A Country of Waiters: The Economic Consequences of Tourism Specialization

Ghizlen Ouasbaa
Universitat de Barcelona

February 2023

PRELIMINARY DRAFT- Please do not circulate & comments are welcome!

Abstract

This paper analyzes the long-term localized effects of tourism specialization on income levels in Spain and examines the mechanisms behind this impact. I use two different empirical strategies to exploit cross-sectional variation in tourism exposure. The first strategy focuses on the initial period of tourism development in the 1960s, while the second strategy examines the second wave of tourism development in the 1990s, driven by exogenous changes. I document that municipalities that were more attractive to tourism before the 1960s or the second development wave in the 1990s have around 8% less income level in 2019. This effect can be attributed to several factors, including lower job stability, a higher proportion of temporary contracts, lower employment rates in industry, and reduced levels of education.

Keywords: Tourism specialization, local economic growth, long-term effects, local labor markets

1 Introduction

Tourism industry is a major source of income for many countries and is often put forth as a potentially important source of economic growth for others. According to the World Travel & Tourism Council (WTTC), prior to the COVID-19 pandemic, tourism accounted for one in four of all new jobs created worldwide, equivalent to 10.3% of all jobs (333 million), and 10.3% of global GDP.[1] As a result, both developed and developing countries invest heavily in infrastructure and promotion of their tourism sector and become increasingly reliant on its revenues.

The existing evidence on the impact of tourism on economic development is mixed, with both positive and negative effects on the local economy and the wider country. On the one hand, tourism can facilitate structural change in economies that depend on the primary sector by reallocating economic activity across manufacturing and services. Additionally, tourism can bring foreign exchange, generate employment, spur local investments, exploit economies of scale, and diffuse technical knowledge (Song et al., 2012).

On the other hand, some studies suggest that the tourism sector has low productivity and may displace other sectors characterized by more intense technological progress, ultimately hindering economic development (the so-called ”Dutch Disease”; Copeland, 1991; Holzner, 2011). Moreover, other studies (Parrilla et al., 2007; Sequeira and Maçãs Nunes, 2008; Arezki, 2009; Brida et al., 2016) suggest that the employment generated by tourism is often temporary and low-skilled, which can result in low wages, limited investment in human capital, or demand volatility (e.g., due to terrorist attacks or pandemics).

This paper brings new evidence to the debate on the long-term impact of tourism on the local economy by providing new empirical evidence on the localized effects of tourism specialization on income levels in Spain. Additionally, this paper aims to examine the mechanisms behind this impact, including the effects on education levels, job stability, and sector displacement. Spain is an ideal setting for this study, as it has experienced an unprecedented increase in tourism since the 1960s, and the industry accounts for a large share of employment (13% in

[1] https://wttc.org/Research/Economic-Impact (consulted on June 3, 2022)

2019). Spain received over 83 million international arrivals and generated 71.2 billion euros in tourism revenue in 2019, making it one of the world's top tourist destinations. The country's weather and beaches are the main attractions for tourists, resulting in a concentration of tourism in warmer and coastal regions. In fact, over 70% of all tourism activity in 2019 was located in coastal municipalities, according to the National Statistics Institute.

To assess the long-term effect of tourism on the local economy in Spain, I employ two different empirical strategies that exploit cross-sectional variation in tourism exposure. The first strategy focuses on the initial period of tourism development in the 1960s, while the second strategy examines the second wave of tourism development in the 1990s, which was driven by exogenous changes such as improvements in air accessibility, transportation, and new forms of accommodation. By exploiting these two distinct periods, I am able to capture the long-term and medium-term effects of tourism on economic outcomes across various Spanish municipalities.

One of the main challenges in studying the causal impact of tourism on economic outcomes is disentangling the effects of tourism from other factors that may be driving overall economic growth in a tourist destination. For instance, local investments and infrastructure development may contribute to economic growth in a tourism destination, but they may be independent of the tourism industry itself. This makes it difficult to isolate the specific effects of tourism on economic outcomes, as tourism may be confounded with other factors that are also contributing to economic growth.

To address the endogeneity issue between tourism and local economic development, I adopt two different approaches. For the first wave of tourism development in the 1960s, I follow the methodology developed by Fabert and Gaubert (2019), and draw on insights from the tourism literature that argues that the quality of local natural amenities significantly impacts tourism activities. Using this, I construct an instrumental variable for tourism attractiveness, I leverage the variation in beach amenities along the Spanish coastline. Specifically, I use aerial photos from the PNOA histórico dataset (1956-57, Instituto Geográfico Nacional) to measure the surface area of sandy beaches across Spanish municipalities. I use the existence of a beach, the fraction of the onshore coastline covered by beach, and other weather features as proxies for tourism attractiveness. By comparing municipalities with beach (treated municipalities) with those in interior municipalities (control municipalities) in coastal provinces, I estimate the causal impact of tourism specialization on the long-term economic outcomes of these municipalities.

The second approach I employ to study the impact of tourism on local economic outcomes focuses on the late 1990s and the changes that occurred during this period in the tourism market. Since sun-and-beach tourism is the primary attraction in Spain, I restrict the sample to 437 municipalities with beaches. This period witnessed changes in policy regulations that improved air accessibility, led to the arrival of low-cost airlines, and the introduction of Airbnb, which significantly impacted tourism inflows to Spain. To exploit this, I document that municipalities with a higher number of residents from tourist countries in 1996 experienced a greater inflow of tourism and a subsequently higher number of overnight stays in hotels after the exogenous changes in the late 1990s. This strategy involves distributing the positive increase in tourism inflows across municipalities, based on their pre-established composition of residents by nationality, and is similar to the shift-share analysis in Bartik (1991). I focus on the overnight stays in hotels-induced increases in beach municipalities between 1997 and 2019 and compare them based on the number of overnight stays predicted by the share of residents by nationality in 1996 and the growth of overnight stays by nationality at the national level.

I employ two different strategies to study the impact of tourism on local economic outcomes because of data availability limitations and the differences in the development of the tourism industry across the two periods. Specifically, data on residents by nationality was only available from 1996 onwards, and the two periods had different levels of tourism development: the 1960s marked the early stages of tourism in Spain, while by the mid-90s there were already well-established tourist areas. By using these two strategies, I can capture the long-term and medium-term causal effects of tourism on local economic outcomes.

Using both research designs, I find strong evidence that higher tourism specialization is associated with lower municipality income per capita in the long and medium-term compared to less touristic areas. Specifically, the results suggest that tourism development decreased the municipalities' income per capita by 8%. I also investigate the potential channels through which tourism negatively affects income per-capita. I find that tourism led to a shift in the local economy, resulting in a greater specialization in the service sector, as well as an increase in demand-related industries such as construction, real estate, hotels, and food-service. However, tourism also resulted in a decrease in manufacturing and agricultural employment, lower educational attainment, and ultimately, higher job instability, since the tourism sector often relies on a relatively large portion of part-time and temporary contracts.

This paper contributes to several strands of literature. First, more broadly, this work is related to a large literature that studies the impact of structural transformation. Most of papers focus on the impact of labor reallocation from agriculture to non-agriculture as a consequence of growth (Makarski et al., 2022; Hjort and Poulsen, 2019; Gollin et, al. 2016; Herrendorf et al., 2014; McMillan and Rodrik, 2011; Cimoli, 2005; Katz, 2000). These studies focus on productivity gains in a large set of poor countries since for most of their countries income data from population censuses are not available.[2] Instead, I focus on a high income country and I use labor income data to study the within country impact of structural transformation.

In terms of the causal impact of tourism on economic development, this analysis complements Faber and Gaubert (2019) as they study the long-term impact of tourism on the development of local economies in Mexico, with a specific focus on general equilibrium effects. This analysis differs both in terms of economic context (a developing country in their case, and on a developed country in my case) and in terms of outcomes studied. Also, another relevant papers in this strand of literature are González & Surovtseva (2020) they focus on the short-term impact of tourism shocks on employment at the regional level in Spain and Nocito et al. (2021) who provide evidence of the impact of tourists' expenditure on municipality income in the short-run. Also, several relevant papers provide estimate of the spatially heterogeneous welfare impact of tourism on locals throughout a

[2] The evidence is mixed, for some countries (in Latin America and Africa) structural change had a negative contribution to overall growth while in Asia had a positive contribution. According to McMillan and Rodrik (2011), structural change, with labor moving predominantly from manufacturing to service industries, in advanced economies have (on its own) has made little difference to productivity overall.

city, on consumption amenities and housing markets (Allen et al.,2020; Almagro and Dominguez-Lino, 2019; Garcia-Lopez et al., 2019). While these papers and most of previous empirical evidence at the subnational level refers to regions or provinces and short-term (with the exception of Faber and Gaubert, 2019), I look at smaller geographical units, the 2658 Spanish municipalities in coastal provinces. Given the high concentration of tourists in a specific municipalities within a region or province, this level of analysis should capture the effect of tourism on local outcomes more accurately.[3] Additionally, the focus on the long term is relevant since the short term effect may be positive and more clear but the long term it is not obvious and the impact of the income level less clearer.

And finally, the paper also relates to the literature that studies possible "Dutch disease" effects associated with natural resource booms by comparing either regional outcomes within countries or providing theoretical models (e.g. Allcott and Keniston 2018, Caselli and Michaels 2013, Corden 1984 and Corden and Neary 1982). The focus on tourism as a special kind of natural resource boom differ from the existing literature, but the economic questions are closely related.

To sum up, the main contributions are the following ones. Firstly, it examines the impact of tourism on a developed country with a high dependency on tourism, where many municipalities derive their revenue from tourism. Secondly, it utilizes unique municipality-level datasets to track localized changes in income and other outcomes, which would not be possible with regional or country-level data. Thirdly, the study provides direct evidence of the mechanisms through which tourism affects socioeconomic outcomes, such as an increase in temporary contracts and changes in the sectoral composition of municipalities. Fourthly, this is the first study to estimate the long-term effects of tourism on a tourism-oriented economy, which is important given the potential differences between short and long-term effects. For example, while tourism may generate high levels of economic activity in the short-term and increase GDP, the long-term adverse effects are unknown. Overall, the analysis suggests that tourism-oriented municipalities in Spain have lower income in the long-term.

[3] For example, in 2019 the municipality of San Bartolomé absorbs the 67% of the total tourist presences of Gran Canaria (and the 12,7% of Spain) and Benidorm the 64% of Alicante (11,6% of the national total).

The paper is structured as follows: Section 2 presents the theoretical expectations behind the results. Section 3 develops in details the Spanish institutional context, highlighting the role of tourism in the economy. Section 4 presents the empirical strategy. Section 5 describes the data and summarizes the main variables used in the analysis. Sections 6 and 7 present the main findings and robustness checks. Finally, Section 8 concludes the paper.

2 Theoretical expectations

The Spanish tourist boom in the 1960s had a significant impact on municipalities that had accessible natural resources such as beaches, high temperatures, and sunny weather. The arrival of tourism in these areas led to the discovery of new uses for these natural resources and a subsequent redistribution of the economy, resulting in specialization in the service sector. According to the natural resource theory (Corden, 1984; Corden & Neary, 1982), the availability of a natural resource leads to an increase in marginal labor productivity within the tourist sector, resulting in a rise in the demand for labor. Since the tourist and non-tradable sectors are labor-intensive, employment in these municipalities is mainly concentrated in the service sector, with fewer jobs in other sectors such as industry and agriculture. Theoretically, this process initially results in a decline in the production of tradable commodities, followed by a likely increase in the production of non-tradable commodities, services, construction work, and other service sectors closely linked to the tourist industry.

Although tourism can bring short-term benefits to a country's economy by improving its economic scenario, its long-term impact is still unclear. Specializing in tourism changes the composition of exports, resulting in a decline in the manufacturing sector, which is more receptive to technological development and innovation (Capó et al. 2007). This affects the economy's growth potential by limiting a crucial source of human development and productivity growth, according to Torvik (2001). Additionally, Gylfason (2001a, 2001b) found evidence that natural resource discovery could hinder the growth of an economy in the long run due to the lower training levels associated with the exploitation of natural resources, which generally have a higher proportion of unskilled workers.

The tourism sector usually employs a large portion of low-skilled, part-time workers or those on temporary contracts paid at minimum wage. While work may be readily available during the high season, there is often high unemployment between peaks. Human capital accumulation models suggest that the substantial increase in productivity of low-skill jobs resulting from the tourist boom may lead some individuals to stop their education, at least in the short term (Becker, 1964; Black et al., 2005b; Charles et al., 2015). Low-skill jobs become more attractive to young adults, increasing the opportunity cost of finishing high school or going to college, and reducing the perceived returns of education.

3 Institutional context

This study focuses on the empirical context of Spain, a country where tourism has become a significant economic force. Compared to other OECD countries, tourism in Spain accounts for a larger percentage of GDP and employment. Specifically, tourism accounts for an average of 12.4% of GDP and 12.9% of employment in Spain, compared to the OECD average of 4.4% of GDP and 6.9% of employment.

The bulk of tourism activity in Spain is driven by coastal tourism, with 70% of tourism activity concentrated in coastal regions, particularly in the 464 coastal municipalities. Despite representing only 16% of the total population, these coastal municipalities generate a significant amount of tourism economic activity.

The development of tourism in Spain, particularly beach-and-sun tourism, is widely acknowledged to have started with the 'Plan de Estabilización' (Stabilisation Plan) of 1959. This plan, which included a significant devaluation of the national currency, marked Spain's return to the international markets after the dictatorship and the end of extreme autarchic policies. The effects of this plan on European tourists seeking sun and beach tourism in Spain were immediate and contributed to a sharp increase in tourism during the 1960s (Sánchez-Sánchez, 2001).[4]

[4]The Stabilization Plan of 1959 refers to a set of economic measures that were implemented by the Spanish government from 1959 onwards. The main objective of the plan was to liberalize the Spanish markets and mark a shift away from the country's previous policies that aimed at achieving autarky. The plan was aimed at stabilizing the economy, reducing inflation, and increasing foreign investment. It involved the devaluation of the national currency, which helped

Between 1997 and 2007, Spanish tourism experienced a second strong development period characterized by continuous growth in the number of international tourists (Albaladejo et al., 2020). This growth was primarily driven by exogenous changes in European regulations, which led to significant changes in the tourism industry in Spain including improvements in air accessibility, transportation, and the emergence of new forms of accommodation. Furthermore, this led to the introduction of low-cost airlines and the opening of new air routes. Thus, travel to Spain became more accessible and affordable for tourists from all over Europe. The number of international tourists visiting Spain increased significantly, particularly in coastal regions where sun and beach tourism are popular. Despite the increase in the number of tourists starting in the late 1990s, the real jump in growth did not occur until 2002, continuing until 2007, with the number of exits also increasing significantly in 2003 (Mason et al., 2016).

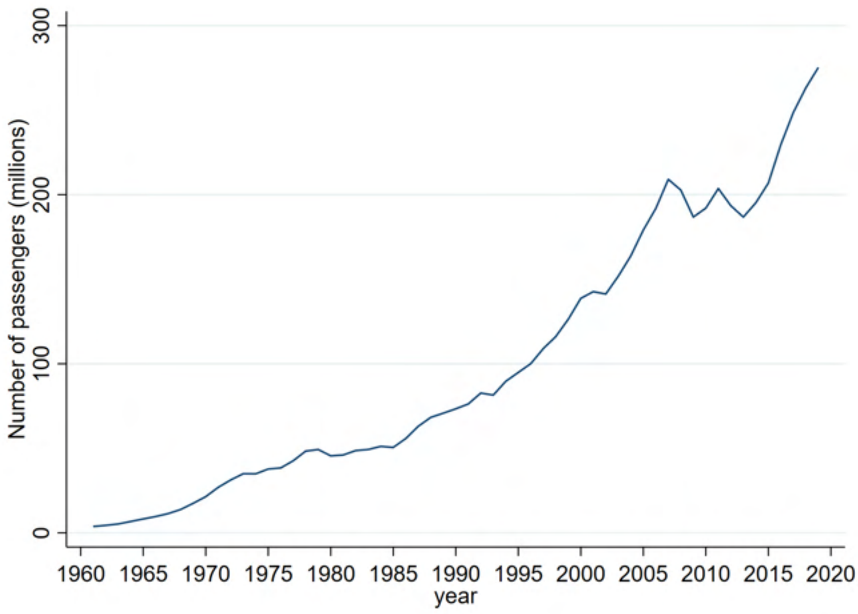
Specifically, the data in Figure 1 shows the numbers of international tourists arriving in Spain from 1961 to 2019, indicating a consistent growth in international arrivals since the 1960s, with several growth phases. The figure reveals two distinct periods of intense growth, the first of which was driven by the tourism boom of the 1960s. The second period was characterized by an accelerated growth at the end of the 1990s and the first years of the 21st century.

By 2019, Spain received 83 million foreign visitors. According to the Statistics on Tourist Movements at the Borders (FRONTUR), this number was less than 5 million foreign visitors in the 1960s and almost no visitors in 1950. The majority of these visitors are from Europe, accounting for 75 percent of foreign tourists, followed by US Americans (4 percent) and Russians (1.5 percent).[5] International visitors make up roughly 70 percent of the tourism activity in Spain.

to make Spanish exports more competitive on the international market.

[5]Britons account for 21%, Germans 13%, French 13%, Italians 5,4%, Dutch 4,4% and Belgium 3%.

Figure 1. Evolution of the number of international passengers arriving in Spain



Notes: (i) This figure illustrates the evolution of international tourist arrivals in Spain spanning from 1961 to 2019. (ii) Source: National Statistic Institute (INE).

4 Empirical Analysis

The development of the tourism industry likely goes hand in hand with overall economic growth in a touristic destination. The appeal of a given tourist destination is a function of fixed factors such as geographic characteristics, climate and historical relevance. To accurately estimate the causal effect of tourism, it is crucial to identify the factors that drive tourist inflows to Spain. To estimate the long-term causal impact of tourism on municipalities' outcomes, I use two distinct empirical approaches. The first strategy covers a broader period from 1960 to 2019, leveraging the early phase of tourism development. The second approach spans from 1997 to 2019, taking advantage of the second wave of tourism growth in the late 1990s. I discuss each of these strategies in the following sections.

4.1 Long-Term Analysis from 1960-2019: Beach amenities

In this subsection, I present the first empirical strategy aimed at capturing the long-term economic consequences of tourism on local economic outcomes across Spanish municipalities. To do so, I exploit cross-sectional variation in tourism municipalities. I begin by documenting the effects of tourism on municipality-level income in the current cross-section of Spanish municipalities, and estimate the following regression:

$$y_{it} = \beta \Delta Tourismpc_{it} + \delta X_{it} + \gamma_p + \epsilon_{it} \quad (1)$$

where y_{it} is the outcome of interest in per-capita terms (e.g. log per capita gross income in 2019) in municipality i, year t. The variable $Tourismpc_{it}$ measures the absolute per capita tourism growth in municipality i, and in year t. The vector, X_{it}, includes a number of municipal-level controls that I describe in the next section and γ_p are province fixed effects. Per capita tourism may be endogenous in this case because it is likely to be correlated with the economic conditions of the municipality, which can affect the level of tourism. For example, municipalities with higher income levels or better economic growth may be more attractive tourist destinations, leading to higher levels of per capita tourism. At the same time, tourism can also impact the local economy, potentially causing reverse causality, where higher levels of tourism lead to higher economic growth. This would lead to a biased OLS estimation of the equation (1).

To address the issue of endogeneity, I use an instrumental variable to control for potential bias in the estimation of the effect of per-capita tourism on the outcome variable. Since the Spanish tourist boom in the 1960s primarily affected municipalities with easily accessible natural resources such as beaches, high temperatures, and sunny weather, I use the existence of a beach in a municipality as an instrument for per-capita tourism growth. This approach exploits plausibly exogenous variation in tourism attractiveness across Spanish municipalities and follows the argument made by Fabert and Gaubert (2019) that local natural amenities, particularly the presence of a beach, are a key determinant of tourism activity.

I compare municipalities that had amenities for tourism at the founding moment of tourism in the 1960s with those that did not. Since almost the entire

Spanish coast have beach (except for 2%), I choose inland municipalities as the control group in this case. I control for several socioeconomic characteristics prior to the beginning of tourism, described in the next section, to make both inland and coastal municipalities comparable. The main advantage of this approach is that it allows to capture the long-term effects of tourism on the local economy, starting from the founding moment.

In the first stage, I estimate the following equation:

$$\Delta Tourismpc_{it} = \beta \mathbb{1}(Beach_i) + \delta X_{it} + \gamma_p + \epsilon_{it} \quad (2)$$

Where $\mathbb{1}(Beach_i)$ is a dummy variable that takes on a value of one if a municipality has a beach, and zero otherwise. Finally, I am able to estimate the second stage equation to measure the causal relation of interest:

$$\Delta y_{it} = \beta \Delta \widehat{Tourismpc}_{it} + \delta X_{it} + \gamma_p + \epsilon_{it} \quad (3)$$

The identifying assumption in specification (3) is that the presence of a beach affects municipality-level economic outcomes relative to other municipalities only through its impact on local tourism activity. The main coefficient of interest, β, captures the marginal effects of an increase in per capita tourism growth on the outcomes of interest. To ensure a meaningful interpretation of this coefficient, it is essential to control for the determinants of long-run development that correlate with tourism.

In Table A.2, I examine the correlates of tourism and report the OLS estimate of regressing per capita tourism growth on each variable separately. I also report standardized-beta coefficients and corresponding p-values in columns 3 and 4, respectively. The results indicate that higher per-capita tourism growth is correlated with a variety of geographic and demographic characteristics measured before the tourism boom. On average, municipalities with beach tended to have larger populations, lower levels of illiteracy, and lower agricultural production than inland municipalities. They also had higher temperatures and less rugged land. In my main specifications, I incorporate these covariates at baseline to mitigate endogeneity concerns. Also, I select the control municipalities using “entropy balancing”, this stem implies reweighting municipalities to improve the covariate balance between the treatment and control group such that the treatment variable becomes closer to being independent of the background characteristics (Hainmueller, 2012).

Furthermore, I create a measure (*Beach measure*) to identify tourism intensity which takes into account not only whether a municipality has a beach or not, but also the length and width of the beach, as well as temperature, hours of sun, and precipitation. To do this, I use 1957 aerial photos and GIS databases to obtain this information for each municipality.[6] This variable captures the quality and attractiveness of a beach. To account for this, I include an interaction term between this variable and the dummy variable *Beach* in Equation (2). Table A.3 shows the correlation between tourism and the variables used to construct *Beach measure*. We can observe that higher beach length and width, as well as more hours of sun and higher temperatures, are positively associated with higher levels of tourism.

4.2 Exploiting the Second Wave: Shift-share analysis

An alternative strategy to examine the medium-term causal impact of tourism is to analyze the second wave of tourism development that occurred in Spain during the late 1990s and early 2000s. In this period there were several changes in the tourism market, such as improvements in air accessibility, transportation, and new forms of accommodations, which had a significant impact on the inflow of tourism to Spain. Since the main attraction for tourism in Spain is sun-and-beach tourism, I choose to focus on a sample of 437 municipalities with beaches.

To analyze the impact of an exogenous increase in the number of tourists, an ideal approach is to estimate the effect of tourism activity at the municipality level. This can be done by using the per capita number of overnight stays in tourist accommodations, which is a widely used measure of tourism activity. Ideally, I

[6] Coastline turns into a significant variable for a Mediterranean country, where tourism demand is characterized by the model of sun and beach. Given that most of tourism activities take place outdoors, they depend on the climate variations as for instance sun and beach destinations (Frechtling 2001). I have included the length and width of the beach by municipality, because it captures the potential of attractive beach holidays (Deller et.al 2008). Then, I look at the annual average temperature and hours of sun. It is expected pleasant weather allows to capture all recreational opportunities for outdoor activities, and the tourism satisfaction. In this line, it is included the annual average precipitation because would have effects on the climatic comfort of tourists, and in sightseeing developing.

would use the following regression:

$$y_{it} = \alpha + \beta \Delta OvernightStays_{it} + \delta X_{it} + \gamma_p + \epsilon_{it} (4)$$

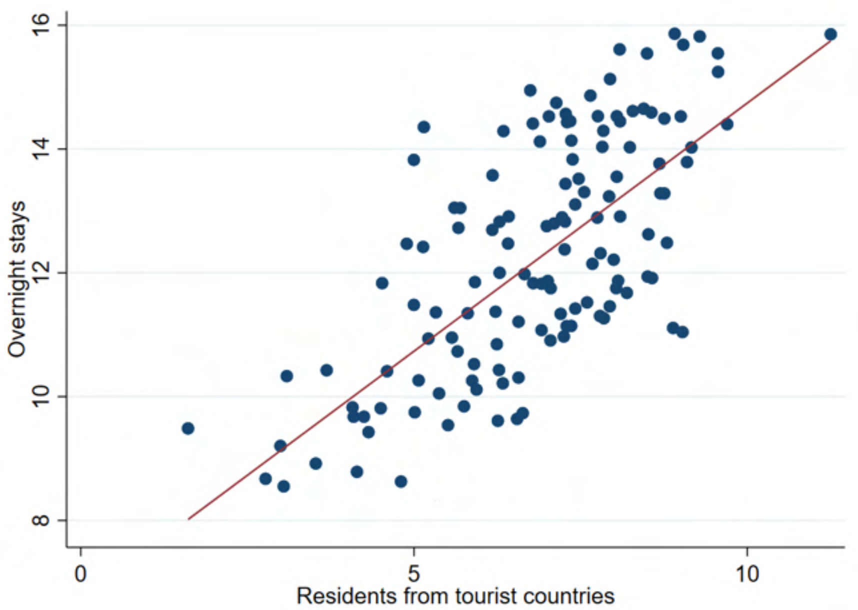
where y_{it} is the outcome of interest in per-capita terms in municipality i at time t. The variable $\Delta OvernightStays_{it}$ measures the change in the per capita number of overnight stays before and after the second wave of tourism development in municipality i. The vector, X_{it}, includes a number of municipal-level controls that I describe in the next section.

This strategy has some limitations, such as the fact that the number of overnight stays is only available for a limited number of municipalities with beaches, which in this case is only 96 out of the 437 municipalities. Additionally, it's important to note that tourists don't randomly choose their destinations but rather are attracted to areas with specific conditions, which means that a simple comparison between high and low overnight stays in beach municipalities may lead to a biased estimate of the impact of tourism.

To address the limitations of using the per capita number of overnight stays as a proxy for tourism activity, I use a shift-share analysis to study the impact of an exogenous increase in the number of tourists. This approach is based on the distribution of residents from tourist countries at the beginning of the second period of tourism development. The basic idea behind the shift-share analysis is that after visiting a municipality, tourists may eventually decide to become residents in that municipality and spend more time there. Furthermore, tourists are more likely to visit areas in Spain where people from their home country live. Thus, the shift-share analysis takes into account the early settlement patterns of residents from tourist countries and uses this information to predict the current distribution of tourists. This type of analysis is closely related to an instrument commonly used in the literature on immigration, which relies on the historical settlement patterns of immigrants by country of origin to predict the current geographic distribution of the immigrant population (Cortes (2008) and Sá (2015)). This is because immigrant networks play an important role in the locational choices of new immigrants, as they facilitate the job search process and assimilation into a new culture (Card, 2003; Jaeger et al., 2018 among others). The same logic applies to tourists, as they may rely on information and recommendations from fellow nationals when deciding where to travel or potentially purchase a property.

Given this, I first analyze the correlation between the number of residents by nationality at the municipality level and the number of overnight stays by nationality for the 96 municipalities for which I have overnight stays data available. Figure 2 shows the correlation between the number of residents from tourist countries in beach municipalities and the number of overnight stays by nationality. We can observe a high and significant correlation between the number of overnight stays by nationality and the presence of residents from tourist countries in a municipality.

Figure 2. Correlation between residents by nationality and overnight stays



Notes: (i) This Figure shows the correlation between the number of residents by nationality and the number of overnight stays at municipal level.

Based on this, I turn to use the share of residents of nationality j from a tourist country in each beach municipality in t_0 to predict where the inflow/number of overnight stays of tourists is going to be more significant in the t. Specifically, I constructed the shift-share in the following way:

$$\widehat{OvernightStays}_{i,t} = \sum_{j=1}^{9} TouristResidents_{i,j,t0} \times (OvernightStays_{j,t} - OvernightStays_{j,t0}) \quad (5)$$

Where $TouristResidents_{i,j,t0}$ is the per capita number of residents of nationality j in each beach municipality in the base year, which I fix to 1996. $OvernightStays_{j,t}$ is the number of overnight stays at national level by nationality in year t and $OvernightStays_{j,t0}$ is the number of overnight stays at national level by nationality in the base year. That is, I multiply the share of residents of each nationality in each beach municipality in t_0 by the change in the total number at national level of tourists from each nationality between t_0 and t. The formula estimates the expected increase in the number of overnight stays for each municipality based on the change in the number of overnight stays at the national level and the distribution of residents from each tourist country in the base year.

In table A.4 in the appendix, we can observe that eleven nationalities explain more than 70% of the tourist inflows in Spain in all the observed years. As argued in Goldsmith-Pinkham et al. (2018), in the shift-share setting identification mostly comes from the 'shares'. The "share" component provides predictive power to the instrument as it exploits the fact that new tourists of a given nationality tend to go to the municipalities where more residents from their country are settled. If the initial importance residents from tourist countries across municipalities is uncorrelated with current changes in outcomes of interest then this identification strategy identifies the causal effect of actual tourism inflows on the variables of interest. By focusing on the period immediately following the exogenous changes in regulation, we can more confidently attribute any changes in tourism inflows to the exogenous shock rather than to pre-existing developments in the municipalities. In order to ensure the validity and robustness of the shift-share analysis, I run several placebo tests in the following section. These tests examine the correlation between the shift-share and local income and population growth prior to the second wave of tourism development. In addition, I follow the method of Goldsmith-Pinkham et al. (2020) and compute Rotemberg weights for each nationality to further assess the validity of the identifying variation. These weights assign a higher weight to nationalities that are more likely to be affected by exogenous shocks. By doing so, I assess whether the shift-share accurately captures the causal effect of tourism inflows on the variables of interest.

5 Data

Hereafter, I first describe the main datasets used in the analysis, and then I shortly describe the results. The data covers 22 coastal provinces and 2,658 municipalities and spans the period 1960-2019. To construct my dataset, I digitized several historical records, including agricultural censuses, education censuses, and the Spanish market yearbook since the beginning of the tourism development. In this section, I describe the main data sources and present some descriptive statistics.

Tourism data

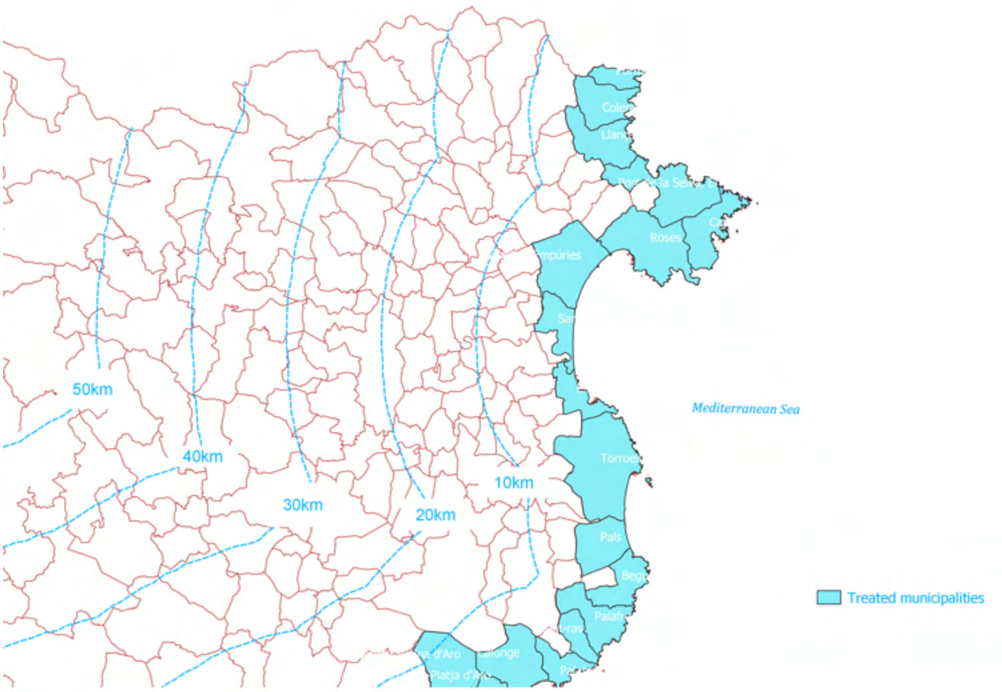
The measure of tourism used in the first empirical strategy is the 'per capita tourism' variable which captures the municipality's development of tourism relative to national levels. This variable is based on the income associated with the number of rooms available and annual occupation of various types of accommodations such as hotels, motels, hotel-apartments, inns, boarding houses, guest houses, camp sites, and apartments, as well as the income from restaurants, cafeterias, and bars. The per capita tourism variable is constructed using data from the tax on economic activities (IAE) levied on the corresponding economic sector. The data for this variable is drawn mainly from the Spanish Market Yearbook for the period 1965-1990 and from "Anuario Economico de España - La Caixa" (Economic Yearbook of Spain) for the period 1995 to 2019 for municipalities with more than 1,000 inhabitants. Table A.1 in the appendix, provides summary statistics for this variable. In addition, to instrument per capita toursim due to endogeneity, I construct *beach* and *beach measure* variables using the aerial photos and GIS data. I use aerial photos data from the National Geographic Institute (PNOA histórico 1956-57) which contains digital aerial orthophotographs of the entire Spanish territory in 1957. This allows to measure the surface, length, width, and area of all the beaches prior to tourism development. Figure 3 provides an extract of the aerial photos data. Additional GIS data layers, including the administrative shape file of municipality boundaries and the position of the Spanish coastline, are obtained from the geo-statistics division of the National Geographic Institute. The average length and width of the beaches in the sample are 3700 meters and 522 meters, respectively, as shown in Table A.1, panel A.

Figure 3. Aerial orthophotograph before the tourism development

Notes: This figure shows an aerial orthophotograph taken prior to the tourism development in 1956, providing a snapshot of the Spanish coast and beaches to the development of the tourism industry. Source: PNOA Americano Serie B for 1956.

The sample that I use in the first empirical analysis, beach amenities, includes all municipalities in coastal provinces with less than 50.000 residents, to exclude big metropolitan areas to address concerns that larger municipalities that are located close to the main economic centers may have both higher tourism attractiveness and more economic activity. In the main analysis the treatment are municipalities with beach and the control municipalities are municipalities without beach. I also exclude municipalities around 10 km of the treatment municipalities to avoid spillovers due to commuting effects, municipalities more than 50 km away from the treatment for comparability reasons and municipalities in the coastfront without beach. In Figure 4, we can observe an example of where are the treated (colored in blue) and the control (colored in white) municipalities located. I do several robustness checks changing the control group modifying the distance cutoff used for the exclusion of nearby municipalities.[7]

Figure 4. Visual map of treated and control municipalities



Notes: This figure shows a sample of treated municipalities with beaches and control municipalities without beaches in coastal provinces, as well as the different distance cutoffs used for excluding nearby municipalities in the robustness section. The treated municipalities are those with beaches, and the control municipalities are those without.

In the second empirical strategy, shift-share analysis, I focus is on the sub-sample of municipalities with beaches. To measure tourism in this sub-sample, I use the number of overnight stays in hotels at the municipal level. This data is obtained from the National Statistics Institute and covers the period from 2003 to 2019. However, it is only available for a limited number of municipalities, which is 96 out of the 461 municipalities with beaches. To address this data limitation, residents nationality data at the municipality level is also used to predict the number of overnight stays in hotels, which is obtained from the Continuous Register

[7]More details in section 6.

Statistics from 1996 to 2019. Further details can be found in section 4.2. The final sample used for this empirical analysis consists of 437 beach municipalities in coastal provinces with less than 50,000 residents.

Additional data

Outcomes variables The main outcome of interest is the gross income per capita in a municipality from the 2019 Household Income Distribution Map of the National Statistics Institute (INE). For the years before 2004, there's no data available regarding the local income at municipal level. Due to this, I follow Parellada (1992) and I estimate local income for the years 1965, 1990 and 1997 at municipal level using several main economic indicators available at the municipality level drawn from Spanish Market Yearbook. For each year and municipality I have data about the number of trucks, commercial vehicles, telephones, commercial licenses except food, and bank offices.[8] Regarding other main outcomes, several variables are coming from SEPE, such as the worker affiliation by economic activity and contract type (permanent or temporary contract) in 2019 and from the population census in 1990. From the population census I also get municipal level years of schooling for the years 1990, 2000 and 2011. Additionally, from the Continuous Register Statistics I get the number of residents by nationality at municipal level from 1996 to 2019 which allows to measure the number residents from tourist countries in a municipality. These variables are summarized in Panel B of Table A.1.

Control variables To account for the possible municipal differences in socioeconomic development prior to the tourism development, I control for the level of agricultural production using data from 1962 agricultural census, from INE for the year 1990, the population level coming from the population censuses and education data for the year 1930 coming from Beltrán Tapia et al. (2011). I also include several controls for geographical variables such as soil quality, altitude, ruggedness, temperature, rainfall and hours of sun coming from Oto-Peralías (2019). The control variables are summarized in Panel C of Table A.1. The average population in 1950 is 5138 inhabitants, with over 47 percent of the adult population unable to read or write. In average, almost 32 percent of the municipality's land was used

[8] Explain more. For more details about the estimation of local income see Parellada (1992) and Alañón (2001).

for agricultural production in 1962 and the agricultural suitability measured with the soil quality index is high.[9] I also present summary statistics for the average annual temperature, precipitations, number of sun hours as well as the average altitude and ruggedness of the municipalities.

6 Preliminary results

6.1 Tourism and municipality income

6.1.1 Beach amenities

In this subsection, I present the results of the impact of tourism on income exploiting the beginning of tourism development in the 1960s and the presence of beach amenity in a municipality. The main results are shown in Table 1. First, we can focus on the association between tourism and beach amenity. Panel B, Columns 5 and 6, show the coefficients from the first-stage regressions, which correspond to the instrumental variable (IV) estimation of Equation (2). To accurately estimate the effect, I take into account various pre-determined socioeconomic and geographic variables, including the population in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, tourism per capita in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. Additionally, I control for province fixed effects and cluster at the province level when indicated.

The results show that having a beach is positively and significantly correlated with higher increase in per-capita tourism. In other words, if a municipality has a beach, the per-capita growth of tourism relative to other municipalities in coastal provinces is approximately 1.6 times higher compared to a municipality without a beach. The standard F-statistic test rule out that the instrument is weak in the context of my specifications.

Having established the correlation between the main measure of tourism attractiveness (existence of a beach) and tourism in Spanish municipalities, I proceed to analyze the impact of per-capita tourism on local income levels. To do this, I

[9] Soil quality is a measure of seven key soil dimensions important for crop production: nutrient availability, nutrient retention capacity, rooting conditions, oxygen availability to roots, excess salts, toxicities, and workability.

estimate specification (3) with per-capita income on the left-hand side and present the results in Panel A of Table 1.

First, in columns 1 and 2, I present the ordinary least squares (OLS) estimates of the effect of per capita tourism growth on income levels in 2019 for the main sample. Subsequently, in columns 3 and 4, I report the reduced-form estimates of the impact of having a beach in 1965 on municipality income levels in 2019. Finally, in columns 5 and 6, I provide instrumental variable (IV) estimates, where per capita tourism growth is instrumented by the presence of a beach in the municipality." . In all specifications I control for population in 1950, income level in 1965, per-capita tourism in 1965, illiteracy rate in 1930, agriculture production in 1960, soil quality, altitude, ruggedness,temperature, rainfall and hours of sun. In column 2, 4 and 6 I add province fixed effects and I cluster the robust standard erros at province level.

The reduced-form results indicate that the presence of a beach is associated with lower per-capita income levels in 2019. The instrumental variable (IV) estimates in columns 5 and 6 are consistent with the OLS estimates, but significantly higher in magnitude. The results in column 6 imply that municipalities with a higher per-capita tourism growth have lower income levels in 2019. Specif-ically, municipalities with higher per-capita tourism have 8,5% less income level in 2019. The average (median) growth of per-capita tourism from 1965 to 2019 for beach municipalities is 2.62, implying these municipalities have 22% less income level in 2019. Hence, it can be conclude that tourism specialization exerts a significant, and negative, effect on municipalities' per-capita income.

Additionally, I also construct the variable *Beach measure* which captures the quality and attractiveness of a beach taking into account the length and width of a beach using the 1957 aerial photos, as well as the temperature, hours of sun and precipitation. I add the variable *Beach measure* interacting with the dummy variable *Beach* in Equation (2). Table A.8 presents the results of incorporating the beach quality/attractiveness measure into the analysis. The table displays the results of five columns, each adding an additional beach feature to the analysis. The first column includes the standardized length of beach in kilometers, the second column incorporates the number of hours of sun, the third column adds the width of the beach in kilometers, the fourth column includes temperature, and the fifth column adds precipitation as a feature. The IV results indicate that not only having a beach, but also having advantageous geographical and climatological characteristics, have a negative impact on municipalities' income level. The first-stage F statistic is slightly lower than when only considering whether a municipality has a beach, suggesting that the presence of a beach is the most important determinant of tourism attractiveness. However, as shown in column 2, kilometers of beach length and hours sun also have a significant impact on tourism attractiveness and thus on income levels.

Table 1. Impact of tourism on local income

	Panel A. Dep. variable: Income level 2019					
	OLS	OLS	RF	RF	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Per-capita tourism growth (1965-2019)	-0.004*	-0.004***			-0.041***	-0.085***
	(0.010)	(0.010)			(0.008)	(0.018)
$\mathbb{1}$(Beach)			-0.071***	-0.137***		
			(0.015)	(0.025)		
Mean dep. var.	9.49	9.49	9.49	9.49	9.49	9.49
	Panel B. First stage: Dep. variable: Per-capita tourism growth					
$\mathbb{1}$(Beach)					2.093***	1.598***
Kleibergen-Paap rk LM F-stat.					38.64	35.890
					[16.38]	[16.38]
Observations	1,196	1,196	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y	Y	Y
Proovince FE	N	Y	N	Y	N	Y

Notes: (1) Panel A reports the estimated effect of per-capita tourism growth on income level using three different regression methods: OLS, reduced-form, and IV. The dependent variable in all cases is the logarithm of income in 2019. Panel B reports the first-stage regression results, where the independent variable is $\mathbb{1}$(Beach), a dummy variable equal to one if the municipality has a beach and zero otherwise. The unit of observation is the municipality. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, per capita tourism in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

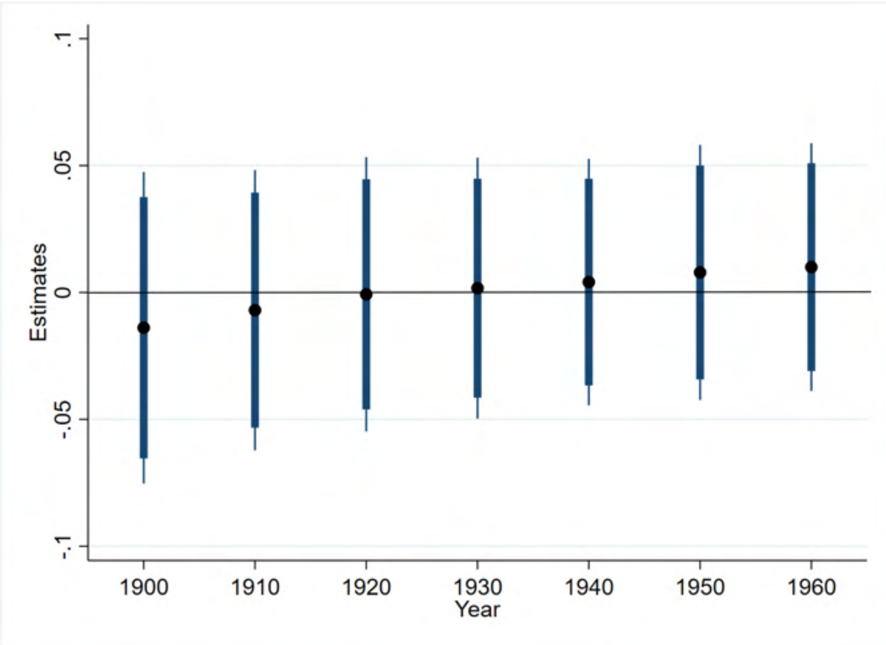
6.1.2 Robustness

Table 1 documents strong negative effect of tourism on local income level. One potential concern is that municipalities with beach and attractive geographical and climatological features (such as extended beach lengths, high temperatures, more sun hours, etc) could affect the local economy not only through their effect on local tourism development, but also by directly influencing the residential choice of Spanish residents relative to other inland municipalities. Even though I sought to be careful in constructing these measures to capture a very particular set of features of the local environment that are arguably specific to tourism attractiveness, it could be the case that they have a significant direct amenity effect on local employment and populations relative to other interior municipalities.

In Figure 4, I further investigate this concern. I run a placebo falsification test on the identical sample of municipalities during a period before beach tourism had become a discernible economic force in Spain. This involves the construction of a long time series of population census data from the years from 1900 to 1960. Due to limited data availability at municipal level before 1960, the historical census only provides municipality population. Results in Figure 4, reports the IV estimates, we can observe insignificant point estimates of the effect on municipality populations before 1960. These results suggest that the presence of a beach that I use to instrument per-capita tourism growth unlikely to capture locational fundamentals that directly enter location choices along the coastline in a discernible way.

Furthermore, the sample used for the analysis in Table 1, excludes from the control (interior) municipalities the ones that are that are located within less than 10 km or more than 50 km away from beach municipalities to address the commuting patterns and ensure similar initial socio-economic and geographical characteristics. To assess the sensitivity of this distance cutoff, in Table A.9 in the appendix, I alternatively report results using different control groups by varying the distance criteria. The IV coefficients are consistently negative and statistically significant across all specifications considered.

Figure 4. Evolution of population. Placebo test



Notes: (1) The graph shows the IV estimation of the impact of per capita tourism growth instrumented by the presence of a beach on population. The dependent variable is the population expressed in logs. (2) In the regression analysis I control for various pre-determined socioeconomic and geographic variables, including population in 1950, illiteracy rate in 1930, income in 1965, agriculture production in 1960, tourism per capita in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3)The standard errors used in the estimation are robust and clustered at the province level.

6.1.3 Shift-share analysis

Next, I turn to present the results when focusing on the second period of tourism development (1997-2019). In this second strategy, ideally, I would like to start showing the first-stage regressions of the relationship between the realized number of overnight stays and the predicted number of overnight stays. But due to data limitations, I only have data regarding the realized overnight stays for a small number of municipalities with beach (96 out of 437 beach municipalities).

Before I turn to show the results of the reduced form for the 437 beach municipalities. I show the first-stage regressions between the predicted number of overnight stays and the realized number of overnight stays for the subsample of the 96 municipalities. The coefficients from these first-stage regressions are in Table A.5 in the appendix. As shown in Table A.5, the predicted growth in overnight stays, as specified in Equation (5), explains the observed change in overnight stays almost perfectly. The coefficients are close to 1 in all specifications and the standard F-statistic test rule out that the instrument is weak in the context of my specifications.

Having established the relationship between the predicted number of overnight stays and the realized number of overnight stays for the subsample of Spanish municipalities with beach, I then analyze themedium-term impact of per capita tourism on local on municipalities' income growth. I estimate specification (4) with log income level in 2019 on the left-hand side. I report directly the Reduced Form results in Table 2. The Reduced Form estimation suggests that the increase in overnight stays after the second period of tourism development is negatively associated with the income level. The result in column (2) indicate that municipalities with higher overnight stays growth have 7% less income level in 2019. The average (median) growth of overnight stays growth from 1997 to 2019 for beach municipalities is 3.03, implying these municipalities have 21,5% less income level in 2019.

Hence, it can also be conclude using the shift-share analysis that tourism specialization exerts a significant, and negative, effect on municipalities' per capita income. Table A.6 in appendix show the impact of tourism measured by the realized number of overnight stays using the subsample of 87 municipalities. The coefficients go in line with the results in Table 2.

Table 2. Impact of tourism on local income 2nd empirical strategy. Reduced Form

	Income level				Population	
	Reduced form		Placebos			
	2019	2019	1990	1990	1990	1990
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Predicted overnight stays	-0.057***	-0.071***	-0.002	-0.001	-0.001	-0.000
	(0.016)	(-0.071)	(0.0003)	(0.0004)	(0.0001)	(0.0002)
Mean dep. var.	9.49	9.49	9.49	9.49	9.49	9.49
Observations	437	437	437	437	437	437
All Controls	Y	Y	Y	Y	Y	Y
Province FE	N	Y	N	Y	N	Y

Notes: (1) The table shows the reduced form estimates of the impact of tourism growth measured by the predicted number of overnight stays and instrumented by the shift-share explained in section 4 on income and population levels. The unit of observation is the municipality. Columns (1) and (2) show the impact of an increase in the number of overnight stays during the period 1997-2019 on income level in 2019. Columns (3) and (4) show the impact of an increase in the number of overnight stays during the period 1997-2019 on income level in 1990. Columns (5) and (6) show the impact of an increase in the number of overnight stays during the period 1997-2019 on population level in 1990. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1990, illiteracy rate in 1990, income in 1990, agricultural production in 1990, per capita tourism in 1990, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

6.1.4 Robustness

Table 2 documents strong negative effect of the increase of the number of overnight stays in hotels on local income for municipalities with beach. One potential concern is the instrument validity. The identification of the shift-share instruments comes from the initial shares. More precisely, I assume that initial shares of residents by nationality measure the differential exogenous exposure to the second period of tourism development. Since the predetermined shares are equilibrium outcomes that are affected by tourism attractiveness, they probably correlate with the income levels in that period. However, the validity of the shift-share analysis hinges on the assumption that the initial shares are exogenous to changes in income, not to the initial level of income. To test this assumption in this framework, I follow Goldsmith- Pinkham et al., (2020) and first, I compute the Rotemberg weights for the different nationalities.[10] These weights indicate which nationalities entering the shift-share are driving the results. In this case, the five most important nationalities are Germans, Britons, French, Belgians and Portuguese. As suggested by Goldsmith-Pinkham et al. (2020), I check the correlation of the five most important initial nationality shares of residents and possible confounders. As confounders, I use the population growth (1997–2019), education level in 1990, unemployment level in 1990, and agricultural and manufacturing production in 1990. I report the results in Table A.7 in the appendix. Reassuringly, this analysis shows that the initial shares of nationalities are not related to possible confounders.

An additional step in evaluating the validity of the shift-share analysis involves examining its correlation with local income growth prior to the second wave of tourism development. Table 2 provides the correlation between the local income growth for municipalities with beaches from 1990 to 1996 and the shift-share from 1997 to 2019. The findings suggest that there is no correlation between the shift-share and the pre-treatment growth of the municipalities' local income. Furthermore, the correlations between the instrument and population growth, as shown in columns (3) and (4), are not statistically significant, which further supports the conclusion that the instrument is uncorrelated with pre-treatment population growth.

6.2 Mechanisms

To better comprehend the mechanisms behind the negative effects of tourism on income levels, I compiled detailed data at the municipal level, including worker affiliation, contract type, and education level. The findings show that an increase in per-capita tourism growth is linked to a rise in temporary contracts and lower education attainment, along with a shift in workers from industry and agriculture

[10] These weights are based on Rotemberg (1983) and Andrews et al. (2017) .

to tourism-related sectors.

6.3 Cross-industry interactions

I analyze the cross-industry interactions at the municipal level as the first channel. Table 3 provides insight into the impact of tourism on employment in various sectors, including manufacturing, construction, real estate, hotels, food-service, and agriculture. Unfortunately, employment data by sectors at municipal level is only available from 1990 onwards. In Table 3, I present the findings using the first empirical strategy, beach amenities, and looking at the impact of per-capita tourism growth on the level of outcomes in 2019. The results from both OLS and IV estimates suggest that an increase in per-capita tourism growth is associated with an increase in construction, real estate, hotels, and food-services employment, and with a decrease in manufacturing and agricultural employment. Thus, we can observe that specialization of tourism reallocates factors of production towards services activities, closely linked to the tourism sector, and away from traded sector production such as industry and agriculture.

More specifically, we can say in that municipalities with higher per-cápita tourism growth in 2019 had 17.9% and 62% less employment in agriculture and manufacturing, respectively, 78% more employment in construction and real estate and 46% more employment in services related to hotels and food-services.

Indeed, the rise in tourism activity has produced a redistribution of resources from industry to services (Copeland, 1991). According to the National Spanish Institute, in 2019 manufacturing workers earn 34 percent more gross annual income than workers in the tourism workers. This is consistent with several studies focusing on developed countries that have found that tourism workers are the lowest paid of any industry (Riley, Ladkin, & Szivas, 2002; Muñoz-Bullón, 2009; Santos & Varejao, 2007; Dogru et al., 2019).

Table 3. Beach amenity analysis: Impact of tourism on the economic structure of the municipality

	Employment in an industry							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Agriculture		Manufacturing		Construction and real estate		Hotels and foodservice	
Per-capita tourism growth	-0.044***	-0.179*	-0.023**	-0.621***	0.109**	0.787**	0.039***	0.464***
	(0.005)	(0.221)	(0.009)	(0.192)	(0.039)	(0.745)	(0.005)	(0.155)
Mean dep. var	14.384	14.384	18.336	18.336	9.077	9.077	15.968	15.968
Kleibergen-Paap rk LM F-stat.	19.343	19.343	19.343	19.343	19.343	19.343	19.343	19.343
	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]
Observations	1,196	1,196	1,196	1,196	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The table shows the estimates of the effect of per-capita tourism growth on employment shares in different sectors using both OLS and IV methods. Per-capita tourism growth is instrumented using $\mathbb{1}$(Beach), a dummy variable equal to one if the municipality has a beach and zero otherwise. which takes the value of 1 if the municipality has a beach and 0 otherwise. The unit of observation is the municipality. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, per capita tourism in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

The above results are also corroborated by the reduced form results when using the second empirical strategy, the shift-share analysis. As we have data on employment by sector at the municipality level from 1990, we can use employment growth by sector between 1990 and 2019 as the outcome variable. Table 4 shows that a 100% increase in the number of overnight stays during this period for the 437 municipalities with beach led to a 27% and 25% decrease in agriculture and manufacturing employment, respectively, as well as a 26% increase in construction and real estate services and an 11% increase in hotels and food-services employment.

Table 4. Shift-share analysis: Impact of tourism on the economic structure of the municipality

	Employment change by industry			
	ΔAgriculture	ΔManufacturing	ΔConstruction	ΔHotels & foodservice
	(1)	(2)	(3)	(4)
Δ Predicted overnight stays 2019-1990	-0.272**	-0.253*	0.262**	0.112**
	(0.106)	(0.09)	(0.102)	(0.052)
Observations	437	437	437	437
All Controls	Y	Y	Y	Y
Province FE	Y	Y	Y	Y

Notes: The table shows the reduced form estimates of the effect of per-capita tourism growth on employment shares growth in different sectors. Per-capita tourism growth is measured by using predicted overnight stays and instrumented by the shift-share explained in section 4. The unit of observation is the municipality. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1990, illiteracy rate in 1990, income in 1990, agricultural production in 1990, per capita tourism in 1990, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

6.4 Job stability

The second channel I exploit to understand the potential reason for the observed negative impact on the overall income level is the contract type. It is possible that when a municipality specializes in tourism, there is a shift towards more temporary contracts and fewer stable jobs due to the seasonal nature of tourism employment.

To test this hypothesis, I show the results of impact of tourism on the share of temporary contracts in 2019 in Table 5, using both empirical strategies. Columns (1) and (2) present results for the first empirical strategy, while column (3) corresponds to the second empirical strategy. In this case, we can observe municipalities with higher tourism growth tend to have a higher share of temporary contracts. This is consistent with González & Surovtseva (2020), who find that the shift towards tourism related employment implies a shift toward fewer permanent and more temporary contracts, resulting in lower employment stability.

Table 5. Impact of tourism on temporary of contracts

	Share of temporary contracts 2019		
	Beach amenity analysis		Shift-share analysis
	OLS	IV	RF
	(1)	(2)	(3)
Per-capita tourism growth	0.034***	0.177**	
	(0.001)	(0.075)	
Δ predicted overnight stays 2019			0.058*
			(0.015)
Mean Dep. Var	0.943	0.943	0.943
Observations	1,196	1,196	437
All Controls	Y	Y	Y
Province FE	Y	Y	Y

Notes: The table shows estimates of tourism on the share of temporary contracts. Columns (1) and (2) use per capita tourism as the independent variable, which is instrumented by the existence of a beach in a municipality. Column 3 uses the independent variable of overnight stays growth. The unit of observation is the municipality. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, per capita tourism in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun.(3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

6.5 Education level

Lower education levels among the workforce could be an additional mechanism contributing to lower income levels, resulting from the opportunity cost of obtaining higher education and decreasing perceived returns of education. In Table 6, we observe that municipalities specializing more in tourism tend to have lower education levels in the years 2011, 2001, and 1991, as shown in columns (1)-(6) using the first empirical strategy. Specifically, a higher increase in per-capita tourism growth is associated with lower tertiary education levels among the 26-35 age population in 2011 and 2001, while a similar effect is observed for secondary education in 1991.[11]

Using the second empirical strategy, column (7) shows the impact of tourism growth on tertiary education growth. We can see that, in municipalities with beach, a 100% increase in the number of overnight stays during 1997-2019 imply that the growth rate of tertiary education is 1% lower. This is consistent with Oliver & Villalonga (2019) who find evidence that Spanish regions specializing in sectors with a high level of unskilled labor crowd students out of the education system, leading to higher school drop-out rates and lower numbers of students that complete the non-compulsory stage of secondary education. In contrast, in regions specializing in sectors with a demand for medium or highly skilled workers (with at least compulsory secondary school studies or higher education studies respectively), young students are driven to remain in the education system because they will have no access to employment unless they achieve a minimum level of education.

[11] Note that in 2011 and 2001, I focus on tertiary education, as secondary education became mandatory in Spain after 1990.

Table 3. Impact of tourism on education level among individuals aged 25-36

	Beach amenity analysis						Shift-share analysis
	OLS	IV	OLS	IV	OLS	IV	RF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2011		2001		1991		1991-2011
Per-capita tourism growth	-0.0008	-0.018***	-0.0004**	-0.013***	-0.0008	- 0.013***	
	(0.000)	(0.006)	(0.000)	(0.004)	(0.000)	(0.015)	
Δ Predicted overnight stays							-0.001**
							(0.004)
Mean Dep. Var	0.150	0.150	0.196	0.196	0.455	0.455	0.05
Kleinbergen-Paap rk LM F-stat.		17.82		14.41		11.92	
Observations	1,196	1,196	1,196	1,196	1,196	1,196	437
All Controls	Y	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y	Y

Notes: The table presents estimates of the impact of tourism on education, using different specifications. Columns (1) and (2) use per capita tourism growth, instrumented by the presence of a beach in a municipality, as the independent variable, while column (3) uses overnight stays growth. The dependent variable in columns (1) and (2) is the tertiary education level in 2011, while in columns (3) and (4) it is the tertiary education level in 2001. Column (5) uses secondary education level in 1990 as the dependent variable. Finally, column (7) examines the growth of tertiary education between 1991 and 2011. The unit of observation is the municipality. (2) All columns control for predetermined socioeconomic and geographic variables, including population in 1950, illiteracy rate in 1930, income in 1965, agricultural production in 1960, per capita tourism in 1965, soil quality, altitude, ruggedness, temperature, rainfall, and hours of sun. (3) The Kleibergen-Paap rk LM F-statistic is used to test for weak instrument validity, with the critical value for the Stock-Yogo weak identification test reported in brackets at the 10% maximal IV size. Robust standard errors are reported in parentheses and clustered at the province level when province fixed effects are included. Significance levels are denoted by *, **, and *** at the 10%, 5%, and 1% levels, respectively.

7 Conclusion

Sun and beach tourism plays a crucial role in the economy of many countries. However, this economic sector is expected to suffer significant damage due to climate change, such as the projected loss of beaches, thermal comfort, water restrictions, and extreme weather events (Aguilles, 2022). Nevertheless, governments continue to invest in and promote tourism as a key sector in the national economy. The economic implications of tourism specialization are mixed and the economic evidence of the long-term impact continues to be scarce. With this paper, I aim to fill this current gap in the literature, by analyzing the local long and medium-term effects of tourism specialization on income.

In this study, I focus on Spain, one of the most popular tourist destinations in the world, where the tourism sector accounts for a substantial share of total employment, especially in coastal regions. To address the endogeneity underlying tourism specialization and economic development, I propose two empirical methodologies that rely on cross-sectional variation to capture the effects of tourism exposure on relative municipal economic outcomes across Spanish municipalities.

The first methodology draws on existing literature to argue that the existence and quality of local natural factors significantly influence tourism activities. I leverage geological variation along the Spanish coastline to construct instrumental variables for tourism specialization, such as the existence of a beach, the fraction of onshore coastline covered by the beach, and other weather features. The results show that accessible natural resources, mainly the existence of a beach, prove to be a strong instrument for identifying the causal impact of tourism on income levels. The second methodology uses shift-share analysis to take advantage of the distribution of residents from tourist countries in beach municipalities, exploiting the second wave of tourism development.

Using these identification strategies, I show that higher tourism activity leads to lower municipality income per capita in the long-run relative to interior municipalities (first strategy) and lower municipality income per capita in the medium-run relative to less touristic regions. Additionally, I investigate the potential channels through which tourism negatively affects income per capita, showing that higher tourism per capita correlates with higher levels of temporary contracts, higher number of workers in sectors related to tourism, and lower number of workers in industry and agriculture, as well as lower levels of education.

This study provides credible empirical evidence of the long-term effects of tourism activity on economic outcomes. Given that most current tourism policies are aimed at investing in the local attractiveness for tourism, these findings reveal the long-term implications of tourism specialization compared to other economic sectors.

8 References

Aguilles, M., Melo-Aguilar, C., & Jordà, G., (2022): "Risk of loss of tourism attractiveness in the Western Mediterranean under climate change." *Frontiers in Climate* 191.

Alañón Pardo, Ángel (2001): "La renta regional en España: análisis y estimación de sus determinantes". Tesis doctoral inédita, Facultad de Ciencias Económicas y Empresariales, Universidad Complutense de Madrid, Madrid.

Albaladejo, I. P., González-Martínez, M. I., & Martínez-García, M. P. (2020): "A double life cycle in tourism arrivals to Spain: Unit root tests with gradual change analysis." *Journal of Destination Marketing & Management* 18, 100497.

Allen, T., Fuchs, S., Ganapati, S., Graziano, A., Madera, R., & Montoriol-Garriga, J. (2020): "Is tourism good for locals? Evidence from Barcelona." Processed Dartmouth College.

Allcott, H., & Keniston, D. (2018): "Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America." *The Review of Economic Studies* 85(2), 695-731.

Almagro, M., & Domınguez-Iino, T. (2019): "Location sorting and endogenous amenities: Evidence from Amsterdam." Processed NYU.

Andrews, I. , Gentzkow, M. , Shapiro, J.M. (2017): " Measuring the sensitivity of parameter estimates to estimation moments." *The Quarterly Journal of Economics* 132 (4), 1553–1592.

Arezki, R., Cherif, R. & Piotrowski, J., (2009): "Tourism specialization and economic development: Evidence from the UNESCO world heritage list." IMF Working Papers: 1-24.

Becker, G. S. (1964): “Human capital theory.” Columbia, New York, 1964.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005): “Why the apple doesn’t fall far: Understanding intergenerational transmission of human capital.” *American Economic Review* 95(1), 437-449.

Brida, J.G., Cortes-Jimenez, I. & Pulina, M., (2016): “Has the tourism-led growth hypothesis been validated? A literature review”. *Current Issues in Tourism* 19(5), pp.394-430.

Capo, J., Font, A. R., & Nadal, J. R. (2007): “Dutch disease in tourism economies: Evidence from the Balearics and the Canary Islands.” *Journal of sustainable Tourism* 15(6), 615-627.

Card, D. (2009): “Immigration and inequality.” *American Economic Review* 99(2), 1-21.

Caselli, F., & Michaels, G. (2013): “Do oil windfalls improve living standards? Evidence from Brazil.” *American Economic Journal: Applied Economics* 5(1), 208-38.

Ciani, Emanuele & Bronzini, Raffaello & Montarul, Francesco (2019). ”Tourism and local growth in Italy.” *Questioni di Economia e Finanza (Occasional Papers)* 509, Bank of Italy, Economic Research and International Relations Area.

Cimoli, M. (2005): “Heterogeneidad estructural, asimetrías tecnológicas y crecimiento en América Latina”. Cepalbid, Santiago de Chile.

Copeland B.R. (1991): ”Tourism, Welfare and De-industrialization in a Small Open Economy” *Economica* 58(232), 515-29.

Corden, M. (1984): “Booming sector and Dutch disease economies: Survey and consolidation.” *Oxford Economic Papers* 36, 359–380.

Corden,M. & Neary, P. (1982): “Booming sector and de-industrialization in a small open economy.” *The Economic Journal* 92, 825–848.

Croes, R., & Semrad, K. (2018): "Destination competitiveness." *The SAGE Handbook of Tourism Management: Applications of theories and concepts to tourism* 77-90

Cortes, P. (2008): "The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data". *Journal of Political Economy* 116(3), 381-422.

Deller, S. C., Lledo, V., & Marcouiller, D. W. (2008): "Modeling regional economic growth with a focus on amenities." *In Review of Urban & Regional Development Studies: Journal of the Applied Regional Science Conferenc* 20, No. 1, pp. 1-21. Melbourne, Australia: Blackwell Publishing Asia.

Dogru, T., McGinley, S., Line, N., & Szende, P. (2019): "Employee earnings growth in the leisure and hospitality industry." *Tourism Management* 74, 1-11.

Faber, B., and Gaubert, C. (2019): "Tourism and economic development: Evidence from Mexico's coastline." *American Economic Review* 109(6), 2245-93.

Frechtling, D. C. (2001): "World conference on the economic impact of tourism." *The International Journal of Tourism Research* 3(3), 253.

Garcia-López, M. À., Jofre-Monseny, J., Martínez-Mazza, R., & Segú, M. (2020): "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona." *Journal of Urban Economics* 119, 103278.

Goldsmith-Pinkham, P., I. Sorkin, & H. Swift (2018): "Bartik Instruments: What, When, Why, and How," NBER Working Paper.

Gollin, D., Jedwab, R., & Vollrath, D. (2016): "Urbanization with and without industrialization." *Journal of Economic Growth* 21(1), 35-70.

González, L., & Surovtseva, T. (2020): "Do more tourists promote local employment?." Processed Graduate School of Economics.

Güell, M. & Petrongolo, B. (2007): "How binding are legal limits? Transitions from temporary to permanent work in Spain". *Labor Economics* 14(2), pp. 153–83.

Gylfason, T. (2001a): "Natural resources, education, and economic development." *European Economic Review* 45(4-6), 847-859.

Gylfason, T. (2001b): "Nature, power and growth." *Scottish Journal of Political Economy* 48(5), 558-588.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2014): "Growth and structural transformation." *Handbook of Economic Growth* 2, 855-941.

Hainmueller, J. (2012): "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies." *Political Analysis* 20(1), 25–46.

Holzner M. (2011): "Tourism and economic development: The beach disease?". *Tourism Management* 32 (4),922-933.

Hjort, J., & Poulsen, J. (2019): "The arrival of fast internet and employment in Africa." *American Economic Review* 109(3), 1032-79.

Inchausti-Sintes, F., Voltes-Dorta, A.,& Suau-Sánchez, P. (2021):" The income elasticity gap and its implications for economic growth and tourism development: the Balearic vs the Canary Islands." *Current Issues in Tourism* 24(1), 98-116.

Katz, J. (2000): "Reformas estructurales, productividad y conducta tecnológica en América Latina." CEPAL.

Jaeger, D. A., Ruist, J., & Stuhler, J. (2018): "Shift-share instruments and the impact of immigration." *National Bureau of Economic Research.*

Mason, K., Morrison, W. G., & Stockman, I. (2016): "Liberalization of air transport in Europe and the evolution of 'low-cost' airlines." *Liberalization in Aviation* 163-177.

Makarski, K., Tyrowicz, J., & Lutynski, J. (2022): "Structural change and inequality in general equilibrium" 67. GRAPE Group for Research in Applied Economics.

McMillan, M., Rodrik, D., Bacchetta, M., & Jensen, M. (2011): ”Making globalization socially sustainable.” *ed. Marc Bacchetta and Marion Jansen, Chapter Globalization, Structural Change, and Productivity Growth. Geneva: International labor Organization and World Trade Organization.*

Muñoz-Bullón, F. (2009): “The gap between male and female pay in the Spanish tourism industry.” *Tourism Management* 30(5), 638-649.

Nocito, S., Sartarelli, M., & Sobbrio, F. (2021): “A Beam of Light: Media, Tourism & Economic Development.” *Tourism & Economic Development.*

Oliver, X., & Villalonga, J. R. (2019):” Does the economic productive structure determine early school leaving rates? Evidence from Spanish regions, 2000-2013.” *Estudios sobre Educación* 36, 181-205.

Oto-Peralías, D. (2019): “Delegation of Governmental Authority in Historical Perspective: Lordships, State Capacity and Development.” *State Capacity and Development.*

Parellada, M. (1992): ”Una aproximación a la estimación de la renta municipal en España”. *Economía española, cultura y sociedad. Homenaje a Juan Velarde Fuentes, tomo II, Eudema, Madrid.*

Parrilla, J.C., Font, A.R. and Nadal, J.R. (2007): “Tourism and long-term growth a Spanish perspective”. *Annals of Tourism Research* 34(3), pp.709-726

Riley, M., Ladkin, A., & Szivas, E. (2002): “Tourism employment: Analysis and planning”. *Channel View Publications* 6.

Rotemberg, J.J. (1983): “Instrumental variable estimation of misspecified models.” Working Paper 1508-83 .

Sá, F. (2015): “Immigration and House Prices in the UK”. *The Economic Journal* 125, 1393-1492.

Sánchez-Sánchez, E. M. (2001): “ El auge del turismo europeo en la España de los años sesenta.” *Arbor CLXX* 669, 201-224

Santos, L. D., & Varejão, J. (2007): "Employment, pay and discrimination in the tourism industry." *Tourism Economics* 13(2), 225-240.

Song H., Dwyer L. & Zhengcao G.L. (2012): "Tourism economics research: a Review and assessment" *Annals of Tourism Research* 39 (3), 1653–1682.

Sequeira, T.N. & Maçãs Nunes, P. (2008). "Does tourism influence economic growth? A dynamic panel data approach". *Applied Economics* 40(18), 2431.

Torvik, R. (2001): "Learning by doing and the Dutch disease." *European Economic Review* 45(2), 285-306.

World Travel and Tourism Council (WTTC) (2021): "Economic Impacts Report", WTTC, Brussels.

9 Appendix

Table A.1. Desciptive statistics

Variables	count	mean	sd	min	max
Panel A					
Tourism per capita 1965	1196	13.82	88.56	0.00	2054.70
Tourist per capita 2019	1196	36.942	202.226	0.000	4454.317
Beach length	367	3691.884	4603.293	0.000	47320.964
Beach width	367	522.190	722.792	0.000	5873.755
Beach area	367	353507.986	668916.524	13.000	6641338
Beach_index	367	3.556	1.049	1.000	5.000
Log Predicted overnight stays in 1997	437	5.047	2.484	1.028	13.816
Share German residents 1996	437	0.007	0.017	0.000	0.178
Share Briton residents 1996	437	0.015	0.043	0.000	0.265
Share Dutch residents 1996	437	0.003	0.008	0.000	0.117
Share French residents 1996	437	0.002	0.005	0.000	0.053
Share Portuguese residents 1996	437	0.001	0.002	0.000	0.020
Share American residents 1996	437	0.001	0.003	0.000	0.065
Share Japanese residents 1996	437	0.000	0.000	0.000	0.001
Share Belgian residents 1996	437	0.001	0.005	0.000	0.068
Share Italian residents 1996	437	0.001	0.002	0.000	0.014
Panel B					
Log Predicted Income per capita 1965	1196	2.977	1.276	1.105	5.029
Log Predicted Income per capita 1990	437	2.710	0.954	1.795	4.818
Log Predicted Income per capita 1997	437	3.750	0.307	1.380	12.27
Log Income per capita 2019	1196	9.474	0.237	8.927	10.409
% Employed in manufacturing 2019	1196	15.551	15.432	0.000	84.215
% Employed in agriculture 2019	1196	18.340	19.707	0.000	80.278
% Employed in construction 2019	1196	9.139	5.694	0.000	49.444
% Employed in real estate 2019	1196	0.507	0.755	0.000	12.500
% Employed in hotels and foodservice 2019	1196	9.969	8.226	0.512	70.760
% Employed in ancillary services 2019	1196	2.688	3.069	0.000	55.479
% Employed in manufacturing 1990	437	1.3	0.9	0.000	9.5
% Employed in agriculture 1990	437	0.127	0.118	0.000	0.573
% Employed in construction 1990	437	12.78	11.7	0.23	85.9
% Employed in hotels and foodservice 1990	437	14.3	0.83	0.39	53.8

(Continued)

Table A.1. Desciptive statistics

Variables	count	mean	sd	min	max
% Secondary education age 25-36, 1991	1196	0.570	0.148	0.012	0.96
% Tertiary education age 25-36, 1991	437	0.071	0.036	0.013	0.261
% Tertiary education age 25-36, 2001	1196	0.171	0.078	0.07	0.486
% Tertiary education age 25-36, 2011	1196	0.22	0.12	0.002	0.856
% of temporary contracts	1196	0.942	0.056	0.381	1.000
% of immigrant workers	1196	0.053	0.052	0.000	0.431
Panel C					
Population 1950	1196	5138	5172	165	58768
Population 1990	437	9492.320	8570.706	57.000	50466
Temperature	1196	14.953	2.108	7.000	19.900
Precipitations	1196	6.751	2.755	1.060	15.570
Anual hours of sun	1196	2649.749	426.967	1600	3500
Altitude	1196	4.334	3.616	0.033	24.060
Ruggedness	1196	1.167	1.106	0.000	8.297
Soil quality	1196	8.324	1.129	0.952	10
Agricultural production 1950	1177	0.318	0.170	0.000	0.954
Agricultural production 1990	437	0.137	0.119	0.000	0.573
Illiteracy rate 1930	1196	47.829	14.202	14.391	88.408
Illiteracy rate 1990	437	1.797	0.991	0.000	4.463

*Notes:*See Section 4 for a description of the datasets. Put in 2 pages

Table A.2. Correlates of tourism

	Dep Var: Per capita tourism growth 1965-2019			
Variables	Coefficient	Std. Error	Std. Beta Coef.	P-vaue
Temperature	0.000	(0.000)	0.017	(0.358)
Precipitations	0.000	(0.004)	0.016	(0.820)
Hour of sun	0.000	(0.000)	0.011	(0.552)
Altitude	0.026**	(0.012)	0.099	(0.042)
Ruggedness	-0.044**	(0.019)	-0.053	(0.024)
Soil quality	-0.013***	(0.004)	-0.054	(0.004)
Share of agricultural production 1962	-0.008	(0.005)	-0.049	(0.120)
Illiteracy Rate 1930	-0.000	(0.001)	0.006	(0.756)
Log Population 1950	-0.020***	(0.004)	-0.110	(0.000)

Notes: Each row reports the OLS estimate of regressing tourism per capita growth on the variable indicated in each row separately. The unit of observation is the municipality. There are 1,196 observations included in each regression. Column 1 shows the point estimate for the regressor of interest. Column 2 shows the corresponding standard error. Column 3 corresponds to the standardized-beta and column 4 to the corresponding p-value.

Table A.3. Measure of tourism intensity

VARIABLES	(1) All sample	(2) Main sample
Beach length	0.591**	0.529*
	(0.270)	(0.291)
Beach width	0.825***	0.991***
	(0.188)	(0.207)
Temperature	0.773**	0.627*
	(0.303)	(0.339)
Precipitation	-0.228	-0.0893
	(0.263)	(0.298)
Hours of sun	0.218***	0.245***
	(0.188)	(0.165)
Observations	461	437

Notes: Each row reports the OLS estimate of regressing tourism per capita in 2019 on the variable indicated in each row separately. The unit of observation is the municipality. Column 1 shows the point estimate for the regressor of interest. Column 2 shows the corresponding standard error. Column 3 corresponds to the standardized-beta and column 4 to the corresponding p-value. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4. International tourist arrivals to Spain by origin

2000		2010		2019	
Germany	0.234	United Kingdom	0.190	United Kingdom	0.185
United Kingdom	0.207	Germany	0.182	Germany	0.129
France	0.092	France	0.108	France	0.106
United States	0.063	Italy	0.071	United States	0.061
Italy	0.062	United States	0.051	Italy	0.057
Netherlands	0.038	Netherlands	0.037	Netherlands	0.037
Belgium	0.035	Portugal	0.035	Portugal	0.029
Portugal	0.035	Belgium	0.027	Belgium	0.024
Japan	0.029	Sweden	0.018	Japan	0.020
Switzerland	0.018	Rusia	0.018	Irlanda	0.019
Sweden	0.016	Japan	0.017	Rusia	0.018
Argentina	0.015	Switzerland	0.016	Sweden	0.017
Cumulative	0.828		0.771		0.703

Notes: The table displays the composition of international tourist inflows in 2000, 2010 and 2019. Only twelve countries with largest tourist inflows are displayed. The data source is the National Statistic Institute data on international arrivals to Spain.

Table A.5. First-stage regression: Correlation between realized and predicted overnight stays

	Realized overnight stays 2019-2003		
	(1)	(2)	(3)
$\Delta \widehat{overnightstays}$	0.978***	0.957***	0.975***
	(0.0802)	(0.0871)	(0.092)
Log population 1990	-0.154	-0.255	-0.396**
	(0.115)	(0.187)	(0.195)
Log income 1990	0.063	0.093	0.143
	(0.215)	(0.229)	(0.250)
Average altitude			-0.364
			(0.472)
Ruggedness			0.355
			(0.457)
Temperature			-3.382
			(3.837)
Precipitations			0.392
			(0.611)
Soil quality			-0.0553
			(0.096)
Illiteracy rate 1990		0.278	1.015
		(0.566)	(0.660)
Agriculture production 1990		-0.251	-0.485
		(0.347)	(0.363)
Observations	96	96	96
F-Stat	49.61	28.97	15.76

Notes: The unit of observation is the municipality. The dependent variable is log growth of realized overnight stays between 2019 and 2003 for the subsample of municipalities of realized overnight stays data available. Each specification adds additional controls as indicated in the table. Predicted overnight stays is calculated as described in section 4. Robust standard errors are in parentheses. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A.6. Impact of tourism on local income. Shift-share analysis for the sub-sample.

	OLS	RF	IV
	(1)	(2)	(3)
Panel A. Dep. variable: Income level 2019			
Realized overnight stays	-0.014**		
	(0.005)		
Predicted overnight stays		-0.029***	-0.029***
		(0.001)	(0.009)
Mean dep. var.	9.49	9.49	9.49
Panel B. First stage: Dep. variable: Predicted overnight stays			
Realized overnight stays			0.966***
Kleibergen-Paap rk LM F-stat.			16.86
			[16.38]
Observations	96	96	96
All Controls	Y	Y	Y
Proovince FE	Y	Y	Y

Notes: (1) Panel A reports the estimates of tourism growth measured by overnight stays as explained in section 4 on income level using the OLS, reduced-form and IV regressions results where the dependent variable is the log of income in 2019 and panel B the First stage where the independent variable is the realized overnight stays. The unit of observation is the municipality; (2) All columns control for pre-determined socioeconomic and geographic variables: population in 1990, illiteracy rate in 1990, income in 1990, agriculture production in 1990, tourism per capita in 1990, soil quality, altitude, ruggedness,temperature, rainfall and hours of sun; (3) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level when adding province fixed effects. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A.7: Correlation of initial nationality shares and potential confounders

Nationality	Germans	Britons	French	Dutch	Portuguese
Population change	0.009	0.0318	0.0033	0.0065	-0.0002
Education level	0.000	-0.0022	-0.0010	-0.0014	-0.0000
Unemployment	-0.0006	0.0018	-0.0005	-0.0004	0.0000
Manufacturing industry employment	0.0105	0.0297	-0.0123	0.0984	-0.0162
Agriculture industry employment	-0.0119	-0.0550	-0.0059	-0.0098	0.0011

*Notes:*Notes: The table shows the correlations of the initial nationality shares across municipalities of the five tourist nationalities with the highest Rotemberg weights with several potential confounders. Changes refer to the period 1997-2019; Education, unemployment and shares by industry employment refer to the year 1990.

Table A.8. Impact of tourism on local income

	(1)	(2)	(3)	(4)	(5)
	Panel A. IV, Dep. variable: Income level 2019				
Yearly per-capita tourism growth	-0.050**	-0.067***	-0.049**	-0.049**	-0.051***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)
Mean dep. var.	9.49	9.49	9.49	9.49	9.49
	Panel B. First stage: Dep. variable: Per-capita tourism growth				
$\mathbb{1}$(Beach)	1.933***	3.480***	3.172***	3.173***	3.172***
Beach measure	1.028***	1.773***	0.150***	0.150***	0.147***
Kleibergen-Paap rk LM F-stat.	27.172	27.596	27.636	27.752	26.987
	[19.93]	[19.93]	[19.93]	[19.93]	[19.93]
Observations	1,196	1,196	1,196	1,196	1,196
All Controls	Y	Y	Y	Y	Y
Proovince FE	Y	Y	Y	Y	Y

Notes: (1) Panel A reports the IV estimates of tourism growth using IV regression, where per-capita tourism is instrumented by $\mathbb{1}$(Beach), a dummy equal to one if the municipality has beach and zero otherwise plus *Beach measure* which measures beach quality and panel B reports the First stage regressions. *Beach measure* in column 1 is the standardized km of beach length, column 2 adds standardized hours of sun, column 3 adds standardized km of beach width, column 4 adds standardized temperatures and column 5 adds standardized precipitations. The unit of observation is the municipality;

(2) All columns control for pre-determined socioeconomic and geographic variables: population in 1950, illiteracy rate in 1930, income in 1965, agriculture production in 1960, tourism per capita in 1965, soil quality, altitude, ruggedness,temperature, rainfall and hours of sun; (3) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.

Table A.9. Sensitivity analysis. Impact of tourism on local income.

	(1)	(2)	(3)	(4)	(5)	(6)
Per-capita tourism	-0.094***	-0.096***	-0.096***	-0.092***	-0.099***	-0.078**
	(0.021)	(0.023)	(0.027)	(0.023)	(0.026)	(0.032)
Kleibergen-Paap rk LM F-stat.	20.070	21.300	19.622	20.611	28.277	19.072
	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]	[16.38]
Observations	791	716	573	585	510	442
All Controls	Y	Y	Y	Y	Y	Y
Province FE	Y	Y	Y	Y	Y	Y

Notes: (1) The table shows the IV estimates of the impact of tourism growth on income level in 2019 using IV regression, where per-capita tourism is instrumented by $\mathbb{1}$(Beach), a dummy equal to one if the municipality has beach and zero otherwise. (2) Column 1 includes in the control group only municipalities that are more than 20 km far away and less than 50 km far away from the treated(beach) municipalities. Column 2 includes in the control group only municipalities that are more than 20 km far away and less than 40 km far away from the treated municipalities. Column 3 includes in the control group only municipalities that are more than 20 km far away and less than 30 km far away from the treated municipalities. Column 4 includes in the control group only municipalities that are more than 30 km far away and less than 50 km far away from the treated municipalities. Column 5 includes in the control group only municipalities that are more than 30 km far away and less than 40 km far away from the treated municipalities. Column 6 includes in the control group only municipalities that are more than 40 km far away and less than 50 km far away from the treated municipalities. The unit of observation is the municipality; (3) All columns control for pre-determined socioeconomic and geographic variables: population in 1950, illiteracy rate in 1930, income in 1965, agriculture production in 1960, tourism per capita in 1965, soil quality, altitude, ruggedness,temperature, rainfall and hours of sun; (4) Kleibergen-Paap rk LM F-stat. is the weak instrument test; in brackets we report the value of the Stock-Yogo weak ID test critical value at 10% maximal IV size. Robust standard errors are in parentheses and clustered at province level. *, ** and *** denote statistical significance at the 10, 5 and 1 % level, respectively.