TOURISM ECONOMICS
The Business and Finance of Tourism and Recreation

Article

Tourism Economics
2021, Vol. 27(6) 1276–1300

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1354816620945440
journals.sagepub.com/home/teu
SAGE

Analyzing online search patterns of music festival tourists

Juan D Montoro-Pons
Universitat de València, Spain

Manuel Cuadrado-García
Universitat de València, Spain

Abstract

Music festivals, as cultural events that induce tourism flows, intermediate both the cultural and travel experience. The present study analyzes online search behavior of potential attenders to a music festival. We hypothesize that the search process reveals latent patterns of behavior of cultural tourists planning to attend music festivals. To this end, information from Google Trends on queries related to three popular music festivals is used to build a network of search topics. Based on it, alternative exponential random graph model specifications are estimated. Findings support the general result of mediated information flows: music festivals induce planning and traveling queries. However, differences relating to the specificities of the cultural event are also found, in particular those regarding what nodes or queries supply the network with more useful information.

Keywords

cultural participation, exponential random graphs models, live music consumption, user-generated data, unstructured data, mediated consumer discovery

Introduction

The increasing relevance of music festivals as cultural events where audiences experience live music has had a twofold impact. Firstly, in the music industry, where curating and bundling cultural content combined with the spatial and temporal concentration of the supply of live music has emerged as a successful business model. Secondly, in tourism-related services, through the linkages and spillovers from cultural engagement into activities that enable such participation. In short, music

Corresponding authors:
Juan D Montoro-Pons, Departamento de Economía Aplicada, Universitat de València, Spain.
Email: Juan.D.Montoro@uv.es
Manuel Cuadrado-García, Departamento de Comercialización e Investigación de Mercados, Universitat de València, Spain.
Email: cuadrado@uv.es

festivals illustrate a type of cultural tourism where cultural participation induces tourism flows. To some extent, music festivals are themselves a tourist destination.

In both cases, music as cultural participation and tourism, consumers face informational problems related to the experiential nature of these activities, which only after consumption can be assessed. Asymmetries of information and unknown quality lead to uncertainty and consumption risks that challenge consumers and affects the extent of their engagement with cultural or touristic activities and, therefore, the performance of the underlying markets. In such cases, online information can be one way of reducing associated risks as it becomes an imperfect substitute for the actual experience (Bei et al., 2004).

The present study analyzes online search behavior about cultural events that, we assume, mediate the process of information acquisition. We use the digital footprint that Internet users generate when searching information about music festivals and propose these are informative and complementary to other standard empirical techniques. The data set of googled terms we use can be interpreted as a set of manifest variables about underlying behavioral patterns of cultural tourists planning to attend music festivals.

Specifically, we posit that related search topics (i.e. terms queried in the same online search session) and the links they form pinpoint the mediating function that certain cultural events serve in cultural and touristic markets. By looking at what is searched, the existing links between queries, and the (in)existence of indirect means to connect two disjointed search terms, we are able to describe the search space of cultural tourists, the process through which searches are associated and how both enable discovery. In short, this article analyzes the online information gathering that cultural events trigger and how it is related to the formation of the choice set of the cultural tourist. Methodologically, we follow a quantitative case study approach (Eisenhardt and Graebner, 2007) and draw on user-generated unstructured data that, after wrangling, allows us to map the process of planning cultural participation to a directed graph.

We collect information on online search activity for three reputed popular music festivals in Spain. The data set identifies the different aspects that are related to the cultural tourist decision-making, from obtaining basic information about the event to planning attendance, traveling, and complementary leisure activities. We process it using a graph layout (topics are nodes and links connect topics that are searched together) and estimate the underlying structure that channels the information acquisition of cultural consumers when planning attendance to a live music event. In so doing, we aim at characterizing what factors act as facilitators of the diffusion of information through the observed graph of search terms and how these reflect consumer behavior. All this to be understood in a framework where actual behavior is preceded by motivation or intention (Ajzen, 1991).

Findings provide evidence of rational addiction and voracity, as well as of spillovers and linkages into other related activities, mainly those related with travel. We also find evidence of the hierarchical structure of search topics, which reinforces the notion of information acquisition as a mechanism that cultural tourists use as a means to reduce risk and uncertainty. Hierarchies of search topics are pinpointed through a core of connected nodes—those supplying the graph with relevant or authoritative information along with others that enable its diffusion—and less accessible peripheral search topics. Furthermore, there are differences across events and the geographical scope of searches, which can be associated with alternative behavioral patterns.

The main contribution of this research is related to the fields of cultural tourism, cultural mediation theories, and theories of asymmetric information in cultural consumption. Drawing on these, we quantitatively show the relevance music festivals have in channeling information gathering that ultimately helps consumer discovery and market creation. This research, in short, is an analysis on mediated information flows in cultural markets that produce informational spillovers. Additionally, we contribute to the literature on cultural tourism by the use of a nonstandard data set and a novel methodology that, to the best of our knowledge, has not been applied before.[1]

The article is organized as follows. First, a brief and general background on cultural tourism and its relation to cultural events and the festivalization of culture is introduced. Here, we include a discussion on the mediated information acquisition that cultural events trigger, which is contextualized within the information search and planning literature in the field of tourism economics and management. Second, an exploration of the data set, the data gathering process, and the formation and classification of the basic units of analysis (nodes and edges of a network of search terms) and its structure and graphical description are introduced. Then statistical models for graphs (exponential random graph models (ERGM)) and testable hypotheses are laid out, followed by the estimation results. The article concludes with a discussion of the main findings.

Background

Music festivals influence cultural and tourism participation decisions through a combination of (i) selection, curation, and bundling—producing and distributing a temporary cultural resource—, and (ii) the geographical concentration of cultural content. As a result, festivals trigger consumers' information acquisition directed toward the reduction of risk and uncertainty that enables cultural tourists' discovery and decision-making. To analyze the search patterns this process activates, first we outline the nature of cultural tourism and discuss the mediating role certain market institutions, such as music festivals, serve.

Festivals and cultural tourism

As cultural tourism sits at the intersection of two activities, any definition implicitly singles out what drives behavior: cultural participation or tourism. Noonan and Rizzo (2017) point out that cultural tourism can be summarized as the analysis of the interconnections between cultural participation and tourism organization and how the former affects the emergence of specific tourism patterns.

Bonet (2013) takes a more restrictive viewpoint when pinpointing that, as most tourists engage in the consumption of goods and services with some cultural component, one should look at what motivates consumers to differentiate cultural tourism from other traveling experiences. In short, the relevant distinction lies on whether culture is the main drive of the touristic experience or just accessory to it and the role that cultural supply plays in explaining tourism.

From this perspective, the evidence shows an association between tourism flows and cultural resources (Borowiecki and Castiglione, 2014), even though the direction of causality is far from settled. Research supports both: cultural resources induce tourism (Cuccia et al., 2016; Guccio et al., 2017) and vice versa (Cellini and Cuccia, 2013). In this respect, observed behavior could reflect the heterogeneity of motivations of tourists, which ranges from pure cultural engagement to a more recreational attitude toward culture (Brida et al., 2016).

The foregoing discussion highlights the demand-side of cultural tourism, where intrinsic cultural motivations explain decision-making. However, it neglects the market creation role of specific institutions such as festivals, which stand out as an urban planning strategy to attract tourism through cultural consumption. The term festivalization describes the increasing relevance of festivals as a medium for cultural consumption. From a policy standpoint, festivalization has been considered as a strategy that structures and organizes leisure and cultural activities and that helps in the reframing of urban spaces (Karpińska-Krakowiak, 2009). Festivals can be seen as a policy tool that relies on cultural events to market cities—a potential driver of success in the positioning of cities—but that exceed the framework and objectives of city marketing (Hitters, 2007; Richards, 2007).

The supply of temporally and geographically constrained cultural resources has been shown to potentially attract tourism flows (Gergaud and Ginsburgh, 2017; Vecco and Srakar, 2017), a market creation effect that is mediated by the differential motivations that local audiences and tourists exhibit (Báez-Montenegro and Devesa-Fernández, 2017; Faulkner et al., 1999; Herrero et al., 2012). All in all, as festivals enter the choice set of cultural consumers, decisions unrelated to cultural participation emerge as a by-product.

Mediated information acquisition

Although the literature has singled out the potential of cultural events as tourism facilitators, it has overlooked how this function is performed. Cultural consumption draws heavily on actors that select, signal, and legitimate cultural artifacts and in so doing, promote consumer awareness of the cultural supply and facilitate consumers' search and discovery (Janssen and Verboord, 2015). In short, cultural markets rely on organizations, such as music festivals, that induce market creation (Hiller, 2016).

Two aspects are noteworthy vis-à-vis festivals. First, when attendance is mostly driven by the cultural event, that is, culturally motivated, the flow of nonlocal audiences can be described as a prototypical instance of *pure cultural tourism*: the cultural activity is to be seen as the main reason for traveling and not just as a complementary activity to be undertaken at destination.

Second, as a consequence, festivals not only influence cultural consumption decisions but also tourists' decision-making. In this respect, music festivals serve a mediating function that spills over the tourism experience, helping cultural tourists to reduce perceived risk and uncertainty. Information collection and planning are part of the decision-making process tourists undertake. In this respect, by restricting the search space of cultural tourists—a process we define as mediated information acquisition—festivals induce the dissemination of information and reduce consumers' search costs. In short, festivals shape the search mechanisms that drive cultural tourists' decision-making.

Search, information collection and planning as part of consumers' decision-making process are central topics in the tourism literature. Research around these issues has highlighted, among others, the impact of information and communication technologies, how widespread they are, and their uses.

For consumers, the use of digital technologies has reduced costs and facilitated access to information, although it has dramatically increased the complexity in decision-making with the number of choices users face (Buhalis and Law, 2008). Note that, as the complexity of consumers' choices escalates, the value to consumers of actors that restrict and mediate the search space, hence simplifying decision-making, also increases.

Users rely on different online tools to acquire information and simplify the planning of the travel experience (Casaló et al., 2010; Chung and Buhalis, 2009). While these have been found to be widespread across all customer segments, there is some variability in the use of particular resources, such as social media, and the utility derived from and weight placed on them (Llodrà-Riera et al., 2015; Xiang et al., 2015). The planning process has been conceptualized through the sources, search patterns, and dimensions of the search process, the activities involved in each of the steps, and the role that the different travel-related technologies play at each step and toward the experience satisfaction (Ho et al., 2012; Huang et al., 2017; Pan and Fesenmaier, 2006; Papathanassis and Knolle, 2011).

In this regard, a central topic is that of travel planning as a strategy to deal with uncertainty and risk. How is risk assessed and its influence on search behavior have been found to have a potential mediating effect in the information search process at the planning stage of the travel (Björk and Kauppinen-Räisänen, 2015). In this context, risk is associated with the choice of information sources: when planning traveling, the higher the risk perception the more likely travelers move beyond informal information sources, which on the other hand could enhance the role of organizations or actors involved in the process.

Moreover, and from the supply side, technologies that provide consumers with information at the planning stage of the tourism experience could be potentially cost-effective means of marketing destinations (Cox et al., 2009; Litvin et al., 2008; María Munar, 2011; Pan and Li, 2011). To put it differently, from the organizations' perspective, inducing and disseminating information about a destination (or a cultural event for that matter) could raise participation.

To sum up, as collecting information is central in tourism decision-making, the specific actors that help disseminate that information influence market outcomes. Cultural tourists embed travel planning, search, and information gathering within cultural attendance, which changes the decision-making framework. Against this background, we aim at identifying the structural mediating role music festivals play in the diffusion of information and how they influence consumers' choice set through the search processes they bring about.

The data set

The empirical research is based on a quantitative analysis of the Internet search terms (and the relations between them) googled by potential attenders to three music festivals in Spain. The increasing access to online user-generated content has not only enlarged the toolset of empirical research but also shifted its focus (Artola et al., 2015; Blazquez and Domenech, 2018; Jun et al., 2018). In the case of cultural participation, where behavioral complexity can be challenging, it allows to obtain valuable information from individuals that might not be elicited with more standard techniques such as surveys (see Scuderi and Dalle Nogare, 2018; Stephens-Davidowitz, 2014).

Selection of cultural events

The audience share of large music festivals has been rising in Spain in the past few years (SGAE, 2018). From 5% of the total attendance of popular music performances in 2008, it jumped to over 21% in 2017. Furthermore, large festivals represent in 2017 over 50% of total income in the sector, up from 20% in 2008. Interestingly, these figures are at best a lower bound for income and attendance as they refer to a handful of festivals classified as large events. Altogether, data show an increasing concentration in live music led by music festivals. This article analyzes online search activity around three renowned live music events in Spain.

The festivals included in the sample meet two criteria: (i) they attract a significant share of non-local audiences (such that these qualify as cultural tourists) and (ii) generate enough online search activity. Given these constraints, a natural choice is large music live events: Table 1 lists the top 10 music festivals in Spain using data available from the website of the Spanish association for music promoters (http://www.apmusicales.com), which ranks festivals by attendance. While this list includes a heterogeneity of live events, two aspects are worth mentioning.

Table 1. Ranking of music festivals in Spain (2018).

Festival name	Attendance[a]	Days	Concerts	Location	Since
Arenal Sound	300	6	98	Burriana	2010
Medusa Sunbeach	300	6	159	Cullera	2014
Mad Cool	240	3	143	Madrid	2016
Primavera Sound	220	4	261	Barcelona	2001
Viña Rock	210	3	122	Villarrobledo	1996
Rototom Sunsplash	208	7	247	Benicassim	1994
Festival Internacional de Benicassim	170	4	136	Benicassim	1995
Dreambeach	155	5	112	Cuevas del Almanzora	2013
Weekend Beach	140	4	140	Torre del Mar	2015
Sónar	126	3	138	Barcelona	1994

[a]Figures in thousands for the whole duration of the event.

First, the ranking shows a mix of long-standing reputed festivals, such as *Festival Internacional de Benicassim* (*FIB*) or *Sónar*, along with newcomers such as *Arenal Sound* and *Medusa Sunbeach* (both set up in the 2010s). As the purpose of this research is to identify cultural consumer traits that emerge from online searches, we need to constrain the variability of consumer uncertainty to that related to planning attendance and (maybe) the lineup and not to the quality of the event itself. Established festivals are therefore a natural choice: they have an accumulated reputation that is reflected on its brand, which provides valuable information about cultural supply for consumers.

Second, festivals in Table 1 can be classified, in general terms, across two dimensions: location and lineup proposal. As for the former, two groups can be identified: those that take place in an urban environment (*Primavera Sound* and *Sónar* in Barcelona, and *Mad Cool* in Madrid) and those that are held in a rural location (rest of the festivals in Table 1). It is noteworthy that with one exception (*Viña Rock*), all of the latter take place in seaside resorts, which reinforces the link of festivals to touristic destinations.

The lineup proposal produces a more heterogeneous classification of live events. However, by looking at how broad is the scope of the music genres festivals portray, one could identify two types. On the one hand, niche or specialized festivals such as *Sónar* (with an experimental/electronic/avant-garde focus) or *Rototom* (a reggae music festival). On the other hand, festivals that have a more general appeal by including in their lineup a wide array of music genres. Here one can include festivals such as *Primavera Sound* or *FIB*, which, although generally classified as alternative rock festivals, have actively pushed their boundaries by enlarging the scope of genres in its lineup to include urban music and hip-hop along with the more expected mix of rock, pop, and R&B.

Taking everything into consideration, three festivals are selected: *FIB*, *Primavera Sound*, and *Sónar*. These are reputed festivals that portray a mix of location (urban/beach resort) and genres (specialized/broader appeal). Furthermore, the selected events rank on top vis-à-vis the share of nondomestic audiences they attract. Using figures from the official report of the music promoters association (APM, 2018), *FIB*, *Primavera Sound*, and *Sónar* are the only three music festivals in Spain whose share of nondomestic attenders exceeds 50%. This choice allows us to identify general search patterns emerging from large established music festivals that attract significant flows of tourists but also differential ones based on location and genres.

Table 2. Time frame of top searches (all dates refer to 2018).

	Festival			Search queries	
	Announced	Starts	Ends	Starting	Ending
Festival Internacional de Benicassim	24 January	19 July	22 July	23 January	22 July
Primavera Sound	29 January	28 May	4 June	5 January	4 June
Sónar	25 January	14 June	16 June	23 January	16 June

Data gathering

To collect the data set, we use data from Google Trends (GT) on queries most frequently performed when searching for web information on the three music festivals selected. GT main functionality is that of providing an index of the volume of queries of any given web search term in a particular geography. Internet search indexes have been widely used in tourism economics to improve the predictive power of time series models. Particularly, the literature abounds with applications of web search indexes to predict inflows of tourists (Artola et al., 2015; Bangwayo-Skeete and Skeete, 2015; Bokelmann and Lessmann, 2019; Padhi and Pati, 2017; Yang et al., 2015).

We draw on the complementary information that GT provides in relation to a search query. Particularly, we make use of top searches: these are terms most frequently queried along with the term entered in the same search session, within the chosen category and country (or region). By using top searches, we produce a database of terms that Google users most frequently queried along with the festivals analyzed. Furthermore, two geographies are selected: world searches and searches performed in Spain. Overall, we produce a database of search terms across three events and two geographies.

Data gathering was automated through the use of an application programming interface (API) for GT.[2] The data collection proceeded in a stepwise fashion. First, the initial or entry search terms—*FIB*, *Primavera Sound*, and *Sónar*—were disambiguated using the *suggestions* functionality of GT. It allows us to unambiguously select only those queries that refer to the festival name. This is relevant as running a query for a search term like *Sónar*, returns several results each one with a unique code identifying its meaning. Among them one finds: {“Sonar”; “Topic”}, {“Sonar”; “Mobile application”}, and {“Sónar”; “Music Festival”}. In this case, only top searches emerging from the latter are of interest. We performed this procedure on the three search terms (*FIB*, *Primavera Sound*, *Sónar*) selecting the codes that uniquely identify each music festival.

Second, for each search term, we collect the list of top searches. Note this list is contingent on the geography of the searches (world and Spain) and the time period considered. As for the latter, Table 2 shows the details of the search. All search queries collected span from (roughly) when the (preliminary) lineup is officially announced (around the end of January in all three cases) until the festival ends. The reason for the starting date is apparent: as a festival announces its lineup, it starts to generate buzz around it and (we suspect) online activity. It should be noted that this produces a somewhat longer time period for *FIB* as it takes place later in the year (mid/end July). Furthermore, this time period queries is split into five equal time intervals. Then, top searches are retrieved for each query term given the geography and time interval.

To find cross-links between top searches (or back-links to the festival search term), we proceed one step further by looking at top search terms related to those found in the first stage. We retain

Table 3. Cross-tabulation of nodes by categories, event, and geographical scope of the search (world/Spain).

	FIB		*Primavera Sound*		*Sónar*	
	World	Spain	World	Spain	World	Spain
Event planning	0.04	0.06	0.04	0.13	0.06	0.14
Festival	0.30	0.42	0.37	0.42	0.26	0.50
Leisure	0.02		0.02	0.03	0.06	
Media	0.02	0.03	0.04	0.05	0.06	0.09
Miscellanea	0.14	0.23	0.17	0.13	0.22	0.09
Music and musicians	0.12	0.13	0.12	0.16	0.14	0.14
Travel	0.36	0.13	0.25	0.08	0.20	0.05

only information on searches that are linked to any of the initial search (festival) or top searches in the first round. As a result, we gather a data set composed of search keywords and the connections between them. Let a graph N be defined as a collection of nodes V, and the ties between them L, that is, $N = (V, L)$. Then, the data set of festival-related searches can be represented as a directed graph, in which nodes are search terms and edges appear when a node is in the list of top searches of another node.[3]

The resulting graph shows queried terms and connections between them as cultural consumers gather information surrounding a cultural event when planning attendance. It also describes the scope of interests that emerge as potential attendees search about and become aware of the cultural supply surrounding a music festival. This process is assumed to have an impact in terms of market creation as it defines the choice set of consumers planning to attend a cultural event.

Classification of nodes

While the data set itself has been transformed into a structured graph, information about the content of the searches (the specific queries users make) is textual and as such unstructured. To process this information, nodes have been classified into categories according to the query they refer to. By looking at categories, we reduce the inherent complexity of the analysis and are able to model and quantify the patterns that emerge within and between these categories.

After revising all the queries produced by each node retrieved, the following categories are considered: (i) music festivals (a category that includes not only the festivals analyzed but other related festivals individuals search for); (ii) event planning (here we include all the information surrounding the festival, such as lineup, ticketing, stages, or locations of venues among others); (iii) travel and travel-planning searches (accounts for information including airlines, accommodation, currency, language, and geographical searches such as cities or countries); (iv) music and musicians (including queries that refer to performers, albums, genres, etc.); (v) leisure activities (mainly searches for cultural activities and cultural institutions, sports and sport events, and nightlife); (vi) media (those queries involving traditional and online media); and (vii) miscellanea as a residual category.

Table 3 shows the cross-tabulation of category of the search nodes and music festival. A breakdown for the geographical scope of searches has been included, showing the relevance of some categories (e.g. travel) when considering the intention to attend.

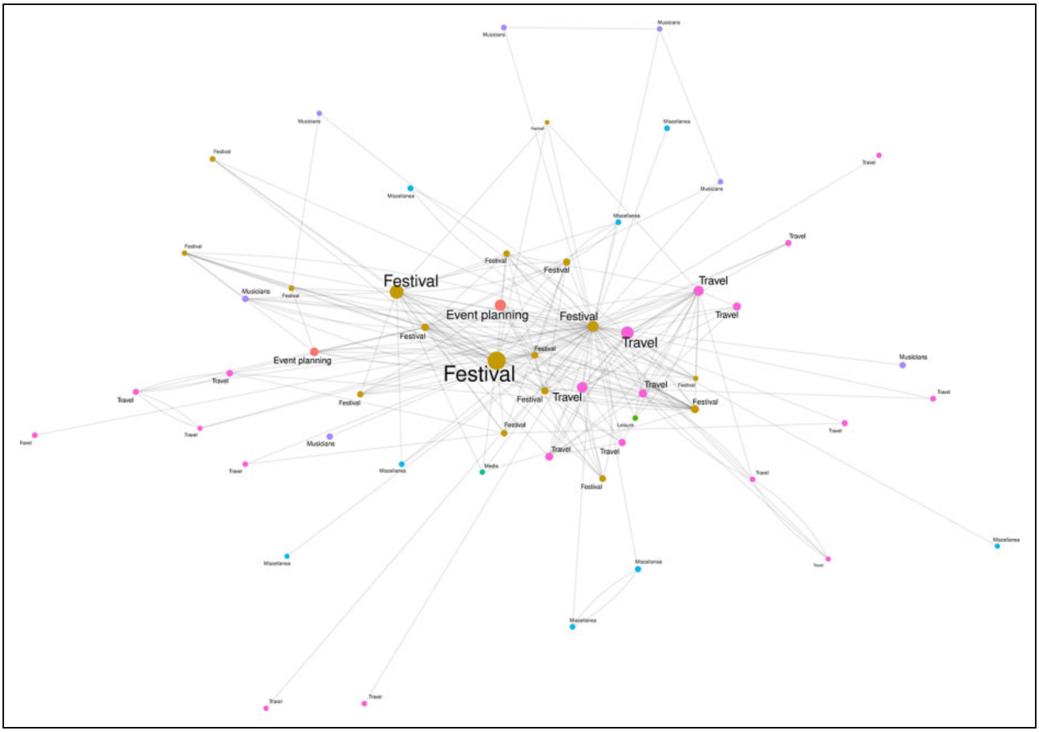

Figure 1. Network of search terms: *FIB*. Geographical scope of searches: World.

The full list of nodes (or search terms) for the events are listed in the Online Appendix. For each node, the table includes full name of the query (node label), the topic in which GT classifies it (Google classification), and the category we assign to each node. From it, it becomes apparent that a small fraction of all search terms are indeed unrelated to the events under consideration. These nodes could be seen as noise (most likely mistakes while entering a search term) that anyhow we retain in the data set.

Structure and description of the data set

After processing the raw data using software tools (Butts, 2008), the resulting networks of search terms are displayed in Figures 1 to 6. We use the category label to identify each node to simplify the display and interpretation of the data set. Furthermore, the size of each node and its label are proportional to its centrality in the graph. Centrality measures summarize the prominence, importance, or popularity of the different nodes (search terms) in the graph. We use authority centrality, whose rationale goes as follows: a node is important when it contains valuable content and hence receives links from other important nodes; conversely, nodes with fewer incoming links have low centrality. Authority centrality measures the extent to which nodes contain relevant information and thus are reached by other nodes that either contain valuable information themselves or facilitate (bridge) connections.

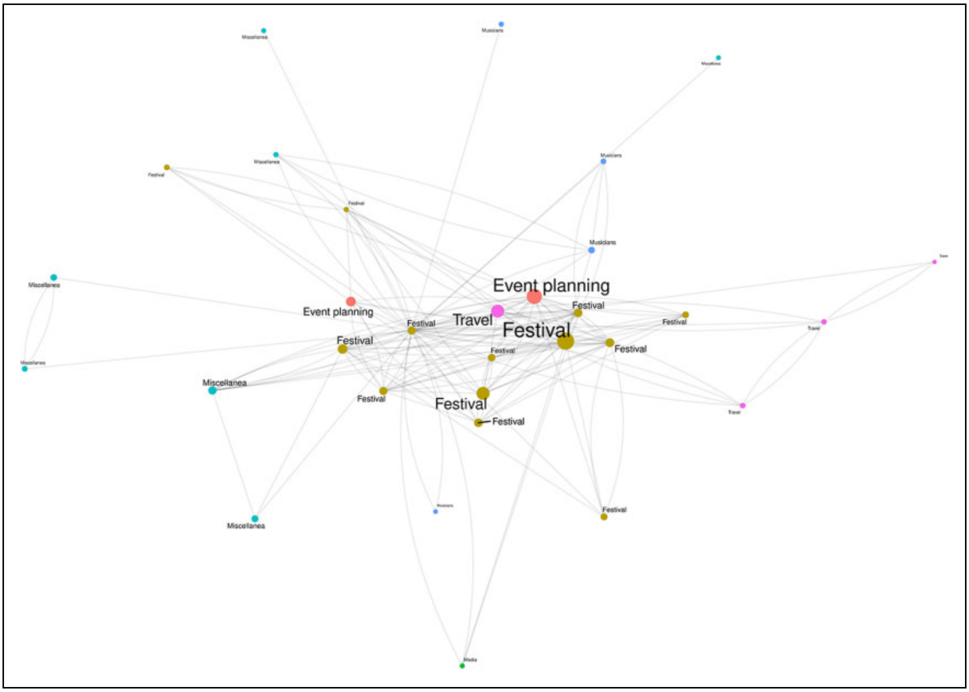

Figure 2. Network of search terms: *FIB*. Geographical scope of searches: Spain.

Three findings are evident from the graphical inspection. Firstly, in all cases, graphs contain a core of densely connected nodes and a periphery of (relatively) isolated end points. Furthermore, search terms that refer to music festivals are overrepresented in this core. This suggests that, in planning participation, live music consumers include and consider a diversity of cultural events in their choice set.

Secondly, and not unexpectedly, gathering travel information (planning the trip by searching about the country/region/city, how to reach there, and where to stay) has a differentiated role depending on the geographical scope of the search. When considering searches from all over the world, and from the perspective of their centrality in the network, traveling comes second in importance to festivals. On the contrary, searches within Spain give prominence to other search categories, such as collecting practical information about the event (planning) or about the musicians in the lineup.

Thirdly, the distribution of the importance or popularity of the different search terms seems to depend on the scope of the festival. Recall that a node's centrality is given by its size. Then, the distribution of centrality is more egalitarian in the more specialized event (Figures 5 and 6) than in the two festivals with a broader appeal (Figures 1 to 4). This suggests patterns of Internet information acquisition that exhibit a more hierarchical structure for *FIB* and *Primavera Sound*—fewer nodes are relevant sources of information in the search process—compared with the more egalitarian distribution of the prominence of nodes for *Sónar*—where most nodes carry equally valuable information.

Nevertheless, we should note that the complexity of the relations emerging between nodes makes the graphical representation of a network of limited value when it comes to obtain additional

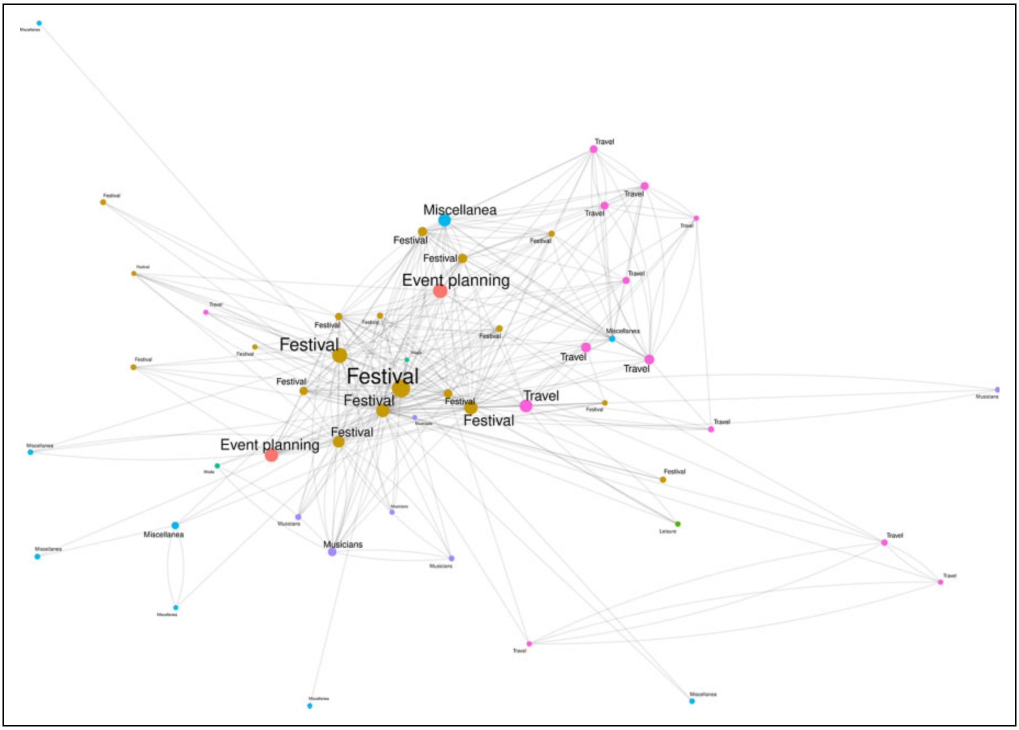

Figure 3. Network of search terms: *Primavera Sound*. Geographical scope of searches: World.

information about its structural properties and how these influence the formation of edges. To do so, we propose a statistical model to make inference on the emergence of patterns in graphs.

Methods

Next, we discuss the empirical approach that allows to analyze the determinants of the emergence of links or edges between the nodes (i.e. search terms) in the network. The statistical analysis produces two applied results. Firstly, it allows us to identify the function that the explicit structural properties of the network serve in triggering consumer planning and discovery patterns. Secondly, and based on the different structural properties of the three cases undertaken, it pinpoints the differential traits in information search of potential attenders to different cultural events.

Exponential random graph models

The empirical analysis draws on a probabilistic representation of graphs, where observed edges are one realization of a random variable. Define a graph $N = (V, L)$ and denote $\mathbf{Y}$ as its adjacency matrix. Let Y be a random variable such that

$$Y_{ij} = \begin{cases} 0 & \text{there is no tie from } i \text{ to } j \\ 1 & \text{there is a tie from } i \text{ to } j \end{cases}$$

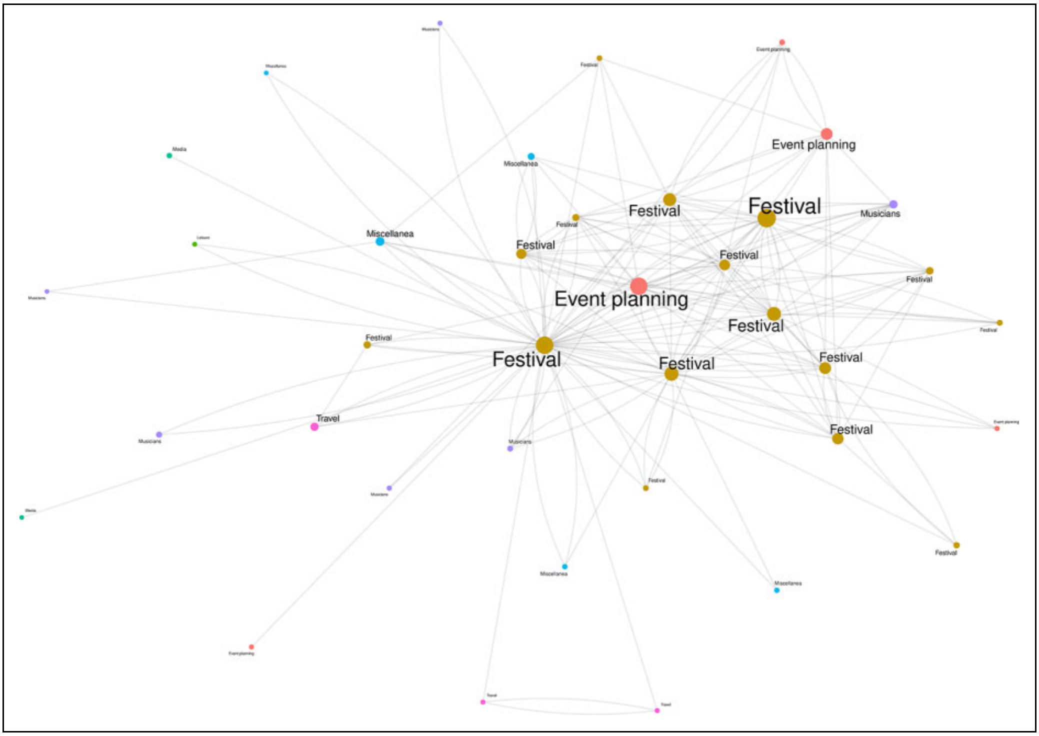

Figure 4. Network of search terms: *Primavera Sound*. Geographical scope of searches: Spain.

Let $\mathbf{y}$ be a realization of this random variable. Then, an ERGM defines the probability of observing a specific network as

$$P(\mathbf{Y} = \mathbf{y}) = \frac{\exp\{\eta^t g(\mathbf{y}, \mathbf{X})\}}{\kappa(\eta)}, \tag{1}$$

where $g(\mathbf{y}, \mathbf{X})$ represents any possible network statistic, which depends on structural features of the network as well as covariates $(\mathbf{X})$ describing node or link properties that are hypothesized to affect the probability of this network forming; η represents the parameters defining the formation of ties, and $\kappa(\eta)$ is a normalizing constant that ensures probability adds up to one. The goal is to obtain maximum likelihood estimators for the parameter vector η from the observed network $\mathbf{y}$. This implies using Markov chain Monte Carlo (MCMC) maximum likelihood estimation for models in which edge formation processes are endogenous (the case under consideration).

Covariates

To estimate an ERGM for the network of search terms, we specify $g(\mathbf{y}, \mathbf{X})$ in expression (1) to include measures of the structure of the network and nodal and dyad-related covariates. These account for the heterogeneity in the formation of ties due to node, interaction, and structural effects.

Node effects determine the *sociality* or propensity of nodes to form links with other nodes. This propensity is measured through the degree or popularity of the node (i.e. search term). Note that, as

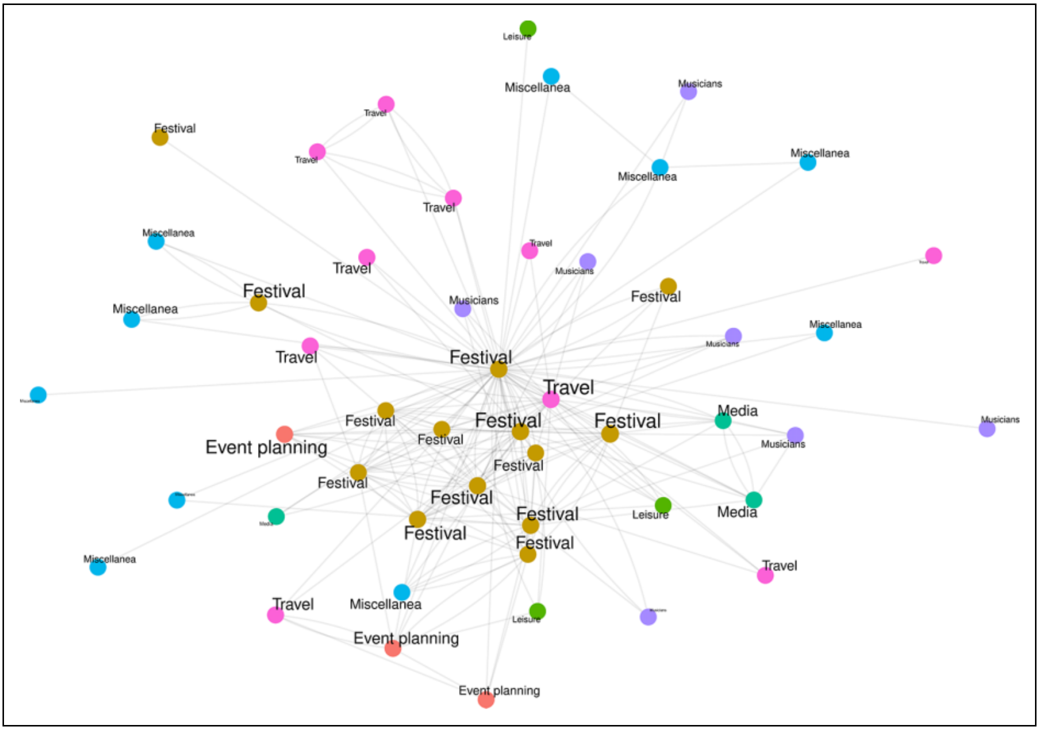

Figure 5. Network of search terms: *Sónar*. Geographical scope of searches: World.

the graph is directed, degree has multiple meanings: out-degree or number of edges originating from a given node; in-degree or number of edges incident to it; and total degree as the sum of both. These are grouped by category of nodes to produce a meaningful interpretation of the effect of search topics within a category on the structure of the graph. When possible, to measure how a particular category increases or decreases the odds of the formation of an outgoing/incoming edge, two terms will be included for each node category. However, in some specifications, due to the sparseness of the resulting graph, total degree is used.

Interaction effects identify the tendency of two nodes to form an edge based on some measure of proximity. The underlying notion is that of nodes clustering together according to their similarity or dissimilarity. Here one distinguishes between assortative mixing or homophily, when links tend to emerge between similar nodes, and dissortative mixing or heterophily, when dissimilarity between nodes induces the formation of ties.

We use a node's category to determine the existence of either homophily or heterophily in the web-search patterns of cultural consumers. Assortative mixing or homophily occurs when nodes in a search category tend to cluster and be segregated from nodes within other categories. If this is the case, reaching a node or search term within the category increases the likelihood of it leading to other searches within the same category, increasing the depth of the information acquisition on that topic. Heterophily, on the other hand, implies a search category creating spillovers on other categories, expanding the breadth of the information acquisition. Consequently, the model includes the inclination of nodes within categories to form ties with similar/dissimilar nodes.

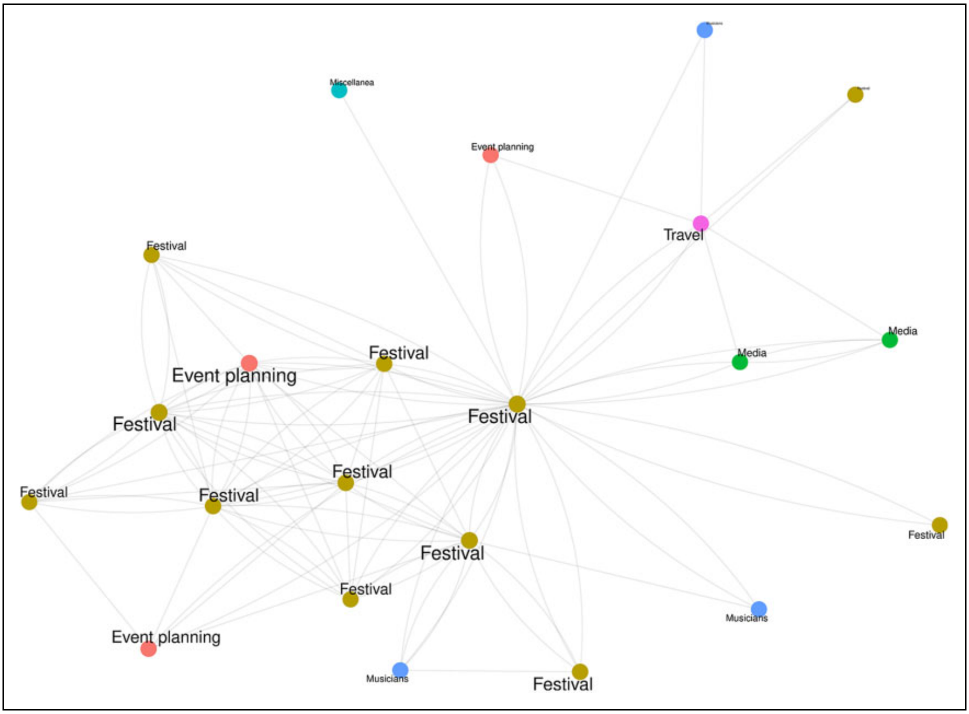

Figure 6. Network of search terms: *Sónar*. Geographical scope of searches: Spain.

Table 4. Mechanisms driving edge formation: List of network effects and covariates.

Mechanism	Description	Covariate
Node attribute*	*Tendency to form (out/in/any) ties*	node(oi∅).Festival
		node(oi∅).Planning
		node(oi∅).Travel
		node(oi∅).Music
		node(oi∅).Media
Interaction effects*	*Tendency homogeneous/heterogenous match*	nodematch.Festival
		nodematch.Planning
		nodematch.Travel
		nodematch.Music
		nodematch.Media
Structural properties	*Density of network edges*	edges
	Tendency to reciprocate	mutual
	Edgewise shared partner distribution	gwesp

*A residual category is included.

Structural properties of the graph are also incorporated in the model. First, we include a measure of the density of the network as a function of a homogenous edge probability (the control variable *edges*, analogous to the intercept in a linear regression) and a term for the reciprocity of ties, *mutual*. Second, an expanded measure of transitivity is also included through the geometrically weighted edgewise partner distribution (*gwesp*), which accounts for the tendency of two nodes that share partner(s) to be also connected. If positive, increasing the number of shared partners increases the likelihood of tie formation. However, link formation exhibits decreasing marginal returns to additional shared nodes. To some extent, *gwesp* measures the effect of local clustering on the likelihood of the dissemination of ties.[4]

Table 4 lists all covariates. Note that the choice of covariates for the different models is determined by the convergence of the estimation process and/or model nondegeneracy.

Finally, two models were estimated: (i) a time-collapsed model in which all ties formed within the sampling period are included regardless of the period(s) of time they are active; and (ii) a separable temporal exponential random graph models (STERGM), which explicitly model the dynamics of tie formation and persistence over time.

Hypotheses

Based on the proposed empirical model, a set of hypotheses in relation to the mechanism that drives connections between searches are formulated. The mechanism driving information acquisition is the observed outcome of festivals' mediation in cultural and tourism markets: it channels the dissemination of information and, in so doing, reduces uncertainty and perceived risk.

First, the central mediating role of music festivals is hypothesized. Mediation is observed through the centrality in the graph of web search activity.

Hypothesis 1: Festival-related searches induce link formation.

Second, we assume that risk and uncertainty of traveling to the festival increases for nonlocal audiences. Therefore, it should be reflected on the connectivity of traveling nodes which are expected to form more ties when the scope of searches is unrestricted (i.e. worldwide searches).

Hypothesis 2: Travel-related terms have a greater tendency to form links in worldwide searches.

Furthermore, interaction effects between searches are expected to emerge. In this respect, it is hypothesized that searches on music festivals spill over other search categories, facilitating discovery.

Hypothesis 3: Festival-related searches exhibit a bias in favor of connections with dissimilar network nodes. This effect broadens information acquisition.

As for searches within other categories, we expect them to be subordinated and lack expansiveness. To the contrary, they are assumed to deepen information acquisition within the category. This case is expected to be relevant in the case of traveling and planning activities: reaching these nodes means that searches will be circumscribed to this domain, allowing users to intensify the information gathering on these particular subjects.

> **Hypothesis 4:** Nonfestival searches exhibit a bias in favor of connections with similar network nodes. This effect deepens information acquisition.

Finally, structural features such as the connectivity derived from clustering have an impact on how knowledge flows through the network, which, in turn, facilitates consumers' discovery process as nodes sharing partners will be very likely linked. In other words, indirect connections induce direct links.

> **Hypothesis 5:** Network structures (local clusters) facilitate the dissemination of links.

Results

We estimate a static (ERGM) and dynamic (STERGM) specification for the different observed graphs. Results are provided in Tables 5 to 7. Each specification includes covariates and the geographical scope of the searches analyzed (world and Spain). Estimates include standard errors and statistical significance, and coefficients are to be interpreted as the conditional log-odds ratio of a tie. In the static framework, a positive (and significant) coefficient implies that the covariate increases the probability of an edge (a negative sign having the opposite interpretation). When time is explicitly considered, then two equations (one for tie formation and one for tie dissolution) are estimated for each graph, and coefficients reflect the probability of a tie forming or its persistence: a positive and significant coefficient increases the likelihood of tie-formation or its persistence over time.

To evaluate the quality of the models, two strategies have been undertaken. First, the achievement of convergence of MCMC has been verified in all models. In this respect, no correlation in the time pattern of the MCMC chain is apparent; moreover, histograms of the difference of observed and simulated sample statistics were roughly bell-shaped and centered at 0. Next, estimated models have been used to simulate networks. Based on these simulations, the 95% confidence interval for the degree distribution is produced and compared to the degree distribution of the observed network. Note that the goodness-of-fit of an ERGM is based on its ability to produce simulations whose distribution (for specific network statistics) include the values of the observed network. In our cases, only small parts of the degree distribution of the observed graphs sit outside the confidence-range bands provided by the models. This allows us to conclude with the adequacy of the estimated models.

ERGM estimation results

Table 5 shows estimation results for the different static graphs. Structurally, ties are less likely than expected by chance (negative sign of *edges*) while reciprocity of ties is a feature of these graphs.

At the node level, we find search terms related to music festivals category tend to generate and attract more edges than those expected under a pure random process. Furthermore, intention to attend (gathering travel and event-specific information) affects information flows. Covariate *node(io).Planning* is found to be unambiguously related to incoming edges in every graph describing search behavior (negative impact on outgoing edges in model 3). As for *node(io).Travel*, estimates are consistent with alternative accounts of its impact, increasing the likelihood of incident links for the broader geographical scope (i.e. world searches) plus model (6); on the other hand, the odds of a travel-related node being a sender of a link are substantially reduced for models (1) and (3)—unrestricted geographical searches for the two alternative rock festivals analyzed. Search terms classified as media were significant sender of edges in models (5) and (6), which could be related to

Table 5. Exponential random graph estimation results.

	Dependent variable: observed network					
	FIB		Primavera Sound		Sónar	
	(1)	(2)	(3)	(4)	(5)	(6)
Structural properties						
edges	−6.482***	−5.048***	−4.357***	−5.190***	−5.031***	−5.439***
	(0.796)	(0.523)	(0.360)	(0.578)	(0.396)	(1.078)
mutual	1.655***	2.727***	2.183***	2.883***	1.864***	3.935***
Node-level effects						
	(0.289)	(0.366)	(0.238)	(0.324)	(0.313)	(0.575)
nodeo.Festival	3.843***	2.399***	2.029***	2.999***	2.757***	2.556**
	(0.760)	(0.530)	(0.316)	(0.588)	(0.381)	(1.120)
nodei.Festival	3.059***	2.065***	1.540***	1.391**	1.310***	0.668
	(0.826)	(0.576)	(0.410)	(0.573)	(0.397)	(0.939)
nodeo.Planning	−1.209	0.416	−1.910**	0.207	0.652	0.963
	(1.080)	(0.510)	(0.744)	(0.535)	(0.535)	(1.169)
nodei.Planning	2.390***	2.552***	2.463***	1.164***	1.322***	1.554*
	(0.440)	(0.477)	(0.364)	(0.378)	(0.367)	(0.886)
nodeo.Travel	−1.516**	−0.249	−2.437***	−1.111	−0.790	−0.832
	(0.626)	(0.564)	(0.554)	(1.105)	(0.589)	(1.529)
nodei.Travel	0.706**	0.249	0.975***	−0.183	0.833***	2.341**
	(0.335)	(0.516)	(0.271)	(0.570)	(0.286)	(0.959)
nodeo.Music	−0.422		−0.121	1.159**	0.588	1.246
	(0.620)		(0.349)	(0.501)	(0.449)	(1.208)
nodei.Music	0.337		0.299	−0.780*	−0.139	0.022
	(0.429)		(0.334)	(0.472)	(0.378)	(0.966)
nodeo.Media			0.053		1.479***	2.078*
			(0.520)		(0.455)	(1.220)
nodei.Media			−0.600	−0.927	−0.029	−0.301
			(0.669)	(0.768)	(0.494)	(1.057)
Interaction effects						
nodematch.Festival	−1.855**	−0.947*	−0.829**	−0.948*	0.057	0.252
	(0.777)	(0.530)	(0.340)	(0.514)	(0.369)	(0.467)
nodematch.Planning				0.950	1.020	
				(0.841)	(1.107)	
nodematch.Travel	4.683***	4.127***	3.856***	4.632***	2.537***	
	(0.889)	(0.840)	(0.592)	(1.309)	(0.624)	
nodematch.Musicians	3.409***	2.092**	1.517**			
	(1.094)	(1.013)	(0.654)			
nodematch.Media					2.192**	
					(0.918)	
nodematch.Miscellanea	3.067***	1.997***	1.243**		0.954	
	(1.022)	(0.729)	(0.583)		(0.790)	
Akaike Inf. Crit.	1230.968	594.877	1549.511	803.767	1047.112	350.624
Bayesian Inf. Crit.	1315.426	652.899	1643.640	877.246	1145.777	404.386
Geographical scope	World	Spain	World	Spain	World	Spain

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

Table 6. STERGM estimation: Network formation.

	Dependent variable: Observed network					
	FIB		*Primavera Sound*		*Sónar*	
	(1)	(2)	(3)	(4)	(5)	(6)
Structural properties						
edges	−6.689***	−5.807***	−5.731***	−6.709***	−6.086***	−4.702***
	(0.469)	(0.462)	(0.296)	(0.615)	(0.339)	(0.477)
mutual	−0.695*	1.154***	−0.016	1.019***	0.429	2.042***
	(0.408)	(0.331)	(0.245)	(0.359)	(0.338)	(0.376)
gwesp	0.725***	0.398***	0.734***	0.482***	0.479***	0.418***
	(0.082)	(0.080)	(0.058)	(0.089)	(0.086)	(0.095)
Node-level effects						
node.Festival	1.846***	1.719***	1.094***	2.100***	1.511***	0.584
	(0.444)	(0.453)	(0.293)	(0.606)	(0.321)	(0.448)
node.Planning	0.831**	0.867***	0.003	0.350	0.803***	−0.135
	(0.360)	(0.293)	(0.291)	(0.303)	(0.281)	(0.343)
node.Travel	−0.313	−0.124	−0.246	−1.017	−0.052	0.511
	(0.275)	(0.340)	(0.204)	(0.742)	(0.268)	(0.404)
node.Media					−0.480	
					(0.447)	
Interaction effects						
nodematch.Festival	−0.992*	−1.002*	−0.962***	−1.705**	−0.829**	−0.046
	(0.532)	(0.525)	(0.366)	(0.676)	(0.420)	(0.555)
nodematch.Planning					0.893	
					(1.085)	
nodematch.Travel	2.569***	4.076***	2.617***		0.955	
	(0.611)	(0.795)	(0.444)		(0.812)	
Akaike Inf. Crit.	1106.229	808.223	1510.928	700.874	1024.369	500.325
Bayesian Inf. Crit.	1165.396	857.471	1568.699	746.933	1095.935	538.328
Geographical scope	World	Spain	World	Spain	World	Spain

Note: STERGM: separable temporal exponential random graph model.
*p < 0.1; **p < 0.05; ***p < 0.01.

the experimental and avant-garde nature of the event, where cultural mediation by specialized actors could prove to be relevant in consumer decision-making.

As for interaction effects, two findings emerge. First, when significant, estimates are negative for festival-related web searches, which supports dissortative mixing or heterophily for search terms within the festival category (models 1 to 4). Namely, festival-related searches tend to form edges with search terms within different categories, that is, spillover, hence broadening the scope of queries.

Second, for other search categories, segregation emerges: a tendency to link to searches within the same category. This is specially so for travel—models (1) to (5)—, which hints at individuals increasing the depth of the information on travel-related topics when planning for attendance. While

Table 7. STERGM estimation: Network dissolution.

	Dependent variable: Observed network					
	FIB		*Primavera Sound*		*Sónar*	
	(1)	(2)	(3)	(4)	(5)	(6)
Structural properties						
edges	−2.973***	−2.448***	−2.323***	−2.037***	−2.366***	−1.920***
	(1.071)	(0.343)	(0.581)	(0.589)	(0.306)	(0.522)
mutual			0.465	1.227***		
			(0.411)	(0.384)		
gwesp	1.574***	1.277***	1.062***	0.942***	1.579***	1.074***
	(0.139)	(0.120)	(0.086)	(0.105)	(0.143)	(0.124)
Node-level effects						
node.Festival	0.555	0.016	−0.040	−0.121	−0.300	−0.074
	(0.697)	(0.206)	(0.579)	(0.320)	(0.194)	(0.314)
node.Planning	0.609		0.999***	0.173		
	(0.644)		(0.372)	(0.402)		
node.Travel	0.515		0.214	1.896**		
	(0.524)		(0.351)	(0.776)		
Interaction effects						
nodematch.Festival	−0.826		−0.012			
	(0.619)		(0.727)			
nodematch.Planning			−0.713			
			(0.843)			
Akaike Inf. Crit.	230.603	178.799	425.480	275.777	200.038	146.690
Bayesian Inf. Crit.	252.412	189.190	459.149	298.020	211.371	155.990
Geographical scope	World	Spain	World	Spain	World	Spain

Note: STERGM: separable temporal exponential random graph model.
**p* < 0.1; ***p* < 0.05; ****p* < 0.01.

similar findings emerge for other covariates, they appear to be less general as they are linked to specific events. In this regard, searches on music and musicians exhibit homophily in both festivals with a broader base models (1), (2), and (3). Likewise, evidence supports assortative mixing of media-related searches in model (5).

Formation and persistence of links

Tables 6 and 7 provide estimation results when the temporal dimension of searches is taken into account. A STERGM includes one equation for the formation and another for the dissolution of links. Positive and significant estimates in Table 6 are interpreted as usual: they increase the probability of an edge formation. As for Table 7, positive coefficients are associated with factors that increase the persistence of an edge.

Starting with Table 6, density effects are found to agree with those already discussed while reciprocity only drives tie formation when searches are restricted to Spain. This could point to the more hierarchical and structured search process of nonlocal audiences as compared to the bidirectional connections found in local ones. As for the impact of *gwesp*, it is found to be positive and significant.[5] Therefore, local clustering increases the likelihood of link formation, helping to disseminate connections. This can be interpreted as the association of any two queries through the connections these share. Overall, it broadens the scope of the information search and induces consumer discovery.

Evidence on node effects are consistent across events. Festival searches increase link formation in all models—but equation (6)—supporting the observed tendency of festival-related nodes to generate and/or receive more ties that expected. Collecting information about the event—*node.-Planning*—drives tie formation in equations (1) to (2) and (5), although is nonsignificant in (3) and (4), which correspond to the most popular music festival analyzed.[6] This finding is consistent with planning being more relevant the less known the event.

Fewer covariates affect the persistence of nodes (Table 7). Structurally, local clusters tend to persist, given the positive and significant coefficient of *gwesp* in all equations. As for other covariates, only planning and traveling were found to increase link persistence in equations (3) and (4). Overall, it seems that once formed, the tendency of a tie to dissolve is mainly driven by its relative position in the network.

Overall, the empirical evidence supports most of the formulated hypotheses. Estimation results show that information acquisition is mediated through the analyzed events (hypothesis 1). Both static and dynamic models point to a tendency of cultural events to produce spillovers in searches (hypothesis 3), while planning and/or travel related information acquisition tend to deepen knowledge accumulation (hypothesis 4). Moreover, the topology of online queries around music festivals favors the dissemination of queries (hypothesis 5). However, the evidence in relation to (hypothesis 2) is mixed: while the sociality of travel-related nodes is affected by geography in Tables 5 to 7 provide inconclusive evidence.

Discussion

Cultural tourism is mostly about consuming experiences and information whose value is influenced by market intermediaries. These are actors that serve, among others, the purpose of spreading information across the market, a function that influences consumers' uncertainty, perceived risk, and discovery. In live music, the tendency toward the temporal and geographical concentration of cultural supply, the so-called festivalization, has made a significant impact not only on what and how music is consumed but also on the consumption of related services linked to the nature of these events, such as travel services.

The aim of this research is to provide an insight into an unexplored aspect of the intersection between cultural participation and tourism behavior as the outcome of a process of mediated information acquisition. Attendance to music festivals by nonlocals is one instance of this type of consumer behavior. As potential tourists are culturally motivated, decision-making is mediated, and to some extent simplified, by the event they plan to participate in. This means that particular traits and patterns related to consumer behavior can be inferred from the process through which individuals collect information.

The research looks at the problem at stake through online searches stemming from three festivals. These searches are then modeled as a network, which emphasizes the mechanism that drives the diffusion of information. Which and how queries are connected and how this information spills over related queries are analyzed.

In this respect, some results stand out. First, we find evidence of the voraciousness of live music consumption (see for instance Sullivan and Katz-Gerro, 2006). Using a purely quantitative dimension of cultural consumption, voraciousness is equated with the scope of searches related to music festivals. Web-based searches indicate that users spread the search over different music festivals, besides the one that triggered the search session. This includes not only franchises but also (domestic and foreign) competitors. The centrality of search terms within the *festivals* category and the fact that edges involving these nodes are more likely than predicted by chance is consistent with the addictive nature of cultural consumption as individuals include in their choice set an interest for a diversity of cultural events. We may infer that this interest (an antecedent of intention) stems from a disposition to attend.

Second, searches related to festivals facilitate information diffusion through two means. On the one hand, they have a tendency to create ties, that is, they are central to the graph. On the other hand, they also show a propensity to mix with search terms within different search categories. This means that searches related to music festivals tend to spillover to searches within different domains, enlarging the scope of information acquisition. These are relevant in the case of planning attendance and traveling arrangements, which are central to cultural tourism. The estimated model suggests that the odds of online search leading to any of these two categories are larger than it would have been by pure chance. Festivals, driving online-search behavior, are actually mediating the process of information acquisition which stands as prior to decision-making. Note this finding is robust even when removing the node of the festival generating the network.

Third, other categories, specially planning and travel, either attract links from other searches or exhibit selective mixing. The former is related to certain nodes (planning) being reached due to their holding of relevant information about the event (e.g. ticketing or lineup). The latter stresses the need of individuals to deepen or complement information about a topic (in this case travel). When mixing occurs, the search is more likely to stay within the domain of the search term once it is reached, increasing the depth of the information acquisition. This finding reinforces the informational role of music festivals and the induced nature of information acquisition.

Fourth, graphs reveal differences across events and geographies. As for the former, two different patterns of information acquisition, potentially linked to differences in the degree of complexity and specialization of the analyzed events, emerge. Figures 1 to 4 describe a hierarchy of nodes in terms of the authority centrality displayed: while some nodes supply useful information to the network (those larger in shape), most contribute marginally. Conversely, Figures 5 and 6 show an egalitarian structure where (almost) all nodes convey similarly relevant information. When no evident authoritative source of information exists, individuals should be expected to search across all sources (i.e. nodes) in the network with similar intensity, which leads to the observed egalitarian pattern. Furthermore, ERGM estimates (Table 5, models (5) and (6)) are consistent with the greater complexity associated with a niche event and consumers' need of specific *cultural capital* accumulated through specialized media when planning attendance to an avantgarde festival—that is, *Sónar*—whose lineup shows a scarcity of headliners.[7] This effect, however, has not been captured in the dynamic specification (STERGMs).

As for the geographical differences, world searches show a differential pattern: ERGM estimates (Table 5) provide evidence of travel queries generating more incoming links, while STERGMs (Tables 6 and 7) suggest a greater likelihood of nonreciprocal ties. The evidence, while being partial, stresses the differential risk faced and the more structured approach—more focused search—of queries run by nonlocals.

Overall, the foregoing discussion allows to identify patterns of behavior in tourist decision-planning across different dimensions. Besides differences emerging as a consequence of the specificities of the events considered, heterogeneity has been found in the relative position of search terms in the search space (which points to the relative relevance of the information that is searched), the likelihood and factors that influence the formation of links and the spillovers from the query on the cultural event generating the graph to other search terms, which determines the formation of the choice set of cultural tourists. To conclude, cultural tourists' behavior is reflected on the structure of the search graphs, where the centrality of festival nodes indicates culturally motivated planning and stresses the potential attracting role that such cultural events play.

Acknowledgments

A previous version of this paper was presented at the 10th Vienna Music Business Research Days (September 2019). The authors wish to thank the organizers (specially Peter Tschmuck) and other participants for their helpful comments and the discussion around it. The authors would like to thank two anonymous referees for their careful review and valuable comments that have helped us to improve the quality of the article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Juan D Montoro-Pons https://orcid.org/0000-0002-5180-7360

Notes

1. Abbruzzo et al. (2014) use a network to identify the determinants of tourism expenditure. Nevertheless, their methodology and application significantly differs to the proposed in this article.
2. https://github.com/GeneralMills/pytrends. Alternatively, there is an R package, gTrendsR that offers similar functionalities.
3. Let search terms A and B be in the node set: then an edge from A to B is defined if B is in the list of top searches of A.
4. See Hunter and Handcock (2006) or Goodreau et al. (2009) for a discussion on these terms.
5. After running the regressions with different values, the decay parameter is fixed at 1 in all models.
6. By any measure in Table 1, but also in terms of the online activity it generates.
7. The lineup can be inspected at https://sonar.es/es/2018/artistas-por-dias.

Supplemental material

Supplemental material for this article is available online.

References

Abbruzzo A, Brida JG and Scuderi R (2014) Determinants of individual tourist expenditure as a network: empirical findings from Uruguay. *Tourism Management* 43: 36–45.

Ajzen I (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes* 50(2): 179–211.

APM (2018) IX anuario de la música en vivo (IX Annual report of live music). Available at: http://www.apmusicales.com (accessed 9 October 2019).

Artola C, Pinto F and de Pedraza García P (2015) Can internet searches forecast tourism inflows? *International Journal of Manpower* 36(1): 103–116.

Báez-Montenegro A and Devesa-Fernández M (2017) Motivation, satisfaction and loyalty in the case of a film festival: differences between local and non-local participants. *Journal of Cultural Economics* 41(2): 173–195.

Bangwayo-Skeete PF and Skeete RW (2015) Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management* 46: 454–464.

Bei LT, Chen EY and Widdows R (2004) Consumers' online information search behavior and the phenomenon of search vs. experience products. *Journal of Family and Economic Issues* 25(4): 449–467.

Björk P and Kauppinen-Räisänen H (2015) Contemporary insights to the dynamic pre-trip information sourcing behaviour. *Tourism and Hospitality Research* 15(1): 39–53.

Blazquez D and Domenech J (2018) Big data sources and methods for social and economic analyses. *Technological Forecasting and Social Change* 130: 99–113.

Bokelmann B and Lessmann S (2019) Spurious patterns in Google Trends data—an analysis of the effects on tourism demand forecasting in Germany. *Tourism Management* 75: 1–12.

Bonet L (2013) Heritage tourism. In: Rizzo I and Mignosa A (eds) *Handbook on the Economics of Cultural Heritage*. Edward Elgar Publishing Ltd. ISBN 9780857931009, pp. 386–401. DOI:10.4337/9780857931009.00032.

Borowiecki KJ and Castiglione C (2014) Cultural participation and tourism flows: an empirical investigation of Italian provinces. *Tourism Economics* 20(2): 241–262.

Brida JG, Dalle Nogare C and Scuderi R (2016) Frequency of museum attendance: motivation matters. *Journal of Cultural Economics* 40(3): 261–283.

Buhalis D and Law R (2008) Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management* 29(4): 609–623.

Butts CT (2008) Network: a package for managing relational data in R. *Journal of Statistical Software* 24(2).

Casaló LV, Flavián C and Guinalíu M (2010) Determinants of the intention to participate in firm-hosted online travel communities and effects on consumer behavioral intentions. *Tourism Management* 31(6): 898–911.

Cellini R and Cuccia T (2013) Museum and monument attendance and tourism flow: a time series analysis approach. *Applied Economics* 45(24): 3473–3482.

Chung JY and Buhalis D (2009) Information needs in online social networks. *Information Technology & Tourism* 10(4): 267–281.

Cox C, Burgess S, Sellitto C, et al. (2009) The role of user-generated content in tourists' travel planning behavior. *Journal of Hospitality Marketing & Management* 18(8): 743–764.

Cuccia T, Guccio C and Rizzo I (2016) The effects of UNESCO World Heritage List inscription on tourism destinations performance in Italian regions. *Economic Modelling* 53: 494–508.

Eisenhardt KM and Graebner ME (2007) Theory building from cases: opportunities and challenges. *Academy of Management Journal* 50(1): 25–32.

Faulkner B, Fredline E, Larson M, et al. (1999) A marketing analysis of Sweden's Storsjöyran musical festival. *Tourism Analysis* 4(3-4): 157–171.

Gergaud O and Ginsburgh V (2017) Measuring the economic effects of events using Google Trends. In: Ateca-Amestoy VM, Ginsburgh V, Mazza I, O'Hagan J and Prieto-Rodriguez J (eds) *Enhancing Participation in the Arts in the EU*. Berlin: Springer, pp. 337–353.

Goodreau SM, Kitts JA and Morris M (2009) Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* 46(1): 103–125.

Guccio C, Lisi D, Martorana M, et al. (2017) On the role of cultural participation in tourism destination performance: an assessment using robust conditional efficiency approach. *Journal of Cultural Economics* 41(2): 129–154.

Herrero LC, Sanz JÁ, Bedate A, et al. (2012) Who pays more for a cultural festival, tourists or locals? A certainty analysis of a contingent valuation application. *International Journal of Tourism Research* 14(5): 495–512.

Hiller RS (2016) The importance of quality: how music festivals achieved commercial success. *Journal of Cultural Economics* 40(3): 309–334.

Hitters E (2007) Porto and Rotterdam as European capitals of culture: toward the festivalization of urban cultural policy. In: Richards Greg (ed) *Cultural Tourism: Global and Local Perspectives*. East Sussex: Psychology Press, pp. 281–301.

Ho CI, Lin MH and Chen HM (2012) Web users' behavioural patterns of tourism information search: from online to offline. *Tourism Management* 33(6): 1468–1482.

Huang CD, Goo J, Nam K, et al. (2017) Smart tourism technologies in travel planning: the role of exploration and exploitation. *Information and Management* 54(6): 757–770.

Hunter DR and Handcock MS (2006) Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics* 15(3): 565–583.

Janssen S and Verboord M (2015) Cultural mediators and gatekeepers. In: Wright James D (ed) *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed. Vol 5. Oxford: Elsevier. pp. 440–446.

Jun SP, Yoo HS and Choi S (2018) Ten years of research change using Google Trends: from the perspective of big data utilizations and applications. *Technological Forecasting and Social Change* 130: 69–87.

Karpińska-Krakowiak M (2009) Festivalization of the city. Contemporary examples. *Lidé města* 11(2): 338–350.

Litvin SW, Goldsmith RE and Pan B (2008) Electronic word-of-mouth in hospitality and tourism management. *Tourism Management* 29(3): 458–468.

Llodrà-Riera I, Martínez-Ruiz MP, Jiménez-Zarco AI, et al. (2015) A multidimensional analysis of the information sources construct and its relevance for destination image formation. *Tourism Management* 48: 319–328.

María Munar A (2011) Tourist-created content: rethinking destination branding. *International Journal of Culture, Tourism and Hospitality Research* 5(3): 291–305.

Noonan DS and Rizzo I (2017) Economics of cultural tourism: issues and perspectives. *Journal of Cultural Economics* 41(2): 95–107.

Padhi SS and Pati RK (2017) Quantifying potential tourist behavior in choice of destination using Google Trends. *Tourism Management Perspectives* 24: 34–47.

Pan B and Fesenmaier DR (2006) Online information search. Vacation planning process. *Annals of Tourism Research* 33(3): 809–832.

Pan B and Li XR (2011) The long tail of destination image and online marketing. *Annals of Tourism Research* 38(1): 132–152.

Papathanassis A and Knolle F (2011) Exploring the adoption and processing of online holiday reviews: a grounded theory approach. *Tourism Management* 32(2): 215–224.

Richards G (forthcoming) The festivalization of society or the socialization of festivals? The case of Catalunya. In: Richards G (ed) *Cultural Tourism: Global and local perspectives*. Binghampton: Haworth Press.

Scuderi R and Dalle Nogare C (2018) Mapping tourist consumption behaviour from destination card data: What do sequences of activities reveal? *International Journal of Tourism Research* 20(5): 554–565.

SGAE (2018) Anuario SGAE de las artes escénicas, musicales y audiovisuales. Available at: http://www.anuariossgae.com/home.html (accessed 10 May 2019).

Stephens-Davidowitz S (2014) The cost of racial animus on a black candidate: evidence using Google search data. *Journal of Public Economics* 118: 26–40.

Sullivan O and Katz-Gerro T (2006) The omnivore thesis revisited: voracious cultural consumers. *European Sociological Review* 23(2): 123–137.

Vecco M and Srakar A (2017) Blue notes: Slovenian jazz festivals and their contribution to the economic resilience of the host cities. *European Planning Studies* 25(1): 107–126.
Xiang Z, Wang D, O'Leary JT, et al. (2015) Adapting to the Internet: trends in travelers' use of the web for trip planning. *Journal of Travel Research* 54(4): 511–527.
Yang X, Pan B, Evans JA, et al. (2015) Forecasting Chinese tourist volume with search engine data. *Tourism Management* 46: 386–397.

Author biographies

Juan D Montoro-Pons is an associate professor at the University of Valencia where he researches on digitization and information consumption, the creative industries and cultural consumption from a quantitative standpoint. He has published on the structure of the music industry, consumer behavior in the cultural industries and the asymmetric impact of digitization.

Manuel Cuadrado-García is an associate professor at the University of Valencia, Spain. His main line of research is marketing in the arts sector. He has undertaken research on consumer behavior, market research, and strategic marketing in performing arts, cinema and visual arts. He has collaborated with different cultural organizations.