Empirical Article

TOURISM ECONOMICS
The Business and Finance of Tourism and Recreation

Tourism Economics
2025, Vol. 0(0) 1–29

DOI: 10.1177/13548166251355678
journals.sagepub.com/home/teu
Sage

Putting the periphery on the map: Tourism activity measured with big data

Karol Jan Borowiecki, Maja U Pedersen and Marco Palomeque
University of Southern Denmark, Denmark

Abstract
This paper examines tourism patterns in peripheral regions, analyzing visitor flows and regional variations using big data from a major travel portal. Covering 10 European peripheral regions from 2016 to 2022, the study maps tourist origins and activity levels, offering insights into how tourism functions in these areas compared to less remote destinations. Our approach, validated against international tourism statistics, maps tourist origins and activity levels, comparing peripheral regions with less remote areas. The findings highlight specific patterns of tourist flows, regional disparities in tourism development, and opportunities for promoting sustainable tourism in underdeveloped regions. The findings reveal specific patterns of tourism dynamics in peripheral regions, contributing to a deeper understanding of how tourism develops in these areas and how it differs from more central destinations.

Keywords
big data, cultural heritage, periphery, tourism

JEL classification
L83, O1, Z11, Z3

Introduction

Tourism presents significant environmental and social challenges. It accounted for 8% of global greenhouse gas emissions between 2009 and 2013 (Lenzen et al., 2018). Far from being reduced, emissions are expected to increase by 25% by 2030 compared to 2016 (World Tourism Organization, 2019). Additionally, overtourism exerts pressure on local environments and communities (Santana- Jiménez and Hernández, 2011; Mohamed Reda Khomsi and Rabier, 2020; Bobic and Akhavan, 2022), often leading to declining public support for tourism as destinations become more crowded (Mansfeld and Ginosar, 1994).

Corresponding author:
Karol Jan Borowiecki, Department of Economics, University of Southern Denmark, Odense 5230, Denmark.
Email: kjb@sam.sdu.dk

Peripheral tourism emerges as a pragmatic alternative to the challenges posed by mass tourism in heavily visited destinations. These areas often feature unique attractions that remain relatively undiscovered compared to major tourist hubs. Moreover, tourism in these regions tends to involve activities centered around nature, heritage, and local traditions, which may align more closely with sustainable tourism practices when properly managed.

While addressing challenges in urban tourism management, tourism in peripheral regions can also contribute to local economies by fostering economic activity in less-visited areas (Vu and Turner, 2009). Domestic tourism, in particular, has been identified as a factor that may support broader regional tourism activity due to its more even spatial distribution compared to international tourism (Goh et al., 2015).

Despite the growing interest in tourism beyond major cities, significant research gaps remain. Current studies are constrained by a reliance on highly aggregated statistics, such as those from Eurostat, which obscure the dynamics of smaller, less-visited areas. This limitation makes it difficult to understand how peripheral destinations attract visitors, the characteristics of these visitors, and the broader implications for sustainability and regional development. Furthermore, existing literature rarely addresses how domestic and international tourist flows differ in peripheral areas or how spatial factors, such as distance and accessibility, shape tourism dynamics. Addressing these gaps is critical for designing policies that balance economic development with environmental and social sustainability. (Goh et al., 2015).

This paper builds on established theoretical frameworks in sustainable tourism and regional economics. Sustainable tourism development theory emphasizes managing tourism in ways that minimize environmental degradation while maximizing socio-economic benefits for host communities (Hunter, 1997; Sharpley, 2009). Meanwhile, regional economic theories highlight the spatial dimensions of economic activity, including tourism, and how tourism patterns relate to broader regional development dynamics (Martin and Sunley, 2006; Storper, 1997). By leveraging big data, this study provides new empirical evidence on tourism activity in peripheral areas, contributing to these theoretical perspectives without making direct claims about tourism-induced economic growth.

Our approach leverages a novel dataset of user-generated content from a popular travel portal, covering 10 peripheral regions in Europe from 2016 to 2022. This dataset enables a detailed exploration of tourist origins, spatial distributions, and activity levels, offering unique insights into how tourism unfolds in peripheral areas. Using this data, we present stylized facts that illuminate tourism activity in peripheral regions relative to their neighboring areas and major urban attractions. For example, we examine how international tourists may gravitate toward well- known landmarks, while domestic tourists might prefer lesser-known sites (Cerisola and Panzera, 2024). Additionally, we compare these patterns with national trends, contributing to a deeper understanding of tourism in peripheral regions.

The findings of this study contribute to both tourism literature and policy design in several ways. On the one hand, we present a new tool that allows for big data analysis on peripheral tourism, almost in real time, allowing us to trace the origin of tourists and to granularize the region to be studied, from the level of a tourist attraction to national aggregations. On the other hand, this study examines the influence of cultural and natural attractions on tourism dynamics in peripheral regions, offering insights into how these areas attract visitors. By analyzing patterns of tourist activity, the study enhances understanding of tourism development outside major urban centers. The findings provide valuable perspectives for tourism planning in regions facing demographic shifts, infrastructure limitations, and economic challenges (Botterill et al., 2000).

Building on these insights, this paper also examines how distance influences tourist flows, using a dataset with significant local and within-country variation. By addressing a key gap in prior

research, which often relied on single observations or population-weighted averages (Rosselló-Nadal and Santana-Gallego, 2024; Bergstrand et al., 2015; Yotov, 2012), our findings enhance the understanding of how accessibility shapes both domestic and international tourism dynamics. While this study does not directly assess the redistribution of tourists, it contributes to discussions on how tourism flows differ between major attractions and peripheral areas. These insights serve as a foundation for future research on redistribution, which would require individual mobility tracking to assess whether visiting peripheral areas impacts demand in top destinations.

The rest of the paper is organized as follows: first, a literature review is provided; then, the data is presented; afterwards, the approach pursued is validated; and, finally, results of the analysis are presented. The paper concludes with a final summary of the findings and contributions.

Literature review

We begin this literature review by summarizing some of the attributes of rural and peripheral tourism. Rural areas accommodation is often managed by individuals or small businesses (Leick et al., 2024; Ye et al., 2019). Tourists in the periphery tend to seek peaceful and rush-free holidays (Komppula, 2005; Molera and Pilar Albaladejo, 2007), exhibiting no big socio-demographic difference compared to urban tourists (Bel et al., 2015), and are more interested in culture and nature than actual rural life (Frochot, 2005). An extensive literature review by Calero and Turner (2020) suggests that regional tourism scholarship is still in its infancy. The present paper contributes to this research strand by developing and testing new tools that allow the measuring of tourism in several regions at the same time, using comparable and consistent data.

The destinations studied in this paper contain predominantly cultural and natural activities to attract tourists. The relationship between cultural participation and tourism flows is empirically shown by Borowiecki and Castiglione (2014). Bertacchini et al. (2021) suggest that attractions and major urban museums can stimulate the demand for regional destinations. Cui et al. (2024) find a positive effect on tourists' well-being when adding co-creation activities in cultural heritage sites. Brandano and Meleddu (2021) observe that the exhibition of newly discovered cultural goods increases the interest of both residents and tourists in the region. This relationship also works in the opposite direction: Brandano and Crociata (2023) observe the positive effect of tourism policies in culture and nature, and Panzer-Krause (2019) study the different approaches to fulfill the socio-ecological goals in a high natural and socio-cultural valued cultural area. Our paper shows how the presence of culture and nature can redirect tourism flows. It also provides one of the first analyses where cultural and natural heritage are considered jointly.

Few recent papers explore the possibilities of big data obtained from user-generated content. Borowiecki et al. (2024) use Tripadvisor reviews to study tourism in Denmark, France, and Spain, and estimate the effect of restrictions imposed during the COVID-19 pandemic on tourism. Leick et al. (2024) use Airbnb data to illustrate the distribution of host types across rural regions on Denmark, Iceland and Norway. Mahat and Hanafiah (2020) use Google Forms questionnaires to reveal travelers' behavioral intentions according to their Tripadvisor use. To the best of our knowledge, the present paper is the first one using Tripadvisor reviews to directly measure tourism in the periphery, allowing fast reports on relative quantities and tourist origins. It also highlights the relevance of digital tools for tourism promotion.

Related is also the literature on ratings and social influence, in particular how opinions are influenced by Airbnb ratings (e.g., Chen and Chang, 2018; Cheng and Jin, 2019; Falk et al., 2019; Sainaghi, 2020), Tripadvisor ratings (Donati, 2022; Martin-Fuentes et al., 2020) or social media reviews (Jean et al., 2019). Improvements in the quality of Airbnb offerings may also spur growth in

the tourism sector in less frequented destinations (Leick et al., 2022). Social media presence and quality attract tourists (Kim et al., 2017). Revealing the strong relationship between Tripadvisor use and actual tourism, our paper emphasizes the relevance of social media metrics and—more widely—the value of digitizing the tourism offer.

Our study relates also to the literature on tourism and local economic development. Faber and Gaubert (2019) focus on tourism's economic impact along Mexico's coastline, while Nocito et al. (2023) examine how media exposure can enhance tourism-driven growth. Luna and Surovtseva (2020) explore tourism's effects on local employment in specific regions. Our work complements these studies by providing an international perspective, focusing on peripheral regions across Europe and using big data to capture real-time tourist activity. By offering a broader, cross-country analy- sis, we provide new insights into the global dynamics of tourism in less-explored areas, contributing both a wider scope and more granular data to the discussion on regional economic development.

Given the time period covered in this study, we contribute also to papers on how the COVID-19 pandemic impacted the tourism sector (e.g., Gil-Alana and Poza, 2022; Plzáková and Smeral, 2022; Wang et al., 2022). Vaishar and Šťastná (2022) point to the opportunities that the pandemic has created for rural tourism development. Seraphin and Dosquet (2020) use news media narrative to suggest a possible increase in rural tourism in France based on the placebo role of two forms of tourism: mountain tourism and second-home tourism. We contribute to this strand by introducing and analysing a novel way to obtain disaggregated data that is available in almost real-time.

Finally, our paper contributes to the literature on the effect of distance on tourist flows, particularly in the context of domestic versus international tourism. While previous research, such as Rosselló-Nadal and Santana-Gallego (2024), suggests that as tourists shift from domestic to international tourism, distance becomes less significant—a notion borrowed from international trade (see Bergstrand et al., 2015; Yotov, 2012)—it often relies on a single observation for domestic tourism. In these studies, distance is typically represented as a population-weighted average, which limits the depth of analysis. Our paper addresses this gap by utilizing a more granular dataset with substantial local and within-country variation, allowing for a more nuanced exploration of how distance impacts tourist flows.

Data description

Summary of peripheral regions

This study presents the results of 10 peripheral regions of the European Union, which can be consulted in Table 1. These regions have in common that they are part of the INCULTUM project.[1] We provide descriptive data on these peripheral regions,[2] allowing us to get a first picture of the situation in these regions and to demonstrate their peripheral nature. The Irish case is different here because the pilot took place in multiple places, hence it is not located in a particular peripheral region. We keep it in the summaries to avoid any subjective exclusions, but our interpretation focus on the remaining sites. All statistics are provided by the European Union (Eurostat) at NUTS2 or NUTS3 level, depending on data availability (NUTS3 is preferred where possible). We cannot use more localized geographic levels due to lack of data, as previously discussed. Although there is some variation within a given region, these data provide a important simple, general overview of the type of regions covered.

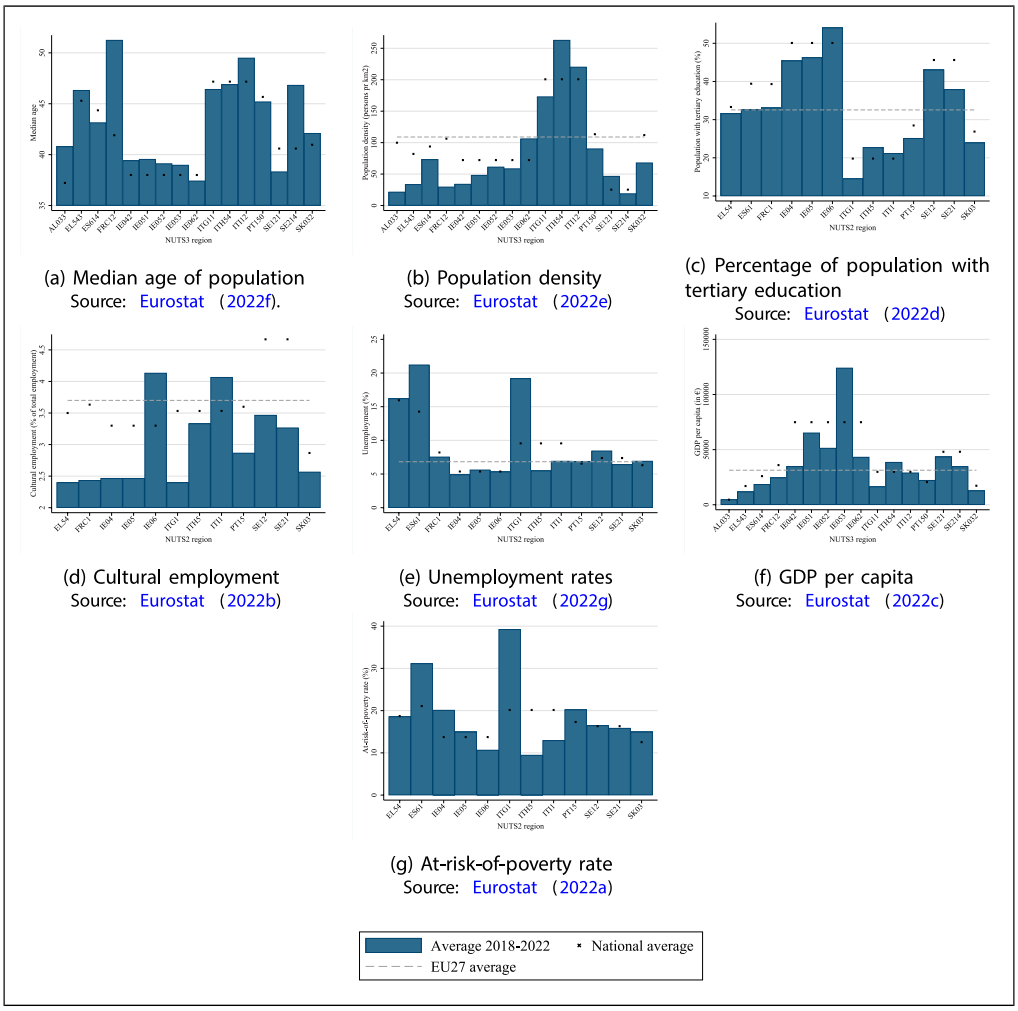
Table 1. Regional classifications of the peripheral areas.

	Country	Location as described by pilot	NUTS3 region name	NUTS3 region code	NUTS 2 region name	NUTS 2 region code
1	Spain	The Altiplano	Granada	ES614	Andalusia	ES61
2	Portugal	Campina de Faro	Algarve	PT150	Algarve	PT15
3	Slovakia	Banská Bystrica, Banská Štiavnica	Banskobystrický kraj	SK032	Central Slovakia	SK03
4	Italy	Monti di Trapani, Calatafimi-Segesta, Custonaci, Buseto Palizzolo	Trapani	ITG11	Sicilia	ITG1
5	Italy	San Pellegrino, Alpe	Modena	ITH54	Emilia-Romagna	ITH5
		Tuscan-Emilian Appennines	Lucca	IT112	Toscana	IT11
6	France	Regional Natural Park	Nièvre	FRC12	Bourgogne	FRC1
7	Greece	Aaos Valley, Konitsa	Ionnina	EL543	Epirus	EL54
8	Albania	Upper Vjosa Valley, Përmet	Gjirokastër	AL033	Southern Albania	AL03
9	Ireland	County Mayo	West region	IE042	Northern and western region	IE04
		County Galway	West region	IE042	Northern and western region	IE04
		County Limerick	Mid-west region	IE051	Southern region	IE05
		County Cork	South-west region	IE053	Southern region	IE05
		County Waterford	South-East region	IE052	Southern region	IE05
		County Wicklow	Mid-east region	IE062	Eastern and Midland region	IE06
10	Sweden	Gotland	Gotlands län	SE214	Småland and the islands	SE21
		Öregrund	Uppsala län	SE121	East middle Sweden	SE12

The following subsections summarize the main descriptors of demographics, labor markets and economic activity. We present averages from the period under review and compare them with national levels and the EU average where available. All indicators are presented in Figure 1.

Demographics. To begin with, we provide a general overview of the population structure in the peripheral regions. We use the median age of the population provided by Eurostat as an indicator of population structure. Figure 1(a), illustrates the median age of the NUTS3 regions compared to the respective national levels at the beginning and end of the study period. In general, the Figure shows that the peripheral regions are older than the national average.

To measure urbanization, we use population density, which is defined as persons per square kilometer. This measure is also available at the NUTS3 regional level for the same years. It is shown in Figure 1(b). Compared to the EU27 average, the majority of the peripheral regions are less densely



(a) Median age of population
Source: Eurostat (2022f).

(b) Population density
Source: Eurostat (2022e)

(c) Percentage of population with tertiary education
Source: Eurostat (2022d)

(d) Cultural employment
Source: Eurostat (2022b)

(e) Unemployment rates
Source: Eurostat (2022g)

(f) GDP per capita
Source: Eurostat (2022c)

(g) At-risk-of-poverty rate
Source: Eurostat (2022a)

Figure 1. Indicators in peripheral regions.

populated. The only regions above the EU27 average are those in Italy, which is not surprising given the country's high population density. Compared with the national average, the majority of peripheral regions tend to be less densely populated. The relatively low population density suggests that the peripheral regions are more rural than metropolitan. This is in line with our expectations, as the regions covered are considered peripheral.

Labour markets. To illustrate the labour markets, we provide an overview of educational attainment, cultural labour and unemployment. All data in this section are presented at NUTS2 regional level, as NUTS3 is not available in Eurostat.

Education is an important determinant of economic development and labour markets, as it determines the supply of labour at different income levels. Figure 1(c), illustrates the share of the population with tertiary education in the peripheral regions. Compared to the EU27 average, there

are several regions (mainly in Southern Europe) where the share of tertiary educated population is below the average. On the other hand, peripheral regions with a higher share of people with tertiary education are mainly from Northern Europe. However, when compared with the national averages, a different pattern emerges: most peripheral regions tend to have a lower share of people with a tertiary education.

Figure 1(d), shows the cultural employment of the NUTS2 regions and compares it with the EU27 and national averages. Almost all peripheral regions are below the EU27 and their respective national averages in terms of cultural employment, especially before the pilot started.

Finally, the unemployment rates of the NUTS2 regions are shown in Figure 1(e). The peripheral regions in Northern Europe more or less follow their respective national unemployment average. However, when looking at the countries of Southern Europe, there seems to be a considerable variation in unemployment across regions. For example, Emilia-Romagna (ITH5) has a significantly lower unemployment rate than the Italian average. On the other hand, Andalusia (ES61) and Sicily (ITG1), together with Epirus (EL54), have unemployment rates well above their respective national averages and also well above the EU27 average.

Economic activity. We measure economic activity in the peripheral regions by using Gross Domestic Product (GDP) per capita at current market prices. Figure 1(f), shows the income levels of the NUTS3 regions. Once again there seems to be a gap between Northern and Southern European regions. Almost all the Northern European peripheral regions have income levels above the EU27 average level of GDP per head, while the majority of the Southern European regions have income levels below it. Compared to national averages, the majority of peripheral regions have income levels below their respective national income levels.

For a more comprehensive view of economic development, we also provide an overview of the at- risk-of-poverty rates.[3] The rates for the NUTS2 regions are presented in Figure 1(g). In the majority of cases, the peripheral regions have at-risk-of-poverty rates at or above their respective national averages. This suggests that most of them have a higher proportion of low-income people than their national averages.

This section has shown that the peripheral regions are on average older than the respective national average, less densely populated and have a lower cultural labour capacity than the respective national average and the EU27 average. They also tend to be less educated than their national average. Furthermore, we show that some of them have unemployment rates well above both their national and EU27 averages.

In terms of income levels, the majority of peripheral regions have income levels below the EU27 and their respective national averages. Finally, we have shown that peripheral regions tend to have at-risk-of-poverty rates that are higher than their national at-risk-of-poverty average.

Tourists' reviews

In this paper, we propose to leverage user-generated content collected from Tripadvisor, a popular travel portal, expanding the work from Borowiecki et al. (2024). We collect reviews posted on Tripadvisor for the period January 2016 to July 2022, covering all attractions in the countries: Albania, France, Greece, Ireland, Italy, Portugal, Slovakia, Spain, and Sweden. The data collection concluded in July 2022 in accordance with the timeline of the INCULTUM project, which ran from May 2021 to April 2024. This period covers the COVID-19 pandemic and its immediate aftermath, allowing us to capture the unique dynamics of this critical phase. We include reviews written in English and the native language whenever available (e.g., French for reviews of attractions in

France). The constructed dataset is a combination of three different entities which we combine: a list of attractions, reviews of attractions, and user profiles, each of which is explained below.

List of attractions. This is a complete list of all the attractions in a country that are present on Tripadvisor on the respective country's *Things to Do* webpage. In this context, an "attraction" refers to any location, site, or feature that draws visitors for recreational, cultural, or educational purposes. The list of attractions includes information about the attraction, such as the name, within-country ranking, overall rating, number of reviews, attraction location and attraction type. The type of attraction is based on Tripadvisor's own classification which covers 20 different cate- gories. The classification system is not mutually exclusive, meaning that some attractions may be classified in more than one category at a time. Finally, the module also includes a brief "About" text describing the attraction.

Attraction reviews. This module is a list of all reviews, in English and the language of the attraction country, for all attractions in the attractions module. This module includes the title and text of the review, along with the date the review was posted and the rating of the attraction. In some cases the type of trip is also indicated, i.e., with family, friends, for business, etc. The list also includes a unique and anonymous identifier for the user who posted the review. This latter can be used to link the review to the user profile module to obtain additional information about the user, such as the user's location. In some cases, the reviews module includes the actual month of the tourist's trip. Since this information is not always available and only specifies the month, we use the review's posting date in our analysis to maximize the number of reviews included. However, a comparison of the trip date and posting date, when both are available, shows that the time difference is minimal.

User profiles. The user profile module contains basic information about users who have written at least one review for at least one attraction in our sample of countries. It includes details about the user, such as their location, the date they joined Tripadvisor, and the number of reviews they have posted.

Together, these three modules form the basis of our dataset, where by processing and combining the information in each module, we obtain information about both the user and the attraction associated with each review. The data is at the individual and daily level, and is therefore highly disaggregated. To obtain additional variables, we use OpenStreetMap (OSM) data to identify the latitude and longitude of locations (user and attraction) in R using the geocode function in the tidygeocoder package Cambon et al. (2021). All attractions include location data, while approximately 68% of users provide this information, allowing us to identify latitude and longitude for about 99.9% of these user locations. Using this data, we can calculate the distance traveled by each user to reach an attraction. Summary statistics are presented in Appendix Table B1. In total, the dataset covers 321,055 attractions, 12,185,780 reviews, and 6,981,081 unique users.

Validating the tripadvisor data

Before turning to the analysis of tourism activity in peripheral areas, we present a series of validation tests of our data. The aim of this section is to demonstrate the validity of the approach pursued here and reliability of the results. To do this, we compare the number of Tripadvisor reviews aggregated at the national level for each of the countries with official tourism statistics from Eurostat. This is in some analogy to Borowiecki et al. (2024), but here we cover a much bigger set of countries and show the results also for domestic and foreign reviews separately.

A visual inspection

We start with a visual inspection of our monthly aggregated data and compare it to the number of arrivals provided by Eurostat. In Figure 2 we show the total number of Eurostat arrivals and Tripadvisor reviews for each of the countries studied. It becomes rather clear that the time series follow each other closely in almost all countries. To ensure that the results are valid when considering only domestic reviews or only foreign reviews, we show these separately in Appendix Figures B1 and B2. We also compare our monthly aggregated data with the occupancy rates in the Appendix Figure B3. The conclusions are similar to above, with some deviations in the case of domestic tourism in Albania and Slovakia.[4]

We estimate also binned scatterplots of Eurostat arrivals or occupancy rates and Tripadvisor reviews. The idea of these plots is to visually illustrate how closely the variables are correlated. The

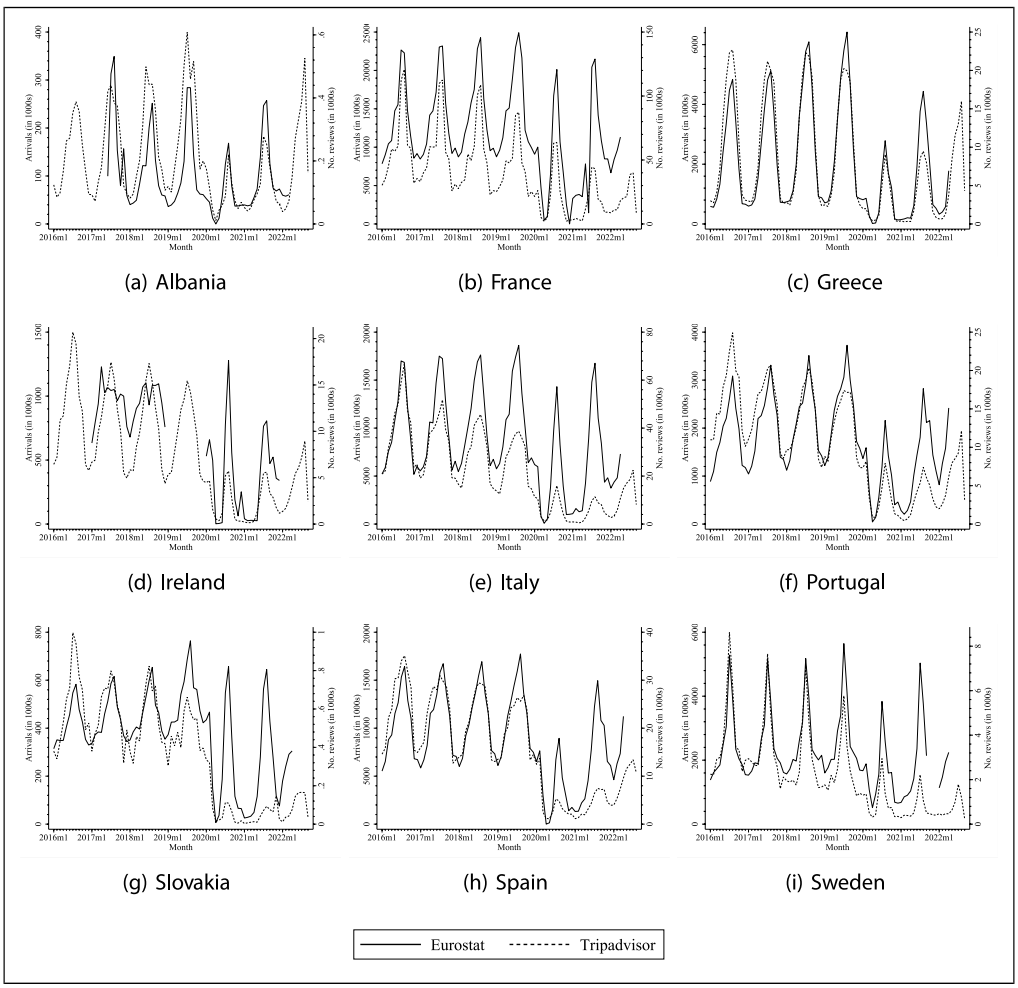


Figure 2. Total number of arrivals and reviews over time. *Notes:* This Figure shows number of Tripadvisor reviews and Eurostat arrivals by country and month. *Source:* (Eurostat, 2023) and own data collected from Tripadvisor.

closer the dots follow a straight line, the more closely the variables are correlated. In Figure 3 we show the correlation between total arrivals and total reviews, while Appendix Figures B4 and B5 show the binned scatterplots between domestic arrivals and domestic reviews and foreign arrivals and foreign reviews, respectively. Finally, Appendix Figure B6 shows the binned scatterplots between Eurstat occupancy rates and Tripadvisor reviews. In all figures, it is very clear that they are well aligned.

Formal validity tests

The visual inspections are complemented by two more formal tests. First, we compute simple correlation coefficients between the variables can be seen in Appendix Table C1. Appendix Table C1, Panel (a), shows the correlation coefficients between total number of Eurostat arrivals and Tripadvisor, Panel (b) between domestic arrivals and domestic reviews, Panel (c) between foreign arrivals and foreign reviews and Panel (d) between Eurostat occupancy rates and Tripadvisor

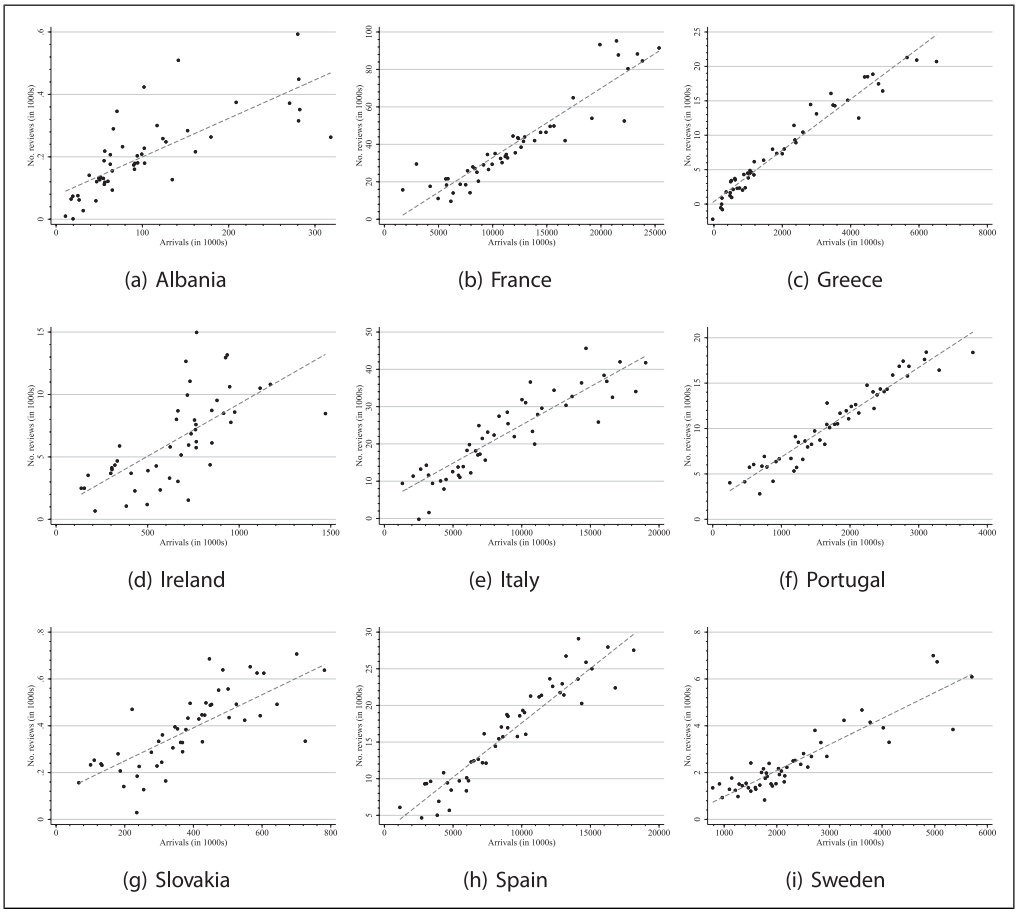


Figure 3. Monthly correlation between tourist arrivals and number of reviews. *Notes:* This Figure shows binned scatter plots of the number of Eurostat arrivals and the number of Tripadvisor reviews by country. *Source:* Official tourism statistics from (Eurostat, 2023) and own data collected from Tripadvisor.

reviews. The correlation coefficients are quite high and also highly significant, with the exception of Albania and Slovakia in Panel (b), which consider domestic measures.

Second, we regress the number of reviews on the number of arrivals or occupancy rate, which allows us to control for country, year, and month fixed effects. In Table 2 we present the results using observations from all countries together. In columns 1, 3, 5, and 7, we present the results without any controls, while in columns 2, 4, 6, and 8, we include country, year, and month fixed effects. In columns 1-2 we show the results using *ln*(*Arrivals*) as the explanatory variable and *ln*(*Reviews*) as the outcome of interest. The estimate is highly significant and large in magnitude.

In Table 2, columns 3-4, we present the results using domestic arrivals and domestic reviews and find that a 1% increase in the number of domestic arrivals corresponds to a 0.61% increase in the number of domestic reviews. In columns 5-6 we present the results using foreign arrivals and foreign reviews and again find a highly significant estimate. Finally, in columns 7-8 we use the occupancy rate, where a 1% increase in the occupancy rate implies a 0.4% increase in the number of reviews. The Eurostat measures in all columns demonstrate substantial explanatory power, as evidenced by the high adjusted R^2 values. These results highlight the strong alignment between variables derived from our web-scraping methodology and official tourism statistics, reinforcing the validity and robustness of our approach. The high explanatory power could, however, also suggest potential issues with multicollinearity. Therefore, to assess the generalizability and ensure that we are not overfitting the data, Appendix Table C2 shows the adjusted R^2 values using a 20-fold cross-validation for all models in Table 2. The results in Table C2 indicate that the R^2 values are generally high and hence suggest no issues with multicollinearity but instead are a result of the nature of the data.[5]

Table 2. Validity test: regression results for full sample.

	ln(Reviews)		ln(Domestic reviews)		ln(Foreign reviews)		ln(Reviews)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(Arrivals)	1.019*** (0.034)	0.673*** (0.051)						
ln(Domestic arrivals)			1.160*** (0.030)	0.610*** (0.060)				
ln(Foreign arrivals)					1.015*** (0.026)	0.688*** (0.046)		
Occupancy rate							0.078**** (0.003)	0.036*** (0.002)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
N	632	632	623	623	630	630	635	635
R^2	0.764	0.964	0.718	0.963	0.786	0.960	0.508	0.946

Notes: Regression results when estimating the number of Tripadvisor reviews on the number of arrivals or the occupancy rate from Eurostat. Columns 1-2 use all arrivals and reviews, columns 3-4 use domestic arrivals and domestic reviews, columns 5-6 use foreign arrivals and foreign reviews, and columns 7-8 use the occupancy rates and number of reviews. Bootstrapped standard errors in parentheses. ***$p < .01$ **$p < .05$ *$p < .10$. *Source:* Official tourism statistics from (Eurostat, 2023) and own data collected from Tripadvisor.

In Appendix Tables C3 - C6 we show the results by country. Showing the results separately helps to illustrate whether our data are reliable in each country separately. These estimations allow us to conclude that in all cases the estimates are highly significantly different from zero and large in magnitude, and that all models have high explanatory power.[6] Therefore, we are confident that our data provide a valid alternative to the official tourism statistics and proceed to our analysis.

Tourism activity and trends in the peripheral regions

Methodology and considerations

We provide detailed insights into tourism activity in the peripheral and neighboring regions by looking at different time trends and comparing their respective tourism activity. We also present several detailed maps showing the location of both attractions and users, as well as maps illustrating travel patterns for four different travel categories: local, domestic, Europe, and world. The analysis will therefore consist of a visual inspection of the trends, after which we will be able to draw conclusions about the impact of the peripheral action.

The decision to use visual inspection of maps and time trends and to exclude regression analysis is based on the consideration that the tourism sector faced massive repercussions with the outbreak of the COVID-19 pandemic and the associated lockdown measures. Restrictive measures on internal and international travel generally began in March/April 2020, and COVID-19 was declared a pandemic by the WHO on March 11, 2020 (World Health Organization, 2020). Most countries had internal movement restrictions and additional controls for international travel when the pilot started in May 2021. The entire tourism sector experienced a record decline due to the COVID-19 pandemic, and tourism-dependent economies were disproportionately affected (Behsudi, 2020). This means that baseline data in all regions (both peripheral and neighboring) are affected by this major event to varying degrees, and therefore future trends in tourism activity are likely to be influenced by this further. Consistent and reliable estimates from a regression analysis would require detailed information on several economic indicators which, to the best of our knowledge, do not exist at the necessary level of aggregation and for all countries. Comparison with neighboring regions could help determine whether the visitor trends observed at the peripheral level are consistent with broader trends in tourism.

We present the results for each peripheral region and their neighboring regions. In some parts of the analysis we consider both peripheral and neighboring regions together to show tourism trends around the area, and in other parts of the analysis we consider peripheral and neighboring regions separately to compare their trends. Additionally, we provide comparisons with national trends.

Describing tourism in the peripheral regions

Location of attractions. We start by showing the location of all the attractions in the regions studied. Figure 4 indicates the location of the attractions in both peripheral and neighboring regions with blue dots. The red squares indicate the approximate location of a pilot site. In Appendix Figure D1 we show more detailed maps for each country, including information about the number of reviews for each attraction location, determined by the size of the dots. Comparing attractions in the peripheral and neighboring regions is useful to ensure that they are not too different in terms of tourist activity.

Considering the countries separately, we can observe different patterns in the location of attractions and the number of reviews.

Albania: The peripheral regions in Albania look quite similar to the selected neighboring regions. Attractions seem to be fairly well distributed across all regions, with the exception of the coastal neighboring regions where there are more attractions. Attractions closer to the coast generally have more reviews.

France: The distribution of attractions is similar in all regions. The attractions around the pilot site have a higher concentration of reviews, while the central part of the region seems to have attractions. The neighboring regions all look quite similar, with some attractions receiving more reviews than others.

Greece: The peripheral regions seem to have a higher concentration of attractions closer to the pilot site, but the attractions are still similarly distributed in the rest of the regions. Peripheral and coastal regions also seem to have larger attractions.

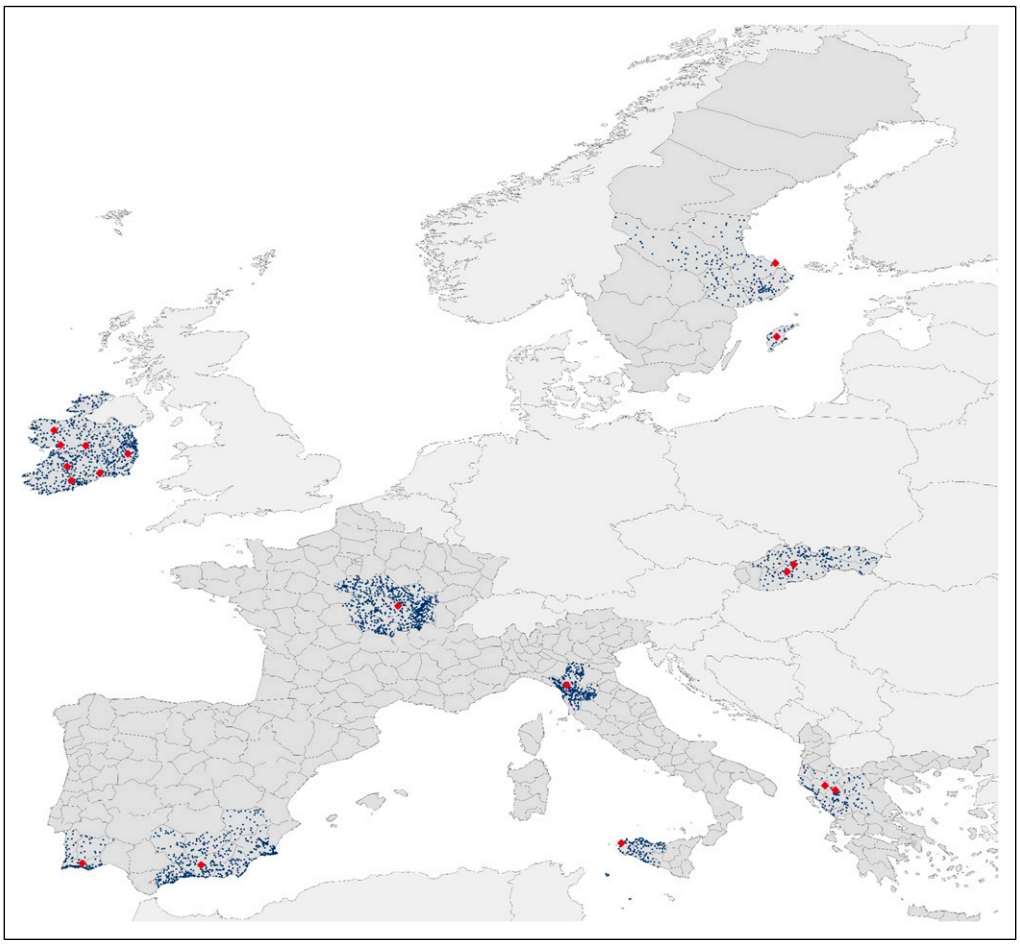


Figure 4. Location of attractions in peripheral and neighboring regions. *Notes:* This Figure shows the location of all attractions on Tripadvisor and located in both peripheral and neighboring regions. The blue dots indicate attractions and the red dots represent the approximate location of a pilot site. In the case of the Irish pilot, the dots indicate locations where visitor surveys have been distributed. *Source:* Own data collected from Tripadvisor.

Ireland: Ireland differs from the other regions in that the pilot action takes place all over the country. Therefore we do not explicitly consider neighboring regions, but Figure 4 shows a similar distribution across regions. All regions have a similar distribution of reviews.

Italy (Sicily): The peripheral region has a larger number of attractions located near the coast, while the neighboring regions have a more even distribution throughout the region. The coastal areas again receive more reviews and all regions have similar patterns.

Italy (Tuscany-Emilia): All regions look similar in their distribution of attractions. The same is true for reviews, except that one of the neighboring regions has more attractions with a greater number of reviews.

Portugal: The peripheral region has a higher number of attractions, mainly located near the coast, while the neighboring regions have fewer, more evenly distributed attractions. Additionally, a large number of all reviews in these regions refer to attractions in the peripheral region close to the coast.

Slovakia: The distribution of attractions is similar in all regions, but the number of reviews is not. The attractions in the peripheral region receive fewer reviews than most of the other attractions.

Spain: Again, the coastal areas seem to have a higher concentration of attractions, but overall, all regions look quite similar. Attractions close to the pilot receive a large number of reviews and, as in other countries, the coastal attractions also receive more reviews.

Sweden: The attraction distributions are also similar for the Swedish regions. All regions seem to have attractions that receive more reviews. The region with the most reviews is the region where the Swedish capital is located.

Travel patterns. To understand where visitors to attractions in the peripheral and neighboring regions come from, we look at the travel patterns of visitors. We classify visitors based on their origin: local, domestic, from elsewhere in Europe, and from other parts of the world. This allows us to identify patterns specific to each geographical category. We classify visitors from different locations, i.e. local, domestic, Europe and world, to look for different patterns in each country. A local visitor is defined as a visitor from the same NUTS3 region as the attraction visited. In Figure 5 we provide an exemplary visualization of travel patterns to Sicily (Italy). The other regions are presented in the Appendix Figures E1 - E9.[7] In each figure, Panel (a) shows where international visitors come from, Panel (b) shows European visitors, Panel (c) national visitors, and Panel (d) local visitors.

Overall, when looking at users from around the world, the patterns are very similar for all regions, with a large number coming from the United States. In terms of European visitors, there is a larger number of visitors from the United Kingdom. In general, it also appears that there are more visitors from Northern Europe, and many visitors from Southern Europe are located closer to the coasts. Some noticeable patterns also emerge from individual countries. In France, it appears that a large proportion of European visitors come from southern Germany and Switzerland, while in Ireland, Portugal and Spain, a large proportion of visitors come from England. In the other countries, the origin of visitors is more evenly distributed across Europe. In terms of domestic and local tourism, there are also different patterns in each country:

Albania: Both domestic and local tourism is largely concentrated on the coast and nearby regions.

France: Domestic visitors come from all locations, but a larger number of visitors come from the Paris area and other major cities. There are no clear patterns of local visitors.

Greece: The origin of domestic tourists is also distributed throughout the country, with no clear pattern as to the choice of destination. Locally, short distances seem to dominate.

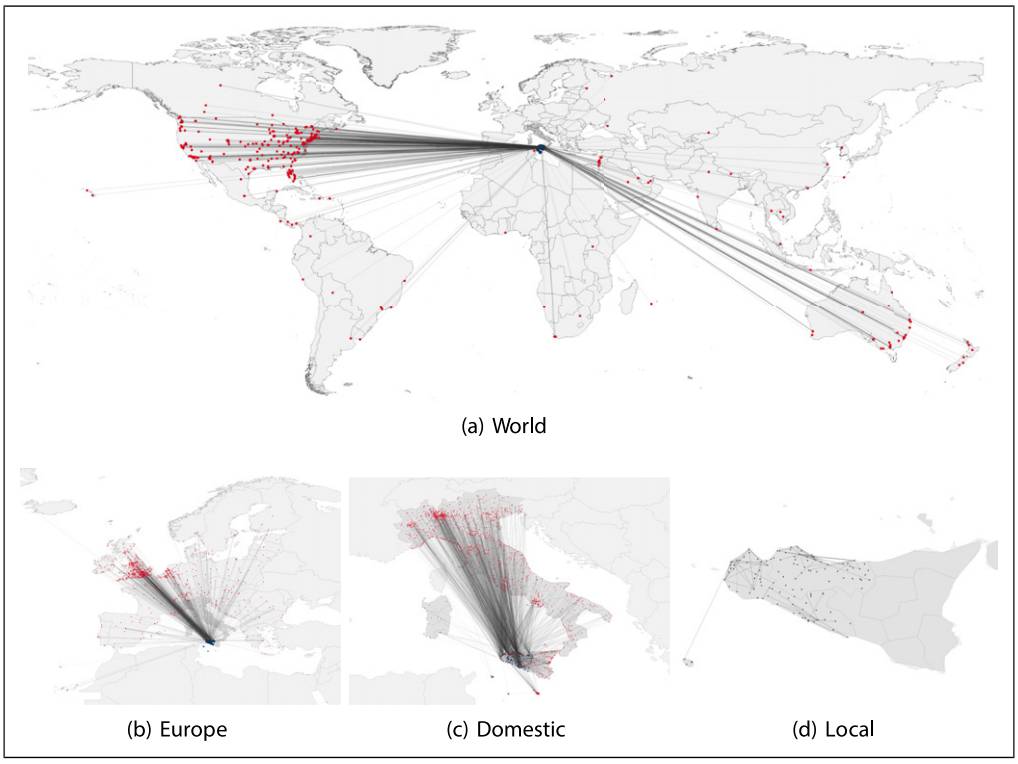


Figure 5. Travel patterns of visitors to peripheral and neighboring regions - Italy (Sicily). *Notes:* Travel patterns of visitors to peripheral and neighboring regions. Panel (a) shows patterns from all over the world, Panel (b) from Europe, Panel (c) domestic travel and Panel (d) local travel. The red dots represent the location of users and the blue dots the location of attractions. *Source:* Own data collected from Tripadvisor.

Ireland: There is a larger number of domestic visitors originating from Dublin, but with no specific pattern in the choice of attractions. The same is true at the local level, where visitors come from all over the region.

Italy (Sicily): Domestic visitors come from all over the country, while a large number of visitors choose attractions closer to the coast. At the local level, it also seems that most visitors themselves come from coastal areas.

Italy (Tuscany-Emilia): There are no clear patterns for domestic and local visitors, apart from some indications that local visitors also come from nearby places.

Portugal: There is a clear pattern where both domestic and local visitors overwhelmingly chose coastal attractions in the peripheral region.

Slovakia: The pattern is less clear. However, there is some evidence that both domestic and local visitors also travel longer distances within the country/region to reach an attraction. There is also some evidence that domestic visitors often choose an attraction in central Slovakia.

Spain: There is a clearer pattern where domestic tourists from all over the country, but often from larger cities and the coast, often choose attractions near the coast. A larger number of domestic visitors also come from places closer to the attraction. Locally, there is more activity in the coastal regions, but visitors from all regions also visit attractions throughout the region.

Sweden: A larger number of domestic visitors come from the south, but with no clear pattern in the choice of attractions. There is no clear pattern for domestic visitors.

In summary, in some peripheral areas there are quite clear patterns of where tourists come from and where they go to, e.g. in Portugal and Italy (Sicily), where most visitors visit the coast, while in other peripheral areas patterns are difficult to generalize.

Time trends of tourism activity. Next, we look at the time trends in tourism flows for the peripheral and neighboring regions. In Figure 6 we illustrate how the number of Tripadvisor reviews in the peripheral regions and in the peripheral and neighboring regions changes over time for each peripheral region. The number of Tripadvisor reviews is a proxy for the number of visits in the regions, and therefore the figure illustrates how tourism activity changes. The vertical line indicates the start of the pilot INCUL- TUM action. Figure 6 illustrates a large drop in the number of reviews in all regions around the beginning of 2020 with the outbreak of the COVID-19 pandemic. Furthermore, the drop seems to last for a few years, after which there is a rebound. Figure 6 also shows a seasonal pattern with peaks during the peak season each summer.

Looking at each Panel of Figure 6 individually, there are some differences when comparing the peripheral regions with the total number of reviews in the peripheral areas. In Panel (g) it is clear that the Portuguese pilot region dominates in terms of reviews. The Sicilian periphery in Panel (e) also has a high proportion of all reviews, while in most other places the peripheral regions receive fewer reviews. In Panel (d), the Irish regions receive all the reviews, as we do not include any neighboring regions.

In Figure 7 we break down the total number of reviews for both the peripheral and neighboring regions into the different travel categories. The different shades of grey show the share of reviews for each of the four travel categories out of the total number of reviews in the peripheral area. Again, there are some interesting differences between the areas.

Albania: Only a small portion of the reviews of the Albanian peripheral area come from local and domestic visitors. There is a slight increase around the outbreak of the pandemic in 2020, but it does not last. On the other hand, European visitors seem to dominate in the Albanian peripheral area.

France: Local and domestic visitors are largely dominant, showing that only about 20% of visitors are from outside France. This pattern is probably not reflective of tourism in France as a whole, but shows the nature of the peripheral area, which is always attractive to domestic visitors.

Greece: There is a predominance of visitors from outside Greece. Local visitors represent a very small share, while domestic visitors represent about 40% during the high season. There is a slight change after the outbreak of COVID-19, but the levels seem to be restored from 2022. In the Greek peripheral area, there is an even distribution between European and non-European visitors.

Ireland: There is a clear increase in domestic travel from 2020 and a decrease from 2022. Domestic visitors are an important part of domestic travel, while international visitors are evenly distributed between European and non-European visitors.

Italy: The Sicilian region in Panel (e) has a higher share of domestic visitors, while the Tuscan region in Panel (f) has a higher share of international visitors, especially visitors from outside Europe. The Tuscan region also seems to have a greater change in composition as a consequence of the COVID-19 pandemic, with a sudden increase in the share of domestic visitors in 2020.

Portugal: There is a very high percentage of European visitors, most of them from abroad, except for a short period following the COVID-19 pandemic. On the other hand, the share of non-European visitors is quite low.

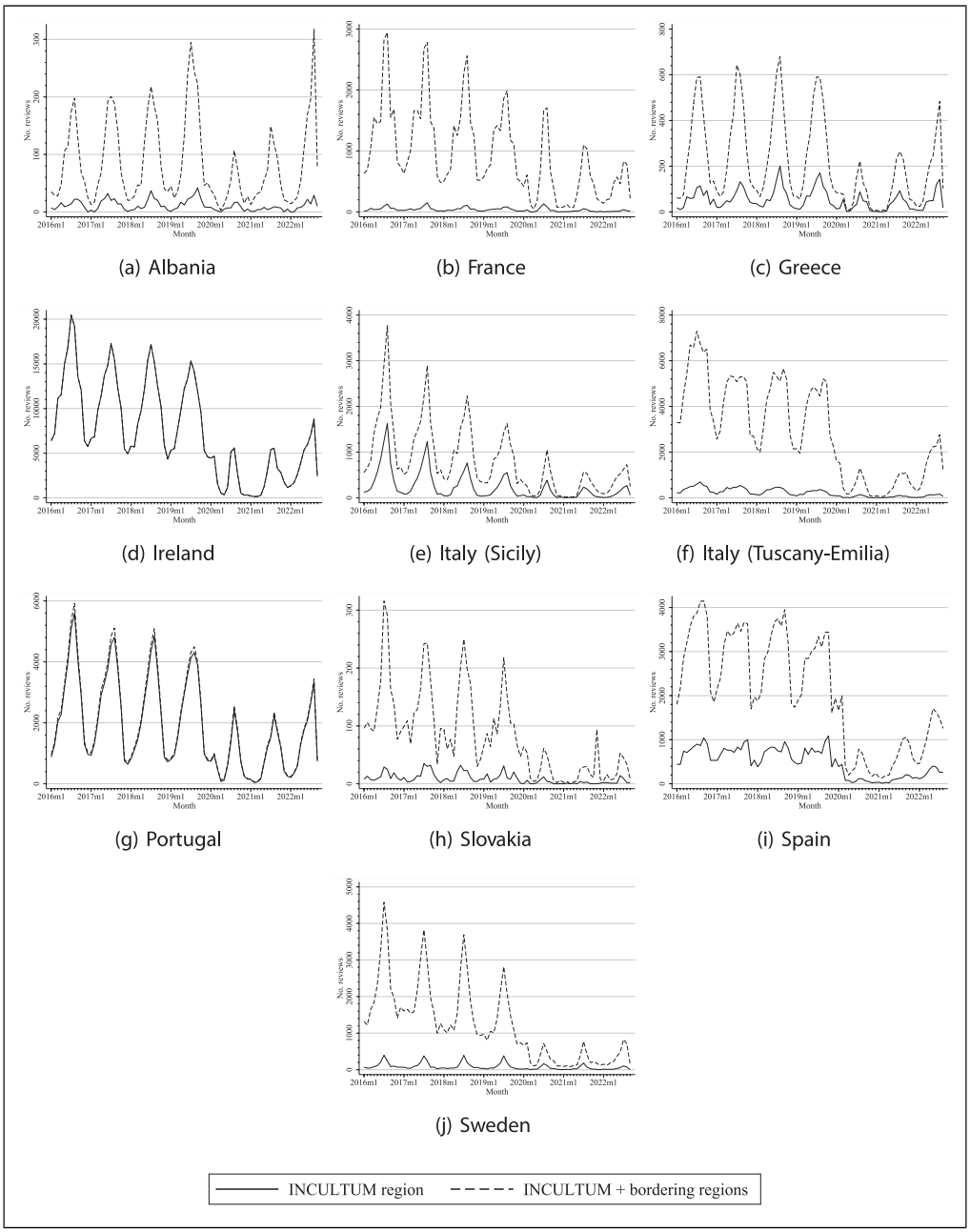


Figure 6. Number of reviews over time for peripheral and neighboring regions. *Notes:* This Figure shows number of Tripadvisor reviews in peripheral and neighboring regions. *Source:* Own data collected from Tripadvisor (see Data Description for details).

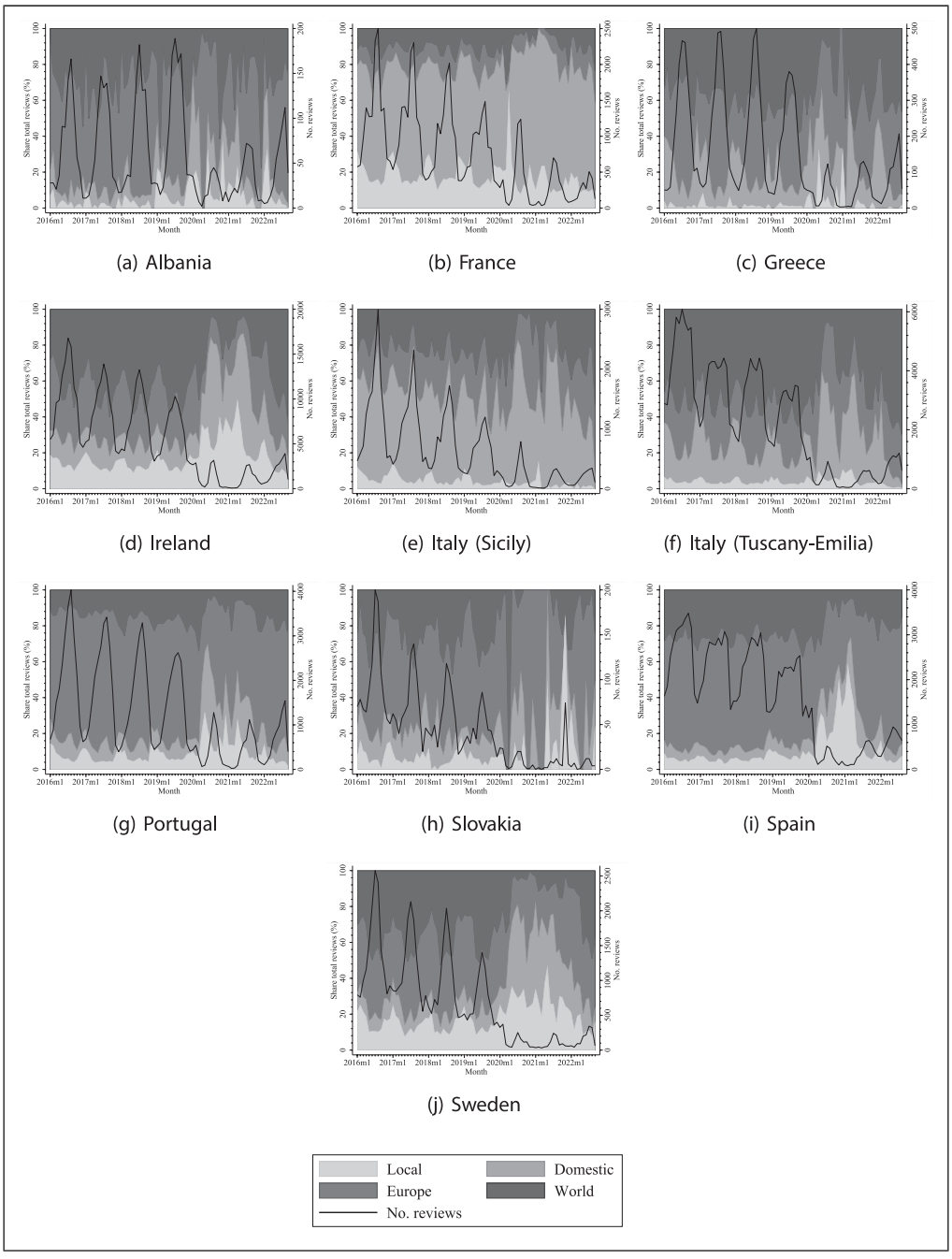


Figure 7. Share of reviews for different travel categories over time for peripheral and neighboring regions. *Notes:* This Figure shows the share of reviews out of the total in the following four travel categories: local, domestic, Europe and world. Regions refer to peripheral and neighboring regions. *Source:* Own data collected from Tripadvisor (see Data Description for details).

Slovakia: The level of both domestic and especially local visitors is quite low, with two major peaks in 2020 and 2021 where they reach almost 50% of all reviews. European visitors dominate, also reaching almost 100% of all reviews for shorter periods.

Spain: European travelers also dominate, followed by other visitors from outside Europe. Domestic visitors remain mostly below 20%, with the exception of a period from 2020 to 2021, when the COVID-19 pandemic was at its peak. From 2022 onwards, the figures are again below 20%.

Sweden: Similar to several other peripheral areas, international visitors dominate, with both European and non-European at similar levels of around 30% each. From 2020, local and domestic visitors reach levels of almost 80% of all reviews, although the levels drop again around 2022.

Overall, international travel appears to dominate in most peripheral areas, with the exception of France and Italy (Sicily). Furthermore, the areas with the highest levels of international travel are also the areas most affected by the outbreak of the COVID-19 pandemic. The higher the level of international travel before the pandemic, the greater the shift to domestic travel after the pandemic.

In Figure 8, we present the various travel categories for each country, focusing on the NUTS3 region receiving most reviews on Tripadvisor. The figure indicates that these highly attractive regions generally attract a larger share of international tourists compared to the peripheral areas considered in our study. Additionally, the impact of the COVID-19 pandemic appears to have had a more significant effect on the composition of tourists in these popular destinations, with the fraction of foreign tourists decreasing while the role of domestic and regional tourists grew significantly.

Another way to look at tourism flows is to look at the distances that visitors traveled to reach an attraction. In Appendix Figure F1 we show the average travel distance over time for each peripheral area. In almost all areas, except the Albanian peripheral area, there is a decrease around the beginning of 2020, followed by an increase in 2021 around the start of the pilot. In the Albanian peripheral area, Panel (a), there are no significant changes in travel distance. This could be due to the smaller number of reviews and, according to the conclusions from Figure 7, the fact that international travel in Albania remains quite high throughout the period. The decrease in travel distance indicates a shift from visitors coming from far away to more local visitors or a decrease in long-distance travel.

The peripheral regions in perspective. This final section illustrates how tourism flows have changed in the peripheral regions compared to national averages, more densely populated regions and the neighboring regions. To get an idea of the overall movements in the peripheral areas, we start by illustrating the fraction of reviews in the peripheral area out of the national level in Figure 9.[8] In areas such as France, Greece and Italy (Sicily), the fraction of reviews out of the national level is very low, reaching shares below 10% of all reviews in the country. On the other hand, in Sweden, Albania and Slovakia, the proportion of reviews in the peripheral areas is quite high, surpassing levels well above 50% of all reviews. Portugal, Spain and Italy (Tuscany-Emilia) are somewhere in between, with levels around 15%–25%. In some cases, e.g. in France and Slovakia, the fractions increase after the outbreak of COVID-19, while in other areas such as Sweden and Italy (Tuscany-Emilia) the opposite is happening.

Figure 10 compares the number of reviews in the peripheral regions with the number of reviews in the most densely populated NUTS3 region, while Figure 11 contrasts the number of reviews in the peripheral regions with the average number of reviews in the neighboring regions. The overall

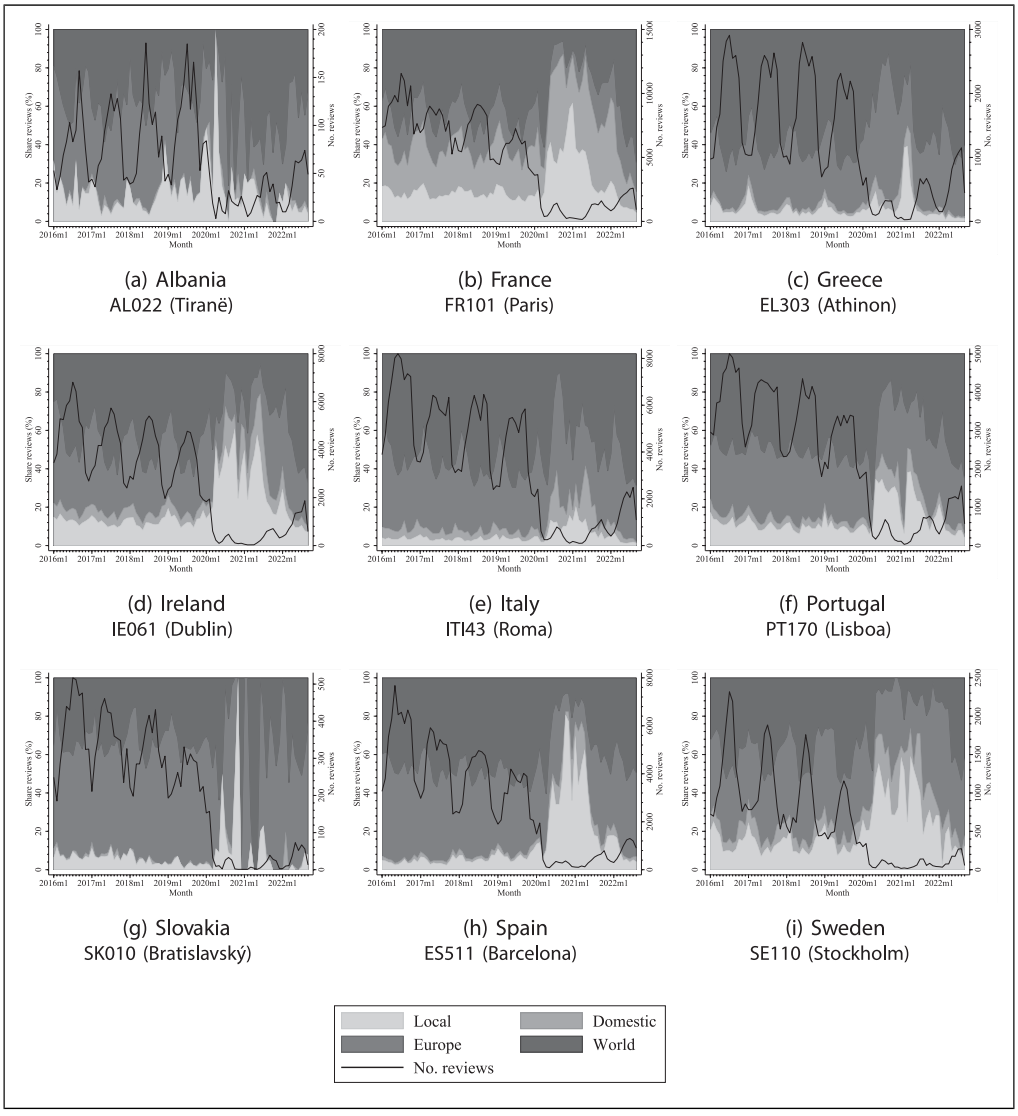


Figure 8. Share of reviews for different travel categories over time for the NUTS3 region with most reviews in each country. *Notes:* This Figure shows the share of reviews out of the total in the following four travel categories: local, domestic, Europe, and world, considering the NUTS3 region in each country receiving most reviews. *Source:* Own data collected from Tripadvisor (see Data Description for details).

trends are similar to what we saw previously in Figure 9. From Figure 10 it appears that in Portugal, the number of reviews in the peripheral region becomes greater than the number of reviews in the most densely populated region after 2020. In all other regions the share of reviews is lower in the peripheral regions spanning from well below 10% in France and Greece to shares above 30% in

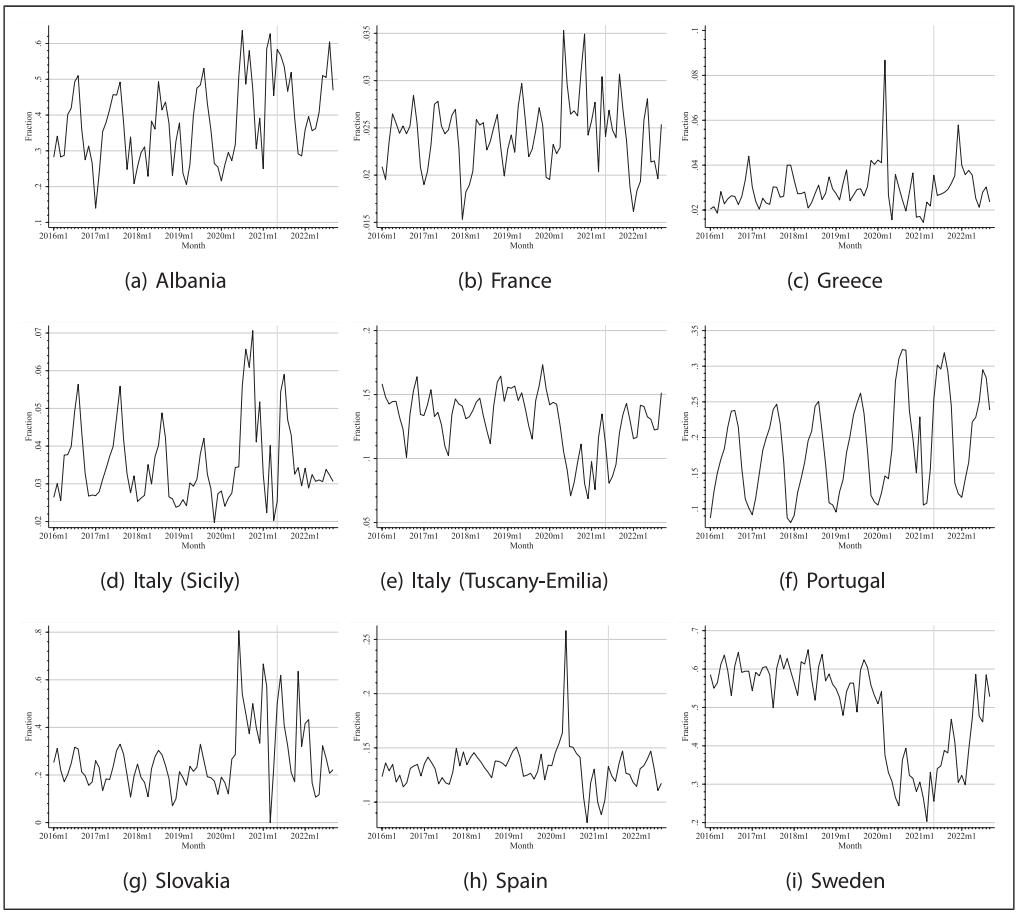


Figure 9. Fraction of reviews in peripheral and neighboring regions out of all reviews in country over time. *Notes:* This figure shows the fraction of Tripadvisor reviews referring to attractions in peripheral and neighboring regions out of the total number of reviews in each country. The vertical line indicates when the pilot started. In this figure, the Irish case is excluded, as we consider the entire country to be part of the pilot area. *Source:* Own data collected from Tripadvisor.

Spain and Sweden. Italy and Albania are somewhat in between. In Slovakia, the share is below 10% before 2020, where after it peaks. In Figure 11 when it comes to magnitude, in Greece, Italy (Tuscany-Emilia), Slovakia and Spain, the number of reviews received in the peripheral regions is greater than the average in the neighboring regions. The Portuguese peripheral region receives more than 20 times as many as the neighboring regions, while in the other cases, the numbers are lower but still indicate that they receive more at all times. In the remaining peripheral regions, the fractions vary a bit, with periods where the number of reviews in the peripheral regions is lower than in the neighboring regions and other periods where the number is higher (mainly in the period after 2020).

In general, the fractions follow the same trend in the period before the outbreak of the COVID-19 pandemic, where there is a sudden increase in the proportion of reviews in the peripheral regions, followed by a decrease below the pre-2020 level. Around the beginning of the pilot, several

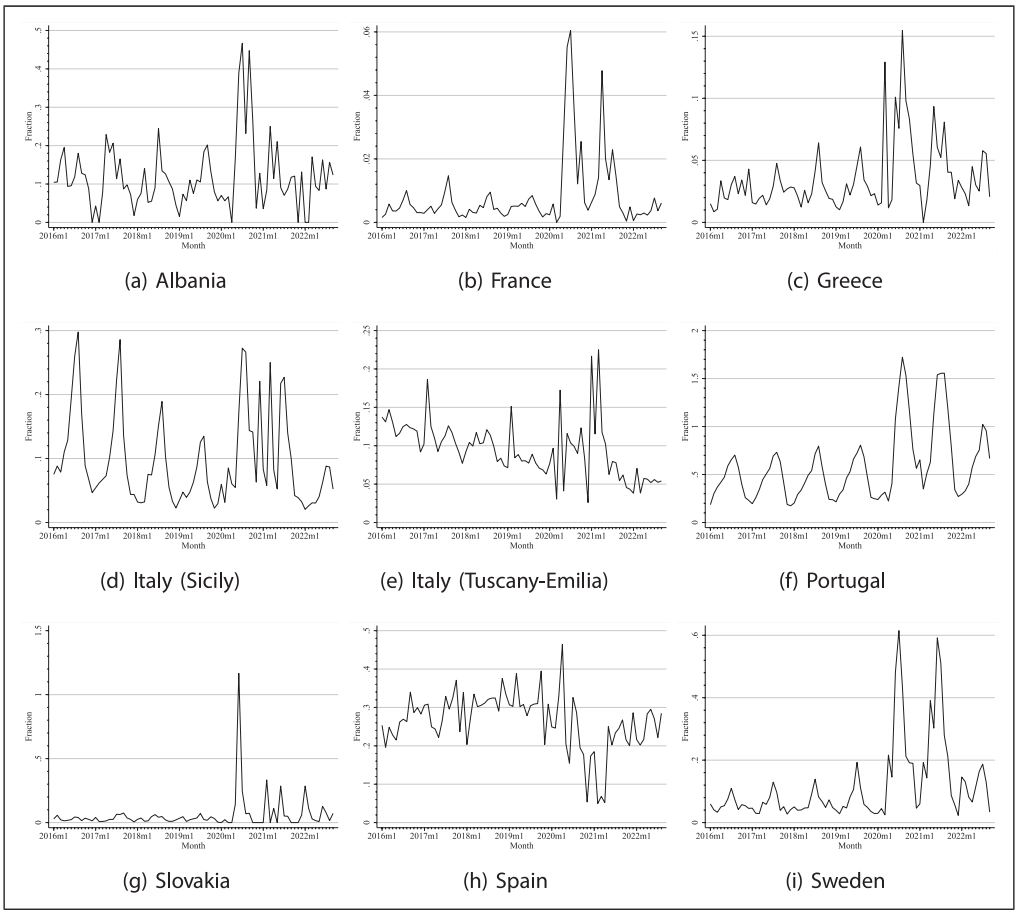


Figure 10. Fraction reviews in peripheral regions out of reviews in the most densely populated region in each country over time. *Notes:* This figure shows the fraction of Tripadvisor reviews referring to attractions in peripheral regions out of the total number of reviews in the most densely populated NUTS3 region in each country. The following NUTS3 regions have the highest population density in each country: AL022 (Tiranë), EL303 (Kentrikos Tomeas Athinon), ES300 (Madrid), FR101 (Paris), ITF33 (Napoli), PT170 (Área Metropolitana de Lisboa), SE110 (Stockholms län), and SK010 (Bratislavský kraj). In this figure, the Irish case is excluded, as we consider the entire country to be part of the pilot area. *Source:* Own data collected from Tripadvisor.

peripheral regions experience a new period with an increase in the share of reviews, interrupting the negative trend. One explanation for this increase is that the pilot measures implemented help the peripheral regions to return to their pre-pandemic levels. Whether this recovery would have happened without the pilot is hard to say given the pre-pandemic trends, but Figure 11 suggests that something is happening around this point. In Figure F2 in the appendix we also show the share of reviews in the peripheral regions out of the number of reviews in the NUTS3 regions receiving most reviews on Tripadvisor with very similar results.

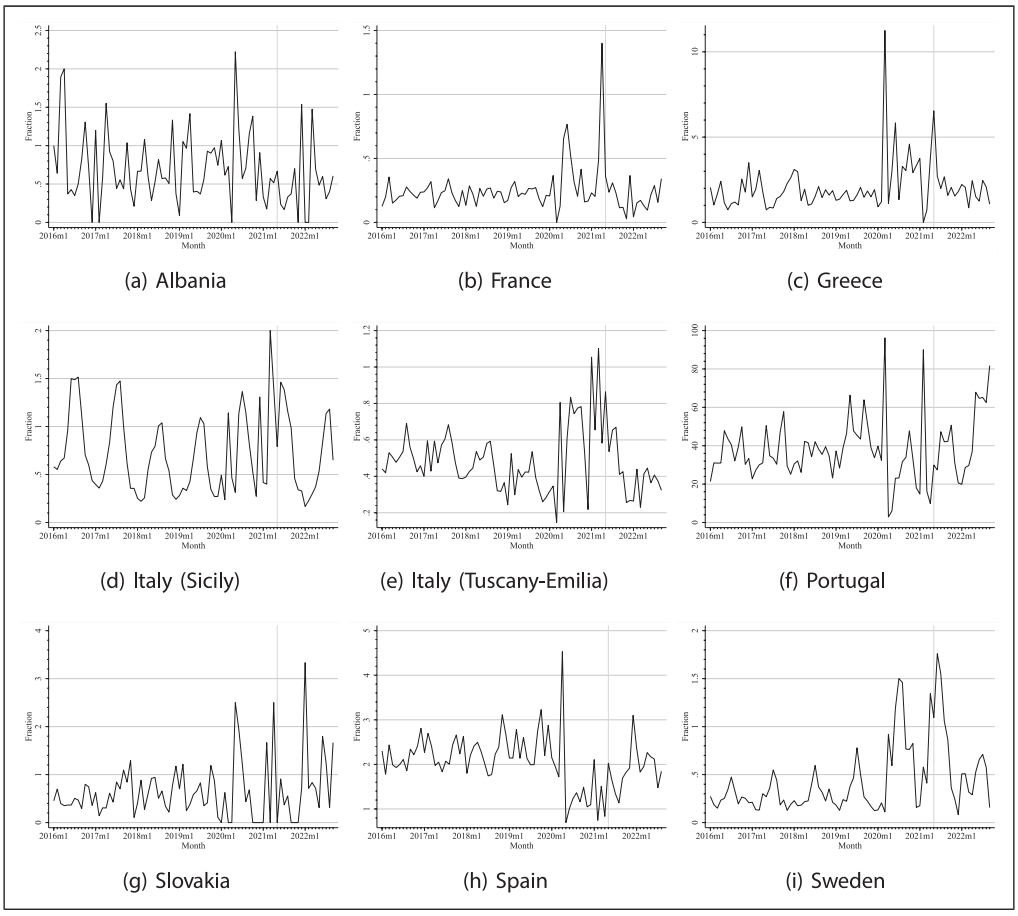


Figure 11. Fraction of number of reviews in peripheral regions out of average number in neighboring regions over time. *Notes:* This figure shows the fraction of Tripadvisor reviews referring to attractions in peripheral and neighboring regions. The vertical line indicates when the pilot started. In this figure, the Irish case is excluded, as we consider the entire country to be part of the pilot area. *Source:* Own data collected from Tripadvisor (see Data Description for details).

From the above analysis, we cannot say with certainty whether the pilot action has had an impact on tourism in the peripheral regions. Given the difficulties explained in the validity process, a longer period without the influence of COVID-19 would help to properly judge the impact of the pilot.

Conclusion

This paper has two main objectives: to evaluate the use of tourist reviews as a tool to measure tourism activity in peripheral regions and to use this novel data to report tourism development across 10 peripheral areas in Europe. Our analysis leverages user-generated content collected from a major travel portal, providing granular insights into peripheral tourism dynamics.

The findings provide applied evidence for the theoretical frameworks of sustainable tourism development and regional economic development, as outlined in the introduction. Sustainable

tourism development theory emphasizes the need to balance economic growth with ecological preservation and cultural heritage protection. This study applies these principles by demonstrating how granular, real-time data can reveal travel patterns in less-explored peripheral areas, offering insights into tourism distribution. By providing a tool to track tourism flows and their origins, the study under- scores the role of data-driven strategies in advancing sustainability goals in tourism management.

In terms of regional economic development, the results highlight the role of tourism in stimulating growth in less-visited areas. Peripheral tourism creates economic opportunities by attracting visitors, supporting local businesses, and fostering regional economic activity. Our analysis of tourism flows, particularly the shift toward domestic tourism during the COVID-19 pandemic, illustrates how these regions can benefit from increased local travel and spending. These findings offer empirical insights into how peripheral regions attract tourists and how their role within national tourism networks differs from that of more prominent destinations.

The study further reveals that peripheral regions, often rich in cultural and natural attractions, demonstrated resilience during the COVID-19 pandemic. The shift toward domestic tourism highlights the potential of these areas as alternatives to overcrowded urban destinations. By diversifying tourism activity and spreading its benefits more evenly, peripheral tourism can play a crucial role in advancing sustainable practices and fostering regional development. However, further research is needed to understand the long-term implications of these shifts, particularly whether tourism can be more evenly distributed across regions in a way that is both economically viable and environmentally sustainable.

From a policy perspective, the study offers practical contributions. Policymakers can use these findings to identify tourist types and tailor strategies accordingly. For example, domestic tourists, who often gravitate toward lesser-known attractions, present an opportunity to promote hidden cultural and natural assets, while international tourists may require a focus on iconic sites. Additionally, the ability to monitor tourism flows in real time can help optimize resource allocation, improve infrastructure planning, and enhance the visitor experience. These insights align with the theoretical perspectives discussed in the introduction by showing how tourism in peripheral regions follows distinct patterns that can inform targeted policy interventions. These data-driven approaches support sustainable and inclusive tourism development while preserving the unique heritage of peripheral areas.

While this study highlights the value of user-generated content for understanding peripheral tourism, it also has limitations. The dataset, reliant on a single travel platform, may not fully capture tourism dynamics in areas with low digital engagement. Furthermore, the selection of regions, guided by their inclusion in the INCULTUM project, may limit the generalizability of findings to other con- texts. Future research could expand this work by exploring causal impacts of policy interventions, investigating the appeal of different types of attractions, or employing text mining techniques to understand tourist motivations more deeply.

In conclusion, this paper demonstrates the applied potential of peripheral tourism in promoting sustainable and balanced regional development. By linking theoretical principles from sustainable tourism and regional economic development to empirical evidence, we emphasize the importance of leveraging cultural and natural assets for regional growth. Through a data-driven approach, this study enhances understanding of tourism activity in peripheral areas, offering valuable insights into their distinct dynamics. These findings provide a foundation for future research on regional tourism patterns and sustainability strategies. By shedding light on peripheral tourism, this study underscores its role in diversifying economic activity, preserving cultural and natural heritage, and fostering a more resilient and inclusive tourism sector.

Acknowledgements

We would like to thank Anne Møller Madsen and Sofus Hesseldahl Laubel for their research assistance in collecting the Tripadvisor data, and Martin Hørlyk Kristensen for his research assistance in preparing the data for the analysis. We also thank the participants at the INCULTUM International Conference and those who attended a seminar at the University of Southern Denmark for many helpful comments.

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) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the H2020 Societal Challenges: 101004552.

ORCID iDs

Karol Jan Borowiecki https://orcid.org/0000-0003-4959-181X
Maja U Pedersen https://orcid.org/0000-0001-8978-3010
Marco Palomeque https://orcid.org/0000-0002-1084-8027

Supplemental Material

Supplemental material for this article is available online.

Notes

1. INCULTUM (Innovative Cultural Tourism in European Peripheries) is a 3-year project funded by the European Union under the H2020 programme. Its primary objective is to investigate the obstacles and tap into the latent opportunities associated with under-rated destinations administered by local communities and stakeholders. For more details, visit: https://incultum.eu/. The original INCULTUM project included specific peripheral locations that, as is often the case with large-scale applications, were selected somewhat arbitrarily. In our analysis, we provide more precise maps showing the exact locations alongside other attractions across a broader area, specifically the NUTS3 regions where they are situated. These maps allow us to differentiate between tourism directed to the exact locations and tourism within the larger NUTS3 regions. In this section, we present a socio-economic overview of these regions using Eurostat data, aiming to explore the common characteristics of the selected peripheral areas and to familiarize the reader with the regions studied. Although we are not explicitly interested in measuring the effect of the INCULTUM project, we use it to have an exogenous peripheral region selection. We will refer to it simply as the "pilot" for now on if we have to say something about its action.
2. In the following we use three definitions: peripheral region, neighboring region and peripheral area. The peripheral region refers to NUTS3 regions where an INCULTUM pilot site is located. The neighboring regions refer to NUTS3 regions adjacent to a peripheral region. These neighboring regions are also peripheral in many cases, and they are included in order to provide a wider overview. The peripheral area refers to the total area of NUTS3 regions included in the analysis, i.e. both peripheral and neighboring regions.
3. The at-risk-of-poverty rate is defined as the share of individuals with an equivalized disposable income (after social transfers) below a threshold of 60% of the national median equivalized disposable income.
4. This is expected, as domestic reviews in these two countries, typically in the national language, could not be incorporated into our data.

5. The variance inflation factor (VIF) tests also indicate no issues with multicollinearity, with all VIF values well below the standard threshold of 10. These results are available upon request.
6. Once again, to illustrate that the high R^2 values are not an indication of issues with multicollinearity, we have performed a 20-fold cross-validation for each model. These results also indicate no issues with multicollinearity and are available upon request.
7. For computational reasons the figures represent smaller samples of actual users in cases where the number of users in a region exceeded 10,000.
8. This section excludes the Irish case, as we consider the entire country to be part of the pilot area.

References

Behsudi A (2020) Tourism-dependent economies are among those harmed the most by the pandemic. *IMF Finance & Development, Washington, DC*.

Bel F, Lacroix A, Lyser S, et al. (2015) Domestic demand for tourism in rural areas: insights from summer stays in three French regions. *Tourism Management* 46(C): 562–570.

Bergstrand JH, Larch M and Yotov YV (2015) Economic integration agreements, border effects, and distance elasticities in the gravity equation. *European Economic Review* 73: 307–327.

Bertacchini E, Nuccio M and Durio A (2021) Proximity tourism and cultural amenities: evidence from a regional museum card. *Tourism Economics* 27(1): 187–204.

Bobic S and Akhavan M (2022) Tourism gentrification in mediterranean heritage cities. The necessity for multidisciplinary planning. *Cities* 124: 103616.

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

Borowiecki KJ, Pedersen MU and Mitchell SB (2024) Using big data to measure cultural tourism in Europe with unprecedented precision. *Tourism Economics* 31(3): 477–503.

Botterill D, Owen E, Emanuel L, et al. (2000) Perceptions from the periphery: the experience of Wales. In: Brown F and Hall D (eds) *Tourism in Peripheral Areas, Aspects of Tourism*. United Kingdom: Channel View Publications, 7–38.

Brandano MG and Crociata A (2023) Cohesion policy, tourism and culture in Italy: a regional policy evaluation. *Regional Studies* 57(4): 763–779.

Brandano MG and Meleddu M (2021) Together or not? Spill-over effects of cultural goods displacement. *Tourism Economics* 27(6): 1202–1220.

Calero C and Turner LW (2020) Regional economic development and tourism: a literature review to highlight future directions for regional tourism research. *Tourism Economics* 26(1): 3–26.

Cambon J, Hernangómez D, Belanger C, et al. (2021) tidygeocoder: an R package for geocoding. *Journal of Open Source Software* 6(65): 3544, R package version 1.0.5.

Cerisola S and Panzera E (2024) Heritage tourism and local prosperity: an empirical investigation of their controversial relationship. *Tourism Economics* 1–20.

Chen C-C and Chang Y-C (2018) What drives purchase intention on Airbnb? Perspectives of consumer reviews, information quality, and media richness. *Telematics and Informatics* 35(5): 1512–1523.

Cheng M and Jin X (2019) What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management* 76: 58–70.

Cui M, Cheng L and Shang Y (2024) The influence of experiencescape of home-based accommodation on tourists' subjective well-being at cultural heritage sites: the role of value co- creation. *Journal of Destination Marketing & Management* 31: 100845.

Donati D (2022) The end of tourist traps: a natural experiment on the impact of tripadvisor on quality upgrading. *SSRN Electronic Journal* Working Paper No. 9834.

Eurostat (2022a) At-risk-of-poverty rate by NUTS regions. Online data code: ILC_LI41. https://ec.europa.eu/eurostat/databrowser/view/ILC_LI41/default/table?lang=en&category=livcon.ilc.ilc_ip.ilc_li Last data update: 29/09/2023.

Eurostat (2022b) Cultural employment by NUTS 2 regions. Online data code: CULT_EMP_REG. https://ec.europa.eu/eurostat/databrowser/view/cult_emp_reg/default/table?lang=en Last data update: 02/03/2023.

Eurostat (2022c) Gross domestic product (GDP) at current market prices by NUTS 3 regions. On- line data code: NAMA_10R_3GDP. https://ec.europa.eu/eurostat/databrowser/view/nama_10r_3gdp/default/table?lang=en Last data update: 21/02/2023.

Eurostat (2022d) Population by educational attainment level, sex and NUTS 2 regions. Online data code: EDAT_LFSE_04. https://ec.europa.eu/eurostat/databrowser/view/edat_lfse_04/default/table?lang=en Last data update: 14/09/2023.

Eurostat (2022e) Population density by NUTS 3 region. Online data code: DEMO_R_D3DENS. https://ec.europa.eu/eurostat/databrowser/view/DEMO_R_D3DENS/default/table?lang=en Last data update: 19/04/2023.

Eurostat (2022f) Population structure indicators by NUTS 3 region. Online data code: DEMO_R_PJANIND3. https://ec.europa.eu/eurostat/databrowser/view/demo_r_pjanind3/default/table?lang= Last data update: 28/09/2023.

Eurostat (2022g) Unemployment rates by sex, age, educational attainment level and NUTS 2 regions (%). Online data code: FST_R_LFU3RT. https://ec.europa.eu/eurostat/databrowser/view/LFST_R_LFU3RT/default/table?lang=en Last data update: 14/09/2023.

Eurostat (2023) Eurostat tourism statistics. https://ec.europa.eu/eurostat/web/tourism Last data update: 05/01/2023.

Faber B and Gaubert C (2019) Tourism and economic development: evidence from Mexico's coastline. *The American Economic Review* 109(6): 2245–2293.

Falk M, Larpin B and Scaglione M (2019) The role of specific attributes in determining prices of Airbnb listings in rural and urban locations. *International Journal of Hospitality Management* 83: 132–140.

Frochot I (2005) A benefit segmentation of tourists in rural areas: a Scottish perspective. *Tourism Management* 26(3): 335–346.

Gil-Alana LA and Poza C (2022) The impact of COVID-19 on the Spanish tourism sector. *Tourism Economics* 28(3): 646–653.

Goh C, Li H and Zhang Q (2015) Achieving balanced regional development in China: is domestic or international tourism more efficacious? *Tourism Economics* 21(2): 369–386.

Hunter C (1997) Sustainable tourism as an adaptive paradigm. *Annals of Tourism Research* 24(4): 850–867.

Jean L, Rozaini A, Radzol M, et al. (2019) The impact of social media influencers on purchase intention and the mediation effect of customer attitude. *Asian Journal of Business Research* 7: 19–36.

Kim S-E, Lee KY, Shin SI, et al. (2017) Effects of tourism information quality in social media on destination image formation: the case of Sina Weibo. *Information & Management* 54(6): 687–702. Smart Tourism: Traveler, Business, and Organizational Perspectives.

Komppula R (2005) Pursuing customer value in tourism–a rural tourism case-study. *Journal of Hospitality & Tourism* 3(2): 83–104.

Leick B, Kivedal BK, Eklund MA, et al. (2022) Exploring the relation- ship between Airbnb and traditional accommodation for regional variations of tourism markets. *Tourism Economics* 28(5): 1258–1279.

Leick B, Beth Mitchell S, Jan Borowiecki K, et al. (2024) Professionalisation and performance of Airbnb hosts in rural regions. *International Journal of Hospitality Management* 118: 103680.

Lenzen M, Sun Y-Y, Faturay F, et al. (2018) The carbon footprint of global tourism. *Nature Climate Change* 8: 522–528.

Luna LG and Surovtseva T (2020) Do More Tourists Promote Local Employment? Economics Working Papers 1746. Department of Economics and Business, Universitat Pompeu Fabra.

Mahat NZD and Hanafiah MH (2020) Help me TripAdvisor! Examining the relationship between TripAdvisor e-WOM attributes, trusts towards online reviews and travellers' behavioural intentions. *Journal of Information and Organizational Sciences* 44(1): 83–112.

Mansfeld Y and Ginosar O (1994) Determinants of locals' perceptions and attitudes towards tourism development in their locality. *Geoforum* 25(2): 227–248.

Martin R and Sunley P (2006) Path dependence and regional economic evolution. *Journal of Economic Geography* 6(4): 395–437.

Martin-Fuentes E, Mateu C and Fernandez C (2020) The more the merrier? Number of reviews versus score on TripAdvisor and Booking.com. *International Journal of Hospitality & Tourism Administration* 21(1): 1–14.

Mohamed Reda Khomsi LF-A and Rabier L (2020) A prospective analysis of overtourism in Montreal. *Journal of Travel & Tourism Marketing* 37(8-9): 873–886.

Molera L and Pilar Albaladejo I (2007) Profiling segments of tourists in rural areas of South- Eastern Spain. *Tourism Management* 28(3): 757–767.

Nocito S, Sartarelli M and Sobbrio F (2023) A beam of light: media, tourism and economic development. *Journal of Urban Economics* 137: 103575.

Panzer-Krause S (2019) Networking towards sustainable tourism: innovations between green growth and degrowth strategies. *Regional Studies* 53(7): 927–938.

Plzáková L and Smeral E (2022) Impact of the COVID-19 crisis on European tourism. *Tourism Economics* 28(1): 91–109.

Rosselló-Nadal J and Santana-Gallego M (2024) Toward a smaller world. The distance puzzle and international border for tourism. *Journal of Transport Geography* 115: 103809.

Sainaghi R (2020) Determinants of price and revenue for peer-to-peer hosts. The state of the art. *International Journal of Contemporary Hospitality Management* 33(2): 557–586.

Santana-Jiménez Y and Hernández JM (2011) Estimating the effect of overcrowding on tourist attraction: the case of Canary Islands. *Tourism Management* 32(2): 415–425.

Seraphin H and Dosquet F (2020) Mountain tourism and second home tourism as post COVID- 19 lockdown placebo? *Worldwide Hospitality and Tourism Themes* 12: 485–500.

Sharpley R (2009) *Tourism Development and the Environment: Beyond Sustainability?* London: Earth Scan.

Storper M (1997) *The Regional World: Territorial Development in a Global Economy*. New York: Guilford Press.

Vaishar A and Šťastná M (2022) Impact of the COVID-19 pandemic on rural tourism in Czechia preliminary considerations. *Current Issues in Tourism* 25(2): 187–191.

Vu JC and Turner L (2009) The economic structure of world tourism. *Tourism Economics* 15(1): 5–21.

Wang C, Meng X, Siriwardana M, et al. (2022) The impact of COVID-19 on the Chinese tourism industry. *Tourism Economics* 28(1): 131–152.

World Health Organization (2020) WHO virtual press conference on COVID-19. https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3_2

World Tourism Organization (2019) *Transport-related CO_2 Emissions of the Tourism Sector – Modelling Results*. Madrid, Spain: World Tourism Organization.

Ye S, Xiao H and Zhou L (2019) Small accommodation business growth in rural areas: effects on guest experience and financial performance. *International Journal of Hospitality Management* 76: 29–38.

Yotov YV (2012) A simple solution to the distance puzzle in international trade. *Economics Letters* 117(3): 794–798.

Author biographies

Karol J. Borowiecki, Professor of economics at the University of Southern Denmark, is renowned for his innovative research methodologies and societal impact. He published >40 items, including in the Journal of Political Economy, a textbook with Cambridge University Press, and a co-edited volume on cultural heritage. He sits on the editorial boards of Tourism Economics and the Journal of Cultural Economics. As President of the Association for Cultural Economics International and a top cultural economist, Karol collaborates with premier European institutions, shaping policy and advancing cultural and tourism economics.

Maja U. Pedersen, PhD degree in economics, currently Assistant Professor of economics at the University of Southern Denmark. Published 7 items on different topics, including economic growth, financial history and globalization. All published in various peer-reviewed journals.

Marco Palomeque, currently postdoc at the University of Southern Denmark. Specialized in Cultural Economics, focuses his research on innovative methodologies and creative research questions, applying Data Science techniques into economic problems. Besides his short career, his research has demonstrated its societal impact with several interviews in international media. His research also highlights in the scientific community, with relevant recognitions such as the President's Prize of the 2023 Conference of the Association for Cultural Economics International.