include the country’s lobbying power in the UNESCO Committee and the size of the cultural and natural capital stock. If the results of the estimates do not vary notwithstanding the substitution of the dependent variables, the number of criteria satisfied proves to be an equally valid proxy of the quality of a site as the Baedeker’s evaluation. Equation 2.7 is specified as follows: Q_i are the four evaluations of quality from the Baedeker guidebooks, N_i are three measures of the increasing number of sites, i.e. *Tenure*, *Site_{t-1}* and *Year*. We also include AV_Q_i , the country average sites’ quality, to capture the marginal variation of Q_i for each country.

$$Q_i = \beta_0 + \beta_1 N_i + \beta_2 AV_Q_i + \beta_3 X_i + u_i \quad (2.7)$$

Since *ln_Citation*, *ln_Icitation* and *sentiment* are not linearly distributed and a considerable part of the sample is equal to zero, a Tobit model is estimated, together with the standard linear regression. In a similar way, *Paragraph* is also estimated in its categorical version, i.e. *Paragraph_c*. In this specification, equation 2.7 is estimated using an Ordered Logit Model, which should be more precise than the simple OLS.

5 Results

Tables 2.1 and 2.2 show the results of the estimates of equation 2.5. The coefficients of the linear model are all positive and significant, at least at the 5% level

¹³The only exception is *Unpaid_contr*, which is removed from the list of control variables in the new specification. Indeed, its inclusion significantly reduces the sample size.

(models from 1 to 8). For all our variables (except *ln_Icitation*), including the country fixed effects increases the magnitude of the coefficients of interest, but reduces the goodness-of-fit of the regression. Since *Criteria* is an integer number, however, the magnitude of the coefficients of the linear models is hard to interpret. To facilitate the interpretation of the results, the coefficients of Table 2.2, presenting the estimates of the Ordered Logit model, are expressed as the 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 receiving one more criterion increases between 37 and 43 percentage points (models 9 and 10). This percentage decreases to around 21-23% when we consider *ln_Icitation*, but it is still positive and significant (models 11 and 12). The odds of receiving one more criterion is 0.1% when the average length of the paragraph increases by one word, regardless if we consider the model with or without fixed effects (models 13 and 14). On the other way, the likelihood of receiving one additional criterion increases by between 8.5% and 12.2% when *sentiment* increases by 1 unit. Contrary to the OLS results, the McFadden's R^2 and the AIC values suggest that adding country fixed effects does not only increase the magnitude of the coefficients, but also the goodness of fit of the regressions. In addition, for both the linear and the logit estimates, the models containing *Paragraph* fit the data better than the other ones.

Table 2.3 reports a summary of the estimations of Equation 2.7¹⁴ The coefficients of *Year* are always negative and significant for the models from 17 to 20, and no relevant differences emerge when the linear or the Tobit versions are considered. In addition, the coefficient of model 22 is also positive and significant. While interpreting this model, we should bear in mind that the impact of *Year* on *Paragraph_c* is negative when the odds ratios are less than 1. Therefore, when *Year* increases by 1 unit, the likelihood to expand the average size of the paragraph decreases by 15%, controlling for the average quality of each country. The coefficient for *sentiment*, even if not significant, is coherent with the whole trend and shows a negative sign. The average value of the dependent variable considered, *AV_Q*, is always positive and significant, as we would expect.

¹⁴The complete estimations can be found in the Appendix A.

Table 2.1: Determinants of Criteria: Linear Regressions

	<i>Dependent variable:</i>							
	Criteria							
	1 Linear	2 Linear	3 Linear	4 Linear	5 Linear	6 Linear	7 Linear	8 Linear
ln_Citation	0.156** (0.074)	0.166** (0.078)						
ln_Icitation			0.100** (0.045)	0.100** (0.046)				
Paragraph					0.0005*** (0.0002)	0.001*** (0.0002)		
sentiment							0.045** (0.021)	0.059** (0.023)
Constant	2.345*** (0.091)	1.741*** (0.155)	2.294*** (0.102)	1.751*** (0.150)	2.272*** (0.093)	1.675*** (0.145)	2.343*** (0.088)	1.723*** (0.148)
Country FE	NO	YES	NO	YES	NO	YES	NO	YES
Observations	247	247	247	247	247	247	247	247
R ²	0.017	0.013	0.019	0.016	0.033	0.031	0.018	0.014
Adjusted R ²	0.013	0.009	0.015	0.012	0.030	0.027	0.014	0.010

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.2: Determinants of Criteria: Ordered Logit Regressions

	<i>Dependent variable:</i>							
	Criteria							
	9 OLR	10 OLR	11 OLR	12 OLR	13 OLR	14 OLR	15 OLR	16 OLR
ln_Citation	1.369** (0.156)	1.435** (0.165)						
ln_Icitation			1.222** (0.097)	1.2316** (0.103)				
Paragraph					1.001** (0.0003)	1.001*** (0.0004)		
sentiment							1.085* (0.044)	1.122** (0.05)
Country FE	NO	YES	NO	YES	NO	YES	NO	YES
Observations	247	247	247	247	247	247	247	247
Log-Likelihood	-311.016	-298.463	-310.904	-298.799	-309.815	-296.989	-311.325	-298.203
McFadden's R^2	0.007	0.047	0.007	0.046	0.01	0.051	0.006	0.047
AIC	634.032	626.925	633.808	627.598	631.63	623.98	634.65	626.406

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.3: Decreasing marginal quality of UNESCO WHL

	<i>Dependent variable:</i>							
	ln.Citation		ln.Icitation		<i>Paragraph</i>	<i>Paragraph_c</i>	<i>sentiment</i>	<i>sentiment</i>
	17 Linear	18 Tobit	19 Linear	20 Tobit	21 Linear	22 Logistic	23 Linear	24 Tobit
Year	-0.011** (0.006)	-0.017** (0.008)	-0.022** (0.010)	-0.028** (0.012)	-3.725 (2.717)	0.985*** (0.0004)	-0.024 (0.021)	-0.034 (0.024)
AV_Q	1.169*** (0.323)	1.510*** (0.474)	0.956*** (0.248)	1.065*** (0.356)	1.281*** (0.290)	1.008*** (0.002)	1.182*** (0.221)	1.330*** (0.280)
Committee	-0.023 (0.041)	-0.028 (0.051)	-0.013 (0.067)	-0.008 (0.077)	-36.900* (23.525)	-0.078 (0.124)	-0.198 (0.178)	-0.226 (0.172)
Exp_contribution	0.062 (0.124)	0.078 (0.144)	0.033 (0.185)	0.034 (0.218)	2.442 (63.408)	0.922 (0.275)	0.034 (0.514)	-0.051 (0.456)
Area	-0.010 (0.048)	-0.016 (0.060)	0.018 (0.065)	0.022 (0.080)	7.072 (14.139)	1.025 (0.095)	0.025 (0.114)	0.029 (0.157)
Population	-0.0001 (0.001)	-0.0001 (0.001)	-0.0002 (0.001)	-0.0002 (0.001)	-0.139 (0.221)	0.999 (0.001)	-0.001 (0.002)	-0.001 (0.002)
Constant	22.478* (11.452)	32.062** (15.552)	43.061** (19.237)	54.523** (23.368)	46.774 (15.394)		66.166 (42.341)	65.603 (48.290)
Observations	232	232	223	223	232	232	223	223
R ²	0.084		0.095		0.161		0.195	
Adjusted R ²	0.059		0.070		0.138		0.173	

Standard errors in parentheses.

Significant levels are: * p < 0.1, ** p < 0.05, *** p < 0.01.

6 Conclusions

In this chapter, we show that the number of criteria of UNESCO WH sites is correlated with four different measures of tastes derived from Baedeker’s travel guidebooks. These evaluations have the advantage not only of being authoritative, but they avoid problems of endogeneity and reverse causation since the 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 construct an original data set based on guidebooks edited at the beginning of the XXth century. These results empirically prove that the number of criteria of UNESCO WHS is an informative indicator of quality¹⁵. This legitimizes the choice of the number of criteria of OUV that a site satisfies as an indicator of the quality of the site.

¹⁵We also sustain our analysis by verifying if the positive correlation between *Criteria* and the other measures of quality is somehow linked to the presence in our sample to both natural and cultural sites. Natural sites, indeed, receive systematically fewer criteria: cultural sites could receive at most 6 criteria, while this maximum is 4 for natural sites. On the other way, the Baedeker guides are mostly focused on cities and monuments; while parks and landscapes could be harder to identify. For this reason, we test our model on a subsample of UNESCO WHL containing only cultural sites. These additional tests do not change our results in any relevant way.

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., and Prayag, G. (2018). World heritage as a placebo brand: A comparative analysis of three sites and marketing implications. *Journal of Sustainable Tourism*, 26(3):399–415.
- Benchimol, J., Kazinnik, S., and Saadon, Y. (2022). Text mining methodologies with r: An application to central bank texts. *Machine Learning with Applications*, 8:100286.
- Bertacchini, E. E. and Saccone, D. (2012). Toward a political economy of world heritage. *Journal of cultural economics*, 36(4):327–352.
- Bruce, D. M. (2010). 7 baedeker: the perceived “inventor” of the formal guidebook-a bible. *Giants of tourism*, page 93.
- Buckley, R. (2004). The effects of world heritage listing on tourism to australian national parks. *Journal of sustainable tourism*, 12(1):70–84.
- Cellini, R. (2011). Is unesco recognition effective in fostering tourism? a comment on yang, lin and han. *Tourism management*, 32(2):452–454.
- Correa, R., Garud, K., Londono, J. M., and Mislant, N. (2021). Sentiment in central banks’ financial stability reports. *Review of Finance*, 25(1):85–120.
- Fraiberger, S. P. (2016). News sentiment and cross-country fluctuations. *Available at SSRN 2730429*.
- Frey, B. S. and Steiner, L. (2011). World heritage list: does it make sense? *International Journal of Cultural Policy*, 17(5):555–573.
- Heston, S. L. and Sinha, N. R. (2017). News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal*, 73(3):67–83.
- Hutto, C. and Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225.

- Jokilehto, J. (2006). World heritage: Defining the outstanding universal value. *City & time*, 2(2):1–10.
- Keane, M. J. (1997). Quality and pricing in tourism destinations. *Annals of tourism research*, 24(1):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(3):323–340.
- Labadi, S. (2007). *World heritage: challenges for the millennium*. UNESCO.
- Musitelli, J. (2002). World heritage, between universalism and globalization. *International journal of cultural property*, 11(2):323–336.
- Nielsen, F. Å. (2011a). Afinn. informatics and mathematical modelling, technical university of denmark. available at : <http://www2.compute.dtu.dk/pubdb/pubs/6010-full.html>.
- Nielsen, F. Å. (2011b). A new anew: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv:1103.2903*.
- Nyman, R., Kapadia, S., and Tuckett, D. (2021). News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127:104119.
- Patuelli, R., Mussoni, M., and Candela, G. (2016). The effects of world heritage sites on domestic tourism: a spatial interaction model for italy. In *Spatial Econometric Interaction Modelling*, pages 281–315. Springer.
- Rakic, T. (2007). World heritage: issues and debates. *Tourism: an international interdisciplinary journal*, 55(2):209–219.
- Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2020). Measuring news sentiment. *Journal of econometrics*.
- Shapiro, A. H. and 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(5):2768–2805.

- Sharma, N. et al. (2018). Sentiment analysis using tidytext package in r. In *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pages 577–580. IEEE.
- Silge, J. and Robinson, D. (2016). tidytext: Text mining and analysis using tidy data principles in r. *Journal of Open Source Software*, 1(3):37.
- VanBlarcom, B. L. and Kayahan, C. (2011). Assessing the economic impact of a unesco world heritage designation. *Journal of Heritage Tourism*, 6(2):143–164.
- Yang, C.-H., Lin, H.-L., and Han, C.-C. (2010). Analysis of international tourist arrivals in china: The role of world heritage sites. *Tourism management*, 31(6):827–837.
- Yang, Y., Xue, L., and Jones, T. E. (2019). Tourism-enhancing effect of world heritage sites: Panacea or placebo? a meta-analysis. *Annals of Tourism Research*, 75:29–41.

Appendix A

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

Table A.1: Decreasing marginal quality: $Q = \ln_Citation$ and $N = Tenure$

	<i>Dependent variable:</i>							
	ln_Citation							
	A-1 Linear	A-2 Linear	A-3 Linear	A-4 Linear	A-5 Tobit	A-6 Tobit	A-7 Tobit	A-8 Tobit
Tenure	-0.006 (0.005)	-0.008* (0.005)	-0.011* (0.006)	-0.011* (0.006)	-0.009 (0.006)	-0.011* (0.006)	-0.015** (0.008)	-0.015* (0.008)
AV_Citation		1.118*** (0.244)	1.192*** (0.248)	1.180*** (0.324)		1.420*** (0.357)	1.540*** (0.383)	1.521*** (0.476)
Committe			-0.013 (0.040)	-0.014 (0.041)			-0.012 (0.050)	-0.015 (0.052)
Exp_contribution			0.064 (0.071)	0.052 (0.125)			0.089 (0.093)	0.061 (0.143)
Area				0.001 (0.049)				0.001 (0.060)
Population				0.0001 (0.001)				0.0001 (0.001)
Constant	1.127*** (0.086)	-0.278 (0.313)	-0.409 (0.318)	-0.409 (0.385)	1.041*** (0.112)	-0.745 (0.464)	-0.955* (0.505)	-0.967* (0.588)
Observations	232	232	223	223	232	232	223	223
R ²	0.008	0.076	0.082	0.082	0.008	0.076	0.082	0.082
Adjusted R ²	0.003	0.068	0.065	0.057	0.003	0.068	0.065	0.057

Note:

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

Table A.2: Decreasing marginal quality: $Q = \ln_Citation$ and $N = Sites_{t-1}$

	<i>Dependent variable:</i>							
	ln_Citation							
	A-9 Linear	A-10 Linear	A-11 Linear	A-12 Linear	A-13 Tobit	A-14 Tobit	A-15 Tobit	A-16 Tobit
Sites _{t-1}	-0.002 (0.004)	-0.007 (0.004)	-0.009* (0.005)	-0.010* (0.006)	-0.004 (0.005)	-0.010* (0.005)	-0.014** (0.007)	-0.014** (0.007)
AV_Citation		1.217*** (0.258)	1.339*** (0.271)	1.267*** (0.336)		1.563*** (0.370)	1.753*** (0.413)	1.648*** (0.487)
Committe			-0.027 (0.040)	-0.030 (0.041)			-0.031 (0.051)	-0.036 (0.052)
Exp_contribution			0.069 (0.074)	0.049 (0.122)			0.099 (0.095)	0.061 (0.142)
Area				0.015 (0.050)				0.020 (0.061)
Population				0.0001 (0.001)				0.0002 (0.001)
Constant	1.068*** (0.089)	-0.392 (0.312)	-0.573* (0.332)	-0.528 (0.390)	0.978*** (0.117)	-0.904* (0.462)	-1.190** (0.525)	-1.138* (0.596)
Observations	232	232	223	223	232	232	223	223
R ²	0.001	0.076	0.082	0.083	0.001	0.076	0.082	0.083
Adjusted R ²	-0.003	0.068	0.065	0.057	-0.003	0.068	0.065	0.057

Note:

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

Table A.3: Decreasing marginal quality: $Q = \ln_Citation$ and $N = Year$

	<i>Dependent variable:</i>							
	ln_Citation							
	A-17 Linear	A-18 Linear	A-19 Linear	A-20 Linear	A-21 Tobit	A-22 Tobit	A-23 Tobit	A-24 Tobit
Year	-0.012** (0.005)	-0.010** (0.005)	-0.011** (0.005)	-0.011** (0.006)	-0.016** (0.007)	-0.015** (0.007)	-0.015** (0.007)	-0.017** (0.008)
AV_Citation		1.050*** (0.244)	1.110*** (0.251)	1.169*** (0.323)		1.330*** (0.358)	1.425*** (0.379)	1.510*** (0.474)
Committe			-0.025 (0.040)	-0.023 (0.041)			-0.030 (0.051)	-0.028 (0.051)
Exp_contribution			0.033 (0.063)	0.062 (0.124)			0.046 (0.082)	0.078 (0.144)
Area				-0.010 (0.048)				-0.016 (0.060)
Population				-0.0001 (0.001)				-0.0001 (0.001)
Constant	24.006** (9.895)	20.206** (9.641)	20.584** (10.119)	22.478* (11.452)	33.782** (13.544)	28.867** (13.145)	29.729** (14.334)	32.062** (15.552)
Observations	232	232	223	223	232	232	223	223
R ²	0.021	0.081	0.083	0.084	0.021	0.081	0.083	0.084
Adjusted R ²	0.016	0.073	0.066	0.059	0.016	0.073	0.066	0.059

Note:

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

Table A.4: Decreasing marginal quality: $Q = \ln_Icitation$ and $N = Tenure$

	<i>Dependent variable:</i>							
	ln_Icitation							
	A-25 Linear	A-26 Linear	A-27 Linear	A-28 Linear	A-29 Tobit	A-30 Tobit	A-31 Tobit	A-32 Tobit
Tenure	-0.011 (0.008)	-0.017** (0.008)	-0.020** (0.009)	-0.020** (0.010)	-0.015 (0.009)	-0.022** (0.009)	-0.026** (0.011)	-0.025** (0.012)
AV_Icitation		1.085*** (0.211)	1.070*** (0.219)	0.963*** (0.246)		1.234*** (0.308)	1.216*** (0.322)	1.076*** (0.360)
Committe			0.010 (0.066)	0.004 (0.067)			0.021 (0.076)	0.013 (0.077)
Exp_contribution			0.057 (0.115)	0.013 (0.184)			0.068 (0.137)	0.008 (0.217)
Area				0.041 (0.065)				0.052 (0.079)
Population				0.0001 (0.001)				0.0002 (0.001)
Constant	2.276*** (0.132)	-0.480 (0.541)	-0.505 (0.548)	-0.365 (0.571)	2.200*** (0.170)	-0.936 (0.801)	-0.973 (0.835)	-0.790 (0.871)
Observations	232	232	223	223	232	232	223	223
R ²	0.009	0.087	0.089	0.092	0.009	0.087	0.089	0.092
Adjusted R ²	0.004	0.079	0.072	0.067	0.004	0.079	0.072	0.067

Note:

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

Table A.5: Decreasing marginal quality: $Q = \ln_Citation$ and $N = Sites_{t-1}$

	<i>Dependent variable:</i>							
	ln_Icitation							
	A-33 Linear	A-34 Linear	A-35 Linear	A-36 Linear	A-37 Tobit	A-38 Tobit	A-39 Tobit	A-40 Tobit
Sites _{t-1}	-0.004 (0.007)	-0.016** (0.007)	-0.017* (0.009)	-0.019** (0.009)	-0.007 (0.008)	-0.020** (0.008)	-0.022** (0.010)	-0.024** (0.010)
AV_Icitation		1.198*** (0.227)	1.208*** (0.235)	1.063*** (0.260)		1.384*** (0.325)	1.397*** (0.347)	1.208*** (0.371)
Committe			-0.015 (0.066)	-0.025 (0.068)			-0.011 (0.077)	-0.024 (0.078)
Exp_contribution			0.048 (0.118)	0.016 (0.184)			0.060 (0.137)	0.016 (0.216)
Area				0.071 (0.066)				0.090 (0.081)
Population				0.0001 (0.001)				0.0001 (0.001)
Constant	2.188*** (0.140)	-0.747 (0.557)	-0.818 (0.571)	-0.624 (0.588)	2.109*** (0.177)	-1.285 (0.818)	-1.384 (0.873)	-1.129 (0.891)
Observations	232	232	223	223	232	232	223	223
R ²	0.002	0.087	0.086	0.093	0.002	0.087	0.086	0.093
Adjusted R ²	-0.003	0.079	0.070	0.068	-0.003	0.079	0.070	0.068

Note:

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

Table A.6: Decreasing marginal quality: $Q = \ln_Icitation$ and $N = Year$

	<i>Dependent variable:</i>							
	ln_Icitation							
	A-41 Linear	A-42 Linear	A-43 Linear	A-44 Linear	A-45 Tobit	A-46 Tobit	A-47 Tobit	A-48 Tobit
Year	-0.023*** (0.008)	-0.023*** (0.008)	-0.022*** (0.008)	-0.022** (0.010)	-0.029*** (0.010)	-0.029*** (0.010)	-0.028** (0.011)	-0.028** (0.012)
AV_Icitation		0.973*** (0.208)	0.980*** (0.220)	0.956*** (0.248)		1.090*** (0.299)	1.097*** (0.316)	1.065*** (0.356)
Committe			-0.013 (0.066)	-0.013 (0.067)			-0.009 (0.076)	-0.008 (0.077)
Exp_contribution			0.011 (0.104)	0.033 (0.185)			0.009 (0.122)	0.034 (0.218)
Area				0.018 (0.065)				0.022 (0.080)
Population				-0.0002 (0.001)				-0.0002 (0.001)
Constant	47.836*** (16.173)	44.849*** (15.778)	42.978** (16.737)	43.061** (19.237)	60.737*** (20.366)	56.627*** (19.664)	54.539** (21.528)	54.523** (23.368)
Observations	232	232	223	223	232	232	223	223
R ²	0.032	0.097	0.095	0.095	0.032	0.097	0.095	0.095
Adjusted R ²	0.028	0.089	0.078	0.070	0.028	0.089	0.078	0.070

Note:

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

Table A.7: Decreasing marginal quality: $Q = \text{Paragraph}$ and $N = \text{Tenure}$

	<i>Dependent variable:</i>							
	Paragraph <i>OLS</i>				Paragraph_ch <i>ordered logistic</i>			
	A-49 Linear	A-50 Linear	A-51 Linear	A-52 Linear	A-53 OLR	A-54 OLR	A-55 OLR	A-56 OLR
Tenure	1.759 (2.261)	-4.412** (2.271)	-3.594 (2.536)	-3.828 (2.566)	1.011 (0.012)	0.978* (0.014)	0.986 (0.015)	0.984 (0.015)
AV_Paragraph		1.121*** (0.151)	1.327*** (0.253)	1.360*** (0.278)		1.007*** (0.001)	1.008*** (0.002)	1.009*** (0.002)
Committe			-35.374* (23.636)	-36.331* (23.585)			0.854 (0.125)	0.849 (0.125)
Exp_contribution			-14.962 (47.695)	1.974 (61.455)			0.792 (0.233)	0.918 (0.308)
Area				11.260 (14.098)				1.043 (0.096)
Population				-0.123 (0.217)				0.999 (0.001)
Constant	474.644*** (33.321)	-126.814 (78.680)	-211.380* (114.702)	-225.447* (116.261)				
Observations	232	232	223	223	232	232	223	223
R ²	0.003	0.162	0.160	0.162				
Adjusted R ²	-0.002	0.154	0.144	0.138				

Note:

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

Table A.8: Decreasing marginal quality: $Q = Paragraph$ and $N = Sites_{t-1}$

	<i>Dependent variable:</i>							
	Paragraph				Paragraph_ch			
	<i>OLS</i>				<i>ordered logistic</i>			
	A-57 Linear	A-58 Linear	A-59 Linear	A-60 Linear	A-61 OLR	A-62 OLR	A-63 OLR	A-64 OLR
Sites _{t-1}	0.834 (1.916)	-3.257* (1.858)	-3.073 (2.254)	-3.962* (2.520)	1.005 (0.011)	0.982 (0.011)	0.988 (0.013)	0.984 (0.014)
AV_Paragraph		1.077*** (0.148)	1.328*** (0.250)	1.313*** (0.281)		1.007*** (0.001)	1.008*** (0.002)	1.008*** (0.002)
Committe			-38.606** (23.423)	-40.105** (23.304)			0.841 (0.125)	0.835 (0.125)
Exp_contribution			-15.542 (48.842)	5.203 (62.177)			0.800 (0.234)	0.933 (0.307)
Area				19.796 (15.786)				1.080 (0.103)
Population				-0.118 (0.214)				0.999 (0.001)
Constant	485.481*** (36.349)	-105.600 (79.241)	-204.126* (114.740)	-211.214* (117.168)				
Observations	232	232	223	223	232	232	223	223
R ²	0.001	0.159	0.160	0.165				
Adjusted R ²	-0.004	0.151	0.145	0.142				

Note:

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

Table A.9: Decreasing marginal quality: $Q = \text{Paragraph}$ and $N = \text{Year}$

	<i>Dependent variable:</i>							
	Paragraph <i>OLS</i>				Paragraph_ch <i>ordered logistic</i>			
	A-65 Linear	A-66 Linear	A-67 Linear	A-68 Linear	A-69 OLR	A-70 OLR	A-71 OLR	A-72 OLR
Year	-4.406* (2.382)	-4.311* (2.235)	-3.502 (2.539)	-3.725 (2.717)	0.977*** (0.0001)	0.979*** (0.0003)	0.987*** (0.0004)	0.985*** (0.0004)
AV_Paragraph		0.978*** (0.145)	1.221*** (0.279)	1.281*** (0.290)		1.006*** (0.001)	1.008*** (0.002)	1.008*** (0.002)
Committe			-35.746* (23.567)	-36.900* (23.525)			0.852 (0.124)	0.846 (0.124)
Exp_contribution			-16.588 (47.814)	2.442 (63.408)			0.785 (0.203)	0.922 (0.275)
Area				7.072 (14.139)				1.025 (0.095)
Population				-0.139 (0.221)				0.999 (0.001)
Constant	9,297.301* (4,750.955)	8,503.597* (4,474.440)	6,795.189 (5,115.807)	7,224.809 (5,464.011)				
Observations	232	232	223	223	232	232	223	223
R ²	0.014	0.161	0.159	0.161				
Adjusted R ²	0.010	0.154	0.144	0.138				

Note:

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

Table A.10: Decreasing marginal quality: $Q = \textit{sentiment}$ and $N = \textit{Tenure}$

	<i>Dependent variable:</i>							
	sentiment							
	A-73 Linear	A-74 Linear	A-75 Linear	A-76 Linear	A-77 Tobit	A-78 Tobit	A-79 Tobit	A-80 Tobit
Tenure	0.017 (0.017)	-0.028* (0.016)	-0.023 (0.019)	-0.024 (0.020)	0.007 (0.021)	-0.041** (0.020)	-0.034 (0.023)	-0.035 (0.023)
AV_sentiment		1.115*** (0.133)	1.230*** (0.201)	1.222*** (0.211)		1.258*** (0.181)	1.411*** (0.249)	1.387*** (0.278)
Committe			-0.194 (0.178)	-0.189 (0.179)			-0.218 (0.171)	-0.213 (0.172)
Exp_contribution			-0.070 (0.335)	0.018 (0.497)			-0.142 (0.315)	-0.068 (0.447)
Area				0.047 (0.115)				0.060 (0.157)
Population				-0.001 (0.002)				-0.0004 (0.002)
Constant	3.331*** (0.253)	-0.892* (0.532)	-1.184* (0.661)	-1.149* (0.663)	3.113*** (0.378)	-1.654** (0.770)	-2.039** (0.895)	-2.004** (0.903)
Observations	232	232	223	223	232	232	223	223
R ²	0.004	0.205	0.195	0.196	0.004	0.205	0.195	0.196
Adjusted R ²	-0.0001	0.198	0.180	0.173	-0.0001	0.198	0.180	0.173

Note:

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

Table A.11: Decreasing marginal quality: $Q = \textit{sentiment}$ and $N = \textit{Sites}_{t-1}$

	<i>Dependent variable:</i>							
	sentiment							
	A-81 Linear	A-82 Linear	A-83 Linear	A-84 Linear	A-85 Tobit	A-86 Tobit	A-87 Tobit	A-88 Tobit
\textit{Sites}_{t-1}	0.005 (0.014)	-0.021 (0.013)	-0.019 (0.016)	-0.024 (0.019)	-0.003 (0.018)	-0.031* (0.017)	-0.026 (0.019)	-0.033 (0.021)
AV_sentiment		1.077*** (0.134)	1.216*** (0.201)	1.153*** (0.224)		1.203*** (0.174)	1.391*** (0.249)	1.292*** (0.283)
Committe			-0.209 (0.176)	-0.204 (0.177)			-0.240 (0.171)	-0.234 (0.171)
Exp_contribution			-0.071 (0.344)	0.051 (0.511)			-0.156 (0.315)	-0.037 (0.449)
Area				0.109 (0.135)				0.142 (0.169)
Population				-0.001 (0.002)				-0.0004 (0.002)
Constant	3.486*** (0.272)	-0.764 (0.546)	-1.098 (0.671)	-0.997 (0.687)	3.282*** (0.393)	-1.466* (0.777)	-1.928** (0.913)	-1.802* (0.928)
Observations	232	232	223	223	232	232	223	223
R ²	0.001	0.203	0.195	0.197	0.001	0.203	0.195	0.197
Adjusted R ²	-0.004	0.196	0.180	0.175	-0.004	0.196	0.180	0.175

Note:

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

Table A.12: Decreasing marginal quality: $Q = \textit{sentiment}$ and $N = \textit{Year}$

	<i>Dependent variable:</i>							
	A-89 Linear	A-90 Linear	A-91 Linear	sentiment		A-94 Tobit	A-95 Tobit	A-96 Tobit
				A-92 Linear	A-93 Tobit			
Year	-0.031* (0.018)	-0.026 (0.016)	-0.022 (0.019)	-0.024 (0.021)	-0.047** (0.023)	-0.039* (0.020)	-0.032 (0.023)	-0.034 (0.024)
AV_sentiment		1.012*** (0.137)	1.166*** (0.222)	1.182*** (0.221)		1.109*** (0.169)	1.318*** (0.258)	1.330*** (0.280)
Committe			-0.203 (0.177)	-0.198 (0.178)			-0.231 (0.171)	-0.226 (0.172)
Exp_contribution			-0.098 (0.326)	0.034 (0.514)			-0.184 (0.305)	-0.051 (0.456)
Area				0.025 (0.114)				0.029 (0.157)
Population				-0.001 (0.002)				-0.001 (0.002)
Constant	66.020* (35.617)	51.585 (32.636)	42.059 (37.921)	46.774 (42.341)	97.555** (45.446)	76.043* (40.783)	60.870 (46.001)	65.603 (48.290)
Observations	232	232	223	223	232	232	223	223
R ²	0.012	0.203	0.194	0.195	0.012	0.203	0.194	0.195
Adjusted R ²	0.008	0.196	0.180	0.173	0.008	0.196	0.180	0.173

Note:

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

Chapter 3

Is there a pro-European bias in the selection of UNESCO sites? A new test based on colonial sites¹

Contents

1	Introduction	84
2	Literature Review	86
3	Nomination and evaluation process	88
4	Empirical Strategy	90
5	Results of the baseline estimates	95
6	Robustness checks	98
7	Conclusion	119
	References	121

¹This chapter is an adapted version of a working paper coauthored with Proff Fabio Padovano and Yvon Rocaboy

”È bello qualcosa che, se fosse nostro, ci rallegrerebbe, ma che rimane tale anche se appartiene a qualcun altro.”

Storia della bellezza, Umberto Eco

1 Introduction

As of 2022, Europe has inscribed 545 sites in the World Heritage List (WHL), almost half of the world total (1154). Several public choice and cultural economics scholars have pointed out such an (alleged) disparity, claiming that Europe is over-represented with respect to the other continents. Two reasons are usually brought forward: first, the superior lobbying power of European member countries (Steiner and Frey, 2012, Frey et al., 2013, Bertacchini et al., 2017, 2016); second, their soft power in imposing aesthetic and cultural standards (Meskell, 2002, Musitelli, 2002). For example, the Convention is said to be written from a “white, middle-class, male’s perspective [...] which favours Western ideas of heritage” (van der Aa, 2005). Such issues would allegedly “bias” the selection of the sites to be inserted in the WHL.

The analysis of this chapter verifies whether this issue receives empirical support, by exploiting a natural experiment. We focus the analysis of the distribution of WHL on a restricted set of countries whose colonial history allows the sites to be either indigenous or European-like (colonial). To verify if either type has an advantage in the process of inscription in the WHL, we consider the evaluations of the candidates to the list that the experts of the International Council on Monuments and Sites (ICOMOS, an independent NGO) deliver to the UNESCO Committee. By that, we focus on a stage of the selection process where the effects of lobbying are minimal. Indeed, lobbying is more likely to appear during the ensuing stage of the Committee’s decision, where the delegates of the member countries play the main role (Bertacchini and Saccone, 2012, Bertacchini et al., 2016). The analysis thus examines the possible biases in the evaluation of sites due to aesthetic and cultural criteria allegedly more favourable for European sites. This aspect has never been empirically examined in the literature so far.

As in Chapter 1, we consider the quality of sites, instead of quantity, because it makes the assessment of the experts’ evaluation of heritage more informative

and avoids the serious shortcomings of using the number of sites as an indicator of the “European selection bias”, the method usually adopted in the literature so far. A non-European country may in fact have a lower number of sites enlisted in the WHL not because the UNESCO has a pro-European bias, but because its cultural capital is limited or it went destroyed over the centuries. In our approach, instead, we start by assembling a new dataset from 2007 to 2016, which includes all the countries that have experienced a period of European colonization. We compare the estimated quality of both the colonial and the indigenous (pre-colonial) sites. If colonial sites receive systematically better evaluations, we interpret it as evidence that European aesthetic and cultural values are favoured in the ICOMOS evaluations. If the opposite result comes out, we take it as evidence of a desire to promote native heritage. If neither type receives systematically better evaluations, we conclude that the ICOMOS experts do not have systematic biases in terms of European (or non-European) aesthetic and/or cultural values. This is indeed the evidence that emerges from our analysis.

We develop a comprehensive analysis to account for the main issues that could lead to non-significant correlation other than the absence of cultural bias. First, to ensure that the results of the analysis are not driven by any measurement error, we compute two measures of the perceived quality of the sites, both based on the ICOMOS evaluations. The first indicator, *Criteria*, is calculated as the number of criteria of Outstanding Universal Value assigned in the evaluation (as in Chapter 1); the second, *Eval_doc* is constructed using a textual analysis of the evaluation documents. Since the hypothesis under test is intuitive and not directly supported by a theoretical model, we corroborate the empirical findings through a battery of robustness tests, each identifying factors that potentially affect the preferences of the ICOMOS experts, such as the type of site, its geographic location, its cultural proximity to the origins of the evaluator, and more. The search for all the possible sources of bias in the preferences of the ICOMOS experts is the goal of the empirical strategy of our analysis.

The rest of the paper is organized as follows: Section 2 includes the review of the literature; in Section 3 we recall the process of the inscription in the WHL and describe the ICOMOS procedure of evaluation; in Section 4 the empirical strategy

is described. Section 5 discusses the results of the baseline analysis; while Section 6 presents the robustness checks. Finally, in Section 7 we draw the main conclusions.

2 Literature Review

The issue of the European “over-representation” is a leitmotiv in the literature regarding the UNESCO WHL. Public choice and cultural economic scholars have stressed that such alleged imbalance raises a question of inequality between member countries (Steiner and Frey, 2012, Bertacchini et al., 2010). The main argument is that all countries that signed the UNESCO Convention should have the same importance in the WHL in purely quantitative terms (Steiner and Frey, 2012, Frey et al., 2013). To prove that UNESCO has pro-European bias, Frey et al. (2013) compare the existing distribution with an unobservable balanced one, which is assumed to be proportional to some indicators, such as population, area and GDP per capita. Steiner and Frey (2012) for instance assume that the number of heritage sites in the list should be correlated with the country’s population, as each person is assumed to have the same capacity to produce cultural goods. However, one can criticize this assumption as not relying on a coherent theoretical approach: as in Cheng (2006) the production of cultural goods actually depends on a “cultural atmosphere” that the consumer-artist absorbs; where cultural goods are more abundant, people tend to be more creative and are more likely to produce new cultural capital. Even the hypothesis that World Heritage sites should be proportioned to the country’s area, despite being more reasonable as far as natural sites are concerned, is not convincing for cultural capital, since its existence does not need a large territory. Moreover, countries that are small today, such as the Holy See or Austria, were quite large at the time of the creation of the cultural capital.

This literature attributes deviations from the hypothetically “fair” distribution of sites to rent-seeking activities implemented by the countries represented in the UNESCO Committee, where the rent is the increase in tourism determined by the inclusion in the WHL of as many sites as possible (Bertacchini and Saccone, 2012, Bertacchini et al., 2017, 2016, Frey et al., 2013). This view is so widely accepted that some authors have started to talk about “world heritage brand” as a tourism advertising strategy (Labadi and Long, 2010). Moreover, the UNESCO recognition

of a national heritage site is an opportunity to reinforce the country's international prestige, improve its public image and build up soft power inside the UNESCO itself (Labadi and Long, 2010).

Some scholars have also pointed out that discrepancies emerge not only with respect to the geographical location but also to the type of heritage that the site represents. Historic towns and cities, religious monuments, and 'elitist' architecture dominate the list, while non-western and traditional cultures seem to be neglected (Bertacchini et al., 2010). Even when new categories of world heritage have been created specifically to extend the representation of countries with a smaller stock of cultural capital, e.g. the "cultural landscapes", European countries have made extensive use of these too (Siririsak and Akagawa, 2007). Again, this alleged fault is attributed to the euro-centrism of the cultural values and the aesthetic categories embedded in the UNESCO Convention. Such discrepancies may prove a hidden but systematic obstacle for non-Western countries to have their potential sites included into the list, although it must be noted that these countries have not shown the intention to change the categories of cultural heritage for a long time (Fowler, 2003). The same European bias argument is used to justify the lack of "negative heritage sites" (Meskell, 2002, Rico, 2008, Trifu, 2018), i.e., sites that testify the vileness of humanity, such as Auschwitz and the Gorée Island, which are often associated with European history.

Member countries, especially non Western ones, have also resorted to the argument of a European cultural bias in the evaluation criteria to sustain positions that openly diverge from the suggestions of ICOMOS experts. These experts have often been accused to be mostly white and European (Meskell, 2012), and to shape their evaluations in favour of European architecture (Siririsak and Akagawa, 2007). Some scholars point out that in recent times the UNESCO Committee have increasingly reverted the recommendations of the ICOMOS experts and that inscriptions into the WHL are increasingly determined by political alliances that serve the goals of emergent nations, like the BRICS. (Meskell, 2012, Meskell et al., 2015, Bertacchini et al., 2017, 2016). Such overturns have been interpreted as a critical, postcolonial

act against an organization perceived to be European-led (Meskell, 2012)². In addition, in the last 15 years, the diversity of the list has significantly increased, as the newly inscribed sites have been mostly industrial or agricultural. All these can be taken as signs of an anti-European attitude toward the UNESCO. Be that as it may, an analysis of whether ICOMOS experts' have specific cultural preferences is missing – and wanting - in the literature.

3 Nomination and evaluation process

3.1 Nomination process

Recalling Section 3.1, the selection process to enter the WHL in which the ICOMOS evaluation is included consists of four steps: 1) setting up of the “tentative list”, 2) the nomination, 3) the evaluation of the nomination by the experts and 4) the final decision of the UNESCO Committee. Only the first three steps are relevant to the scope of this chapter and will be described in detail. First, the “tentative list” is an “inventory” of important natural and cultural heritage sites located within a country. It provides a forecast of the attractions that a country may decide to submit for inscription in the next five to ten years. The country is the only actor responsible for the choice of including a specific site in such a list; the redaction of this list is mandatory before its nomination.

Secondly, the country can plan when to file a nomination. By that, the country should demonstrate that the proposed heritage meets the minimum requirement for the inscription, including the satisfaction of a certain number of criteria.

²For example, delegates during the 2013 meeting from Brazil and India suggested changing the Operational Guidelines, bringing in independent evaluators, reducing the role of the advisory bodies such as the ICOMOS and inscribing a larger quota of sites per country.

3.2 ICOMOS evaluation process

The participation of two separate groups of experts characterizes the ICOMOS evaluation process. The first group certifies the Outstanding Universal Value of the nominated site, namely whether it deserves to be included in the WHL. To be considered of OUV, a site should receive a positive evaluation concerning a number of features: a comparative analysis (i.e., the site must be original with respect to the sites already in the list), the condition of integrity and authenticity, and a number of criteria that consider the historical relevance, the representativeness, the aesthetic beauty and so on. Experts evaluate separately each feature and assign a synthetic judgement to each of them. At this stage, experts can add any missing criteria that the country was not able to identify but is relevant for the evaluation. An anecdotal example was the evaluation of the Khangchendzonga National Park in 2016: India nominated the property under criterion (iii); while ICOMOS experts consider that the also criterion (vi) was satisfied. Typically, for each characteristic, four alternative judgements are rendered (in decreasing order of satisfaction): “the condition has been met”; “it is partially met” or “met despite limitations”; “it could be met, but it has not been proven yet”; “the condition has not been met”.

The second group of experts evaluates characteristics related to the management of the site, i.e., they verify that boundaries, protection, conservation and management are adequate. A draft appraisal is the synthetic outcome of these two sets of evaluations. It is presented to the ICOMOS World Heritage Panel, which, based on these separate judgements, expresses a recommendation regarding the inscription. Ultimately, this recommendation can be of four types: “inscription”, “referral”, “deferral” or “not to inscribe”. A recommendation for “not inscription” implies that the country cannot propose that site ever again. The “referral” and “deferral” evaluations encourage the country to provide minor changes (in the case of “referral”) or substantial revisions (in the case of “deferral”) and then resubmit the candidature at a later session. Finally, a recommendation of “inscription” implies that the site is evaluated to be meritorious to enter the list as it is.

4 Empirical Strategy

4.1 Dependent variable

A quantitative assessment of the quality of the sites is needed to represent the ICOMOS' preferences with respect to the various types of sites. To this aim, we focus on the characteristics that concur in attesting that the site is, indeed, of OUV. Two measures are actually computed: the first, named *Criteria*, is the number of criteria of OUV that, according to the ICOMOS experts' evaluations, the site satisfies. It has the great advantage of being simple and widely available, but it cannot distinguish between sites that are nearly accepted ("referred" ones) and sites that are nearly rejected ("deferred" ones). The second measure, named *Eval.doc*, exploits the information available in the ICOMOS evaluation document. We associate a score from 3 to 0 to every type of evaluation of the site, i.e. from "the condition has been met" (3 points), to "it is partially met" or "met despite limitations" (2 points); "it could be met, but it has not been proven yet" (1 point); "the condition has not been met" (0 points). We consider the features that are crucial to determine the OUV, i.e., the comparative analysis, the integrity, the authenticity and the criteria. We then linearly aggregate these values in a single score. *Eval.doc* is still based on the hypothesis that the higher the number of the criteria satisfied, the higher the quality of the site; however, it overcomes the limits of *Criteria* as it is more complete and precise. Unfortunately, before 2007 the evaluation document does not explicitly report judgements for each feature, making it impossible to collect comparable data before this date. Therefore, our main dataset contains information from 2007 to 2016 for both indicators.

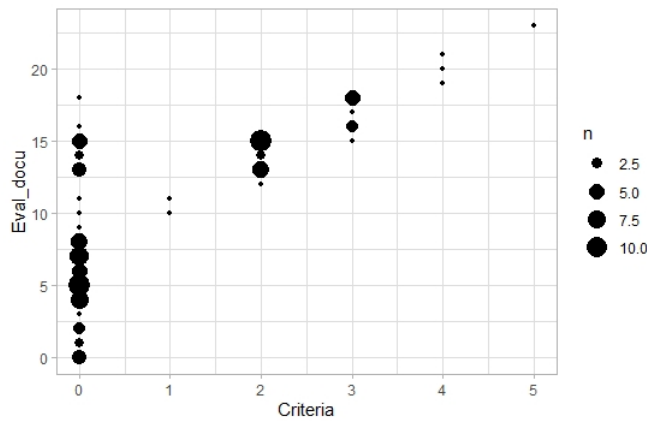
Table 3.1 shows the descriptive statistics for the two measures, *Eval.doc* and *Criteria*. The latter is also analyzed with respect to the subsample of sites inscribed at the first attempt (third column). The last row of the table, where the number of observations is reported, shows that more than 60 % of sites are rejected at first, since only 36 sites out of 99 receive a positive judgement (and consequently satisfy a certain number of criteria) at the first attempt.

Table 3.1: Descriptive Statistics

<i>Eval_doc</i>	<i>Criteria</i>	<i>Criteria - inscribed</i>	<i>Statistics</i>
10.14	0.90	2.47	Mean
5.66	1.30	0.84	SD
32.02	1.68	0.71	Variance
0.00	0.00	1.00	Min
23.00	5.00	5.00	Max
99.00	99.00	36.00	Observations

Figure 3.1 shows that *Eval_doc* and *Criteria* are positively correlated, but *Eval_doc* is more informative, due to its greater variability: for the majority of rejected sites, i.e. those with a value of *Criteria* equal to zero, the value of *Eval_doc* is well below the minimum of the same variable for the inscribed ones. Still, a non-negligible number of sites are rejected with an *Eval_doc* score higher than the mean (between 13 and 15 points), this shows that some sites have shown to meet OUV standards, but still have received a negative recommendation. A possible explanation is that these sites fail to satisfy the standards of conservation, leading to a negative evaluation on the first review.

Figure 3.1: Correlation between *Eval_doc* and *Criteria*



Note: the size of the black dots represents the number of sites with the same combinations of *Eval_doc* and *Criteria*. Most of the rejected sites, i.e. with *Criteria* = 0, are associated with positive values of *Eval_doc* and show high variability.

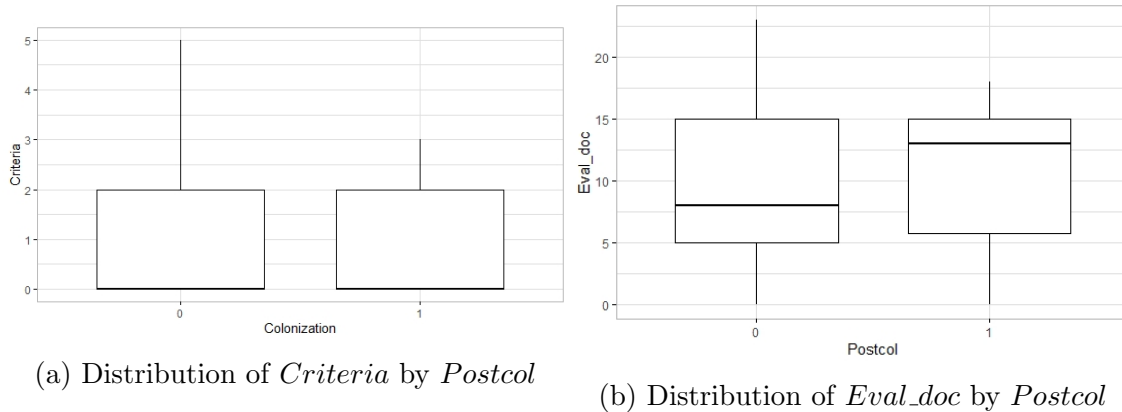
4.2 Independent variable: Colonization

We focus our analysis on non-European countries that were a colony of a European one. This allows for the potential presence of both colonial and native candidate sites in the country and enables the detection of systematic preferences for either type of site by ICOMOS experts. The list of countries included in the sample is reported in Appendix B.1.

Evaluating the degree of colonization of a site is a complex task. For example, there is no doubt that the city of Teotihuacan in Mexico was built before the Spanish invasion, while the many Franciscan missions in the American continents are the work of European architects. However, a city like Rabat, shaped by the French Protectorate but with significant Arabo-Muslim influence, is harder to classify. To deal with this problem, we compute three alternative measures of colonization. The first one, *Postcol*, is constructed using the information contained on the first page of the ICOMOS evaluation document, i.e. a brief historical description of the site. It is a dummy variable with a value of 1 if the site was constructed after the colonial invasion of the country, and 0 otherwise. This indicator is site-specific, as it is computed for all the sites ever considered by UNESCO. Its main advantage is that it shows a within-country variance; the main disadvantage, however, is that it relies on the discretion of the researcher, especially for the sites whose history is partly indigenous and partly colonial and must be therefore classified either way³. As plotted in Figure 3.2a, the maximum number of *Criteria* reached by colonial sites is equal to 3, while it is 5 for native sites. The median, however, is close to 0 for both kinds of sites, showing that there are no differences in the distribution. This intuition is consistent with the Wilcoxon test that confirms no differences in mean (p-value = 0.6768). As in Figure 3.2b, colonial sites seem to have a higher median with respect to *Eval_doc*, but the difference in means is not statistically significant, based on the Wilcoxon test (p-value = 0.7043).

³Recalling the example above, the Franciscan Missions in the Sierra Gorda are considered a colonial site, with a value of 1, while we assign a value of 0 to the city of Rabat, as a large number of buildings were constructed before the French invasion. The opposite is true for contemporary heritage, such as the Universidad Nacional Autónoma de México, which is associated with strictly colonial sites with a value of *Postcol* equal to 1 as well.

Figure 3.2: Distribution of quality by *Postcol*



The second measure, *Domination*, measures the duration of the colonization period; it is computed as the number of years spanning from the first invasion of the Europeans to the date of the country's independence. We assume that the longer the domination, the higher the likelihood that any site in the country is colonial, since destroying the traditional heritage and/or building a new one is a process that requires time. *Domination* has the advantage of reducing the subjectivity of the analyst, but the shortcoming of being country-specific, with a null within-country variance. The correlation between *Domination* and *Criteria* appears slightly negative in Figure 3.3a, while the curve is almost flat when *Domination* is correlated with *Eval_doc* (3.3b). In addition, in both cases the correlation coefficient is not statistically significant (p-value = 0.3127 and p-value = 0.914), suggesting that no differences exist between a native and a colonial site, in terms of evaluations by the ICOMOS' experts.

The third variable, *Distance*, calculated as the distance of the site from the coast as the crow flies, is continuous and different for each site. It is based on the historical evidence that the colonial invaders first arrived from the sea and then penetrated the inner territory encountering a certain resistance; hence the assumption is that the closer is a site to the coastline, the more likely it is that it has been built by colonizers. Again, the correlation between *Distance* and *Criteria* is not statistically significant (p-value = 0.2162), nor is the one between *Distance* and *Eval_doc* (p-value = 0.3312; Figures 3.4a and 3.4b).

Figure 3.3: Correlation between quality and *Domination*

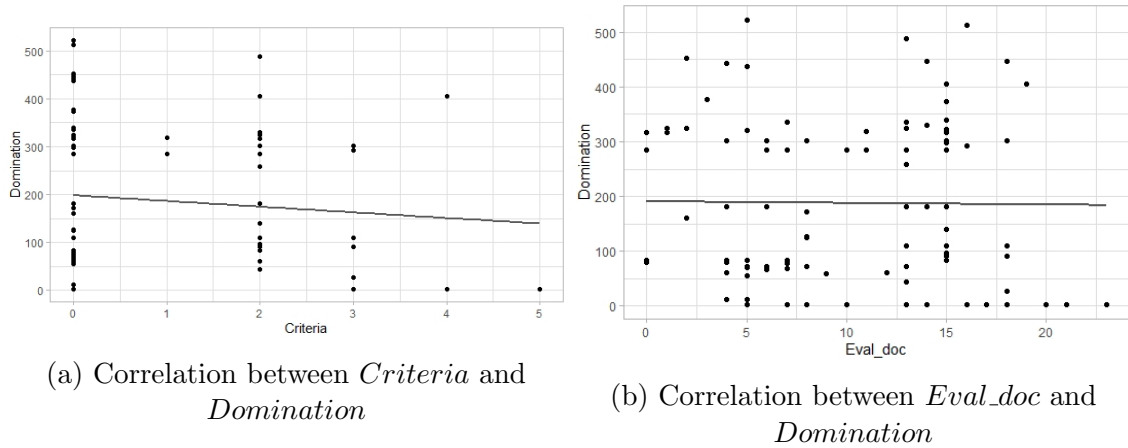
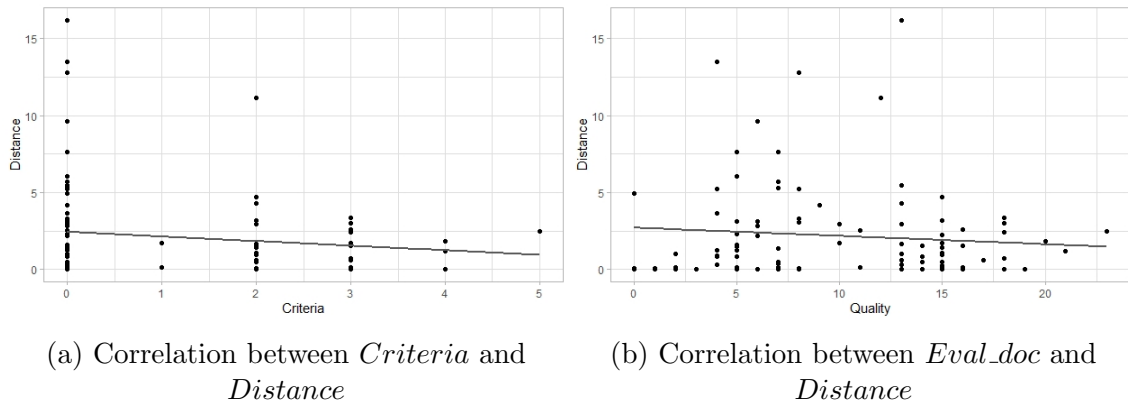


Figure 3.4: Correlation between Quality and *Distance*



Interestingly, The Wilcoxon test confirms that colonial sites identified by *Postcol* have a higher average value for *Domination* (p-value = 4.804e-08) and a lower one for *Distance* compared to native ones. In addition, the probit estimations (Appendix B.2) prove that these two variables are good predictors of *Postcol*. We can therefore conclude that all our measures of colonization are consistent between them.

Equation 3.1 specifies the baseline model to detect the impact of colonization on the perceived quality of the site by the ICOMOS experts:

$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 Year_i + u_i \quad (3.1)$$

This regression is estimated using the two measures of quality (Y_i corresponds to *Criteria* and *Eval.doc*) and the three measures of colonization (the vector *Colonial* includes the variables *Postcol*, *Domination* and *Distance*) described above, on a data set which includes 99 observations. As both *Eval.doc* and *Criteria* are non-negative integers, we consider econometric estimators apt for count variables. Since the variance is greater than the mean for both variables, we assume a negative binomial distribution. We test the model with a year fixed effect ($Year_i$), where the reference year is the first one, 2007.

5 Results of the baseline estimates

The results of the baseline regression of equation 3.1 are reported in Tables 3.2 and 3.3. In line with the previously presented descriptive analysis, the estimated models do not give any evidence of differences in the evaluation of colonial and native sites: *Colonial* is never statistically significant in any of its three alternative specifications. In addition, their p-value is far from the 10%: the average p-value of *Criteria* is 0.4, while it is 0.653 for *Eval.doc*, meaning a pooled average of 0.527. In addition, the magnitude of the estimated coefficients is generally close to zero and the statistical power of the regressions is strong enough to conclude that there are no relevant differences between the treatment and the control group⁴.

The AIC differences analysis, reported in table 3.4, shows that model 5 and model 11, where *Colonial* is proxied by *Distance*, should be preferred, as the AIC is minimum for those models. Following the rule of thumb by Burnham and Anderson (2002), the regression containing *Postcol* and *Domination* are equally correctly specified, since the AIC differences, Δ_i , are less than 2, showing that the results obtained do not rely on econometric biases. The introduction of the year fixed effects does not affect the estimates of the model; on the contrary, it significantly raises the AIC differences to values close to 10. We can then exclude that introducing a year fixed effect improves the quality of our regressions. This is true for both variants of the dependent variables, *Criteria* and *Eval.doc*.

⁴Given the number of observations considered, the statistical power of our regression in the case of a medium effect size (Olivier et al., 2017) is greater than 0.9 for all our predictors.

All in all, the estimates clearly point out that no relation exists between the experts' evaluations of the quality of sites and their being either colonial or native. In other words, we can exclude a pro-European bias in the experts' evaluations of the sampled sites once their quality is accounted for.

Table 3.2: Baseline model. Dependent variable: *Criteria*

	<i>Dependent variable:</i>					
	<i>Criteria</i>					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>Postcol</i>	0.012 (0.380)	-0.244 (0.389)				
<i>Domination</i>			-0.001 (0.001)	-0.002 (0.001)		
<i>Distance</i>					-0.091 (0.067)	-0.093 (0.067)
<i>Constant</i>	-0.110 (0.217)	0.324 (0.481)	0.070 (0.278)	0.567 (0.513)	0.063 (0.215)	0.501 (0.509)
Year FE	NO	YES	NO	YES	NO	YES
Observations	99	99	99	99	99	99
Log-Likelihood	-129.214	-126.311	-128.864	-125.590	-128.401	-125.649
θ	0.494*** (0.166)	0.603*** (0.221)	0.506*** (0.173)	0.633*** (0.235)	0.518*** (0.177)	0.633*** (0.237)
AIC	262.429	274.622	261.728	273.180	260.803	273.298

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3: Baseline model. Dependent variable: *Eval_doc*

	<i>Dependent variable:</i>					
	<i>Eval_doc</i>					
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
<i>Postcol</i>	0.020 (0.135)	-0.073 (0.138)				
<i>Domination</i>			-0.00004 (0.0004)	-0.0002 (0.0004)		
<i>Distance</i>					-0.019 ((0.021)	-0.015 (0.021)
<i>Constant</i>	2.310*** (0.077)	2.442*** (0.180)	2.324*** (0.101)	2.455*** (0.194)	2.356*** (0.077)	2.482*** (0.193)
Year FE	NO	YES	NO	YES	NO	YES
Observations	99	99	99	99	99	99
Log-Likelihood	-315.639	-310.784	-315.646	-310.825	-315.247	-310.696
θ	3.373*** (0.690)	3.941*** (0.857)	3.3729*** (0.690)	3.939*** (0.857)	3.4158*** (0.702)	3.953*** (0.861)
AIC	635.278	643.568	635.291	643.649	634.494	643.392

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.4: AIC differences analysis

<i>Criteria</i>	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6
	1.626	13.819	0.475	12.377	0	12.495
<i>Eval_doc</i>	Δ_7	Δ_8	Δ_9	Δ_{10}	Δ_{11}	Δ_{12}
	0.784	9.074	0.797	9.155	0	8.898

Note: This table reports the AIC differences expressed as $\Delta_i = AIC_i - AIC_{min}$. Such differences estimate the relative expected K-L differences between f and $g_i(x|\theta)$ (Burnham and Anderson, 2002). The model estimated to be best has $\Delta_i = \Delta_{min} = 0$.

6 Robustness checks

6.1 Type of sites

To verify the generality and robustness of these results, we have performed a battery of additional tests, controlling for additional factors that may affect the correlation between Y and *Colonial*, thus influencing the preferences of experts. First, we test if the perception of the quality of the sites by the ICOMOS experts depends on the type of site. The UNESCO Convention establishes three different categories: *Site*, *Buildings* and *Monument*⁵, that we treat as dummy variables. Sites that include two or more unconnected attractions of different types are called *Serial*⁶. *Site*, the most numerous type in the sample, is considered as a benchmark.

In a second specification, we consider a vector of dummy variables that identify the types mentioned in the Operational Guidelines (OG). They comprise a larger set of variants with respect to the Convention: *canal*, *cultural landscape*, *historic town*, *cultural route* and *historic centre*. A large set of nominated sites, however, do not specify the OG type in the ICOMOS evaluation document. They are then classified as *residual*. To complete the information extracted from this document, we consider the name of the site as a signal of the classification. For example, “Rice Terrace Cultural Landscape” is considered as *cultural landscape*, even if it is not specified in the evaluation document. Using the same reasoning, sites such as “Thimlich Ohinga Archaeological Site” allow us to add another specification, the type *archaeological*.

Traditionally, *Buildings* and *Monument* are associated with European architecture, while *cultural landscape* was introduced to favour non-Western traditions. The

⁵Specifically, monuments are architectural works, works of monumental sculpture and painting, elements or structures of an archaeological nature, inscriptions, cave dwellings and combinations of features, which are of outstanding universal value from the point of view of history, art or science; groups of buildings are groups of separate or connected buildings which, because of their architecture, their homogeneity or their place in the landscape, are of outstanding universal value from the point of view of history, art or science; sites are works of man or the combined works of nature and man, and areas including archaeological sites which are of outstanding universal value from the historical, aesthetic, ethnological or anthropological point of view (UNESCO, 1972).

⁶For example, the site “Pearling, Testimony of an Island Economy” in Bahrain includes 4 *Site*, 9 *Monument* and 2 *Buildings*.

distribution of sites is plotted in Figures 3.5a and 3.5b. Contrary to Bertacchini et al. (2010), *Monument* and *Buildings* are far from being dominant; indeed the generic classification *Site* represents almost 3/4 of all nominations; the *cultural landscape* category counts three times more inscription than *historic centre* and *historic town* combined together. This is not surprising, however, since our sample contains only non-European sites. The proportion of sites without any label, classified as *residual*, remains a quantitatively important group and is taken as a reference.

The literature on UNESCO WHL has also pointed out that, as Christianity is one of the main pillars of European history and identity, religious sites could receive a preferred treatment by the ICOMOS experts, thus generating a bias. The sample does not allow controlling for Christian sites, however, as only two observations are classified under this label. Instead, to verify whether a general religious bias affects the evaluation of the quality of a site, the dummy *Religious* is included in the model. It assumes the value of 1 if the word “sacred”, “religious” or the name of a religion is present in the short description of the site, and 0 otherwise.

The regression equations are therefore specified as follows:

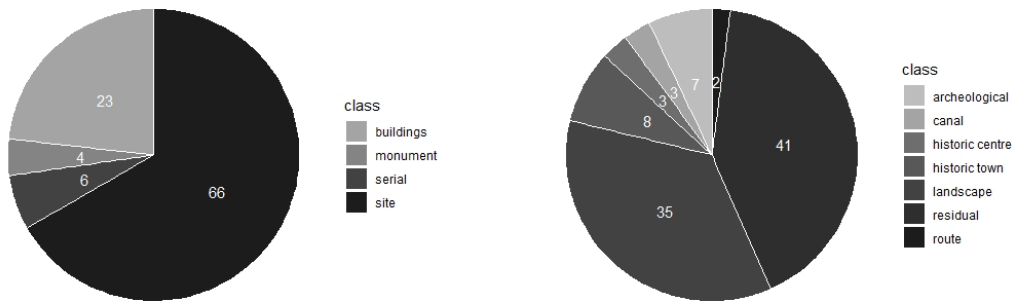
$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 Type_i + u_i \quad (3.2)$$

$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 Religious_i + u_i \quad (3.3)$$

Excluding model 18, where *Distance* is barely significant at the 10% level, *Colonial* is never statistically significant, confirming the result of Section 5. The average p-value of *Colonial* estimated by regression 3.2 is 0.59, similar to our baseline analysis. In addition, the average p-value for *Distance* is 0.2915, lower than the other variables of interest but still far from being significant. The AIC differences suggest that the level of empirical support is considerably lower than our reference model, where equation 3.2 is estimated using *Criteria* as the dependent variable. Still, the models estimated with respect to *Eval_doc* are more reliable, with an average Δ_i of 4, which becomes 3 for models 23 and 24. Therefore, these models are preferred to models 17 and 18.

The type of site is generally not significant, with the only exception of *route* in Table 3.6 that displays a negative sign with a p-value lower than 0.10. However, this fact on its own does not allow us to conclude that there is a systematic disparity in the evaluations, since the effect of all the other types is equal to zero. In the same way, the religious connotation does not seem to have any distinct effect.

Figure 3.5: Distribution by type



(a) Distribution by type (Convention)

(b) Distribution by type (OG)

Is there a pro-European bias in the selection of UNESCO sites?

Table 3.5: Robustness check: Type of site. Dependent variable: *Criteria*

	<i>Dependent variable:</i>					
	<i>Criteria</i>					
	Model 13	Model 14	Model 15	Model 16	Model 17	Model 18
<i>Postcol</i>	0.078 (0.393)	0.051 (0.394)				
<i>Domination</i>			-0.001 (0.001)	-0.001 (0.001)		
<i>Distance</i>					-0.093 (0.067)	-0.115* (0.068)
<i>Buildings</i>	0.089 (0.433)		0.037 (0.417)		0.139 (0.411)	
<i>Monument</i>	0.105 (0.887)		-0.041 (0.893)		0.039 (0.871)	
<i>Serial</i>	-1.035 (0.943)		-0.940 (0.928)		-0.995 (0.906)	
<i>canal</i>		0.936 (0.874)		0.994 (0.857)		0.921 (0.841)
<i>landscape</i>		-0.287 (0.387)		-0.253 (0.387)		-0.282 (0.384)
<i>archeological</i>		-0.588 (0.746)		-0.507 (0.735)		-0.614 (0.746)
<i>historic centre</i>		0.247 (0.942)		0.361 (0.954)		0.719 (0.897)
<i>historic town</i>		-0.738 (0.743)		-0.667 (0.730)		-0.800 (0.710)
<i>route</i>		-26.366 (2.2e+05)		-23.193 (4.9e+04)		-35.366 (2.0e+07)
<i>Constant</i>	-0.118 (0.233)	0.013 (0.279)	0.070 (0.305)	0.140 (0.305)	0.068 (0.254)	0.228 (0.280)
Observations	99	99	99	99	99	99
Log Likelihood	-128.532	-126.246	-128.329	-126.025	-127.697	-124.985
θ	0.515*** (0.176)	0.603*** (0.223)	0.525*** (0.182)	0.615*** (0.230)	0.544*** (0.190)	0.641*** (0.239)
AIC	267.064	268.492	266.657	268.051	265.395	265.971

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.6: Robustness check: Type of site. Dependent variable: *Eval_doc*

	<i>Dependent variable:</i>					
	<i>Eval_doc</i>					
	Model 19	Model 20	Model 21	Model 22	Model 23	Model 24
<i>Postcol</i>	0.019 (0.139)	0.037 (0.141)				
<i>Domination</i>			0.0001 (0.0004)	0.0001 (0.0004)		
<i>Distance</i>					-0.018 (0.021)	-0.021 (0.020)
<i>Buildings</i>	0.065 (0.155)		0.071 (0.149)		0.063 (0.149)	
<i>Monument</i>	0.257 (0.313)		0.267 (0.313)		0.240 (0.310)	
<i>Serial</i>	-0.317 (0.279)		-0.315 (0.278)		-0.326 (0.276)	
<i>canal</i>		0.305 (0.345)		0.316 (0.342)		0.306 (0.339)
<i>landscape</i>		-0.196 (0.139)		-0.202 (0.139)		-0.191 (0.137)
<i>archeological</i>		-0.179 (0.248)		-0.189 (0.248)		-0.185 (0.245)
<i>historic centre</i>		0.235 (0.348)		0.237 (0.351)		0.284 (0.343)
<i>historic town</i>		-0.349 (0.246)		-0.340 (0.241)		-0.357 (0.238)
<i>route</i>		-0.934* (0.509)		-0.919* (0.504)		-0.904* (0.495)
Constant	2.299*** (0.083)	2.401*** (0.101)	2.293*** (0.110)	2.400*** (0.112)	2.346*** (0.091)	2.454*** (0.101)
Observations	99	99	99	99	99	99
Log Likelihood	-314.470	-311.406	-314.470	-311.424	-314.087	-310.883
θ	3.499*** (0.726)	3.862*** (0.833)	3.499*** (0.726)	3.860*** (0.833)	3.543*** (0.739)	3.937*** (0.857)
AIC	638.940	638.811	638.939	638.849	638.175	637.766

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.7: Robustness check: Religion

	<i>Dependent variable:</i>					
	Model 25	<i>Criteria</i> Model 26	Model 27	Model 28	<i>Eval_doc</i> Model 29	Model 30
<i>Postcol</i>	0.071 (0.387)			0.035 (0.138)		
<i>Domination</i>		-0.001 (0.001)			-0.00001 (0.0004)	
<i>Distance</i>			-0.109 (0.071)			-0.022 (0.021)
<i>Religious</i>	0.228 (0.383)	0.183 (0.375)	0.353 (0.381)	0.060 (0.137)	0.051 (0.135)	0.087 (0.137)
<i>Constant</i>	-0.208 (0.267)	-0.003 (0.315)	-0.028 (0.237)	2.285*** (0.094)	2.302*** (0.114)	2.334*** (0.084)
Observations	99	99	99	99	99	99
Log Likelihood	-129.041	-128.744	-127.978	-315.544	-315.576	-315.049
θ	0.500*** (0.169)	0.511*** (0.175)	0.531*** (0.184)	3.383*** (0.693)	3.380*** (0.692)	3.435*** (0.708)
AIC	264.082	263.487	261.955	637.088	637.152	636.099

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Heterogeneity across continents

The distribution of sites is not homogeneous around the world, even when Europe is excluded. Figure 3.6 shows that Asia is the first continent for number of sites, followed by America, Africa and the Arab countries. Moreover, Figure 3.7 shows that not in all continents the ratio of colonial to native sites is the same: the majority of colonial sites are located in Latin America, while they are rare in Arab countries. Such an unequal geographical distribution is the consequence of different colonization campaigns and styles of colonial administration in the various continents. The Spanish exploitation of American resources for instance involved the conversion of the Indian population; thus, great monasteries were established in the conquered districts; in Africa instead, colonizers concentrated more on the extraction of resources and what they constructed was less likely to become potential cultural heritage (Bolton, 1917). An appearance of a pro-European bias in Latin America might, therefore, arise only because there are relatively more colonial sites. In general, these historically generated asymmetries must be taken into consideration, as it might be the case that no systematic differences in evaluation emerge at the global level, but a bias might exist in continents where the number of colonial and native sites is similar. Estimating equation 3.1 separately for each continent would however lead to inefficient estimates, as the statistical power would be insufficient. We, therefore, introduce dummy variables for continents and interaction terms between those dummies and the colonial variables. The aim is to disentangle the average quality of the continent's sites, which is captured by the dummies, from the impact of *Colonial* on quality, which is captured by the interaction terms.

The following equation is estimated:

$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 Continent_i + \beta_3 Colonial_i * Contints_i + u_i \quad (3.4)$$

where *Continent* is a vector of dummy variables that identify the continent of the site. Asia, the continent with the higher value for Y and the largest number of sites, is taken as a benchmark.

Table 3.8 reports the estimates of equation 3.4. In the majority of the regressions, *Colonial* is not significant; yet model 32 seems to suggest that countries with a

longer *Domination* present also sites that receive a lower Y today. In other words, native sites seem to be preferred. The AIC of this regression, however, is far from the optimal value (260.803 in Model 5), to the point that we can exclude that this model interprets the distribution of our data better than the baseline analysis. On the contrary, Δ_{34} , Δ_{35} and Δ_{36} lay in an acceptable range. Models 34, 35 and 36, indeed, give some evidence that African and Arabic sites are generally undervalued with respect to their Asian counterparts, since the continent's dummy variables are negative. African sites, in particular, show a negative and highly significant coefficient, which likely points out that African countries fail to efficiently promote their proposed sites, regardless of their being colonial or indigenous. American nominations, on the contrary, seem to receive better evaluations, since the continent dummy variable is positive. None of the interactive terms can be considered different from zero. This confirms that, even when each continent is examined separately, no pro-European bias emerges. If a difference in experts' perception exists with respect to continents, it does not rely on the greater availability of "European-like" types of sites and heritage.

Figure 3.6: Distribution of sites by continents

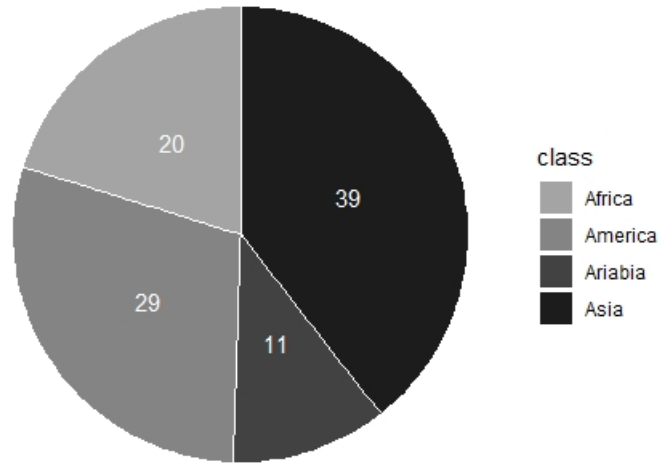


Figure 3.7: Geographical distribution of the sites



Table 3.8: Robustness check: Continents

	<i>Dependent variable:</i>					
	<i>Criteria</i>				<i>Eval_doc</i>	
	Model 31	Model 32	Model 33	Model 34	Model 35	Model 36
<i>Postcol</i>	-0.386 (0.826)			0.259 (0.272)		
<i>Domination</i>		-0.004** (0.002)			-0.0003 (0.001)	
<i>Distance</i>			-0.109 (0.098)			-0.012 (0.028)
<i>Africa</i>	-0.738 (0.547)	-1.327** (0.672)	-0.748 (0.624)	-0.391** (0.184)	-0.607*** (0.233)	-0.249 (0.209)
<i>America</i>	0.163 (0.735)	-0.677 (1.952)	-0.301 (0.515)	0.127 (0.262)	0.021 (0.757)	0.354* (0.187)
<i>Arabia</i>	-0.951 (0.662)	-0.791 (0.890)	-0.614 (0.812)	-0.395* (0.211)	-0.343 (0.295)	-0.395 (0.256)
<i>Africa * Postcol</i>	0.268 (1.360)			-0.003 (0.432)		
<i>America * Postcol</i>	0.386 (1.125)			-0.373 (0.386)		
<i>Arabia * Postcol</i>						
<i>Africa * Domination</i>		0.005 (0.003)			0.002 (0.001)	
<i>America * Domination</i>		0.004 (0.006)			-0.001 (0.002)	
<i>Arabia * Domination</i>		-0.010 (0.0171)			-0.003 (0.004)	
<i>Africa * Distance</i>			-0.04 (0.149)			-0.037 (0.044)
<i>America * Distance</i>			0.069 (0.241)			0.069 (0.083)
<i>Arabia * Distance</i>			-0.640 (0.873)			-0.043 (0.115)
<i>Constant</i>	0.163 (0.2781)	0.545 (0.336)	0.262 (0.187)	2.463*** (0.100)	2.546*** (0.130)	2.529*** (0.114)
Observations	99	99	99	99	99	99
Log Likelihood	-252.835	-248.168	-250.345	-311.073	-310.565	-310.407
θ	0.562*** (0.201)	0.672*** (0.262)	0.591*** (0.212)	3.918*** (0.852)	3.976*** (0.869)	3.996*** (0.875)
AIC	268.830	266.170	268.340	636.147	637.130	636.814

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Experts' personal bias

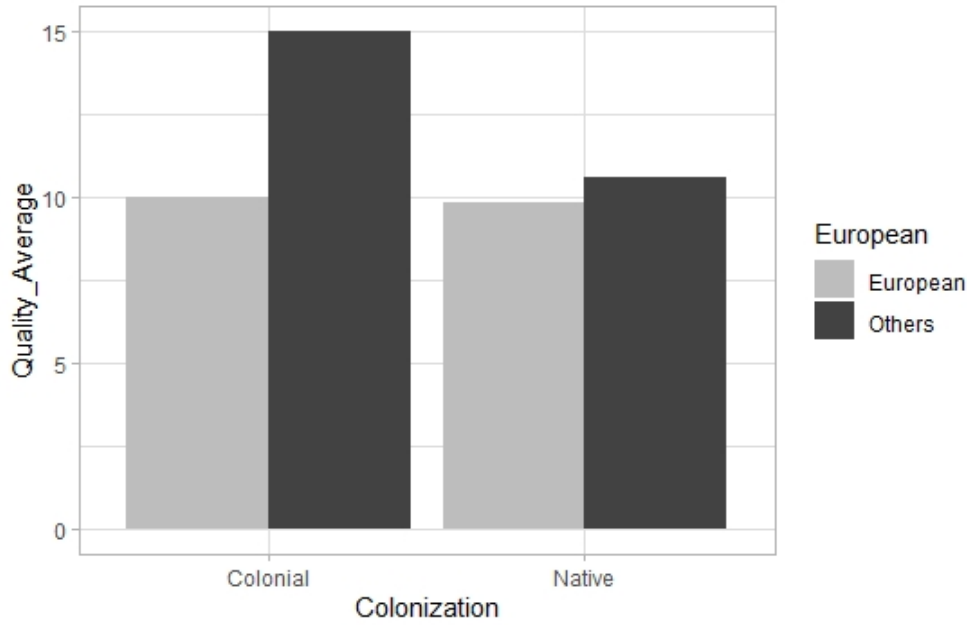
So far the estimates have failed to detect any systematic differences in the evaluation of the quality of native and colonial (European) sites by the experts of ICOMOS. Yet we have not accounted for possible biases that the personal cultural background of the individual expert may cause. As experts come from all around the world, the lack of bias emerging from the estimates presented so far may be due either to the fact that ICOMOS indeed matches experts and sites so to avoid “pro-local” preferences; alternatively, it may be due to the impartiality of the experts themselves. Unfortunately, no information is available about the process of selection and assignment of experts to the site subject to review. The only information available in the evaluation document is the name of the expert and the country in charge of its identification. We exploit this information to verify whether the expert’s background somehow affects the evaluation of the site.

Specifically, we test the hypothesis that experts chosen by European countries evaluate colonial sites more favourably than native ones, because of a conflict of interest and/or culturally generated pro-European bias. To this end, we develop two proxies: the first, *Eu_dummy* is a qualitative variable equal to 1 if the selected expert is from a European country and 0 otherwise. Figure 3.8, where the quality of sites is proxied by *Eval_doc*, provides some preliminary insights on that variable: non-European selected experts tend to award higher scores, than European ones and this effect is larger for colonial sites. There is no evidence, on the other hand, that European experts privilege colonial sites since the average quality assessment that they render is close to 10 for both colonial and native sites.

Eu_dummy has the advantage of being simple and straightforward, but it does not take into account factors such as immigration, language proximity and cultural roots of the individual expert. For example, a French expert with Cameroonian parents could be “culturally close” to the African culture. We then develop a second proxy, *Eu_surname*, which is a continuous variable reporting the relative diffusion of the expert’s surname in Europe, i.e., how many European citizens carry the expert’s surname with respect to people with the same surname in the rest of the world. This second variable is especially useful to verify whether a European bias might emerge,

for instance, in the case of a U.S. expert of Spanish descent is assigned to evaluate a colonial Mexican site.

Figure 3.8: The impact of a European mission on quality



Following the same reasoning, two are the proxies taking into account the experts' "cultural proximity" to a site: *Prox_dummy* is computed as a dummy variable with the value equal to 1 if the nation in charge of the mission belongs to the continent where the site is located; while *Prox_surname* is the percentage of the incidence of the expert surname in the same continent of the site, i.e., how many people have the expert's surname in the continent where the site is located⁷.

The econometric models that control for the experts' origin are the following:

⁷To exemplify the difference between those variables, consider this case taken from the dataset: the Mexican delegation asked to Dr. Niklas Schulze to evaluate a pre-Columbian site in Costa Rica. Then *Eu_dummy* is equal to 0, since the country in charge of the evaluation is non-European; *Eu_surname* is instead equal to 0.88, since "Schulze" is likely to be a European surname; *Prox_dummy* is equal to 1, since Mexico and Costa Rica belong to the same geographical area; finally, *Prox_surname* is equal to 0.11, as the cultural proximity between the authors and the site is likely to be low, based on the local frequency of the expert's surname.

$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 European_i + \beta_3 Colonial_i * European_i + u_i \quad (3.5)$$

$$Y_i = \beta_0 + \beta_1 Colonial_i + \beta_2 Proximity_i + u_i \quad (3.6)$$

They are estimated for *Criteria* and *Eval_doc* and all the colonial variables. An interaction term is added to capture the combined effect of being European and evaluating a colonial site. Since two out of three colonial variables and one European are continuous, we construct the regression in such a way that the interaction always contains one continuous variable and one dummy.

Tables 3.9, 3.10 and 3.11 report the estimates of equations 3.5 and 3.6. Again, *Colonial* is not significant with an average p-value of 0.57, coherently with the previous tests; only in model 45, *Distance* is negative and significant at a 10% level. There is no systematic pro-European bias since no one interaction term in Table 3.9 is statistically significant; this is true for both *Criteria* and *Eval_doc*. On the other hand, cultural proximity has no impact on *Criteria*, but shows a negative sign in Table 3.11. This suggests that experts in fact tend to award lower scores to sites from their own area. This is quite the opposite to what the “pro-European bias” hypothesis would lead us to expect and it would suggest that the competition for listing would be local rather than global. Model 51 shows the lowest AIC value of all the estimates that consider *Eval_doc* as dependent variable, meaning it should be preferred to the baseline model.

Table 3.9: Robustness check: Experts' personal bias

	<i>Dependent variable:</i>					
	<i>Criteria</i>				<i>Eval.doc</i>	
	Model 37	Model 38	Model 39	Model 40	Model 41	Model 42
<i>Postcol</i>	-0.320 (0.578)			-0.043 (0.195)		
<i>Domination</i>		-0.0001 (0.001)			0.0002 (0.0005)	
<i>Distance</i>			-0.063 (0.079)			-0.011 (0.025)
<i>Eu_surname</i>	0.728 (0.603)			0.316 (0.218)		
<i>Eu_dummy</i>		0.633 (0.566)	0.900* (0.521)		0.196 (0.211)	0.220 (0.189)
<i>Interaction1</i>	1.180 (1.363)			0.300 (0.493)		
<i>Interaction2</i>		-0.002 (0.004)			-0.001 (0.001)	
<i>Interaction3</i>			-0.169 (0.167)			-0.038 (0.045)
<i>Constant</i>	-0.403 (0.305)	-0.224 (0.362)	-0.132 (0.240)	2.196*** (0.106)	2.247*** (0.129)	2.312*** (0.086)
Observations	98	99	99	98	99	99
Log Likelihood	-124.871	-128.227	-127.024	-310.528	-315.195	-314.584
θ	0.547*** (0.198)	0.534*** (0.187)	0.566*** (0.201)	3.545*** (0.745)	3.421*** (0.704)	3.484*** (0.721)
AIC	257.743	264.454	262.047	629.056	638.390	637.169

Table 3.10: Robustness check: Experts' personal bias

	<i>Dependent variable:</i>					
	<i>Criteria</i>					
	Model 43	Model 44	Model 45	Model 46	Model 47	Model 48
<i>Postcol</i>	0.141 (0.401)			0.213 (0.390)		
<i>Domination</i>		-0.001 (0.001)			-0.001 (0.001)	
<i>Distance</i>			-0.118* (0.071)			-0.097 (0.067))
<i>Prox_surname</i>	-0.583 (0.477)	-0.477 (0.470)	-0.682 (0.462)			
<i>Prox_dummy</i>				-0.616 (0.395)	-0.497 (0.391)	-0.576 (0.376)
<i>Constant</i>	0.085 (0.278)	0.198 (0.319)	0.386 (0.320)	0.236 (0.315)	0.325 (0.345)	0.457 (0.347)
Observations	98	98	98	99	99	99
Log Likelihood	-126.091	-126.030	-124.881	-128.013	-128.057	-127.212
θ	0.496*** (0.171)	0.498*** (0.171)	0.541*** (0.193)	0.541*** (0.190)	0.539*** (0.190)	0.574*** (0.208)
AIC	258.182	258.060	255.761	262.027	262.115	260.424

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.11: Robustness check: Experts' personal bias

	<i>Dependent variable:</i>					
	<i>Eval_doc</i>					
	Model 49	Model 50	Model 51	Model 52	Model 53	Model 54
<i>Postcol</i>	0.079 (0.139)			0.068 (0.140)		
<i>Domination</i>		0.0001 (0.0004)			0.0001 (0.0004)	
<i>Distance</i>			-0.028 (0.021)			-0.021 (0.021)
<i>Prox_surname</i>	-0.283* (0.165)	-0.273* (0.164)	-0.310* (0.163)			
<i>Prox_dummy</i>				-0.172 (0.145)	-0.159 (0.145)	-0.162 (0.140)
<i>Constant</i>	2.415*** (0.099)	2.413*** (0.114)	2.511*** (0.115)	2.417*** (0.119)	2.414*** (0.130)	2.474*** (0.131)
Observations	98	98	98	99	99	99
Log Likelihood	-311.060	-311.181	-310.340	-314.942	-315.041	-314.571
θ	3.484*** (0.727)	3.469*** (0.722)	3.569*** (0.752)	3.448*** (0.712)	3.438*** (0.709)	3.492*** (0.725)
AIC	628.120	628.363	626.680	635.884	636.083	635.142

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6.4 ICOMOS recommendation

Recalling Figure 3.1, in our sample 14 out of 99 sites were rejected despite a value of *Eval_doc* higher than the sample mean. A possible reason could be that this variable fails to include all the relevant information for the decision of inscription. In other words, we have been able to exclude a pro-European bias at the stage of the evaluation of the sites. Still, something might happen at the stage of the recommendation of being inscribed or not, which might result in a bias. In order to be able to exclude that, we should check that colonial sites have the same probability of receiving a recommendation of being “inscribed” as native sites, holding constant the evaluation received. If that is not the case, a pro-European bias might arise. In addition, we analyze the probability that a site is “deferred”, i.e. its OUV is not demonstrated at that stage. Sites that receive this recommendation could still be included in the list, but this would require more effort than “inscribed” or “referred back” nominations.

Inscribed and *Deferred* are treated as dummy variables and substitute Y as the dependent variable. Figures 3.9 and 3.10 report some descriptive statistics. First, most of the deferred sites are native, while the majority of sites inscribed are colonial (Figure 3.9). Moreover, deferred sites have a higher value for *Distance* and a lower one for *Domination* with respect to the sites inscribed, meaning that a native site is more likely to be deferred than a colonial one. The standard errors, however, are too large to support any conclusion (Figure 3.10) and regression analysis is required.

The estimated model for ICOMOS recommendation takes the following form:

$$Y_i = \beta_0 + \beta_1 Eval_doc_i + \beta_2 Colonial_i + \beta_3 Year_i + u_i \quad (3.7)$$

Where Y_i represents the ICOMOS recommendations *Inscribed* and *Deferred*. Since they are both dichotomous variables, they are estimated using a probit model. Equation 3.7 verifies if, for the same level of quality, a difference exists in the probability of inscription for colonial and native sites. As long as no correlation exists between *Eval_doc* and *Colonial*, we do not expect that the specification of the regression introduces any bias. Conversely, we cannot estimate the same equation by

adding *Criteria* as a control variable because, by definition, it is equal to 0 when a site is rejected, which would induce an orthogonality issue.

Table 3.12 shows the estimates of equation 3.7, when Y_i corresponds to inscribed sites. The results show that *Eval_doc* is positively correlated with *Inscribed*, which confirms that sites that receive a better evaluation are more likely to be inscribed. *Postcol*, *Domination* and *Distance* are never statistically significant, even after controlling for the year fixed effects. Models 55, 56 and 57 should be preferred as they have the lowest AIC in this class of models.

Table 3.13 instead shows the estimates of equation 3.7, when Y_i corresponds to deferred sites. It also shows that, as expected, the higher the quality, the lower the probability that a site is deferred. *Postcol* is negative and statistically significant, showing that native sites are more likely to be deferred, with respect to colonial ones. This result, however, is not confirmed by *Distance* and *Domination*.

Figure 3.9: Site distribution by recommendation and type

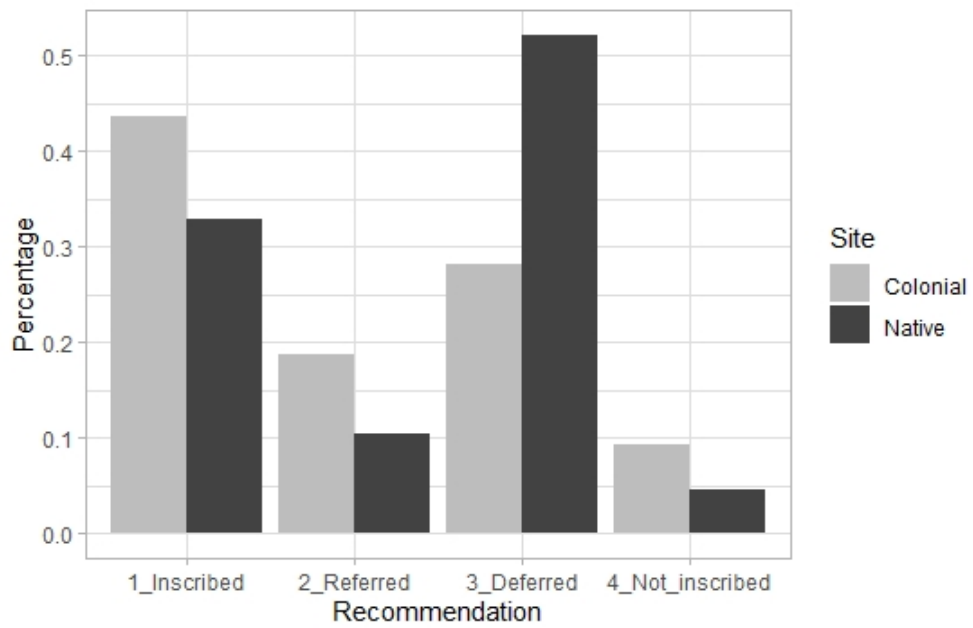
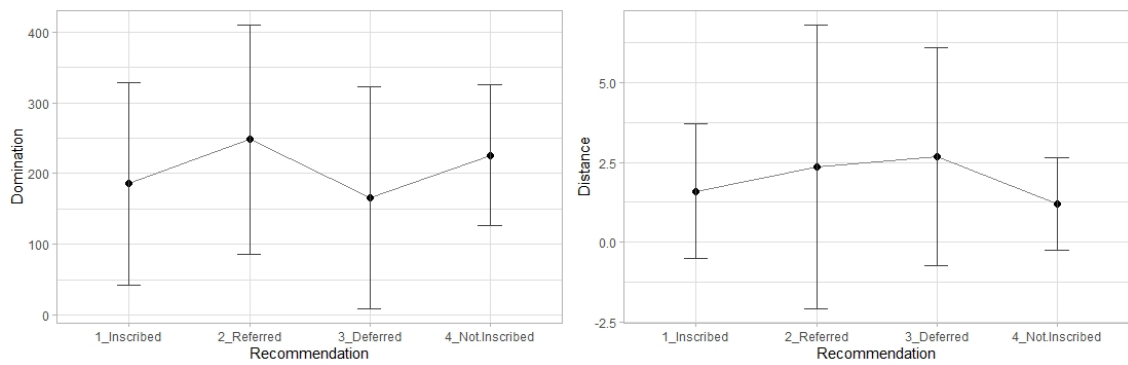


Figure 3.10: Average *Colonial* by recommendation



(a) Average *Domination* by recommendation

(b) Average *Distance* by recommendation

Is there a pro-European bias in the selection of UNESCO sites?

Table 3.12: Robustness check: recommendation. Dependent variable: *Inscribed*

<i>Dependent variable:</i>						
<i>Inscribed</i>						
	Model 55	Model 56	Model 57	Model 58	Model 59	Model 60
<i>Eval_doc</i>	0.296*** (0.056)	0.315*** (0.060)	0.300*** (0.057)	0.386*** (0.083)	0.408*** (0.088)	0.386*** (0.085)
<i>Postcol</i>	0.156 (0.381)			0.236 (0.460)		
<i>Domination</i>		-0.001 (0.001)			-0.001 (0.002)	
<i>Distance</i>			-0.001 (0.059)			-0.053 (0.069)
<i>Constant</i>	-3.918*** (0.750)	-3.850*** (0.769)	-3.910*** (0.814)	-4.345*** (1.098)	-4.666*** (1.239)	-4.193*** (1.320)
Year fix effect	NO	NO	NO	YES	YES	YES
Observations	99	99	99	99	99	99
Log Likelihood	-31.096	-30.743	-31.181	-26.357	-26.036	-25.938
AIC	68.192	67.486	68.363	74.713	76.073	75.876

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.13: Robustness check: recommendation. Dependent variable: *Deferred*

	<i>Dependent variable:</i>					
	<i>Deferred</i>					
	Model 61	Model 62	Model 63	Model 64	Model 65	Model 66
<i>Eval_doc</i>	-0.206*** (0.035)	-0.196*** (0.033)	-0.198*** (0.033)	-0.208*** (0.036)	-0.196*** (0.034)	-0.198*** (0.036)
<i>Postcol</i>	-0.961** (0.376)			-1.028** (0.421)		
<i>Domination</i>		-0.001 (0.001)			-0.001 (0.001)	
<i>Distance</i>			0.059 (0.048)			0.077 (0.053)
<i>Constant</i>	2.069*** (0.383)	1.981*** (0.397)	1.587*** (0.360)	1.998*** (0.511)	1.498*** (0.632)	0.917 (0.654)
Year fix effect	NO	NO	NO	YES	YES	YES
Observations	99	99	99	99	99	99
Log Likelihood	-38.303	-40.954	-41.144	-37.569	-39.495	-39.250
AIC	82.606	87.909	88.288	97.139	102.991	102.500

Standard errors in parentheses. Significant levels are: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7 Conclusion

This study contributes to the existing political economy literature about the UNESCO by empirically investigating the process of evaluation and inscription of sites into the UNESCO WHL using two proxies of the quality of sites that have never been investigated before: the number of criteria of OUV assigned to each site and an index based on the textual analysis of evaluation documents. We focus the analysis on a sample where the alleged bias is most likely to emerge, i.e. countries that are former European colonies where both European-like and indigenous sites are present. The paper critically revises the claim of European “over-representation” in the WHL, discarding the hypothesis that an “even” distribution of sites may or should exist. Our analysis fails to detect any clear and systematic difference in the evaluation of native and colonial sites by the ICOMOS experts, suggesting that sites that satisfy aesthetics canons that are typically European and/or that carried European values are not systematically favoured over those that instead represent a native heritage, at least in the sample period 2007-2018.

This result holds for both our measures of quality and for the recommendation regarding the inscription. Quite importantly, this result seems robust even after controlling for time trends, the type of the site, the continent where the site is located and the cultural background and origin of the experts⁸. The estimates provide little evidence that cultural proximity matters. If anything, they seem to suggest that ICOMOS experts do not favour the sites that belong to their own continent; to the contrary, they seem to reward diversity. The extensive set of robustness checks corroborates the conclusion of neutral, unbiased evaluations.

In conclusion, contrary to the predominant view in the literature, the analysis does not find any empirical support for the idea that ICOMOS experts have priv-

⁸In addition to the test presented, we reproduce our main analysis on each *Eval.doc* component using an Ordered Logit Model and adjusting the p-values for multiple hypothesis testing. We obtain some evidence of opposite effects of *Postcol*: colonial sites tend to receive better judgments for criterion 3 and less laudatory judgments for criterion 4. This result is, however, not corroborated by the other two control variables. We also extract data about *Criteria* for a longer period, namely from 1999 to 2017. Replicating the analysis and the main robustness checks on this second dataset does not significantly change our results. Again, we conclude there are no differences between the control and treatment groups.

CHAPTER 3

ileged European-like sites in the last 20 years. This suggests that the increasing criticism that ICOMOS experts are facing is more likely to be generated by political interests of the UNESCO member countries than by an actual bias in the judgement of the experts.

References

- Bertacchini, E., D.Saccone, and W.Santagata (2010). Embracing diversity, correcting inequalities: Towards a new global governance for the unesco world heritage. *International Journal of Cultural Policy*, 17(3):278–288.
- Bertacchini, E., Liuzza, C., and Meskell, L. (2017). Shifting the balance of power in the unesco world heritage committee: an empirical assessment. *International journal of cultural policy*, 23(3):331–351.
- Bertacchini, E., Liuzza, C., Meskell, L., and Donatella, S. (2016). The politicization of unesco world heritage decision making. *Public Choice*, 167.
- Bertacchini, E. E. and Saccone, D. (2012). Toward a political economy of world heritage. *Journal of cultural economics*, 36(4):327–352.
- Bolton, H. E. (1917). The mission as a frontier institution in the spanish-american colonies. *The American Historical Review*, 23(1):42–61.
- Burnham, K. and Anderson, D. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. New York, Springer.
- Cheng, S.-W. (2006). Cultural goods creation, cultural capital formation, provision of cultural services and cultural atmosphere accumulation. *Journal of Cultural Economics*, 30:263–286.
- Fowler, P. (2003). World heritage cultural landscapes 1992-2002. Technical report, UNESCO World Heritage Centre.
- Frey, B. S., Pamini, P., and Steiner, L. (2013). Explaining the world heritage list: an empirical study. *International Review of Economics*, 60(1):1–19.
- Labadi, S. and Long, C. (2010). *Heritage and Globalisation*. Key Issues in Cultural Heritage. Taylor & Francis.
- Meskell, L. (2002). Negative heritage and past mastering in archaeology. *Anthropological Quarterly*, 75:557–574.
- Meskell, L. (2012). The rush to inscribe: Reflections on the 35th session of the world heritage committee, unesco paris, 2011. *Journal of Field Archaeology*, 37(2):145–151.

- Meskell, L., Liuzza, C., Bertacchini, E., and Donatella, S. (2015). Multilateralism and unesco world heritage: Decision-making, states parties and political processes. *International Journal of Heritage Studies*, 21.
- Musitelli, J. (2002). World heritage, between universalism and globalization. *International journal of cultural property*, 11(2):323–336.
- Olivier, J., May, W. L., and Bell, M. L. (2017). Relative effect sizes for measures of risk. *Communications in Statistics-Theory and Methods*, 46(14):6774–6781.
- Rico, T. (2008). Negative heritage: The place of conflict in world heritage. *Conservation and Management of Archaeological Sites*, 10(4):344–352.
- Sirisrisak, T. and Akagawa, N. (2007). Cultural landscape in the world heritage list: Understanding on the gap and categorisation. *City & Time*, 2.
- Steiner, L. and Frey, B. S. (2012). Correcting the imbalance of the world heritage list: Did the unesco strategy work? *Journal of International Organizations Studies*, 3(1):25–40.
- Trifu, I. (2018). *Dealing with a Difficult Past: Japan, South Korea and the UNESCO World Heritage List*, pages 197–217.
- UNESCO (1972). Convention concerning the protection of the world cultural and natural heritage.
- van der Aa, B. J. (2005). Preserving the heritage of humanity? obtaining world heritage status and the impacts of listing. Ph. D. Dissertation, Faculty of Spatial Science, University of Groningen, The Netherlands.

Appendix B

Is there a pro-European bias in the selection of UNESCO sites? A new test based on colonial sites

Contents

B.1 List of Countries	124
B.2 Colonization determinants	125
B.3 List of variables	126

B.1 List of Countries

Table B.1: List of Countries

2007_2016	
1	Antigua and Barbuda
2	Argentina
3	Australia
4	Bahrain
5	Barbados
6	Brazil
7	Cabo Verde
8	Cambodia
9	Canada
10	Chad
11	Colombia
12	Costa Rica
13	Côte d'Ivoire
14	Cuba
15	DPR of Korea
16	Ethiopia
17	Fiji
18	Ghana
19	Guinea-Bissau
20	India
21	Indonesia
22	Iran (Islamic Republic of)
23	Iraq
24	Jamaica
25	Jordan
26	Kenya
27	Kyrgyzstan
28	Madagascar
29	Malawi
30	Malaysia
31	Marshall Islands
32	Mauritius
33	Mexico
34	Micronesia (Federated States of)
35	Morocco
36	Myanmar
38	Nicaragua
38	Nicaragua
39	Niger
40	Nigeria
41	Palau
42	Papua New Guinea
43	Peru
44	Philippines
45	Qatar
46	Senegal
47	Singapore
48	South Africa
49	Sri Lanka
50	Sudan
51	Syrian Arab Republic
52	Tajikistan
53	Turkmenistan
54	Uganda
55	United Arab Emirates
56	United States of America
57	Uruguay
58	Vanuatu
59	Viet Nam
60	Zambia

B.2 Colonization determinants

Table B.2: Colonization determinants - Domination

<i>Dependent variable: Postcol</i>	
Domination	0.005*** (0.001)
Constant	-1.655*** (0.289)
Observations	99
Log Likelihood	-46.423
Akaike Inf. Crit.	96.846
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

Table B.3: Colonization determinants - Distance

<i>Dependent variable: Postcol</i>	
Distance	-0.292*** (0.097)
Constant	-0.023 (0.178)
Observations	99
Log Likelihood	-54.845
Akaike Inf. Crit.	113.690
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

B.3 List of variables

Table B.4: List of Variables

Variable	Type	Description
<i>Eval_doc</i>	Integer	Quality of sites extracted from ICOMOS evaluation
<i>Criteria</i>	Integer	Sum of criteria approved by ICOMOS evaluation
<i>Inscribed</i>	Dummy	ICOMOS recommendation
<i>Deferred</i>	Dummy	ICOMOS recommendation
<i>Postcol</i>	Dummy	Equal to 1 if the site is colonial
<i>Domination</i>	Integer	Years between the first colonization and independence
<i>Distance</i>	Float	Distance to the sea as crow flies
<i>Year</i>	Integer	Year of first site nomination
<i>Continent</i>	Dummy	Site continent
<i>Site</i>	Dummy	Site type following UNESCO Convention
<i>Buildings</i>	Dummy	Site type following UNESCO Convention
<i>Monuments</i>	Dummy	Site type following UNESCO Convention
<i>Serial</i>	Dummy	Residual site type
<i>canal</i>	Dummy	Site type following Operational Guidelines
<i>landscape</i>	Dummy	Site type following Operational Guidelines
<i>area</i>	Dummy	Site type following Operational Guidelines
<i>historic town</i>	Dummy	Site type following Operational Guidelines
<i>route</i>	Dummy	Site type following Operational Guidelines
<i>historic centre</i>	Dummy	Site type following Operational Guidelines
<i>archaeological</i>	Dummy	Site type following Operational Guidelines
<i>residual</i>	Dummy	Site type following Operational Guidelines
<i>Religious</i>	Dummy	Equal to 1 if the site is religious
<i>Eu_dummy</i>	Dummy	Equal 1 if the expert mission is european
<i>Eu_surname</i>	Float	Expert surname incidence in Europe (percentage)
<i>Interaction1</i>		Interaction term between <i>Postcol</i> and <i>Eu_surname</i>
<i>Interaction2</i>		Interaction term between <i>Domination</i> and <i>Eu_dummy</i>
<i>Interaction3</i>		Interaction term between <i>Distance</i> and <i>Eu_dummy</i>
<i>Prox_dummy</i>	Dummy	Equal 1 if expert belong to the same continent of the site
<i>Prox_surname</i>	Float	Expert surname incidence in site continent

General Conclusion

This thesis contributes to the existing literature in more than one way. First, by focusing on the quality instead of the quantity of UNESCO WH sites, we raise the argument that countries' lobbying activities to boost the inscription of sites could depend on their intrinsic characteristics. Indeed, high-quality sites may not need much political pressure to enter the WHL. Another merit of the thesis is raising the issue of the reputation of the UNESCO WHL. We point out that the UNESCO WHL reputation is at stake: by increasing the number of sites inscribed the marginal quality of the newly inscribed sites decreases. Thus, the average quality of the list might fall in the long term and the UNESCO label can lose its function to signal the outstanding quality of heritage sites. This result is coherent with the literature pointing out the inability of the WHL label to increase the number of tourist arrivals in recent years (Yang et al., 2019, Cellini, 2011, Adie et al., 2018). A deeper examination is needed in order to properly interpret these results. Indeed, one of these two hypotheses can be true: or these studies fail to find a proper counterfactual, i.e. the sample of potential WH sites, or UNESCO's reputation is falling. Testing these hypotheses might be the object of future research.

The third important contribution of this thesis is the development of three original metrics to evaluate the quality of UNESCO heritage. These indexes apply to two different samples: the quality of sites inscribed in the UNESCO WHL is measured by counting the number of criteria satisfied at the moment of the inscription; while the quality of sites nominated is computed by analysing the evaluations of ICOMOS experts. The latter index improves the former by expanding the number of cultural heritage values considered. Indeed, in addition to the satisfaction of a certain number of criteria of OUV, it also evaluates the fulfilment of the conditions of *uniqueness* (i.e. the comparative analysis), *integrity* and *authenticity* of the sites. Moreover, it captures their intensity, by ranking each characteristic on a scale from

0 to 3. The last measure of quality we employ is derived from the Baedeker guidebooks. It evaluates the author's perception of heritage sites along different statistics (the number of citations, the length of the paragraph and the sentiment of the text) and is shown to be coherent with our first index, i.e. the number of criteria of OUV. Despite being innovative in more than one aspect, this approach has some important limitations. By linearly aggregating the criteria of OUV, we are implicitly assigning the same weight to each of them; while some aspects could be, indeed, more important than others. Future research could encompass this limitation by conducting a survey among ICOMOS experts. This methodology would allow attributing the proper weight to each dimension but also taking into account potential synergies and redundancies of the dimensions that are aggregated to obtain the index.

Another limitation of the thesis concerns the sample of heritage sites considered. As mentioned before, we study the sites included in the list (Chapters 1 and 2) and nominated for inclusion (Chapter 3), but we do not analyse the sample of all potential WH sites. This approach would need a theoretical model assessing countries' incentives to propose a site. In Chapter 1, we suggest that the decreasing marginal quality of the UNESCO WH sites is due to the stock of cultural capital being fixed and the unavailability of high-quality cultural capital. Another possible explanation for the same phenomenon is, however, that high-quality cultural capital is available, but countries do not have sufficient incentives to propose it for inclusion. The empirical test of such a hypothesis would require data on the quality of potential UNESCO WH sites. Our analysis of the Baedeker guidebooks enables building a data frame of all potential UNESCO WH sites, which could be used for policy evaluation. Moreover, the potential sites' quality obtained with this methodology will be coherent with the evaluation of UNESCO experts and can be consistently compared with the quality of UNESCO WH sites.

These and similar future research works will help to identify more clearly the role of UNESCO in the cultural tourism sector and the role played by countries in driving the policies implemented by the organisation. This might be useful to identify new policies to efficiently preserve and promote the cultural heritage of humanity.

References

- Adie, B. A., Hall, C. M., and Prayag, G. (2018). World heritage as a placebo brand: A comparative analysis of three sites and marketing implications. *Journal of Sustainable Tourism*, 26(3):399–415.
- Cellini, R. (2011). Is unesco recognition effective in fostering tourism? a comment on yang, lin and han. *Tourism management*, 32(2):452–454.
- Yang, Y., Xue, L., and Jones, T. E. (2019). Tourism-enhancing effect of world heritage sites: Panacea or placebo? a meta-analysis. *Annals of Tourism Research*, 75:29–41.



Titre : Mesurer la qualité du patrimoine culturelle: conséquences pour la littérature sur la liste du Patrimoine Mondiale de l'UNESCO **Mots clés :** UNESCO, Patrimoine Culturel, Qualité,

Lobbying

Résumé : cette thèse contribue à la littérature en fournissant un aperçu de la manière dont la qualité du patrimoine culturel peut être évaluée. En effet, nous considérons comme indicateur de qualité le nombre de critères de Valeur Universelle Exceptionnelle (VUE) qu'un site remplit au moment de son inscription à la LMP. Grâce à cet indicateur, nous examinons de manière critique certains points de vue établis dans la littérature à la LPM de l'UNESCO. Par exemple, nous montrons que l'augmentation du nombre de sites inscrits sur la LPM réduit la qualité marginale des sites nouvellement inscrits (chapitre 1). En exploitant un ensemble de données original, nous testons la robustesse de cet indicateur et nous montrons qu'il est corrélé avec d'autres mesures possibles de la qualité (chapitre 2).

Ainsi, l'indicateur que nous proposons s'avère être proche de la qualité des sites, telle qu'exprimée dans les guides de voyage Baedeker, source faisant autorité au début du XXème siècle. En outre, cette thèse clarifie le rôle des experts dans l'établissement de ces normes et suggère comment ces mesures pourraient être utilisées pour l'évaluation des politiques (chapitre 3). Cet indicateur nous permet d'évaluer le rôle des experts de l'UNESCO dans la définition et l'application des critères d'inscription. En contrôlant la qualité de cette manière, on constate que les experts de l'UNESCO sont impartiaux et qu'ils attribuent le même nombre de critères au patrimoine indigène et colonial, ce qui exclut les craintes d'un parti pris culturel pro-européen dans la sélection des sites du PM.

Title: Measuring quality of cultural heritage: consequences for the literature on the UNESCO World Heritage List **Keywords:** UNESCO, cultural heritage, quality, lobbying, rent-seeking

Summary: This thesis contributes to the cultural economics literature by providing insights on how the quality of cultural heritage can be proxied. Indeed, we consider the number of criteria of Outstanding Universal Value (OUV) that a site met at the moment of the inscription as an indicator of quality. Thanks to this index, we critically investigate some established views of the literature on the UNESCO World Heritage List (WHL). For example, we show that enlarging the number of sites inscribed in the WHL reduces the marginal quality of newly inscribed sites (Chapter 1). Exploiting an original dataset, we show that this indicator is correlated with other possible measures of quality (Chapter 2).

In detail, our index is found to be in line with the quality of the sites, as expressed in the Baedeker travel guidebooks, authoritative source of the beginning of the XXth century. In addition, this thesis clarifies the role of experts in setting those standards and suggests how these measures could be used for policy evaluations (Chapter 3). Indeed, it allows us to assess the role of UNESCO experts in setting and applying the criteria for inscription. Controlling for quality in such a way, UNESCO experts are found to be impartial and assign the same number of criteria to native and colonial heritage, thus excluding concerns of a pro-European cultural bias in the selection of WH sites.