

MODELLING MUSEUM EFFICIENCY IN PRODUCING INTER-RELIANT OUTPUTS

María José del BARRIO-TELLADO
(ORCID: 0000-0002-4818-932X)
(mjose@emp.uva.es)
(Universidad de Valladolid, Spain)

Luis César HERRERO-PRIETO
(ORCID: 0000-0002-2915-5706)
(herrero@emp.uva.es)
(Universidad de Valladolid, Spain)

Grupo de Investigación Reconocido en Economía de la Cultura. Universidad de Valladolid

Abstract:

The aim of this work is to evaluate the performance of a homogeneous state-run network of museums. Non-parametric models are used to measure relative efficiency in these institutions, and we employ a complex production function embracing a number of inputs and outputs adapted to the various functions which museums fulfil: preservation, research, communication, and exhibition. Our approach considers that managers drive certain outputs, but that others escape their control since they are co-produced by visitors and determined by demand conditions and external factors. Based on this, a network two-stage Data Envelopment Analysis (DEA) approach is applied to evaluate museums' overall performance and to distinguish between efficiency in two stages: internal management and external outcomes. The low levels of performance and gaps in the scores from the first to the second stage suggest there are external factors that might determine museum performance. We therefore apply truncated regression models to analyse how and how much certain environmental variables might shape levels of museum efficiency. In this case, we consider indicators such as accessibility, tourism capacity, cultural appeal, museum age and the institutional management model. The application is performed on a sample taken from a Spanish state-run network of museums. Results show that, in general, good levels of efficiency in terms of management do not guarantee success when attracting visitors, and there seems to be a trade-off between the two goals. Variables such as tourism capacity and heritage endowments in the surrounding area, as well as the museum's management model, may determine museums' efficiency levels. The research findings may prove useful for running these cultural institutions and for those responsible for public resource allocation in cultural policies as well as for scholars, who may find a fresh approach for modelling museum efficiency and for discussing drivers of museum management success.

Keywords: Museum management assessment, technical efficiency, Data Envelopment Analysis, DEA, two-stage performance evaluation, national network of museums in Spain, performance evaluation.

JEL Classification: Z11; Z18; D24

This research was financed by the Regional Ministry of Education of the Regional Government of Castilla y León (Spain) (Project Ref. VA012G19).

Introduction

The goal of our analysis is to delve further into the evaluation of cultural institutions' results and performance, taking the network of publicly owned Spanish museums as a case study. These institutions conduct their activity in the public sector, and are funded through the taxes collected by the state as well as donations from private organisations or individuals, albeit to a much smaller degree. This implies a kind of tacit contract between the state and citizens, who agree to pay taxes in exchange for being provided with a range of cultural goods and services that include cultural facilities (Carnegie and Wolnizer, 1996). This by no means clear social contract has aroused interest as regards gaining an understanding of how museums make use of the funds they are allocated, not only in terms of ensuring management common sense vis-à-vis the use of such funding but also with regard to ascertaining whether the previously established objectives are achieved in an environment of dwindling resources and ever-increasing competition for available resources. As a result, there is a desire to gauge museums' performance and to relate this to the resources used in an effort to ascertain whether museums are efficient in their work.

In this vein, and given the problems inherent in objectivising in market terms many of the services such institutions provide and the repeated difficulties involved in gathering financial information on the functions they carry out, most performance evaluation studies focus on analysing technical efficiency; in other words, exploring optimal use of resources, rather than allocative efficiency, which implies a cost minimisation strategy and the need to know the cost of services (Fernández-Blanco et al., 2013). When undertaking such a task, defining the production function is crucial, in terms of the relation between the resources used and the goods and services procured as well as measuring their performance, at least through comparative references vis-à-vis possible best practices. The main aim is therefore to delimit a frontier of optimal behaviour in the transformation process in order to determine how far away each unit lies from the optimum established by the best case frontier (Farrel 1957; Färe et al., 1985). For this purpose, estimating the efficient frontier may be accomplished by using parametric and non-parametric models (Fernández-Blanco et al., 2013). The former require an explicit definition of the production function whereas in non-parametric models, the efficient frontier is estimated through available data, and requires no prior definition. This approach endows these models with greater flexibility and allows them to be applied to situations involving multiple outputs. One disadvantage is that non-parametric models imply deterministic approaches, since they consider any distance from the frontier to be a cause of inefficiency, and fail to take account of the influence of other external factors not specifically considered in the production function (Dyson et al., 2001)

This proves particularly important when some of the variables involved in the production process are at least partly beyond the control of the service provider. Following De Witte and Geys (2011, 2013), a distinction may thus be drawn between programmed outputs and observed outputs. The former are under the manager's control, since they involve using primary resources to produce institutions' basic services, such as activities involving conservation and the organisation of exhibitions in the case of museums or holding and gathering bibliographical material in the case of libraries. In contrast, observed outputs are partly dealt with in the market through the intensity of demand preferences for this service; in other words, museum visitors, or requests for book loans and consultations in libraries, the intensity of which not only depends on the appeal of what is available but also on the demand characteristics and contextual factors that might be driving it. The role of citizens as "co-producers" of public services has aroused substantial interest in the area of public economy (Whitaker, 1980; Pestoff, 2006; Meijer, 2011) and has led to more wide-ranging evaluation studies of efficiency in the provision of public services in a dual sense. Firstly, studies have sought to break down the production process into two stages (Hammond, 2002; Kao and Hwang, 2008), firstly reflecting the production of outputs programmed by the manager based on primary resources, which are later considered as intermediate inputs in the accomplishment of final outcomes, and partially determined by consumer preferences and other external

variables. Secondly, and as a complementary analytical result, the aim is to test the influence of these contextual variables outside the production function on efficiency outcomes (Banker and Natarajan, 2008; Daraio and Simar, 2005).

This methodological approach has given rise to numerous applications in the field of efficiency evaluation of cultural institutions such as libraries (Vitaliano, 1998; De Witte and Geys, 2011 and 2013; Guccio et al., 2018; Simon et al., 2011); archives (Guccio et al., 2016), entities devoted to protecting and maintaining cultural heritage (Finocchiaro-Castro and Rizzo, 2009; Finocchiaro-Castro et al., 2011; Guccio et al., 2014), and even efficiency analysis of tourist destinations based on territorial production functions (Cuccia et al., 2016; Herrero and Gómez, 2017; Figueroa et al., 2018; Guccio et al., 2017). However, to the best of our knowledge, and bearing in mind that museums are one of the most commonly explored cultural institutions in efficiency evaluation studies, to date there are no analyses that consider this analytical approach for evaluating efficiency in two stages and gauging the impact of external variables in the case of museums. The principal contribution and novelty the present study makes is thus to evaluate efficiency of a public system of museums in Spain from a three-fold perspective: firstly, by examining a full scale multi-output production function in an effort to reflect indicators that are representative of the diversity of functions that museums engage in; secondly, to break down the production process into two stages, assuming that the influx of visitors constitutes an observed and non-programmed output, unlike the remainder of museums' activities that make up the cultural supply, and that efficiency should therefore be evaluated in two stages from such a perspective. Finally, an analysis is made of the influence of various contextual variables on efficiency outcomes in museums.

Our case study deals with a sample of 23 representative museums in the Spanish national network of museums, the main network of national and provincial museums in the country. The empirical strategy is based on applying Data Envelopment Analysis (DEA), one of the techniques most widely used to study the efficiency of institutions, and which is also applied in a three-pronged approach in accordance with the previous method: firstly, by applying a non-radial approach using a DEA SBM (*Slacks-based measure*) (Tone, 2001) under the hypothesis of variable returns to scale applied to the whole production function; secondly using the DEA SBM network model proposed by Tone and Tsutsui (2009) in the evaluation of the two stages of the production process; and finally, the Simar-Wilson (2007) truncated regression model in order to gauge the extent to which the efficiency outcomes obtained are driven by other variables outside the institution.

The main contribution of our research is eminently practical, since we aim to show the utility of our methodological approach for assessing the efficiency of cultural institutions. As far as we are aware, our paper is the first to use network DEA jointly with the evaluation of environmental factors to explore a case study of museum assessment. Given these novelties in the approach for modelling museum efficiency, we also aim to contribute to the discussion concerning drivers of performance in museum management from a theoretical point of view. Additionally, the research findings may also prove useful for running these cultural institutions and for those responsible for public resource allocation policies in the area of cultural heritage. The study we present is structured in four parts. After this introduction, we offer a review of previous studies carried out into museum performance. In section 3, we introduce the empirical problem to be addressed, offering a description of the sample of museums used as a reference and describing the methodology to be applied in the empirical strategy for evaluating museum efficiency. In section 4 we present the main results to emerge from the study. We finish with the main conclusions.

1. Efficiency analysis in museums: *state of the art*

Museums are perhaps one of the most representative institutions of cultural heritage, as they perfectly sum up the desire to preserve the cultural legacy handed down to us, and

reflect the wish to maintain and select those assets that reflect the creativity, beauty and identity of a society accumulated over time. However, museums are not just a lifeless ensemble of artefacts, since they also represent a specific type of cultural creation, a distinctive cultural offer which should be programmed and managed. Indeed, how the various exhibits housed in the museum's collection are displayed or how much of an impact their various activities have, may affect to a greater or lesser extent, the institution's appeal and the number of visitors it attracts. Indeed, how museums are managed with regard to achieving their main purposes might entail an array of different economic, cultural and social implications (Johnson and Thomas, 1992; Bertachini et al., 2018).

According to the International Council of Museums (ICOM), a museum is a permanent non-profit institution at the service of society and its development, open to the public, and which acquires, preserves, researches, communicates, and exhibits the material and immaterial heritage of humanity and its environmental for the purpose of study, education and enjoyment. This definition, generally accepted in the museum world, clearly describes the main functions that museums pursue and which involve preservation, exhibition, research, education and dissemination of their collections. Consequently, at this stage, measuring the performance of museums might prove to be a complicated task mainly for three reasons: firstly, because museum involve a wide range of resources, many of which are not easy to measure due to their qualitative and disperse nature; secondly, because museums' ultimate purpose is to provide a complex and multiple product that is not always tangible or commercial in nature; and thirdly, because these institutions are frequently public or non-profit entities that do not often follow cost minimisation behaviour, such that management success cannot easily be measured in the market (Fernández-Blanco et al., 2013; Gómez-Zapata et al., 2018).

This does not mean that the efficiency of these organisations should not be measured or that we may not posit tools that allow us to reflect the quality of the work done in museums or the degree to which they fulfil their purposes. Hence, a museum's activity may be regarded as one production function, involving inputs such as work, provision of buildings and equipment, together with the museum collection itself, in order to obtain various goods and services, corresponding to the main tasks allocated to a museum. The problem here is to identify and specify representative variables of each function (Mairesse and Vanden Eekaut, 2002) and, therefore, to measure the degree of competence as a distance between the resources employed and to what extent the purposes and functions undertaken are achieved, where optimal cases may at least be used as benchmarks. The primary works in this area seek to measure museum's efficiency by drawing up a series of performance indicators (Ames, 1994; Weil, 1995) as a simple or complex set of ratios concerning the different activities carried out. However, these indicators could never hope to offer an all-inclusive and fully comprehensive description of how well a museum is working and indeed are not suited to comparing institutions and compiling rankings amongst them. Another way of measuring museum efficiency is by calculating each museum's distance from a frontier made up of the best practices of a given group. In this case, it is necessary to define the efficient frontier, for which parametric and non-parametric models may be used (Fernández-Blanco et al., 2013). The former are more rigid since they require a precise definition of the functional form of the production function, although they prove more accurate in estimating the productivity linked to each factor and when dealing with stochastic error. Nevertheless, this approach has not enjoyed widespread success apart from the seminal work of Jackson (1988) for a sample of North-American museums, and Bishop and Brand (2003) who measure the efficiency of a group of British museums. In this instance, a simple production function is estimated and leads to the conclusion that the amount of public funding received and volunteer participation have a negative impact on mean efficiency in terms of visitor numbers.

Non-parametric techniques, particularly DEA and its offshoots, have been the most widely used to study the performance of museums given the flexibility they afford when defining the frontier, and because they are able to adapt to a situation of multi-output

production function, which proved more difficult with the previous approach. Studies that opt for non-parametric methods to evaluate museums include Paulus (1995), who examines the technical efficiency of a group comprising 64 French museums. Analysis of the technical efficiency of a group of Belgian museums is also the subject of the work by Mairesse and Vanden Eeckaut (2002) where the technique applied is FDH (Free Disposal Hull), a variant of DEA that removes the need for convexity in the frontier. This is perhaps one of the most complex evaluation approaches for museums since, drawing on a single input set (jobs and items related to budget and infrastructure), they analyse three service models (conservation, dissemination and impact), with their corresponding output, obtaining, in the order stated, increasingly efficient levels for the models. The work of Taalas (1998) introduces the notion of allocative efficiency for a sample of Finnish museums using also the FDH model¹. Other works based on non-parametric models include Pignataro (2002) and Basso and Funari (2004). The former measures the efficiency and technical change of Sicilian museums, whilst the latter introduces variable returns to scale and offers a detailed comparison procedure using a crossed efficiency matrix for a sample of museums located in three major tourist cities (Bologna, Florence and Venice). Within this area of study, Del Barrio et al. (2009) explore the efficiency of a sample of museums in Spain based on a prior classification thereof using multivariate techniques, and Del Barrio and Herrero (2014) introduce a complex production function with multiple outputs to evaluate the performance of museums belonging to a regional network of museums in Spain. Finally, Taheri and Ansari (2013) assess the technical efficiency of a regional system of museums in Tehran, and Basso et al. (2018) evaluate a group of municipal museums in Venice using a two-stage DEA model based on the Balanced Scorecard (BSC) reference constructed for museums.

All of these studies examine museum efficiency assuming a single process that is controlled by the institution, in other words, the use of a set of resources to provide different goods and services linked to museums' functions. They therefore fail to take account of an analytical innovation that has already been applied in efficiency studies in other cultural institutions, namely the possibility of splitting the production process into two stages, based on distinguishing between programmed outputs and observed outputs (De Witte and Geys, 2011, 2013). Indeed, when providing public services a distinction may be drawn between an initial stage controlled by the institution's manager, which involves merging a set of primary resources to produce programmed outputs related to the entity's basic functions, and a second stage in which the programmed services act as intermediate input to achieve final outputs that are partly determined by consumer preferences, consumers who therefore play the role of co-producers. It is not a matter of conducting two separate efficiency studies, but rather of viewing efficiency analysis as a process involving interrelations between activities, where the second stage, geared towards the public, is partly beyond the manager's control. Libraries have provided the primary field for such analyses, since a clear distinction may be drawn between the functions controlled by the manager (basically, maintaining and managing the bibliographical collections) and those involving loans and consultation services, which ultimately depend on public interest. Numerous studies have been conducted in this regard, from the seminal works of Vitaliano (1998) and Hammond (2002) up to the more recent works carried out on a sample of Belgian (De Witte and Geys, 2011, 2013), Spanish (Simón et al., 2011), and Italian libraries (Guccio et al., 2018). On the basis of this methodological approach, evaluations have also been performed of other cultural institutions such as archives (Guccio et al., 2014) and cultural heritage agencies (Finocciaro Castro and Rizzo, 2009; Finocciaro Castro et al., 2011; Guccio et al., 2014) although, to the best of our knowledge, no works as yet exist exploring museum efficiency adopting this method.

¹ Few works address the analysis of allocative efficiency of cultural institutions. Mention may again be made of Taalas (1997) for the case of Finnish theatres; Fernández-Blanco et al. (2019), who estimate marginal costs for performing arts production of a sample of theatres in Warsaw; and Fernández-Blanco and Rodríguez-Alvarez (2018), who study allocative efficiency for a non-governmental organisation devoted to promoting cultural, humanistic and scientific values.

Finally, and assuming that part of museum management success is not totally controlled internally, the question arises as to what extent external factors might shape final demand or might have impact on the performance of the production process. By applying regression models, analyses to complement efficiency studies have thus been developed in an effort to measure the link between the efficiency ratios obtained and the different contextual variables, related mainly to the socioeconomic features of the area in which the institutions are located. Such analytical techniques have been widely applied in tourist destination efficiency studies, specifically exploring the impact of cultural variables on tourist competitiveness (Cuccia et al., 2016; Herrero and Gómez, 2017; Figueroa et al., 2018; Guccio et al., 2017). Nevertheless, they have been applied to a lesser degree in the case of cultural institutions. To the best of our knowledge this methodology has only been applied in the case of Flemish libraries (De Witte and Geys, 2011) and cultural heritage agencies in Sicily (Finocciaro Castro et al., 2011), although there are still no works in this sense in the area of museums.

3- Evaluation of the Spanish state-run network of museums as a case study

3.1- Data

A country's museum network tends to be wide-ranging and diverse given the disperse nature of the museum collections involved and the differing institutional structure that emerges over time. Nevertheless, efficiency evaluation studies demand that a certain consistency of the sample analysed be assumed, within each institution's diversity, in order to avoid spurious results when applying evaluation techniques (Dyson et al., 2001). The case study involved in this present research consists of an institutional network of museums, the Spanish system of national museums, all of which are owned by the state and which are, on the one hand, made up of provincial museums, comprising the principal archaeological and fine arts collections accumulated during the 19th and 20th centuries at a provincial scale; and on the other, so-called national museums, which emerged as a result of the specific nature and importance of their artistic collection and which embrace a certain thematic diversity². In sum, this is a homogenous group of museums with a long history and classical approach, similar in size and pursuing a parallel mission as museums: gathering, maintaining, studying and disseminating the corresponding museum collections. Nevertheless, as a result of political decentralisation in Spain, provincial museums have been run by regional authorities for the past 25 years, whereas national museums continue to be under the charge of the Ministry of Culture and Sports, giving rise to two groups that are distinguishable merely in terms of their management status, and that have followed different paths. In all, the study is composed of a group of 50 museums³, on which a survey was carried out, requesting information on the resources used and the main activities undertaken between 2008 and 2015⁴. Answers were obtained from 23 representative museums of the whole sample, which can be seen in Table 1. Nevertheless, in an effort to enhance the robustness of the subsequent statistical and econometric analysis and also to avoid measurement or other errors resulting from outlying data, we posited an efficiency analysis over eight years that we consider as a single time period in order to delimit a database panel in which each museum is taken into account as a decision making unit each year. We finally obtained a total of 184 observations from 23 museums over eight years. In this case, the degree of representativeness of the sample provides a sampling error of 5.3% with a 95% significance level.

² For instance, the Prado National Museum, the National Sculpture Museum, the National Museum of Altamira, the National Museum of Ceramics and Sumptuary Arts, etc.

³ So-called *House Museums*, dedicated to the life and work of certain historical figures and whose collections are considered to remain virtually unaltered and to be of mainly ethnographical interest, have been removed. Likewise, also removed from the sample are the Prado National Museum and the Queen Sofia National Museum of Contemporary Art which, due to their having an autonomous management status and given their condition as *star museums* (Frey, 1998), would be outliers in the sample.

⁴ Survey and data gathered are available upon request from the authors of the research.

Based on this information, the main variables that make up the production function of museums have been constructed in an effort to pinpoint representative indicators of the museum's main missions. On the input side, we first have the work or staff factor for the personnel involved in the museum's various activities: management, administration, technical staff, security, maintenance staff, and so on. Secondly, capital resources are specified in two variables: the museum's size in square metres, which gives an idea of the scale and importance of the building that houses the collection, and one indicator for the equipment and services⁵ deemed to be essential for the museum to undertake many of its activities⁶. As regards the variables representing output, we first consider those linked to the museum's exhibition function; namely, visitor numbers, the most basic expression of demand; and the number of temporary exhibitions organised by the museum. Indeed, this second variable is also an output that is specific and representative of one of a museum's most characteristic activities, namely the dissemination and research concerning its own cultural corpus, since many of these exhibitions are linked to the permanent collection and contribute to its dissemination. Continuing with the representative outputs, two new variables were calculated that deal with the museum's communication, education and research purposes, since we first calculated the number of publications issued by the institution (guides, catalogues, artworks, and research articles) and then computed the museums' activities that are more closely related to social commitments, since these refer to organising dissemination actions, such as educational workshops, concerts, seminars, conferences and mini-conferences, open days and so on.

TABLE 1
Museums in the sample

Museum	Name	Foundation	Management
M1	Museum of the National Library of Spain	1995	Central Administration
M2	National Museum of Ceramics and Sumptuary Arts	1947	Central Administration
M3	National Sculpture Museum	1933	Central Administration
M4	Sephardic Museum	1964	Central Administration
M5	Altamira National Museum and Research Centre	1979	Central Administration
M6	Lázaro Galdiano Museum	1951	Public Foundation
M7	Álava Museum of Fine Arts	1941	Regional Administration
M8	Badajoz Provincial Museum of Archaeology	1867	Regional Administration
M9	Burgos Museum	1846	Regional Administration
M10	Caceres Museum	1898	Regional Administration
M11	Cuenca Museum	1974	Regional Administration
M12	Casa de los Tiros Museum in Granada	1929	Regional Administration
M13	Úbeda Archaeological Museum	1972	Regional Administration
M14	León Museum	1869	Regional Administration
M15	Murcia Museum of Fine Arts	1867	Regional Administration
M16	Palencia Museum	1997	Regional Administration
M17	La Rioja Museum	1963	Regional Administration
M18	Seville Museum of Fine Arts	1841	Regional Administration
M19	Soria Numancia Museum	1919	Regional Administration
M20	Tarragona National Archaeological Museum	1844	Regional Administration
M21	Valladolid Museum	1879	Regional Administration
M22	Zamora Museum	1911	Regional Administration
M23	Santa Cruz Museum	1965	Regional Administration

Source: Authors' own

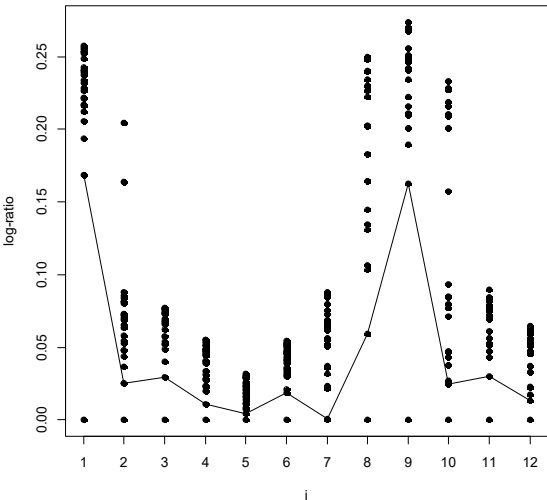
⁵ This indicator specifically calculates the existence of library services, archives, restoration workshops, warehouses, photography workshops, audio-visual facilities, areas for educational activities, environmental control, computerised control, cloakrooms, public car parks for the disabled, areas for rent, tourist guides, audio-guides, webpages, conference rooms, cafeterias and shops.

⁶ This section does not contain any variable reflecting the cultural value of the museum collection itself, since they all enjoy the same official accreditation as "goods of cultural interest". Additionally, a museum's cultural value cannot be confined to the number of exhibits in the collection, given the disperse nature thereof. Nor is it possible to consider qualitative external evaluations, since these tend to be applied to the collection as a whole and fail to draw any distinction between the various pieces. Indeed, quantitative measurement of a museum's cultural value remains one of the challenges facing economic analysis, and is one which might only prove possible by estimating stated preferences through the contingent valuation method or even following tourist valuation standards (TripAdvisor and so on) that would surely tip the balance towards collections that are better known or more accessible to tourists. However, positing any such technique or approach would fall well outside the scope of the present research. Nevertheless, the impact of a museum's cultural value is assumed to have a direct correlation on the remaining variables, such as through museum size, which tends to be linked to the museum's importance or to the historical value of the building where it is housed.

On this full model of the basic functions undertaken by museums, expressed through a multi-output production function, an initial performance evaluation is carried out. We are, however, aware that most of the outputs are controlled and programmed by the manager, whereas the number of visitors is an observed output that depends on the appeal of the supply as well as on visitor preferences and external factors that shape said demand. For this reason, two new analytical contributions are included in the study of the public system of museums; two-stage efficiency analysis using a DEA SBM-Network (Tone and Tsutsui, 2009), and an analysis of the influence of external variables using Simar-Wilson (2007) type truncated regression models.

Before proceeding with the methodological application, certain clarifications should be made concerning the database, since specific efficiency evaluation models are extremely sensitive both to possible measurement errors as well as to potential outliers. The presence of these elements might mean that the estimated frontier is above the real frontier, leading to efficient institutions being deemed inefficient in the estimation. This advocates the application of procedures that remove possible outliers from the data. With this goal in mind, we use the technique developed by Wilson (1993), since it is the one that is best adapted to the problem of efficiency analysis⁷. This procedure is based on measuring the effect caused by excluding K observations (K being a number chosen by the analyst) in the whole sample. With this aim in mind, we calculate the value of the $R_L^{(i)}$ ⁸ statistic for all the possible subsets L of size i (i being arbitrary) subsequent to the successive exclusion of cases. In this way, the presence of an outlier is identified graphically based on the greatest gap between the two smallest values of the statistic for each test.

FIGURE 1
Log-Ratio Plot for the sample of museums



Source: authors' own

In order to be clearer in the graphs, we calculate the logarithm of the quotient between the value of the statistic in each case and its minimum value. Figure 1 thus shows the 25 lowest values of the logarithm of ratios for our case study. The line connects the second lowest value of the ratio for each i, highlighting the gap between the two lowest values. The separation is significant for i=1, 2 and 3, suggesting in Table 2 that observations 61, 84 and 166 might be identified as outliers at this stage. Once again, the separation proves important for observations i=8, 9 10 and 11 suggesting in Table 2 a new group of outliers corresponding to observations 23, 46, 69, 92, 115, 138, 161 and 184. An

⁷ The Wilson procedure does not require OLS residuals and can thus be used with linear programming based models.

⁸ For further information concerning the calculation of the $R_L^{(i)}$ statistic, see Andrews and Pregibon (1978) and Wilson (1993). To apply this procedure, the FEAR 1.15 package for R has been used.

analysis of our data shows that eight of the observations identified as outliers correspond to the Santa Cruz Museum in each of the eight years for which data are available, such that we opted to exclude it from the sample. Indeed, the idiosyncratic nature of its building means that the Santa Cruz Museum is often used to house large-scale exhibitions linked to the historical and cultural wealth of the city of Toledo (El Greco, Charles I, etc.) that occasionally overestimate the impact of their own collections. The three other data identified as possible outliers correspond to different museums in the sample that have witnessed some special event in the year in question and that is reflected in the data gathered but not in its progress over time, on average. Following Wilson (1993), we feel that simply because an observation is unlikely to occur is not enough to claim that it is an error, such that we believe in this case that we are not dealing with an outlier and that it can be kept in the sample.

TABLE 2
Observations and $R^{(i)}_{\min}$ values

i	Observations												$R^{(i)}_{\min}$
1	166	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.7079
2	84	166	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.5930
3	61	84	166	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.4944
4	18	61	84	166	NA	NA	NA	NA	NA	NA	NA	NA	0.4242
5	18	23	61	84	166	NA	NA	NA	NA	NA	NA	NA	0.3677
6	92	69	46	16	123	166	NA	NA	NA	NA	NA	NA	0.3021
7	115	92	69	46	161	23	166	NA	NA	NA	NA	NA	0.2362
8	138	184	115	92	69	46	161	23	NA	NA	NA	NA	0.1609
9	138	184	115	92	69	46	161	23	166	NA	NA	NA	0.1138
10	138	184	115	92	69	46	161	23	84	166	NA	NA	0.0947
11	138	184	115	92	69	46	161	23	61	84	166	NA	0.0784
12	138	184	115	92	69	18	46	161	23	61	84	166	0.0668

Source: authors' own

To conclude this section, Table 3 shows the descriptive statistics of the group of variables considered in the study, and for the 22 museums finally taken into account in the analysis. First, the variables involved in the museums' production function are shown, and then the contextual variables taken for the final part of our study in order to determine the possible influence of external factors on the efficiency outcomes. In this vein, efforts have been made to specify indicators that characterize where museums are located, gathering variables related to tourism potential (accommodation capacity in hotels), complementary cultural endowments (number of protected cultural goods in the province), access (motorways) or wealth (regional GDP) and, secondly, specifying museums' characteristics, such as years in existence and type of managing institution. Based on these premises, our aim is to ascertain whether the most efficient institutions in terms of their cultural programme are also the most efficient in attracting visitors and, in turn, to what extent external non-discretionary economic or sociodemographic variables might be affecting museum performance.

TABLE 3
Variables and descriptive statistics

Variables		Mean	Std.Dev.	Minimum	Maximum	Range
Inputs and Outputs						
Size	Surface area of museum in m ²	2913.66	1651.31	426.00	6646.00	6220.00
Equipment	No. of facilities and equipment	10.65	4.19	0.00	18.00	18.00
Employment	Museum staff	34.22	24.71	10.00	112.00	102.00
Exhibitions	No. of exhibitions scheduled	4.28	3.51	0.00	16.00	16.00
Museum Activities	No. of dissemination and educational activities	161.87	230.92	0.00	1705.00	1705.00
Publications	No. of publications (research and dissemination)	7.13	10.73	0.00	55.00	55.00
Visitors	Visitors	87416.22	87462.87	6647.00	375170.00	368523.00
Environmental variables						
Heritage sites	No. of protected/listed cultural goods in the province	268.73	156.00	52.00	546.00	494.00
Hotels	No. of hotel beds per km ²	2.17	3.47	0.26	13.26	13.00
Motorways	Km of motorway in the province	0.04	0.02	0.01	0.10	0.09
Institutions	Dummy: Central / Regional Government	0.27	0.45	0.00	1.00	1.00
Age	Years since it was opened	89.14	50.71	11.00	181.00	170.00
GDP	GDP per inhabitant in the region	4.33	0.09	4.18	4.51	0.33

Source: Authors' own

3.2- Methodology

Most previous works addressing museum performance apply standard DEA models to obtain an indicator that measures the relative efficiency of these institutions by comparing a set of inputs and outputs (see section 2). In all cases, the models presented fail to take account of the possible relations between the different activities carried out in the museum nor the possibility that the output from one activity might again be involved as an intermediate input for another activity. Pinpointing these possible links between activities makes single stage performance assessment insufficient and perforce leads to the need for models in which the various activities undertaken are broken down into two stages, where the first is under the manager's direct control, with a production of programmed outputs based on primary resources, whereas the second is beyond the manager's control since the outputs are observed. In other words, they are part-produced by visitor demand and the variables that shape it (De Witte and Geys, 2011). In this way, network DEA models (Liang et al., 2008; Tone and Tsutsui, 2009; Kao, 2009 and 2014) allow the internal processes of the entities evaluated and their interrelations to be considered and to provide, in addition to an overall measure of efficiency, consistent measures of efficiency for each stage, facilitating more accurate information on the possibilities of improving the units assessed. As pointed out, this kind of approach has been widely used to evaluate other cultural institutions such as libraries, archives, and so on, although less so in museums. As a result, in our work, starting from a conventional DEA SBM efficiency model known as the *black-box model*, we subsequently apply a network DEA model to evaluate the performance of the group of museums in two stages. More specifically, we apply an SBM network model based on the work of Tone and Tsui (2009) and already used on previous occasions to gauge performance in other sectors such as sport (Moreno and Lozano, 2014) or the banking sector (Avkiran, 2009; Akther et al., 2013)

Conventional DEA CCR (Charnes-Cooper-Rhodes) and BCC (Banker-Charnes-Cooper) models provide radial measures of efficiency under the assumption of proportional changes in inputs or outputs. In contrast, the SBM network model is based on the SBM model proposed by Tone (2001) that provides a non-radial efficiency index and which may prove to be appropriate when inputs and outputs do not change proportionally. Such might be the case for the production function in museums, where capital and work are not fully interchangeable and where there is subsequent diversity of scales and heterogeneity of DMUs. Furthermore, the efficiency index of the non-oriented SBM model includes information related to slacks in both inputs and outputs, providing a *strong* measure of efficiency, as opposed to the concept of *weak* efficiency offered by conventional radial models that fail to take account of slacks information.

Taking the work of Tone and Tsutui (2009) as a reference, in order to describe the non-oriented SBM network model, we assume a set composed of n DMUs ($j = 1, \dots, n$) that carry out K activities ($k = 1, \dots, K$) consuming m_k inputs and producing r_k outputs in each activity k . We call x_j^k ($j = 1, \dots, n$) ($k = 1, \dots, K$) the set of inputs consumed by the DMU j in activity k ; y_j^k ($j = 1, \dots, n$) ($k = 1, \dots, K$) the set of outputs produced by DMU j in activity k , and $z_j^{(k,h)}$ ($j = 1, \dots, n$) the intermediate outputs of activity k included in activity h . We can evaluate the overall efficiency of each DMU by solving the following optimisation problem:

$$\rho_0^j = \min_{\lambda^k, s^{k-}, s^{k+}} \frac{\sum_{k=1}^K w^k \left[1 - \frac{1}{m_k} \left(\sum_{i=1}^{m_k} \frac{s_i^{k-}}{x_{io}^k} \right) \right]}{\sum_{k=1}^K w^k \left[1 - \frac{1}{r_k} \left(\sum_{r=1}^{r_k} \frac{s_r^{k+}}{y_{ro}^k} \right) \right]}$$

subject to:

$$x_0^k = X^k \lambda^k + s^{k-} \quad (k = 1, \dots, K)$$

$$y_o^k = Y^k \lambda^k - s^{k+} \quad (k = 1, \dots, K)$$

$$e \lambda^k = 1 \quad (k = 1, \dots, K)$$

$$\lambda^k \geq 0, s^{k-} \geq 0, s^{k+} \geq 0, \quad (\forall k)$$

$$z^{(k,h)} \lambda^h = z^{(k,h)} \lambda^k \quad (\forall (k,h))$$

where λ^k is the vector of intensity corresponding to activity $k (k = 1, \dots, K)$ and $s^{k-} (s^{k+})$ are the vectors of input (output) slacks indicating excesses of inputs and defects of output, respectively. We consider $\sum_{k=1}^K w^k = 1, w^k \geq 0 (\forall k)$, with w^k being the relative weight allocated to activity k . In our case, we have opted to attach the same importance to each of the DMUs' activities.

The previous model is a non-oriented model following the guideline of the SBM model applied previously in the conventional model (black box) where both input and output slacks are taken into account. Moreover, this model assumes a situation in which the intermediate products may vary (*free link case*), either increasing or diminishing freely in the optimum, as opposed to the option in which the intermediate products are beyond the control of the DMUs (*fixed link case*). This is the option which best fits our case study since we feel that museum managers faithfully programme the museum's activities as internal products of their cultural supply.

In line with Tone (2001), the previous model may be solved through a transformation to a linear model, applying the Charnes and Cooper (1962) transformation. This provides the overall non-oriented efficiency scores for each DMU (ρ_o^k), where input and output slacks are simultaneously taken into account. A DMU is said to be efficient overall when $\rho_o^k = 1$. Using input and output slacks in the optimum (s^{k-}, y, s^{k+}) derived from the previous model, we can calculate the non-oriented efficiency of each activity as follows:

$$\rho_k = \frac{1 - \frac{1}{m_k} \left(\sum_{t=1}^{m_k} \frac{s_t^{k-}}{x_{to}^k} \right)}{1 + \frac{1}{r_k} \left(\sum_{r=1}^{r_k} \frac{s_r^{k+}}{y_{ro}^k} \right)} \quad (k = 1, \dots, K)$$

with ρ_k being the efficiency of each activity that optimises the overall efficiency of institution ρ_o^k . The DMU_o is considered efficient for activity k when $\rho_k = 1$ and will be efficient overall if, and only if, it is efficient for each one of the activities undertaken; in other words, in the two stages of the production process.

The SBM network model just specified operates assuming constant returns to scale (CRS). However, it is possible to modify said specification in order to consider variable returns to scale (VRS), by introducing an additional restriction, set out as follows:

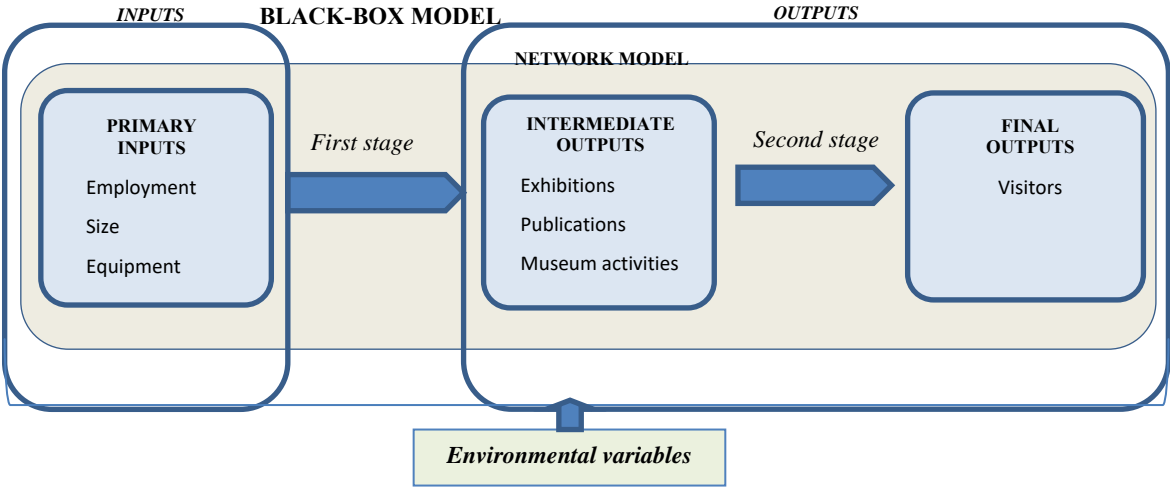
$$\sum_j \lambda_j^k = 1 \quad (\forall \lambda), \quad \lambda_j^k \geq 0 \quad (\forall j, k)$$

Estimating efficiency using variable returns to scale means using a more flexible technology, since it allows sections with varying returns to be present at the frontier. Estimation using constant returns, however, tends to be considered as the long term reference frontier for the units analysed. There are different criteria for identifying the returns to scale in each situation. One of the most widely used criteria is that put forward by Färe et al. (1985) based on calculating –for each institution– the measures of technical efficiency compared to three technologies with differing returns to scale. Banker (1996) suggests semi-parametric statistical tests to evaluate returns to scale. Simar and Wilson (2002 and 2011) posit bootstrap procedures for testing hypotheses regarding returns to scale. We opted for this latter alternative to identify returns to scale in the two previously described models. In this way, taking into account the statistic 48 put forward in Simar

and Wilson (2011), we posited various hypothesis contrasts, taking constant and non-increasing returns to scale as the null hypothesis. In all cases, results lead us to reject the null hypothesis and to confirm the hypothesis of variable returns to scale for our model at a 95% confidence level.

Based on these premises, the empirical strategy in our research work is shown in Figure 2. First, we carry out the efficiency analysis of a conventional model (*black box model*) with three inputs and four outputs. We assume that museums consume work and capital resources, measured in terms of the number of employees in the institution, the size of the museum and the equipment required to generate a series of services that make up the institution’s cultural offer in compliance with museums’ basic functions of conservation, exhibition, education and dissemination. Following this idea, the outputs considered in our model are measured in terms of the number of temporary exhibitions staged, the number of publications from the institution, the number of complementary activities (educational workshops, conferences, social integration, etc.) and the number of visitors. Secondly, in an effort to understand what may underlie possible inefficiencies in the units analysed, we posited a two-stage SBM network model. During the first stage (*managerial stage*), the inputs consumed are related to the cultural supply programmed by the institution. All the output variables in this stage are handled directly by the managers taking account of the budgetary restrictions and technical availability of resources, which is consistent with the previously described model (*free link case*). During the second stage of the SBM network model (*outcomes stage*), the output variables from the first stage become intermediate inputs, and the relation between the institution’s cultural offer and the number of visitors is analysed. In this case, managers may improve efficiency levels by putting together an appealing cultural supply. Nevertheless, this part of the process is not fully under managers’ control since the decision whether to visit or not ultimately depends on the public and there may be other external factors beyond the scope of the museum that can shape said decision. For this reason, the final contribution made by this study seeks to identify some of the external variables that might determine how well museums perform when aiming to achieve their goals, and which are related to the socioeconomic characteristics of the environment where the museum is located (hotel capacity, alternative cultural attractions, economic conditions and ease of access), as some of the differentiating features of museums (age and managing institution)

FIGURE 2
Empirical strategy of the research



Thus, we assume that the efficiency scores can be regressed on a vector of environmental variables in line with the following general specification:

$$e_k = \beta x_k + \epsilon_k$$

where \hat{e}_k represents the efficient scores from each model of the previous stage, $\varepsilon_k \sim N(0, \sigma^2)$ is a vector of error terms, and β is a vector of parameters for the series of independent variables x_k . In order to estimate the parameters of this equation, models such as OLS (ordinary least squares) or Tobit regressions are unsuitable since they fail to prevent correlation between the efficiency outcomes and the error term ε_k because the efficiency scores estimated in the first-stage are serially correlated by construction (Benito, 2014, Cuccia et al., 2016). This is why we adopt the approach suggested by Simar and Wilson (2007) —based on truncated regression and bootstrapping— to explain differences in museum efficiency according to our vector of external variables. This approach allows us to account for the nature of our DEA-based efficiency scores and the unknown serial correlation between them. Basically, it requires simulating a sensitive data-generating process from which to create artificial bootstrap samples, and then constructing standard errors and confidence intervals for the parameters of interest through bootstrapping. We specifically follow the first algorithm in Simar and Wilson (2007: 41-42), since it allows for the introduction of efficiency ratios based on non-radial distances (Reig-Martínez, et al., 2011), such as the SBM *black box* and SBM network models used in the research.

4. Results

This section shows the main efficiency evaluation results of the state-run system of museums in Spain from a triple methodological perspective, in accordance with the empirical strategy of the research: firstly, evaluating the initial *black box* model, which involves applying DEA SBM to the full multi-output production function, and which considers the outputs that are representative of all museums' functions and activities. Second, the application of the SBM network model that enables us to assess overall efficiency as well as two-stage efficiency under the hypothesis that we are able to distinguish between outputs programmed by the managers and observed outputs that are beyond their control, specifically the number of visitors, which is co-produced by those attending. Finally, we aim to identify certain environmental variables that might determine the efficiency results obtained previously, by applying a truncated regression model following algorithm 1 from Simar and Wilson (2007).

TABLE 4
Statistics of efficiency ratio estimates: DEA SBM (*Black Box*) and DEA SBM network

	DEA SBM <i>Black Box</i>		DEA SBM network		
	SBM	SBM Superff.	Overall Score	Stage 1	Stage 2
Average	0.4129	0.4272	0.2130	0.3349	0.2330
Max	1	1.6881	0.6747	1	1
Min	0.0357	0.0357	0.0223	0.1066	0.0177
St. Dev	0.3438	0.3724	0.1615	0.1802	0.2325
No. of efficient museums	35	1	0	3	1
No. of observations	176	176	176	176	176

Source: Authors' own

Consequently, Table 4 shows the descriptive statistics of the efficiency indices obtained in all the evaluation models⁹. The results from the first model, a DEA SBM non-oriented model with variable returns to scale applied to the full production function with three inputs and four outputs, indicates a relatively low average efficiency value of 0.41, with important variations among the units analysed, as shown by the high value of the standard deviation (0.34). This tells us that the museums in our sample are, on average, relatively technically inefficient in terms of accomplishing their functions. There are only 35 optimal observations considered as best practices in the whole period, while the remaining cases and museums evidence ample room for managerial improvement. We have also applied an SBM super-efficiency model (Tone, 2002) to discriminate between efficient DMUs so as then to be able to compare the rank order of the efficiency results

⁹ All the estimates are available from the authors upon request.

among the different models. The efficiency results of this latter approach are obviously very close to the first model and the rank correlation among the two indices is significant and very high (Table 5).

TABLE 5
Spearman and Kendall rank correlation matrix between efficiency results

SPEARMAN	SBM	SBM Superff.	Overall Score	Stage 1	Stage 2
SBM	1				
SBM Supeff	0.9961*	1			
Overall Score	0.3046*	0.3104*	1		
Stage 1	0.4244*	0.4228*	-0.3039*	1	
Stage 2	0.2268*	0.2312*	0.9868*	-0.4201*	1
KENDALL	SBM	SBM Superff.	Overall Score	Stage 1	Stage 2
SBM	1				
SBM Supeff	0.9613*	1			
Overall Score	0.2004*	0.2058*	1		
Stage 1	0.2844*	0.2873*	-0.2043*	1	
Stage 2	0.1509*	0.1553*	0.9219*	-0.2823*	1

***, ** and * stand for statistical significance at 1%, 5% and 10% confidence levels, respectively.
Source: authors' own

In order to pinpoint the reasons for the inefficiencies observed in the previous approach, we subsequently implemented an SBM network model (Tone and Tsutsui, 2009), which was also non-oriented and with variable returns to scale, allowing us to consider the relations between the different activities undertaken by museums, and to split the production process into two stages depending on how much the manager is able to control the outputs. In light of Table 4, this model yields a mean overall score of 0.21, and displays significant variation between a maximum value of 0.67 and a minimum of 0.02. The results obtained are substantially lower than those in the *black box* model¹⁰. Moreover, the mean efficiency level in the first stage is much higher than in the second, indicating that, broadly speaking, Spanish museums in the national system are more efficient when it comes to purely organisational activities and cultural programming than when measuring their capacity to attract visitors. In any case, there are only three efficient units in the first stage and one in the second, whereas there are no museums that may be considered efficient overall, since there are no cases where optimal performance concurs between the two stages.

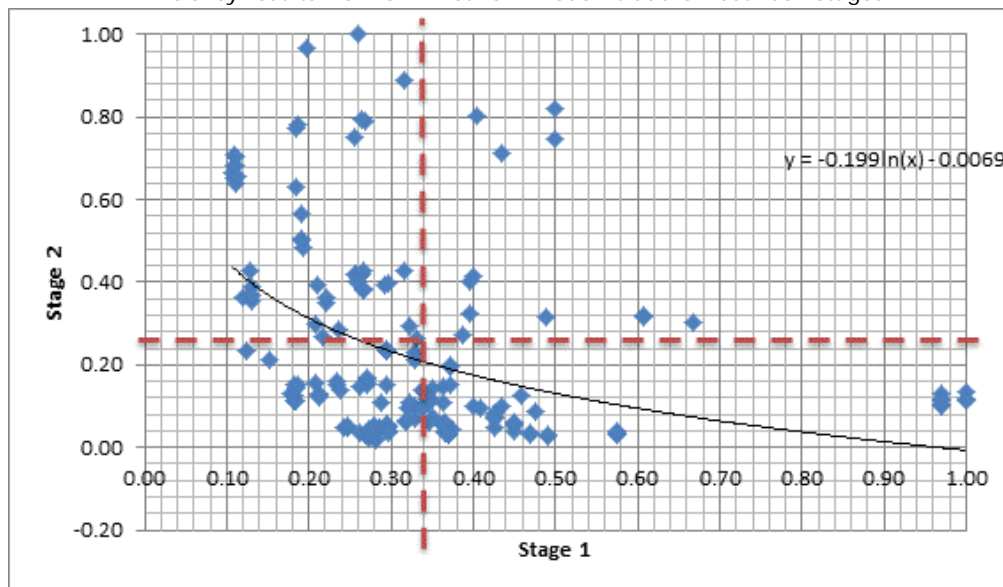
This result leads us to infer three analytical deductions. First of all, we found that the standard DEA SBM seems to fail to identify inefficiencies within the internal processes. Thus, it provides few insights into sources of inefficiency and the operational stages where inefficiency may arise. This also seems to be evident through the Spearman and Kendall rank test (Table 5) since, even though correlations are significant, they reflect a relatively low degree of dependency between the ranking of efficient units in the *black box* model, and the ranking of the SBM network models. Nevertheless, the value of the correlation coefficient between the first and second stage of the model reaches significant and negative values (specifically, -0.42 in the Spearman correlation), thus pointing to a possible inverse relation between the two efficiency indexes. In other words, the units that obtain the best results on the management stage should be expected to perform worst in terms of attracting visitors and vice-versa. This is the second deduction to emerge from our research, since there seems to be a certain trade-off between management and cultural programming functions on the one hand and the exhibition function on the other, expressed in terms of influx of visitors to museums. Consistent with this claim, Figure 3 shows the scatter plot of the observations (museums) for the whole period analysed for stage 1 (management) and stage 2 (attracting visitors) efficiency index values, together with the fit of a logarithmic function with a negative slope (-0.2), showing this interchange between the two goals.

¹⁰ This result is also related to the fact that the efficiency values in each model depend on the number of variables considered: three inputs and four outputs for the traditional model, compared to three inputs and one output for the overall efficiency of the network model. In fact, the reliability of the results from DEA models depends on the number of inputs and outputs in the analysis, for a given number of DMUs (Kneip et al., 1998; Guccio et al., 2018)

This trade-off between management activities and efficiency in attracting visitors seems to suggest that visitor flows to museums remains to a certain degree independent from said museums' cultural programming, and is shaped by other factors such as the appeal of the collection itself, the museum's reputation or other external factors not linked to the management and scope of the museum. This leads us to the third implication to come out of our study where, given the lack of variables that are representative of the museum's value and reputation¹¹, we aim to determine how certain environmental or contextual factors might impact on museums' performance vis-à-vis accomplishing their goals. We thus define a series of environmental variables related to tourist appeal, the accumulation of cultural capital endowments or conditions of accessibility and the economic development of the area where the museum is located, together with other factors linked to the particular features thereof, such as how long it has been open and the type of management (national or regional)¹².

FIGURE 3

Efficiency results from SBM network model: trade-off between stages



As proposed in the methodology, we apply the Simar and Wilson (2007) truncated regression and simple bootstrap model (algorithm 1), since this proves appropriate when considering efficiency values from estimation models with non-radial distances (Reig-Martínez, et al., 2011). However, the applicability of this method stems from compliance with the condition of separability between the space of the inputs-outputs and the contextual variables, which entails assuming that the exogenous factors affect the probability of being more or less efficient, but do not determine the type of frontier. In our case, it is understood that the characteristics of the environment should not *per se* influence museums' resource allocation and cultural programming capacity since, applying the interpretation of the principal-agent theory, the museum's resources as well as the goals it pursues are defined externally by the administration which acts as the principal agent, in other words, the Ministry of Culture or the Regional Culture Departments. Nevertheless, as empirical support for this claim, we calculated the correlations matrix between the variables of the production function and the contextual variables (Table 6), where ratios can be observed that are not always significant or that

¹¹ As pointed out, we do not have any estimations of the value of the museum collections. Nor does the Ministry of Culture draw any kind of distinction in terms of museums' reputation or accreditation. All of them are given the general title of "goods of cultural interest". Due to the disperse nature of museums, neither did we consider the number of exhibits in the artistic or archaeological collection to reflect its estimated value.

¹² These latter two reasons may ultimately be indirectly linked to museums' reputation, both because of what it means to be a museum of long-standing repute and because national museums have a specific relevant theme.

reflect low degrees of dependency, such that it is reasonable to assume the condition of separability¹³, although this should be taken with a certain degree of caution.

TABLE 6
Spearman Correlations Matrix

	SIZE	EQUIP	EMP	VISIT	EXP	ACT	PUB	Heritage sites	Hotels	Motorway	Institution	Age	PIB
SIZE	1												
EQUIP	0.5566*	1											
EMP	0.3655*	0.3326*	1										
VISIT	-0.0035	0.1192	0.6732*	1									
EXP	0.1186	0.2225*	0.0559	0.0246	1								
ACT	0.0707	0.0922	0.3135*	0.1583**	-0.1731**	1							
PUB	0.1653**	0.4192*	0.1528**	0.2875*	0.1394**	0.1958*	1						
Heritage sites	-0.1254***	-0.1268***	0.2487*	0.3904*	0.1005	0.0866	0.2015*	1					
Hotels	-0.0225	0.1679**	0.3953*	0.4431*	0.3095*	0.1803**	0.4114*	0.6345*	1				
Motorway	-0.1110	-0.0483	0.3731*	0.3565*	0.2445*	0.1018	-0.0313	0.4379*	0.3695*	1			
Institution	-0.1287***	0.1986*	0.5439*	0.5924*	-0.1470**	0.4113*	0.2789*	0.3058*	0.4510*	0.2305*	1		
Age	0.2285*	-0.2236*	0.0730	0.0298	-0.0348	0.0658	0.0150	-0.1018	-0.0374	0.0668	-0.2539*	1	
PIB	0.1528**	0.3902*	0.0981	-0.1441**	0.1851**	0.2626*	0.2411*	-0.0524	0.3778*	-0.1510**	0.2148*	-0.1712**	1

*, **, *** Statistically significant at 0.01%, 0.05% and 0.1%

TABLE 7
Results of Simar and Wilson regression

Variables	DEA SBM <i>Black Box</i>		DEA SBM Network		
	SBM	SBM Supereff.	Overall	Stage 1	Stage 2
Heritage sites	0.0013* (0.0003)	0.0013* 0.0004	0.0001 0.0001	0.0004* 0.0001	3.13E-05 0.0004
Hotels	-0.0835* 0.0287	-0.0835* 0.0278	0.0102 0.0103	-0.0158*** 0.0095	0.0749** 0.0324
Motorways	0.0004 0.0003	0.0004 0.0003	0.0001 0.0001	-5.05E-05 0.0001	-1.68E-05 0.0004
GDP	4.17E-05* 1.44E-05	4.17E-05* 1.40E-05	-2.86E-05* 6.45E-06	-1.40E-06 4.80E-06	-0.0001* 2.22E-05
Institution	0.1835** 0.0753	0.1835** 0.0742	0.3304* 0.0395	-0.1039* 0.0315	1.1899* 0.2075
Age	-0.0010 0.0007	-0.0010 0.0007	0.0004 0.0003	-0.0012* 0.0003	0.0027* 0.0010
C	-1.0448** 0.4281	-1.0448** 0.4095	0.5440* 0.1434	0.4329* 0.1279	1.5785* 0.4331
Sigma	0.2493* 0.0333	0.2494* 0.0317	0.1442* 0.0116	0.1504* 0.0098	0.3410* 0.0420
Wald Chi squared	18.07*	19.95*	91.73*	38.56*	41.33*

***, ** and * represent statistical significance at 1%, 5% and 10% confidence levels, respectively. Bootstrap standard errors are in brackets. The number of bootstrap replications has been set at 5000. Source: authors' own

Table 7 thus shows the results of the Simar-Wilson regression under 5000 iterations¹⁴ between the efficiency indices from the different evaluation models, with respect to six external variables, four concerning the environment and two that emerge from within the museum itself, although none are controlled by the museum managers themselves. The results may be interpreted as follows. When we take efficiency calculated using the conventional SBM model (*black box* and *superefficiency*) as the dependent variable, a direct relation can be seen between cultural potential (heritage sites) and the area's level of economic development (GDP) with museums' levels of performance, whereas there is an inverse relation with tourist intensity (concentration of hotel places). There is a positive relation with the type of institution, in the sense that national museums seem to be more efficient, and that variables concerning accessibility (motorways) and the

¹³ The only variable that displays slightly higher weights is the type of institution. Nevertheless, as an indication of the robustness of the results as regards maintaining the conditions of separability, the Simar-Wilson regression analysis presented below was repeated, removing this variable. Results were seen to be very similar, at least in the significance and sense of the rest of the indicators.

¹⁴ Estimations have been made with the Simar-Wilson Stata package.

museum's age do not appear to be significant. Nevertheless, if we look at the efficiency ratios in the SBM network evaluation model, interpretations are clearer and more easily distinguishable between the two stages. In fact, in stage 1, which focuses on management and cultural programming activities, museum efficiency seems to be positively affected by the cultural environment, whereas it is negatively impacted by hotel concentration, the museum's age and the type of management. Successful museums in terms of efficiency are, therefore, relatively young museums, run by regional authorities and located in cultural areas with a high density of heritage attractions but which are not as congested in terms of tourism, given that they might be non-iconic cultural ensembles located in cities that are not so famous. In contrast, in stage 2, which focuses on attracting visitors, museum performance is positively shaped by tourist concentration, which obviously stems from attracting a larger number of visitors, as well as the type of institution and age; in other words, the most efficient museums here are the older national museums located in popular tourist cities with large hospitality facilities. Broadly speaking, the locations' accessibility does not prove to be significant.

The variable related to the museum's management status (national or regional) partly seems to determine a certain group of museums vis-à-vis their performance and the orientation of their management functions. Specifically, there seems to be a distinction between museums that achieve better efficiency ratios in the first stage and in the second. Bearing this in mind, we computed the efficiency means for state-run museums and for regionally managed museums. The results are shown in Table 8, where the gap in efficiency between the two groups of museums is seen to be virtually unnoticeable in the *black box* model but is more conspicuous when observing the outcomes from the SBM network model. National museums thus emerge, in general terms, as more efficient, but particularly when it comes to attracting visitors, whereas provincial museums (run by regional governments) perform better in the first stage, which is geared towards creating cultural supply, but which does not seem to be enough to attract visitors. This result once again underlines two of the arguments set out previously in this study: firstly, the general *black box* evaluation model fails to reveal the causes of inefficiency in a network of museums, and secondly, museums may be efficient at undertaking tasks related to organisation and creating cultural offer, but this is not enough to attract visitors, which depends on other contextual factors such as complementary cultural facilities and the area's tourist appeal as well as the museums' reputation.

TABLE 8
Efficiency indexes per managing institution

Description	DEA SBM <i>Black Box</i>		DEA SBM network		
	SBM	SBM Supereff.	Overall Score	Stage 1	Stage 2
National museums	0.45	0.46	0.35	0.25	0.42
Provincial museums	0.40	0.42	0.16	0.37	0.16

Source: authors' own

5. Conclusions

This work presents a methodological approach to analyse the efficiency of publicly owned Spanish museums using an SBM network model. Museum performance has traditionally been evaluated by examining the relation between inputs and outputs for the various activities carried out simultaneously, without taking into account possible interrelations between such activities. Applying an SBM network model allows these relations to be considered, and provides information concerning the root cause of the inefficiencies to emerge. Nevertheless, the inefficiencies pinpointed in this manner stem from inside the organisation, although they allow for the possibility that other external variables that make up the context in which the museum engages in its activities, or certain intrinsic features of museums that are beyond managers' specific control, might partly be the cause of certain inefficiencies. In order to identify some of these variables, we apply a truncated bootstrap regression procedure following the Simar-Wilson algorithm 1 (2007).

Our model is grounded on the basic functions performed by museums and considers two stages in the development of the related activities. Firstly, the management stage, which corresponds to activities that shape the institution's cultural supply as the principal output, and through which museums seek to obtain, during the second stage, a final outcome measured in terms of their impact through visitor numbers. We conjecture that the first stage is under the manager's control, since it involves the production of programmed outputs in line with available resources, whereas the second stage is reflected through observed outputs, partly determined and co-formed by visitor demand.

The results to emerge from our research reveal low levels of overall efficiency in the national system of museums in Spain, and in which museums perform better in the first stage than in the second, reflecting a certain trade-off between good performance in how museums are organised and an orientation towards potential visitors. With regard to how much contextual variables might determine museums' efficiency rates, it would seem that the area's cultural endowments do not act as competitive factors, but that they do help to improve organisation, offering a complementary supply of culture in the cultural context. There is also a link between the area's tourist potential and museum performance, which has a negative impact on management efficiency but a positive effect on visitor numbers. Finally, the museum's age and management status have significant effects on efficiency but with contrasting signs with regard to the first and second stages, since it would seem that provincial museums (run by regional authorities) are more efficient in the first stage, whereas national museums perform better in the second.

Summing up, it can be said that organisational efficiency is not enough to attract visitors and that, although we identify other factors not controlled by managers and which might impact on the ability to attract visitors, it would be advisable to reduce the gap between how well museums are organised and their potential demand, perhaps by re-focusing their activities. Clearly, the value of the collection and the museum's reputation play a key role in attracting visitors and are factors that our work has not taken into account due to the difficulty involved in allocating a monetary value to museums' holdings. Further inquiry which takes into consideration the scope of the exhibits on display, the value of the art collection or the museum's reputation as input variables, might provide fresh insights into how well these institutions perform. The results to emerge from the research also reveal the suitability of the technique and methodological design vis-à-vis identifying the drivers of museum efficiency, and may also prove useful for running these cultural institutions as well as for those responsible for public resource allocation policies in the area of cultural heritage.

References

- Akther, S., Fukuyama, H. & Weber, W.L. (2013). Estimating two-stage network slacks-based inefficiency: An application to Bangladesh banking. *Omega*, 41, 88-96.
- Ames, P. (1994). Measuring museums merits. In K. Moore (ed): *Museum management*. Routledge, London, 22-40.
- Andrews, D. F., & Pregibon, D. (1978). Finding the Outliers That Matter. *Journal of the Royal Statistical Society*. 40(1), 85-93.
- Arroyo, L. & Hart, J. (2011). *Museum Accreditation Program*. Taylor & Francis. USA.
- Avkiran, N.K. (2009). Opening the black box of efficiency analysis: an illustration with UAE banks. *Omega*, 37(4), 930-941.
- Banker, R.D., (1996). Hypothesis tests using data envelopment analysis, *Journal of Productivity Analysis*, 7(2-3), 139-159.
- Banker, R.D., & Natarajan, R. (2008). Evaluating Contextual Variables Affecting Productivity. *Operations Research*, 56, 48-58.
- Basso, A., Funari, S. (2004). A Quantitative Approach to Evaluate the Relative Efficiency of Museums, *Journal of Cultural Economics*, 28 (3), 195-216.
- Basso, A. Casarin, F. & Funari, S. (2018). How well is the museum performing? A joint use of DEA and BSC to measure the performance of museums, *Omega*, 81, 67-84

- Benito, B., Solana, J. & López, P. (2014). Determinants of Spanish Regions' Tourism Performance: A Two-Stage, Double-Bootstrap Data Envelopment Analysis. *Tourism Economics*, 20(5), 987-1012.
- Bertachini, E., Dalle Nogare, C. & Scuderi, R. 2018. Ownership, organization structure and public service provision: the case of museums. *Journal of Cultural Economics*, 42 (4), pp 619-643
- Bishop, P. & Brand, S. (2003). The Efficiency of Museums: A Stochastic Frontier Production Function Approach, *Applied Economics*, 35 (17), 1853-1858.
- Carnegie, G.D. & Wolnizer, P. W. (1996). Enabling accountability in museums. *Accounting, Auditing & Accountability Journal* 9(5), 84-99.
- Charnes, A. & Cooper, W. (1962). programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 15, 333-334.
- Cuccia, T., Guccio, C. & Rizzo, I. (2016). The effects of UNESCO World Heritage List inscription on tourism destinations performance in Italian regions. *Economic Modelling*, 53, 494-508.
- Daraio, C. & Simar, L. (2005). Introducing Environmental Variables in Nonparametric Frontier Models: a Probabilistic Approach. *Journal of Productivity Analysis*, 24(1), 93-121.
- De Witte, K., & Geys, B. (2011).Evaluating efficient public good provision: Theory and evidence from a generalised conditional efficiency model for public libraries. *Journal of Urban Economics*, 69(3), 319-327.
- De Witte, K. & Geys, B. (2013). Citizen coproduction and efficient public goods provision: Theory and evidence from local public libraries. *Journal of Operational research*, 224, 592-602.
- Del Barrio, M.J., Herrero, L.C. & Sanz, J. A. (2009). Measuring the efficiency of Heritage institutions: A case Study of a regional system of museums in Spain, *Journal of Cultural Heritage*, 10(2), 258-268.
- Del Barrio, M.J. & Herrero, L.C. (2014). Evaluating the efficiency of museums using multiple outputs: evidence from a regional system of museums in Spain. *International Journal of Cultural Policy*, 20(2), 221-238.
- Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V.V., Sarrico, C.S. & Shale, E.A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2), 245-259.
- Färe, R., Grosskopf, S. & Lovell, C.A.K. (1985). The measurement of efficiency of production, Kluwer-Nijhoff, Boston, MA.
- Farrell, M. J. (1957). The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society. Series A (General)*, 120 (3), 253-290.
- Fernández-Blanco, V., L. C. Herrero & Prieto-Rodríguez, J. (2013). Performance of Cultural Heritage Institutions, In I. Rizzo & A. Mignosa (ed): *Handbook on Economics of Cultural Heritage*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing Ltd, 470-490.
- Fernández-Blanco, V., Rodríguez-Álvarez, A. 2018. Measuring allocative efficiency in Cultural Economics: the case of "Fundación Princesa de Asturias" (The Princess of Asturias Foundation). *Journal of Cultural Economics*, 42 (1), pp 91-110
- Fernández-Blanco, V., Rodríguez-Álvarez, A. and Wisniewska, A. 2019. Measuring the efficiency and marginal cost in the performing arts: the case of the municipal theatres of Warsaw. *Journal of Cultural Economics*, 43 (1), pp 97-119
- Figuroa, V., Herrero, L.C., Báez, A. & Gómez, M. (2018). Analysing how cultural factors influence the efficiency of tourist destinations in Chile. *International Journal of Tourism Research*, 20, 11–24.
- Finocchiaro-Castro, M., Guccio, C. & Rizzo, I. (2011). Public intervention on heritage conservation and determinants of heritage authorities' performance: a semi-parametric analysis. *International Tax and Public Finance*, 18(1),1-16.
- Finocchiaro-Castro, M. & Rizzo, I. (2009). Performance Measurement of Heritage Conservation Activity in Sicily. *International Journal of Arts Management*, 11(2), 29-41
- Frey, B. (1998). Superstar Museums: An Economic Analysis. *Journal of Cultural Economics*, 22(2/3), 113-125.

- Gómez-Zapata, J., Espinal-Monsalve, N. & Herrero-Prieto, L.C (2018) Economic valuation of museums as public club goods: Why build loyalty in cultural heritage consumption?, *Journal of Cultural Heritage*, 30 (1), pp 190-198
- Guccio, C., Pignataro, G. & Rizzo, I. (2014). Evaluating the efficiency of public procurement contracts for cultural heritage conservation works in Italy. *Journal of Cultural Economics*, 38 (1), pp 43-70
- Guccio, C., Martorana, M., Mazza, I. & Rizzo, I. (2016). Back to the Future. The effect of digital technology on the performance of public historical archives. *ACEI Working paper series*, AWP-10-2016.
- Guccio, C., Lisi, D., Martorana, M. & Mignosa, A. (2017). On the role of cultural participation in tourism destination performance: an assessment using robust conditional efficiency approach. *Journal of Cultural Economics*, 41 (2), pp 129-154
- Guccio, C., Mignosa, A. & Rizzo, I. (2018). Are public state libraries efficient? An empirical assessment using network Data Envelopment Analysis, *Socio-Economic Planning Sciences*, 64, 78-91.
- Hammond, C. J. (2002). Efficiency in the provision of public services: a data envelopment analysis of UK public library systems. *Applied Economics*, 34(5), 649-657.
- Herrero-Prieto, L.C. & Gómez-Vega, M. (2017). Cultural resources as a factor in cultural tourism attraction: Technical efficiency estimation of regional destinations in Spain. *Tourism Economics*, 23(2), 260-280.
- Jackson, R. (1988). A Museum Cost Function, *Journal of Cultural Economics*, 12(1), 41-50.
- Johnson, P. & Thomas B., Tourism (1992). *Museums and the Local Economy: the Economic Impact of the North of England Open Air Museum at Beamish*, Edward Elgar, Aldershot
- Kao, C. (2009): Efficiency decomposition in network data envelopment analysis: a relational model. *European Journal of Operational Research*, 192, 949-962.
- Kao, C. (2014): Network data envelopment analysis: A review. *European Journal of Operational Research*, 239, 1-16.
- Kao, C. & Hwang, S-N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, 185(1), 418-429.
- Kneip A, Park BU, Simar L. (1998) A note on the convergence of nonparametric DEA estimators for production efficiency scores. *Economic Theory*, 14(6):783-93.
- Liang, L., Wade, D.C. and Zhu, J. (2008): DEA Models for Two-Stage Processes: Game Approach and Efficiency Decomposition. *Naval Research Logistics*, 55, 643-653.
- Meijer, A. J. (2011). Networked Coproduction of Public Services in Virtual Communities: From a Government-Centric to a Community Approach to Public Service Support. *Public Administration Review*, 71, 598-607.
- Moreno, P. and Lozano, S. (2014): A network DEA assessment of team efficiency in the NBA. *Annals of Operational Research*, 214(1), 99-124.
- Paulus, O. (1995). Museums' Efficiency, Paper presented at the Fourth European Workshop on Efficiency and Productivity Analysis, Conference, CORE, Université Catholique de Louvain.
- Pestoff, V. (2006). Citizens and Co-Production of Welfare Services: Childcare in eight European countries. *Public Management Review*. 8(4), 503-519.
- Pignataro, G. (2002). Measuring the Efficiency of Museums: A Case Study in Sicily, In: I. Rizzo & R. Towse, eds., *The Economics of Heritage. A Study in the Political Economy of Culture in Sicily*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing Ltd, pp. 65-78.
- Reig-Martínez, E., Gómez-Limón, J.A., Picazo-Tadeo, A.J. (2011). Ranking farms with a composite indicator of sustainability. *Agricultural Economics*, 42, 561-575.
- Simar, L., Wilson, P. W. (2002). Non-parametric test of returns to scale, *European Journal of Operational Research*, 139, pp. 115-132.
- Simar, L., Wilson, P. W. (2007). Estimation and inference in two-stage, semiparametric models of production processes, *Journal of Econometrics*, 136(1), 31-64.
- Simar, L., Wilson, P. W. (2011). Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis*, 36(1), pp. 33-53.

- Simon, J., Simon, C. & Arias, A. (2011). Changes in productivity of Spanish university libraries. *Omega*, 39(5), 578-588.
- Taalas, M. (1998). Efficiency of Finnish Museums - Free Disposal Hull Method to Measure Cost Efficiency, Paper presented at the Tenth International Conference on Cultural Economics, Conference, Barcelona, Spain.
- Taalas, M. (1997). Generalised Cost Functions for Producers of Performing Arts – Allocative Inefficiencies and Scale Economies in Theatres. *Journal of Cultural Economics*, 21(4), 335-353.
- Taheri, H., Ansari, S. (2013). Measuring the relative efficiency of cultural-historical museums in Tehran: DEA approach, *Journal of Cultural Heritage*, 14(5), 431-438.
- Tone, K. (2001): A slacks-based measure of efficiency in data envelopment analysis. *European Journal of operational Research*, 130, 498-509.
- Tone, K. (2002): A Slacks-Based Measure of Super-Efficiency in Data Envelopment Analysis. *European Journal of operational Research*, 143(1), 32-41.
- Tone, K. & Tsutsui, M. (2009): Network DEA: a slacks-based measure approach. *European Journal of Operational Research*, 197, 243-252.
- Vitaliano, D. F. (1998). Assessing Public Library Efficiency Using Data Envelopment Analysis. *Annals of Public and Cooperative Economics*, 69, 107-122.
- Weil, S. E. (1995). Performance Indicators for Museums: Progress Report from Wintergreen, *The Journal of Arts Management, Law, and Society*, 23 (4), 341-351.
- Wilson, P. (1993). Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business and Economics Statistics*, 11(3), 319-323.
- Wilson, P. (2008). FEAR 1.0: A software package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences*, 42, 247-254.
- Whitaker, G. (1980). Coproduction: Citizen Participation in Service Delivery. *Public Administration Review*, 40(3), 240-246. doi:10.2307/975377