Policy Research Working Paper 9189

What Are the Empirical Determinants of International Tourist Arrivals and Expenditures?

An Empirical Application to the Case of São Tomé and Príncipe

Gabriel Montes-Rojas
Rafael Barroso

WORLD BANK GROUP
Macroeconomics, Trade and Investment Global Practice
March 2020

Abstract

The link between tourism and growth is very important for some countries, especially small island countries like São Tomé and Príncipe. This paper investigates the empirical determinants of tourism outcomes (tourist arrivals and expenditures) and uses the findings to assess the performance of the tourism sector in São Tomé and Príncipe. The paper confirms most of the results found in the literature on the general determinants of tourism. Tourist arrivals increase with the gross domestic product and exports of the host country, as well as with increased air connectivity. Real exchange rate variations affect tourist decisions, and tourism outcomes have a persistent effect. The paper also finds that a positive attitude toward acceptance of lesbian, gay, bisexual, and transgender people increases tourist arrivals. Unfortunately, the relationship between digital presence and tourism outcomes could not be tested. The paper shows that São Tomé and Príncipe can do better in tourist arrivals, but it already has good performance on expenditures per tourist. Improving air connectivity is key to attracting more tourists, and demand for tourism is not very price sensitive, implying that the strategy to focus on high-spending tourists is the correct one.

This paper is a product of the Macroeconomics, Trade and Investment Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at rbarroso@worldbank.org.

What Are the Empirical Determinants of International Tourist Arrivals and Expenditures? An Empirical Application to the Case of São Tomé and Príncipe

Gabriel Montes-Rojas*
Rafael Barroso♦

JEL: C33, F14, L83, O14, Z32
Keywords: Service Trade, Tourism, Panel Data, Small-Island States

* Professor of Econometrics at Universidad de Buenos Aires and Universidad Nacional de La Plata, Argentina. He holds a PhD in Economics from University of Illinois at Urbana-Champaign and works on theoretical and applied econometrics with applications to labor and development economics.
♦ Senior Economist at the World Bank.

I. Introduction

The purpose of this paper is to empirically evaluate the main determinants of tourist arrivals and expenditures that have been found in the literature, incorporating new explanatory variables, and then assess the tourism performance of São Tomé and Príncipe (STP). The economic literature has established that international trade is one of the causes of economic growth.[1] This relationship can also be extended to the tourism sector, which is a service export.[2] STP is a small-island African country in the Gulf of Guinea, in which tourism is a comparative advantage, but it is far from the characterization of a tourism-dependent small economy. The empirical examination of the determinants of tourism aims to help STP and other small islands to understand what can be done to spur tourism growth.

The economic literature shows that tourism demand is affected by not only price and income, but also a host of other factors, such as air connectivity, language, and culture, among others. First, there is an extensive research agenda on measuring price and income elasticities of tourism, which is specific for different types of tourism destinations. Second, there is a myriad of characteristics that are found to be important to the tourism industry, such as remoteness, language, culture, air connectivity, bilateral trade, etc. More recently, there have been many studies emphasizing the role of digital media and digital presence[3] as a key determinant of tourist decisions. Finally, the models were expanded to investigate the link between acceptance of lesbian, gay, bisexual, and transgender (LGBT) people and its impacts on tourism.

Following the literature on determinants of tourism demand, several empirical approaches were applied and yielded general findings applicable to all countries and some specific to STP. This paper constructs a database from different sources to provide a complete picture of tourism statistics and explanatory variables. Different statistical models were used to establish correlations and, whenever possible, to establish causation.

The first set of general findings confirms most of the results found in the literature. Gross domestic product (GDP) and exports of the host country are positively associated with tourism. Higher flight connectivity increases tourist arrivals but does not change average expenditures per tourist. Tourist decisions are influenced by relative price variations, which means that a real exchange rate depreciation has a positive effect on arrivals and a negative effect on expenditures. Finally, tourist arrivals and expenditures per tourist have a persistent effect, meaning that increases in one year tend to be maintained in future years.

The second general finding is that positive attitude towards LGBT population, as measured by decriminalization, is associated with an increase in tourist arrivals. A causal analysis was performed to evaluate the effect of decriminalization of LGBT and large positive effects were observed. Using

---

[1] See Frankel, J. A., & Romer, D. H. (1999). Does trade cause growth? American Economic Review, 89(3), 379-399 and Alcalá, F., & Ciccone, A. (2004). Trade and productivity. Quarterly Journal of Economics, 119(2), 613-646.

[2] See section II for a review of the literature.

[3] Digital presence is how a business, in this case a tourism destination, appears online in all forms possible: website, social network, mobile, review sites, etc.

a difference-in-differences approach of a dummy variable that takes the value of 1 if being LGBT is not a crime in the country and zero otherwise, the dummy is positively associated with tourist arrivals, although this has no relationship with tourism per capita spending. If we assume that, conditional on the countries' observed characteristics, the change in the LGBT legal environment is not correlated with unobserved characteristics, the causal effects are valid and highlight the importance of tolerance and security for attracting tourists.

The data obtained on digital presence indicators is partial and from a short period, which limits the ability for meaningful statistical analysis. Unfortunately, the frequency and nature of the data cannot be used to further investigate the effect of digital footprints on tourist arrivals or related measures. As the literature speculates on positive effects derived from digital presence, this should be encouraged at all levels. As such, investing in digital visibility is likely to have a positive effect on tourist arrivals.

The paper shows first that STP can do better in terms of the number of tourist arrivals but has a good performance in terms of expenditures per tourist. Using a comprehensive list of determinants and a hedonic model of tourist arrivals and expenditures per tourist, the empirical analysis reveals that STP had a lower number of tourists in 2016 than what would be otherwise predicted by the model, but spending by tourists was larger than what the model suggests.

In addition, the empirical analysis confirms the importance of improving air connectivity to the island. STP can potentially receive more tourists because it has a low ratio of tourists-to-population and land area; however, flight demand meets supply as measured by seats sold on incoming flights, which shows that the country needs to increase air connectivity. STP ranks low in centrality indexes in the networks using flight data. Its major flight origin is Lisbon, Portugal, which determines that many of the European tourists need to fly to Lisbon first. As connections decrease utility and propensity to spend at destinations, STP might benefit greatly from connections to central and large hubs. A simple exercise simulating a connection to London reveals major changes in centrality. The tourism elasticity with respect to seat availability on flights is of the order of 0.1, which means that an increase of 10 percent in seat availability will increase tourism arrivals by 1 percent.

Finally, results show that demand for tourism in STP is not very sensitive to price, implying that the strategy to focus on high-spending tourists instead of large volume of tourists is the correct one. The price elasticity of tourism is less than one, which means tourism is price inelastic; moreover, the elasticity of expenditures per tourist is higher than one. As a result, the pricing strategy to maximize revenues is to focus on high-spending tourists. Since arrivals of tourists have a high persistence (as measured by the autoregressive parameter in a dynamic panel data model), the long-term effects are large. STP had a recent trend of currency appreciation with respect to the euro and US dollar. The latest results confirm that the total expenditures of the tourism industry benefited from this as was measured in real terms.

II. Literature review on the determinants of tourism

The link between tourism and development is significant and empirically well-documented. Trade competitiveness is associated with higher GDP growth. Tourism, being an export service, is certainly

important for countries with comparative advantages in tourism. That tourism promotes economic growth is confirmed in Dritsakis (2004) for Greece, Balaguer and Cantavella-Jordá (2002) for Spain, Oh (2005) for the Republic of Korea, Durbarry (2004) for Mauritius, and Tosun (1999) and Gunduz and Hatemi (2005) for Turkey. Cross-country studies by Sequeira and Campos (2006) and Brau et al. (2003) demonstrate that tourism-specialized countries grow more than others. Sinclair's (1998) survey indicates that the main channel through which tourism helps growth is the provision of foreign currency, which can finance imports of capital goods. Culiuc (2014), based on a gravity model, found that the income, real exchange rate and bilateral trade flows of source countries have a strong impact on world tourist arrivals, although the impact is much reduced for smaller countries. Wolfe and Romeu (2011) reached similar conclusions. Lederman and Maloney (2007) also found that tourism exports have less volatility than other natural resource-based export industries. The following sections review the literature on specific determinants of tourism demand.

IIa. Price elasticity of tourism

An important feature of tourism markets is whether tourism services are price elastic or not. The price of tourism is a multivalued concept, as tourism services differ in quantity and quality more than other goods and services. The price of tourism services depends on whether they are denominated in local or foreign currencies, and most of the studies reviewed below use different measures of the real exchange rate. The price of the tourism-related services should be specific to the type of consumption; however, detailed price data are not common in tourism empirical analysis; prices are generally replaced by average spending per tourist and deflated by a real exchange rate index.

In most papers, the real exchange rate has the expected effect: an appreciation of the origin's currency increases bilateral tourism while the appreciation of the destination reduces it, and an increase in income tends to increase demand more than proportionately. If tourism is inelastic; that is, the price elasticity of tourism is less than one, then increasing its price will reduce tourist arrivals and that will result in a higher spending per tourist (total revenues increase more than the reduction in tourist arrivals). If, on the contrary, tourism is elastic; that is, price elasticity is more than one, implying that tourism is a luxury good, then increasing the price will produce a reduction of the average spending per tourist. The estimated elasticity is in general large and is highly sensitive to the estimation techniques and measures of the real exchange rate. Rosensweig (1988) finds elasticities of substitution of Caribbean islands with each other is 1.33 to 2.45, with Mexico 1.85 to 1.00 and with Europe 1.7, depending on whether the origin is restricted to the United States or all countries. Rosensweig's estimates are broadly in line with those of Papatheodoru (1999), who estimates a model in levels for the Mediterranean region for tourists from France, Germany, and the United Kingdom and finds price elasticities greater than 1; however, Garín-Muñoz and Pérez Amaral (2000) estimate a dynamic model in levels and a non-dynamic differences specification on the determinants of tourist flows to Spain and obtain price elasticities of only 0.30 and 0.25 respectively. Song et al. (2000), working in an error correction framework on single country data, obtain price elasticities that range from 0 to 2 for the principal destination countries of English tourists. Estimated income elasticities tend to confirm the intuition of the sector as being a luxury

good although there is also substantial variance. Garín-Muñoz and Pérez Amaral (2000) find elasticities of 0.97 and 0.91 and Song et al. (2000) of around 2. Maloney and Montes-Rojas (2005) study substitution among Caribbean islands by type of tourist origin and find a very high price elasticity of close to 4.9, confirming that tourism is a luxury good with income elasticity above 2.

Small islands may suffer from the problem known as Dutch Disease[4] that results in an unfavorable over-appreciated exchange rate (Bertram, 2009). While most often Dutch Disease refers to natural resource discovery, it can also refer to any development that results in a large inflow of foreign currency. The natural beauty endowment of small islands (warm temperatures, beaches, etc. that result in an abundance of tourism-related amenities) can be considered the natural resource on which tourism develops and brings foreign exchange inflows. Therefore, an expansion of the tourism industry may result in a lack of competitiveness of other exports and may increase the relative price of non-tradable goods (see the discussion in Lederman and Maloney, 2007). This leads to the fact that the macroeconomic consequences of tourism growth should be considered for appropriate policy design.

IIb. Digital media

Information and digital presence play a central role in the tourism industry. Rauch and Trindade (2002, 2003) have emphasized the role that informational barriers play in restricting trade, and how private networks can overcome those barriers. Informational barriers are defined as the barriers that arise because of asymmetric or incomplete information in a given market. Trade is affected by information about the counterpart, and this is relevant for the tourism service exports. For tourism, Gawande, Maloney, and Montes-Rojas (2011) find a similar mechanism at work by involving transboundary business-government networks. They show that the lobbying in the destination countries increases the level of tourism and reduces the price elasticity of demand, suggesting that the destinations have the ability to differentiate their products. Both are of magnitudes that suggest important policy implications for developing economies.

Informational barriers are reduced by digital media. There is a large literature in tourism-related journals (Leung et al., 2013, Standing et al., 2014, Li et al., 2018) that highlights the role of digital media in overcoming informational barriers. They all recognize that social media and digital information play a central role in the shape of the tourism industry, involving both demand (e.g., "The newly released *World Travel Market 2011 Industry Report* announced that more than one-third of all leisure travelers in the United Kingdom choose their hotels on the basis of social media sites like TripAdvisor and Facebook" Leung et al. 2013, p.4) and supply (the market share for intermediation of large websites has increased considerably over traditional travel agencies). Leung et al. (2013) state that "social media have been widely adopted by travelers to search, organize, share, and annotate their travel stories and experiences through blogs and microblogs (e.g., Blogger

[4] The Dutch Disease is the causal relationship between growth of a specific sector (for example, natural resources) and a decline in other sectors (like the manufacturing sector or agriculture). As revenues increase in the growing sector, the nation's currency appreciates compared to currencies of other nations. This results in the nation's other exports becoming relatively more expensive for other countries to buy, and imports becoming cheaper, making those other sectors less competitive.

and Twitter), online communities (e.g., Facebook, RenRen, and TripAdvisor), media sharing sites (e.g., Flickr and YouTube), social bookmarking sites (e.g., Delicious), social knowledge sharing sites (e.g., Wikitravel), and other tools in a collaborative way." (p.4).

Digital information may increase tourism competition among tourism destinations, and among tourism service providers within tourism destinations, which might benefit consumers. Moreover, the Internet may reduce inequality among tourism providers, giving all involved the chance to promote their services on an equal basis. The evidence is scattered.[5] However, most econometric analysis shows that providers (i.e., hotels) do respond to consumer valuation. Ye et al. (2009) find that "results indicate a significant relationship between online consumer reviews and business performance of hotels." (p.1). Xie et al. (2014) state that "The results show that overall rating, attribute ratings of purchase value, location and cleanliness, variation and volume of consumer reviews, and the number of management responses are significantly associated with hotel performance." (p.1). Standing et al. (2014) clearly attribute this shift of power to the information technologies: "Indeed, it is argued the Internet has facilitated a shift of power from travel providers to consumers since they now have access to a wide range of travel providers on the Internet that often compete on price (Law et al., 2010). New travel intermediaries have become a recognized force in the industry and tourism destinations have embraced the Internet for promotion and marketing. The use of Internet purchasing of travel products and the growth in numbers of people with Internet access has led to continued growth in travel consumers (Ramos & Rodrigues, 2010). These consumers are continually developing in terms of their needs for information, and research should evolve in terms of the areas of focus to address practitioner concerns and requirements."

IIc. Transport and connectivity

Geographical proximity and cultural affinity positively influence tourism demand especially for small islands. The literature survey of McElroy and Parry (2010) clearly identifies geographic proximity as the main factor explaining the tourism attractiveness of small islands, followed by dependent, colonial, or political status and uncrowded island ambience. In fact, the same factors are observed by Coscia, Hausman and Neffke (2016) using anonymized and aggregated foreign transaction data from credit card expenditures: distance, a country's reported wealth, and cultural affinity in affecting tourism demand.

Increased airlift connectivity benefits the destination country and the inbound tourists. The evidence on the effect of airlift supply is clear: more airlift benefits any tourism destination (see, for instance, the review and the econometric evidence for the Caribbean of Acevedo et al., 2016). Another issue is that more airlift supply should increase competition among airlines for a destination, and as such, it should be expected that less money would be spent on transportation

[5] Of theoretical relevance is the question of whether, as more islands participate in the informational game, competition among them may be used strategically by agents to capture rents. For instance, Zhang et al. (2016) indicate that expert reviews influence consumer ratings. As such, information might be manipulated by the bigger players, a concern that the recent focus on digital platforms such as Facebook clearly illustrates.

and more will be spent on the destination. There should be ceteris paribus an unambiguous positive effect of having more airlift connectivity for a given destination.

STP lacks appropriate transport connections, which results in difficulties to compete with other tourism destinations. The data show that flight origins with STP as the destination are Angola, Cameroon, Cabo Verde, the Democratic Republic of Congo, Equatorial Guinea, Gabon, Ghana, Côte d'Ivoire, Nigeria and Portugal. The list of countries is not the same across all years of analysis. This illustrates that STP might require flight stops in non-traditional tourism destinations in Africa for receiving tourists from high-income countries. Even in that case, the Portuguese tourism market is not big enough and it serves only as a limited flight connection. A simple market analysis might conclude that tourists prefer more direct flights to reach a destination, since time spent at airports decreases the tourism experience utility.

IId. Protected areas and tourism attractiveness

Since tourism is a relatively heterogeneous good, countries can differentiate on the product offered by establishing protected areas. Conditional on its tourism attractiveness endowments, such as natural resources, countries can promote tourism through investing in infrastructure and legislation. Protected areas are one of the most important investments for tourism. The report by the World Economic Forum (2017) highlights this issue in clear terms: "Data reveals that the environmental strength of a country is directly related to tourism revenue (…). Although this relationship is complex, and there is no evidence of direct causality, the more pristine the natural environment of a country, the more tourists are inclined to travel there, and the more they are willing to pay to access well-preserved areas. Consequently, as the natural capital depletes, destinations lose revenue." (World Economic Forum, 2017, p.6)

Protected areas increase the attractiveness of a country for tourism. Many of the protected areas established in the late 19th and early 20th centuries responded to practical interests such as favoring tourism or preserving iconic landscape features. Baldi et al. (2017) found that the tourism attractiveness of an area is positively related to its level of protection, achieving a top importance in the ranking of variables. A common feature is that tourism and protection are involved in positive feedbacks, as protection itself attracts visitors interested in remarkable natural or cultural landscapes, and visitors drive protection to preserve this quality. Tourism also engages local communities and regional and national governments in the preservation of these landscapes, offering economic revenues that eventually exceed those obtained from traditional land uses (Mulholland and Eagles, 2002; Siikamäki et al., 2015). Given that land resources are scarce, protected touristic areas may compete with other industries. As a result, the effect on the local population's well-being cannot be determined without further analysis. In fact, tourism may have a negative effect on the environment as well: "Although tourism is often negatively impacted by the pollution caused by other human activity, it is important to recognize that processes, mechanisms and activities associated with tourism also damage the environment." (World Economic Forum, 2017, p.6)

III. Data description and econometric analysis

Tourist arrivals is the main outcome variable of interest. Arrivals data measure the flows of international visitors to the country of reference: each arrival corresponds to one inbound touristic trip. If a person visits several countries during the course of a single trip, his/her arrival in each country is recorded separately. In an accounting period, arrivals are not necessarily equal to the number of persons traveling (when a person visits the same country several times a year, each trip by the same person is counted as a separate arrival). Being an island that is not close to other destinations, in STP's case, number of arrivals is closer to the number of visitors within a year. *Arrivals* data should correspond to the *inbound visitors* by including both tourists and same-day non-resident visitors. All other types of travelers (such as border, seasonal and other short-term workers, long-term students and others) should be excluded, as they do not qualify as visitors. As noted by UNWTO (2018), data are obtained from different sources: administrative records (immigration, traffic counts, and other possible types of controls), border surveys or a mix of them. If data are obtained from accommodation surveys, the number of guests is used as an estimate of arrival figures; consequently, in this case, breakdowns by regions, main purpose of the trip, modes of transport used or forms of organization of the trip are based on complementary visitor surveys. The main variable of interest to this note is the log of yearly tourist arrivals (LN_ARR).

Unfortunately, the UNWTO does not report data for STP on arrivals by origin. The data only show aggregate arrivals from all origins, which is a major limitation of the analysis as it does not allow to disaggregate by type of tourist. It could be assumed that tourist origins are significantly correlated with their income, and then as done in the literature, European and American tourists are considered as the representatives of tourist tastes and behavior. As noted above, flight arrivals from developed countries are connected only through Lisbon, and as such, it can be safely assumed that most tourists are from Europe.

Expenditures by tourists is also an outcome variable of interest. Expenditures associated with the activity of visitors have been traditionally identified with the travel item of the Balance of Payments (BOP). In the case of outbound tourism, those expenditures associated with resident visitors are registered as "debits" in the BOP and refer to "travel expenditure". As in the case of *inbound tourism*, BOP data are used. The 2008 International Recommendations for Tourism Statistics consider that "tourism industries and products" include transport of passengers. Consequently, a better estimate of the tourism-related expenditures by resident and non-resident visitors in an international scenario would be, in terms of the BOP, the value of the travel item plus that of the passenger transport item.

A key variable for our analysis is the average spending per tourist. Given information limitations, this is approximated by the ratio between Expenditure and Arrivals, measured as the log of the ratio Expenditure / Arrivals (LN_EA). This is a proxy of prices paid by a representative tourist. As noted above, this aggregate and average measure is not rich enough to study the prices of different tourism services.

Supply-side information is used from airline schedules on numbers of flights and seats. This corresponds to the total number of flights reported by commercial airlines among the destinations (including local flights within countries) and the seats that were reported available from each flight.

The data are then aggregated on a yearly basis and; thus, they correspond to the total flights that occurred during the year between a unique pair of origin and destination, measured as the log of the yearly total number of seats available on the incoming flights (LN_SEATS).

Finally, data on Instagram posts, as a proxy of digital presence, are also collected for studying STP tourism performance. These data consist of the daily counts of posts containing #hashtags following #travel[COUNTRYNAME] pattern or similar, if applicable, from January 2016 to October 2018. These hashtags contain the most relevant posts to this task. Construction is based on Instagram API output following a query of said hashtags on a global scale using a unique data access.

Other country-level information (GDP, population, exports, prices, nominal exchange rates) is used as control variables and taken from the World Development Indicators database. The variables are used in logs and they are given by LN_GDP, LN_POP and LN_X. The real exchange rate indicators are constructed using the exchange rate and prices of the country and with respect to the euro area, the United States and the United Kingdom. These are given by the variables, RER_EURO, RER_US and RER_POUND, respectively.

Over the last two decades, STP experienced major changes regarding its tourism industry. In order to construct a comparison group, STP performance is studied in several dimensions with respect to other "similar" tourism destinations. First, this note uses islands that are at a similar distance with respect to major tourism origins (i.e., Europe, the United States). The list of countries used for this group is Cabo Verde (CPV), Madagascar (MDG), Maldives (MDV), Mauritius (MUS), and the Seychelles (SYC). The econometric analysis below expands this control group to include Caribbean and Pacific islands as well as some other countries with strong tourism sectors.

IIIa. Baseline random-effects model

This first model establishes correlation and association between the tourism the outcome variables and a set of tourism determinants. In order to study the determinants of tourism, consider first a baseline random-effects panel data model of the form:

$$Y_{it} = \beta X_{it-1} + \gamma Z_i + \phi W_{t-1} + \delta_t + \mu_i + \varepsilon_{it},$$

where ***i*** indexes tourism destinations and ***t*** time, ***Y*** is the dependent variable of interest (log of tourist arrivals, or log of expenditures per tourist), **X** is a set of control variables that varies with respect to both i and t, **Z** varies only across i and **W** only with respect to time. $\boldsymbol{\delta}$ contains year effects, $\boldsymbol{\mu}$ a country specific random-effect and $\boldsymbol{\varepsilon}$ the unexplained component of the model. For **X**, the note considers the destination specific information: log of GDP (LN_GDP), log of population (LN_POP), log of total number of seats in incoming flights (LN_SEATS), real exchange rate (with respect to euro, RER_EURO); for **Z**, the notes considers the size of the country (LN_SIZE), and for **W**, it considers the log of GDP of the euro area (LN_GDPEURO). The variables are described in the previous section. This model does not intend to establish causal relationships among the control variables and the dependent variables, but rather to establish correlation and association. The model, thus, produces a hedonic regression, in which a set of characteristics are used as determinants of a given outcome

variable.[6] Then, using this strategy, we can evaluate how STP relates to the tourism determinants used as inputs in the hedonic model.

IIIb. Tourism price elasticity

A dynamic panel data model was regressed to estimate price elasticity. As noted above, a very important determinant of the tourism industry performance is price and, in particular, if tourism can be considered a market with elastic or inelastic price elasticity. Consider now the following dynamic panel data model:

$$Y_{it}=\alpha Y_{it-1} + \beta ER_{it} + \gamma X_{it-1} + \delta_t + \mu_I + \varepsilon_{it},$$

where **Y** is the dependent variable of interest (LN_ARR and LN_EA), **ER** is the real exchange rate with respect to either euro, US dollar or British pound, **X** is a set of control variables that affect tourists' decisions (GDP of destination, population, exports, and total seats in flights for that destination, all lagged one period), δ contains time-specific effects, μ is island-specific effects, and ε is an idiosyncratic error term. Stationary panels require that $|\alpha|<1$, which is assumed**.** The main coefficients of interest are β and γ. In particular, for dynamic models the main interest is in the long-run effects, which mean the ratio: $\beta/(1-\alpha)$.

This model is appropriate for this estimation for two reasons: (i) tourism arrivals and receipts are likely to have a clear temporal persistence. The experience to spend holidays in a certain destination provides information that cannot be achieved by other methods. As such, if a certain destination was visited (provided the experience was positive), it will increase the chances of revisiting and arrivals from friends, family, etc. Nevertheless, the available data set is not long enough to construct VAR models and, therefore, it cannot be evaluated by using island-specific dynamic effects, but rather by using a common autoregressive parameter. The above model is similar to that of Maloney and Montes-Rojas (2005) for Caribbean tourism destinations.

And (ii) data availability only allows the use of total aggregate tourist arrivals from all destinations. Since different islands have a different bundle of tourist origins, the model also has controls for this: the presence of fixed-effects by country and by year. The former captures all unobservable components that affect tourism arrivals (or expenditures) for a particular destination. The aggregation procedure, thus, passes all unobservable characteristics of the tourism destinations to this factor. The latter does the same to all year specific unobservable components. In particular, the main interest is controlling for changes in origin GDP and tastes. For this analysis, this note considers 36 tourism destinations from different locations. The data availability is from 2009 to 2016.

A well-known issue in dynamic panel data models is that the model cannot be estimated consistently by conventional methods such as OLS or FE, the so-called Nickel bias. The reason is that the presence of individual (i.e., island) specific effects produces a bias that affects α and other parameters. This applies to this data set since its shorth length implies that the bias effect cannot be assumed to be

[6] In hedonic regression models, characteristics are used as determinants of a certain outcome variable. The typical application is the house price hedonic price model, where house characteristics are used to infer about the price of a house.

zero. In order to address this problem, the regression estimates a model in differences (to eliminate the fixed-effect) and then instruments the lag of the dependent variable, following the procedure of Anderson and Hsiao (1981,1982). Since there are more instrumental variables (IV) than parameters, it is common to consider the Generalized Method of Moments (GMM) implementation of Arellano and Bond (1991) and Blundell and Bond (1998). In particular, the empirical model used here is the Roodman (2009a,2009b) GMM with collapsed IV.

IV. General Findings

The first model produces results that are in line with the literature reviewed: GDP in the issuing country and the exports of the host country are positively associated with tourism. Table 3 reports the regression coefficients using both tourism arrivals and expenditures per arrivals as dependent variables. The coefficients are in line with the expected results. Richer countries, as measured by GDP controlling for population, are associated with more tourist arrivals, indicating a weak preference for destinations with amenities correlated with destination wealth. Still, the same control variable has a negative and significant effect on expenditures per tourist. Controlling for all other covariates, countries with higher GDP have, on average, tourists with lower expenditures per capita. Total exports are positively associated with more tourists and expenditures per tourist. This suggests that countries that are trade oriented may have cultural attitudes that promote tourist satisfaction.

Higher flight connectivity increases tourist arrivals but does not change expenditures per tourist. Consider now the effect of flight connectivity, as given by LN_SEATS. The results suggest that increasing the availability of seats on incoming flights by 1 percent increases tourism arrivals by 0.12 percent, but it does not affect expenditures per tourist. Note that as analyzed above, the connectivity of a given destination is richer than the number of incoming flights or sold seats, but it depends on the entire network of connections among destinations.

IVa. Tourism Price Elasticity

Tourist decisions are influenced by relative price variations. A real exchange rate depreciation has a statistically significant positive effect on LN_ARR and a negative effect on LN_EA. This relationship is explored in detail in the following paragraphs using a dynamic panel data model. This simple association, however, indicates that, controlling for other amenities, tourist decisions are influenced by relative price variations as expected.

Tourist arrivals and expenditures per tourist have a persistent effect, meaning that increases in one year tend to be maintained in future years. The econometric results appear in Table 5 and Table 6 for arrivals and expenditures per arrival respectively. In both cases, tourism variables are persistent: the autoregressive lag coefficient is higher than 0.7.[7] This means that even after controlling for

[7] Note that the FE and Sys-GMM coefficients are of similar magnitude. This indicates that the Nickel bias is small in the first place and makes the empirical strategy more reliable. The Arellano-Bond test for absence of AR2 correlation of residuals cannot be rejected, while the Hansen test indicates that the instruments are valid. Since the number of instruments is close to the number of countries, these results are reliable.

islands characteristics, increasing tourist arrivals will be maintained into the future (naturally, this applies equally to both positive and negative shocks). This issue also highlights the importance of digital presence to reinforce positive experiences.

A real devaluation of the currency increases tourist arrivals by a smaller margin, and increases total expenditures because expenditures increase more than the decline in tourist arrivals. The euro real exchange rate is statistically significant for both FE and Sys-GMM cases, while the US dollar and UK pound are marginally significant. As a result, a depreciation of the exchange rate results, in general, in an increment in the number of tourists. When the long-run elasticity is taken into account, it is less than one in all cases. Moreover, the test for those elasticities being equal to one is always rejected. As such, tourist arrivals respond relatively less when prices increase, and it determines that a pricing strategy that focuses on high-spending tourists is suitable. When looking at the real exchange rate effects on price, expenditures per tourist arrival, the effects show that an appreciation of the local currency results in an increment of total expenditures (because expenditures increase more than tourist arrivals decline). In this case, the effects are negative and significant for all currency comparison cases. The long-run elasticity is close to -1 in all cases.

IVb. LGBT tolerance and tourism

It has been long hypothesized that tourism destinations could benefit from an inclusive attitude towards LGBT population. The tourism industry can take advantage of the rise of the global phenomenon of the "pink dollar" by creating an inclusive environment for international LGBT tourists. The "pink dollar" is the term used to describe the purchasing power of sexual and gender minorities. Sexual and gender minorities are more likely to spend their money on travel, and when they do, they are more likely to explore new destinations, provided those are safe.

Lack of appropriate data has been an issue to establish causal relationships relating to this, but the recent trend in changing legal environment can help to establish this relationship. Using a data set on the LBGT environment prepared by the World Bank, we explore the effect of changing the legislation with respect to criminalization of LGBT using a difference-in-differences (d-in-d) approach. This paper, in particular, uses a dummy variable that takes the value 1 if the country does not criminalize LGBT and 0 otherwise, and it matches the data in Table 1 of Cortez and Arzinos (2019) to our sample data, resulting in 15 countries that appear in both samples. Since there are some countries that changed their legal attitude towards LGBT (e.g., STP does not criminalize LGBT since 2012), it can explore if the change in legislation is associated with an effect on tourism arrivals. The identification analysis, as it is common in difference-in-difference, is based on the assumption that controlling for observable characteristics, the timing of the implementation of the change in LGBT legislation is not related to other tourism related policy.

Positive attitude towards LGBT population, as measured by not criminalizing LGBT, positively impacts tourist arrivals. Table 6 presents the d-in-d econometric estimators for log of tourism arrivals (columns (1) and (2)) and log of expenditures per arrivals (columns (3) and (4)). Column (1) provides a point estimate of 0.431, statistically significant at 5 percent level, implying that removing the criminalization of LGBT would increase tourism arrivals by 43 percent. Including additional

control variables as in the baseline model in column (2) also produces a positive and significant effect of 0.273. There is a positive but not statistically significant effect for expenditures per arrivals. In sum, the results indicate that tourist arrivals are very responsive to the liberties and democratic attitudes toward minorities of destinations. As shown in columns (3) and (4), however, they affect neither the consumer propensity to consume tourism services nor the economic type of consumer.

V. Specific findings for STP

Va. Arrivals and Tourism receipts

The total number of tourist arrivals has increased, but there seems to be considerable room for further growth. Figure 1 plots total arrival figures for STP and five other countries that serve as a comparison group. Overall, STP has a small market compared to its competitors, but it has experienced a rapid increase in the total arrivals, which grew by three times when compared to 2010 (see Figure 2), and this is very distinctive when compared to the other five countries in the comparison group. Despite this major change, the long-run tourism capacity seems to be far from being saturated. In fact, Figure 3 and Figure 4 show that STP has relatively low ratios of tourist arrivals with respect to total population and land area, which serve as crude indices of potential expansion. Thus, in a relative perspective, it seems that STP can increase its tourist numbers without the risk of over-crowding even after taking into account that the type of tourism marketed in STP will always yield lower ratios than in most of its peers.

Together with this expansion in the number of tourist arrivals, STP also experienced a considerable increase in tourism receipts. Expenditures per capita or per tourist show that there were major changes in STP during the last decade. Figure 5 and Figure 6 show significant increments in the ratio of expenditures per arrivals. As noted above, this is a proxy of the average price of tourism services. This trend is unique to STP: from 1999 to 2009, the average receipts per arrival declined, but it has increased (more than doubled in current USD) since 2009. As a result, using the latest available data, STP has an average spending per tourist that is higher than the Seychelles and Maldives, two of the most exclusive competitors. The increase in expenditure per capita seems to be related to the introduction of the exchange rate peg in 2010, which resulted in appreciating the country's currency. Thus, the increment noted above was not a conscious decision of the country, but rather a result of the macroeconomic policies. Given the estimated price elasticities above, this could be seen as an adequate revenue maximization strategy.

STP can do better in the number of tourism arrivals but not in the expenditure per tourist, given the structural determinants as assessed by the first model. In order to evaluate the overall performance of STP with respect to the comparison group, this note uses the hedonic model above and computes the ratio of the actual value of the model linear prediction of LN_ARR with respect to the actual value of LN_ARR, and then the imputed value of the unobservable component of the model (μ) with respect to the actual value of LN_ARR. The plot of these two ratios for 2016 appears in Figure *28*. The same procedure is used for LN_EA in Figure 29. The analysis for LN_ARR indicates that STP has a predicted value that is above the actual value for 2016, which indicates that given the characteristics corresponding the STP and the coefficients of the joint regression model, STP is in

the group of countries that can do better in terms of number of tourists. Moreover, STP can also do better in terms of the unobservable component. The unobserved component contains all other tourism determinants that have not been considered explicitly in the model, in particular, location, culture and other fixed determinants of the countries in the sample. The fact that STP ranks low in this is associated with one of those. Probably, the main candidate is related to location and connectivity (which is not captured by LN_SEATS, an aggregate measure of arriving flights). The analysis for expenditure per tourist, however, reveals that STP has a better performance than what would be otherwise expected with the hedonic model. This confirms previous empirical analysis indicating that STP is one of the destinations where expenditure per tourist has increased the most in absolute and relative terms.

Vb. Flights and seats arrivals

Since STP is an island, the change in tourist arrivals is accompanied by a marked increase in flight arrivals, but a decrease in seats per flight. Table 2 presents, in detail, all flights that have STP as a destination: as of 2016, Angola, Ghana, and Portugal were the countries of departure from which most tourists came. Moreover, Angola and Portugal are the ones for which the number of connections has been more stable over the past years, a result that indicates the importance of the colonial past in the present configuration. Figure 10 and Figure 11 present the number of flight arrivals both in absolute and in relative terms (with respect to 2010). STP shows the steepest increment over the last years when total flights are 2.5 times the number in 2010. However, when looking at seats (Figure 12 and Figure 13), STP had a similar increase as the five countries in the comparison group, at the same level as Maldives and the Seychelles. As a result, it seems that STP had an increment mostly based on flights with fewer seats. The ration of seats per flight declined from 109 in 2010 to 67 in 2016. Figure 14, indeed, shows that STP is a unique case for this abrupt decline.

The connectivity pattern is the result of its colonial past and cultural affinity with Portugal. Another barrier, as in the case of STP, the language spoken is Portuguese, which is not widely known outside Portuguese-speaking countries, and which does not represent a large tourist-issuing market. An obvious implication of this is that STP workers in the tourism industry should have appropriate knowledge of foreign languages.

The lack of connectivity can be clearly observed when analyzing flight connections as a full network graph. Take all countries with flight data to be a collection of nodes or vertices, where origin and destination of flights define an edge or link. Consider, thus, an unweighted and undirected network analysis of the flight data for 2016, with special emphasis on STP and related tourism destinations. The unweighted network structure is preferred because the main interest is on the connectivity potential of STP and not in the number of flights, which are related to the node size (i.e., tourist capacity), but on its potential for connectivity (i.e., attracting tourists). Consider also an undirected network as most destinations have the same amount of incoming and outgoing flights. Of interest for a given node is a measure of centrality which determines how important each node is for the full network structure and cohesion.

STP has a relatively low measure of centrality, as measured by different methods. The degree of a node corresponds to the number of links that it has. STP had only five incoming connections in 2016 (10, summing both directions) and, as such, it is on the 12.3 percentile of the node degree distribution. Moreover, it ranks lower than the other five island destinations in the peer group. While this may be affected by size, other centrality measures analyze the whole network structure. Three popular centrality measures are betweenness,[8] closeness,[9] and eigenvalue vector centrality.[10] While for betweenness STP is on the lowest 15.5 percentile, it performs worse for closeness: 6.4 percentile, and eigenvalue: 11.8 percentile. Figure 15 illustrates the network structure of STP together with its connections, and the countries connected to those. The figure suggests that STP is not a central part of the network. Figure 16 and Figure 17 clearly illustrate this issue by plotting the centrality measures discussed above. In all cases, STP appears on the lower end of the centrality distributions. In fact, when compared to the five peer countries, it is clearly in a disadvantageous situation.

This analysis suggests that policy actions need to be taken to compensate STP's isolation and remoteness. The network analysis evaluates all nodes in relation to their connected nodes, and through them, to the connections of these, which means it seeks to capture the idea that the more central the neighbors of a vertex are, the more central that vertex itself is. As such, isolation is not only a problem of being far away (in fact the analysis above does not use geographical distance as an input), but also of having poor links mostly. As a result, it shows that being linked to major hubs, other than Portugal (which is not large), is very important.

Adding one weekly flight to a major hub would significantly impact STP's connectivity rankings. As a hypothetical example, consider the effect of adding a link to London, keeping everything else constant. While STP's degree only increases by two - moving STP from the 12.3 percentile to the 14.5 - betweenness increases from 15.5 percentile to the 21.8, closeness from 6.4 percentile to 23.2 and eigenvalue centrality from 11.8 percentile to 18.2. This simple exercise, which can be replicated by adding any other tourist origin or connection, highlights that major gains can be achieved by small increases in connectivity.

---

[8] *Betweenness centrality* measures are aimed at summarizing the extent to which a vertex is located 'between' other pairs of vertices. These centralities are based on the perspective that 'importance' relates to where a vertex is located with respect to the paths in the network graph. If we picture those paths as the routes by which, say, communication of some sort or another takes place, vertices that sit on many paths are likely more critical to the communication process.

[9] *Closeness centrality* measures attempt to capture the notion that a vertex is 'central' if it is 'close' to many other vertices. The standard approach is to let the centrality vary inversely with a measure of the total distance of a vertex from all others.

[10] A good centrality measure should be based on notions of 'status' or 'prestige' or 'rank.' They seek to capture the idea that the more central the neighbors of a vertex are, the more central that vertex itself is. These measures are inherently implicit in their definition and typically can be expressed in terms of eigenvector solutions of appropriately defined linear systems of equations. There are many such *eigenvector centrality* measures. The one used here is Bonacich centrality.

Vc. Exchange rate and price competition

The increasing spending per tourist in STP reflects the REER appreciation and can also be linked to high airfares. In order to evaluate this dynamic, consider the evolution of the real exchange rates; in particular, evaluate the relative trends in prices and exchange rates with respect to the euro-euro zone (RER_EURO), the US dollar-US (RER_US), and the British pound-UK (RER_POUND). This analysis is done for the 1999-2016 period for STP (Figure 18), for the five countries of comparison (Figure 19 to Figure 21), and the 2016 values for a larger comparison group (Figure 22 to Figure 24). All the figures indicate that STP's currency has appreciated the most in the relevant comparison group. Note, however, that the average spending per tourist includes the air ticket; therefore, a high expenditure in STP could be a reflection of comparably more expensive tickets to STP due to low airlift competition rather than actual high-spending tourists. As discussed above, a more detailed data set on how tourists spend their money would be needed to separate expenditures by categories.

Vd. Digital presence and Instagram posts

ICT and digital marketing allow countries to reach potential tourists in a more direct way and can be a boost for niche destinations. Many tourism destinations have taken to the social media and social networks to market their products, reaching markets previously inaccessible to them.[11] This can be particularly more intense in the type of tourism that STP wants to promote, "discovery" tourism, targeting travelers who are looking for lesser known destinations that offer attractions like nature, culture, etc. with strong emphasis on environmental protection and positive social impact.

The collected digital presence data proved to be very volatile, and only Belize and the Maldives showed some persistent improvement in this metric. Following the relatively new literature about the digital presence effects on tourism and other industries, consider now the Instagram posts. As discussed above, these correspond to daily data, aggregated on a monthly basis that mentions a given country. Figure 25 reports the evolution of posts for 2016 to the latest 2018 available data. The series are very volatile and, therefore, it is difficult to extract clear trends and patterns. This may signal that it is difficult to maintain a consistent digital presence or a great influence of one-off factors such as campaigns or events. From the data, some preliminary conclusions can be extracted. First, there is a significant increment in volume for all destinations (but only Belize and Maldives seem to have achieved some persistent digital presence). There is, however, a relative decline for the latest data collected. Second, STP has a considerable increment around 2018, when it reaches the numbers of other locations with many posts. This is not maintained for subsequent periods, however. Unfortunately, the frequency and nature of the data cannot be used to further investigate the effect of digital footprints on tourist arrivals or to relate measures. The digital presence cannot be used in the subsequent regression analysis because the tourism data set only has observations up to 2016, and, therefore, it overlaps with this data in only one year.

[11] The OECD Tourism Trends and Policies 2018 recognizes that the increased use of digital means is a recent trend: "*An increasing focus on digital strategies, with digital platforms opening new partnership opportunities and routes to market with reduced costs compared to traditional marketing approaches*".

VI. Concluding remarks

This note uses a multidimensional approach to estimate the determinants of international tourism and to produce a partial equilibrium analysis of the tourism sector in STP. The tourism industry is analyzed using different determinants highlighted in the literature as important. The main empirical findings, general and specific to STP, can be summarized as follows:

The first set of general findings confirms most of the results found in the literature. GDP in the issuing country and the exports of the host country are positively associated with tourism. Higher flight connectivity increases tourist arrivals but does not change expenditures per tourist. Tourist decisions are influenced by relative price variations, which means that a real exchange rate depreciation has a positive effect on arrivals and a negative effect on expenditures. Finally, tourist arrivals and expenditures per tourist have a persistent effect, meaning that increases in one year tend to be maintained in future years.

A causal analysis is performed to evaluate the effect of decriminalization of LGBT and large positive effects are observed. Using a difference-in-differences approach of a dummy variable that takes the value 1 if being LGBT is not a crime in the country and zero otherwise, the dummy is positively associated with tourist arrivals although there is no relationship with per capita spending.

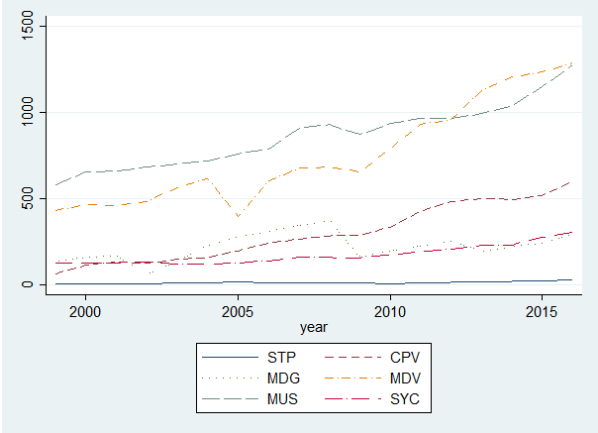
The data obtained on digital presence indicators are partial and from a short period, which limits the ability for meaningful statistical analysis. The preliminary data analysis using Instagram posts reveals high visibility for STP, however with great variability across time. As the literature speculates on the positive effects derived from digital presence, this should be encouraged at all levels. The empirical evidence for other destinations reveals that tourism is positively related to digital presence. As such, investing in digital visibility is likely to have a positive effect on the arrival of tourists.

Using a comprehensive list of determinants, the empirical analysis reveals that STP can do better in terms of the number of tourist arrivals, but it has a good performance in terms of expenditures per tourist. Using a hedonic model of tourist arrivals and expenditures per tourist, STP has a lower number of tourists in 2016 than what would be otherwise predicted with the model, but spending by tourists is larger than what the model finds.

The empirical analysis confirms the importance of improving air connectivity in the island. STP can receive more tourists because it has a low ratio of tourists-to-population and land area, however, flight demand meets supply as measured by seats sold on incoming flights. Therefore, the country needs to increase air connectivity. STP ranks low in centrality indexes in networks using flight data. Its major flight origin is Lisbon, which means that many of the European tourists need to fly to Lisbon first. As connections reduce utility and propensity to spend at the destinations, STP might benefit greatly from having connections to central hubs. A simple exercise simulating a connection to London reveals major changes in centrality. The tourism elasticity with respect to the seat availability on the flights is of the order of 0.1.

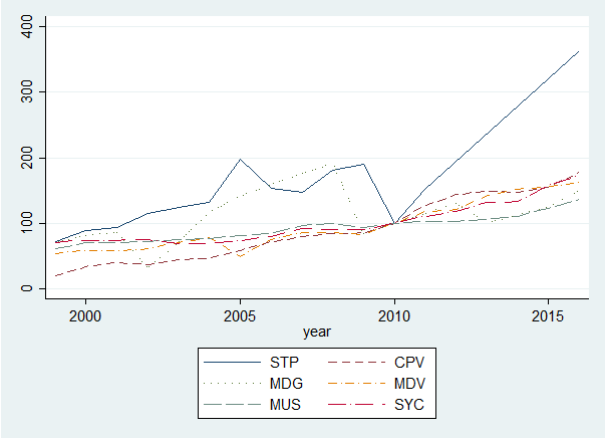
Price effects on tourist arrivals and expenditures per tourist confirm that a pricing strategy, which focuses on high-spending tourism is appropriate. Since the arrivals of tourists have a high persistence, the long-term effects are large. STP recently had a trend of currency appreciation with respect to the euro and US dollar. The latest results confirm that this has been beneficiary for total expenditures in tourism as measured in real terms.

Figure 1: Total Arrivals – thousands of tourist arrivals: 1999-2016



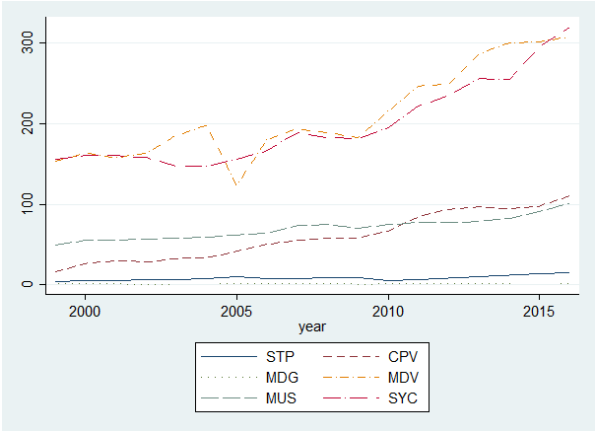
Source: Author´s calculations using UNWTO.

Figure 2: Total Arrivals (2010=100): 1999-2016



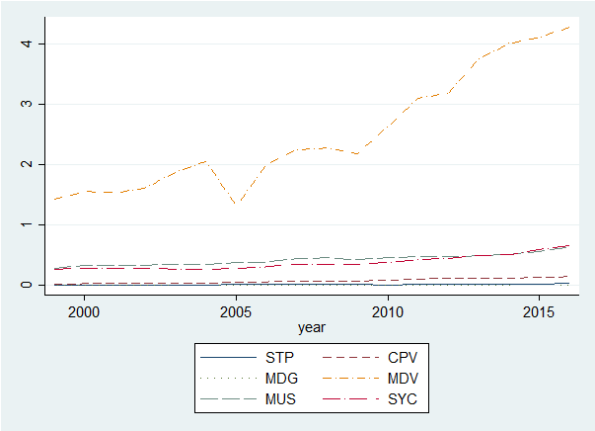
Source: Author´s calculations using UNWTO.

Figure 3: Total Arrivals per Population: 1999-2016



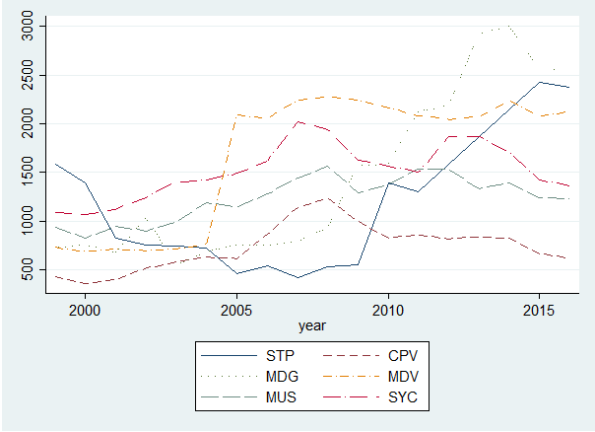
Source: Author´s calculations using UNWTO and WDI.

Figure 4: Total Arrivals per Land: 1999-2016



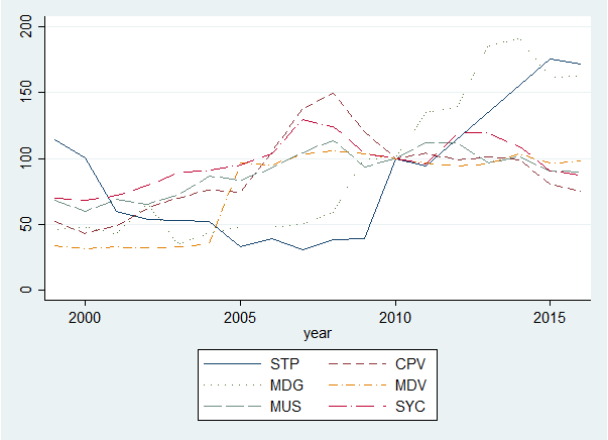
Source: Author´s calculations using UNWTO and WDI.

Figure 5: Expenditures per arrival (US dollars)): 1999-2016



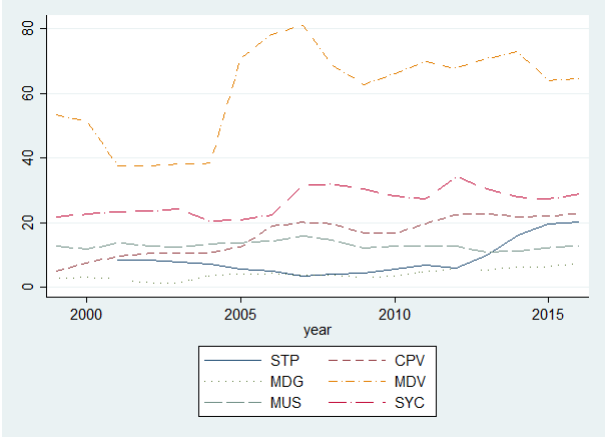
Source: Author´s calculations using UNWTO and WDI.

Figure 6: Expenditures per arrival (2010=100, US dollars): 1999-2016



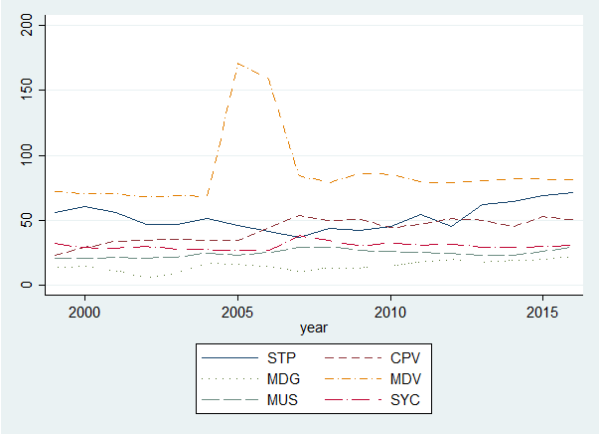
Source: Author´s calculations using UNWTO and WDI.

Figure 7: Tourism receipts / GDP: 1999-2016



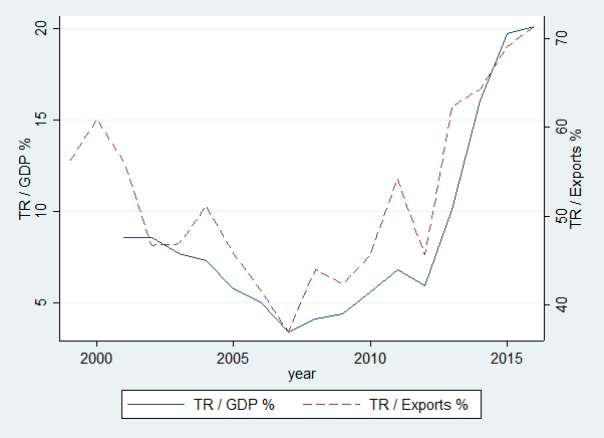
Source: Author´s calculations using UNWTO and WDI.

Figure 8: Tourism receipts / Exports: *1999-2016*



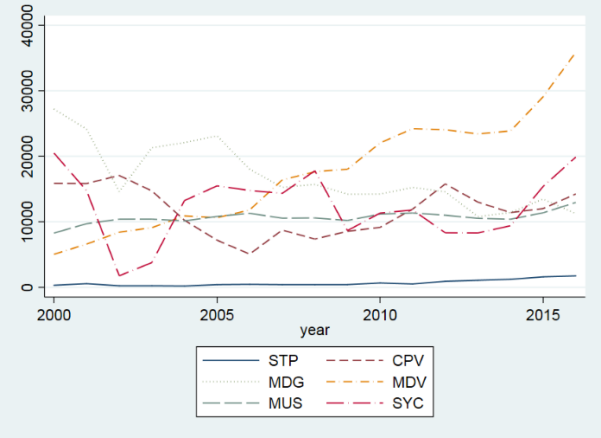
Source: Author´s calculations using UNWTO and WDI.

Figure 9: Tourism receipts / GDP and Tourism receipts /Exports: 1999-2016



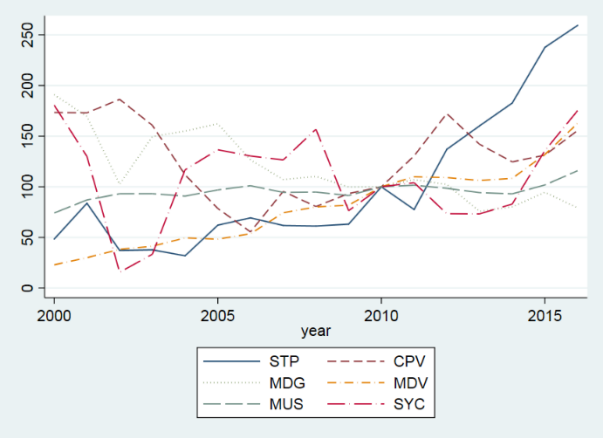
Source: Author´s calculations using UNWTO and WDI.

Figure 10: Total flights arrivals: 2000-2016



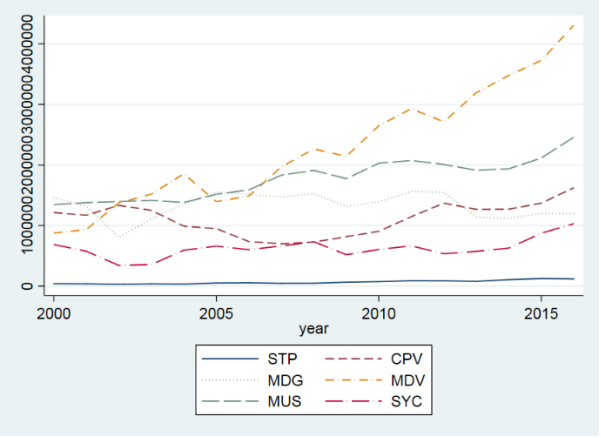
Source: Author´s calculations using World Bank data.

Figure 11: Total flights arrivals (2010=100): 2000-2016



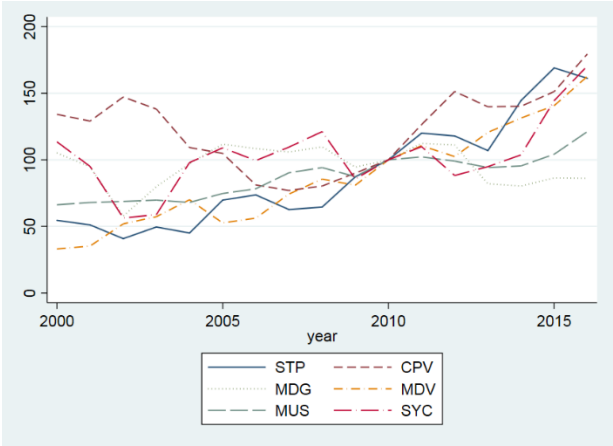
Source: Author´s calculations using World Bank data.

Figure 12: Total seats arrivals: 2000-2016



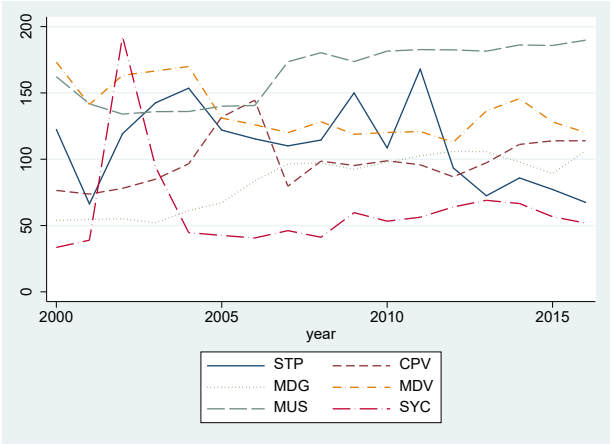
Source: Author´s calculations using World Bank data.

Figure 13: Total seats arrivals (2010=100): 2000-2016



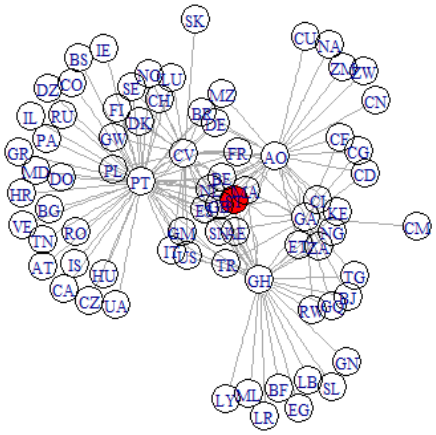
Source: Author´s calculations using UNWTO and WDI.

Figure 14: Seats / Flights: 2000-2016



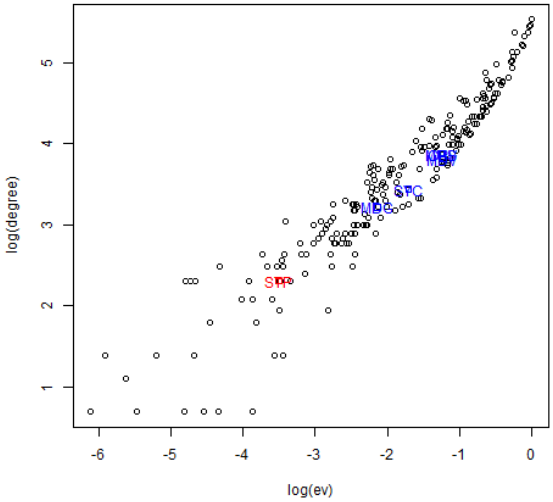
Source: Author´s calculations using World Bank data.

Figure 15: Flights connection network structure of STP: 2016



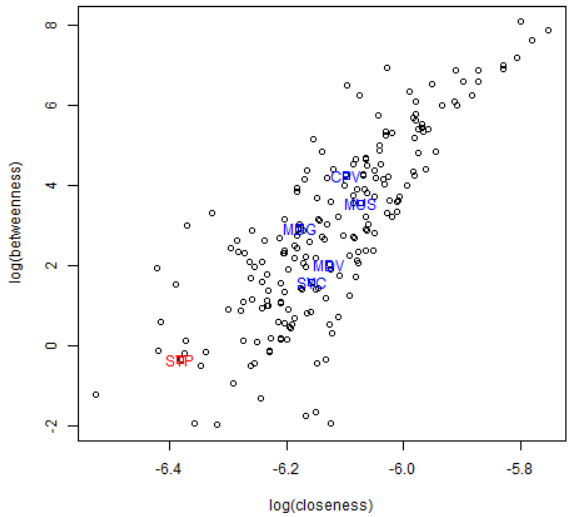
Source: Author´s calculations using World Bank data.

Figure 16: Degree and eigenvalue vector centrality (in logs): 2016



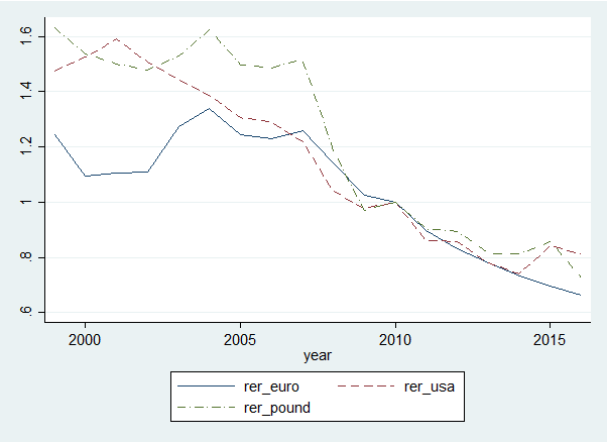
Source: Author´s calculations using World Bank data.

Figure 17: Betweenness and Closeness: 2016



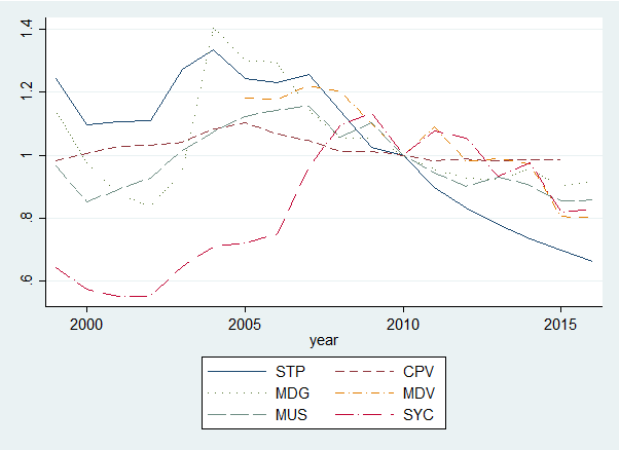
Source: Author´s calculations using World Bank data.

Figure 18: Real exchange rate: 1999-2016



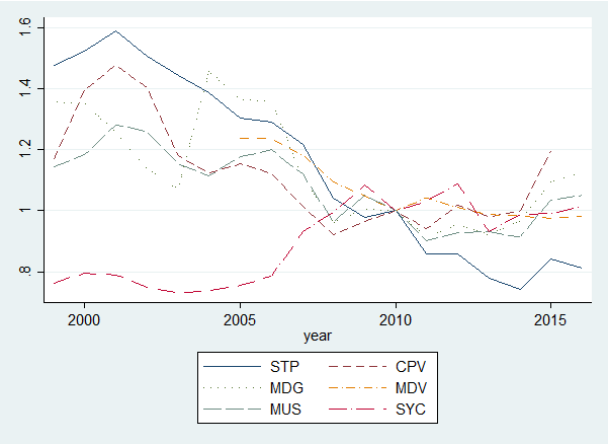
Source: Author´s calculations using UNWTO and WDI.

Figure 19: Real exchange rate (wrt Euro): 1999-2016



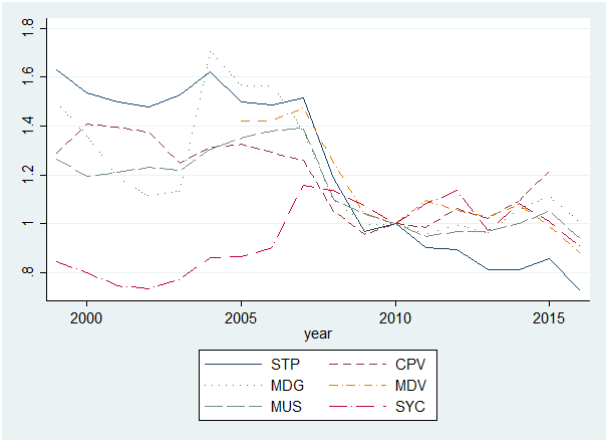
Source: Author´s calculations using UNWTO and WDI.

Figure 20: Real exchange rate (wrt US Dollar): 1999-2016



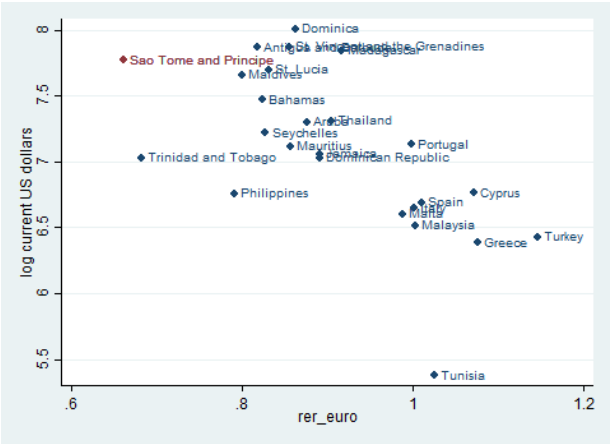
Source: Author´s calculations using UNWTO and WDI.

Figure 21: Real exchange rate (wrt British Pound): 1999-2016



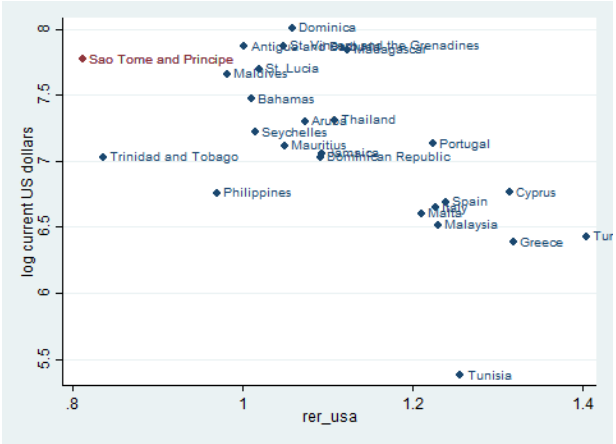
Source: Author´s calculations using UNWTO and WDI.

Figure 22: Real exchange rate (wrt Euro): *2016*



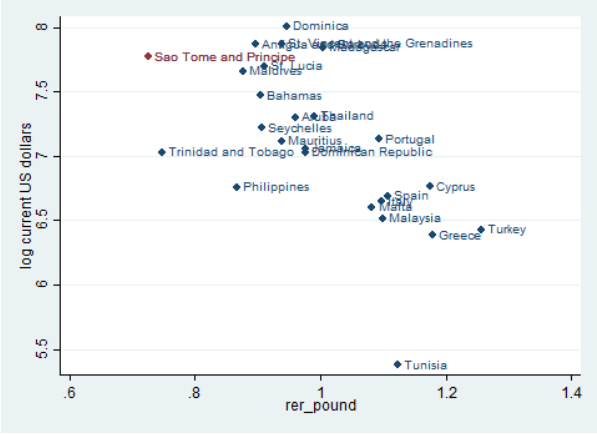
Source: Author´s calculations using UNWTO and WDI.

Figure 23: Real exchange rate (wrt US Dollar): *2016*



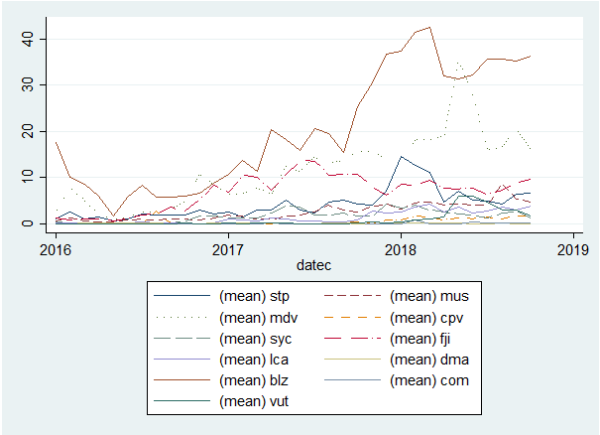
Source: Author´s calculations using UNWTO and WDI.

Figure 24: Real exchange rate (wrt British Pound): *2016*



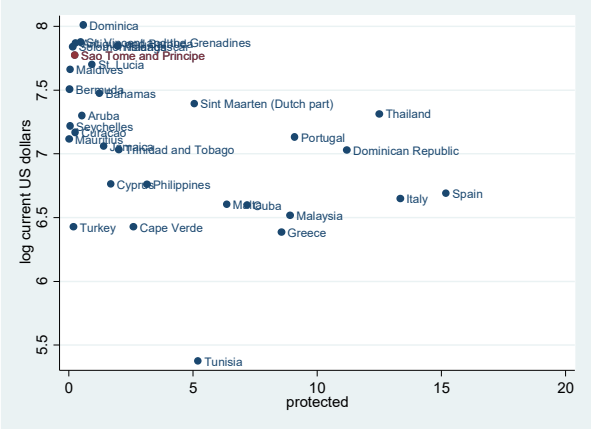
Source: Author´s calculations using UNWTO and WDI.

Figure 25: Instagram posts: 01/2016 - 10/2018



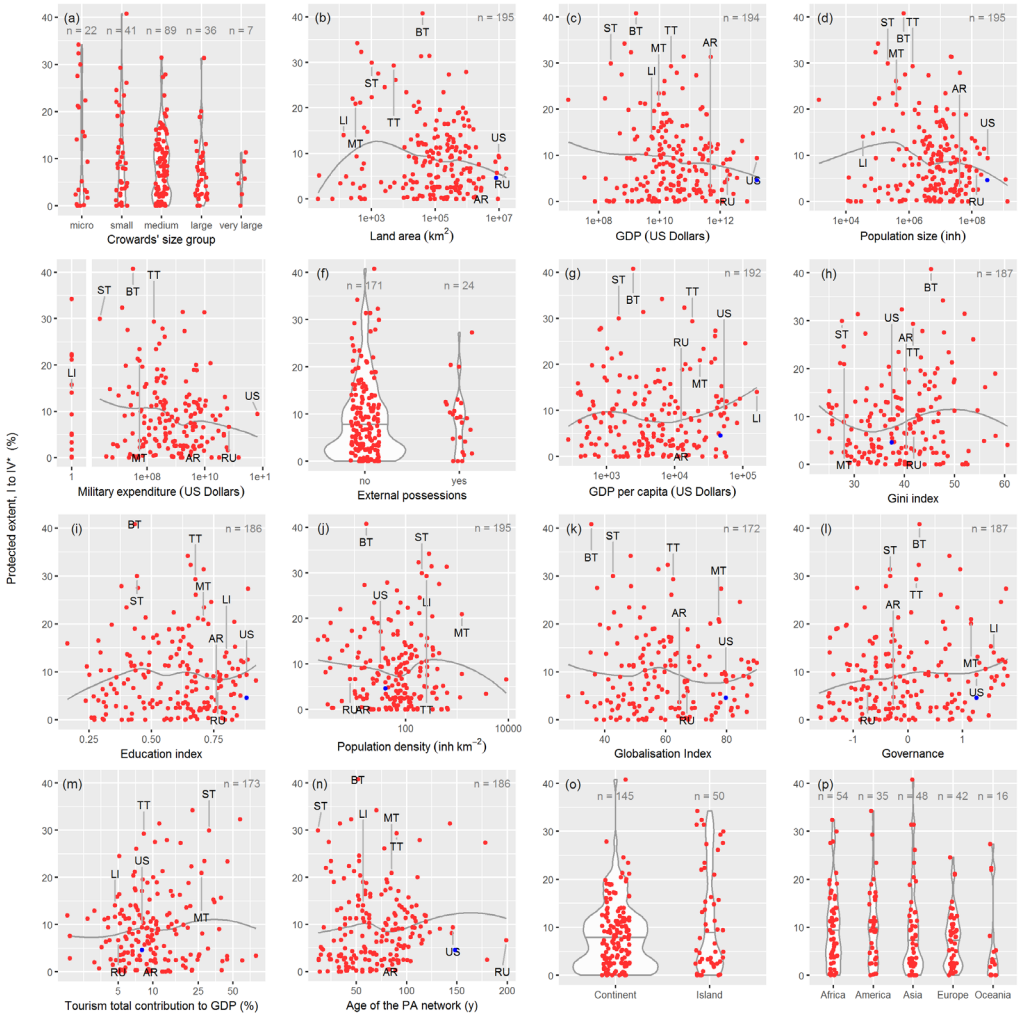
Source: Author´s calculations using World Bank data.

Figure 26: Protected areas: 2016



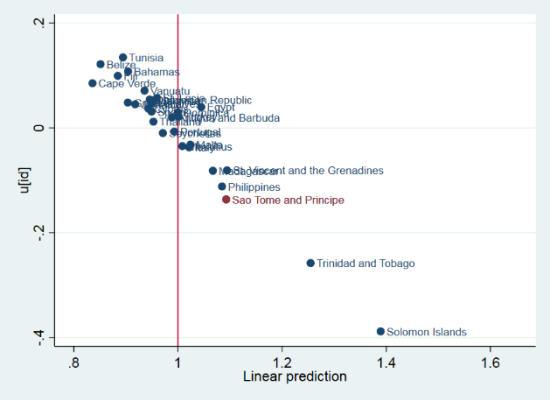
Source: Author´s calculations using UNWTO and WDI.

Figure 27: Protected area and other measures from Baldi et al. (2017)



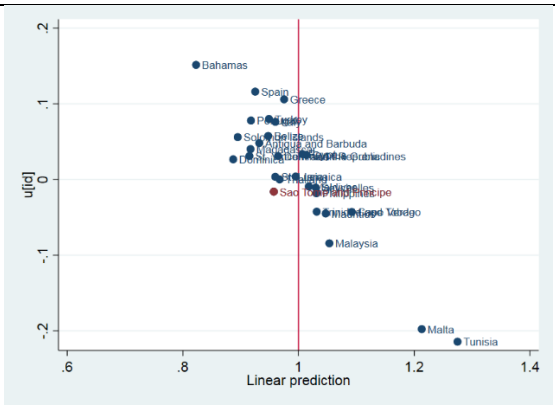
Notes: ST is STP. Source: Baldi et al. (2017).

Figure 28: Linear prediction and unobservable component in the hedonic model for tourism arrivals: 2016



Source: Author´s calculations using World Bank data. Horizontal axis contains the value of the linear prediction of the hedonic model divided by the actual value of LN_ARR in 2016. Vertical axis is the value of the unobservable country-specific component divided by the actual value of LN_ARR in 2016.

Figure 29: Linear prediction and unobservable component in the hedonic model for expenditures per tourist arrivals: 2016



Source: Author´s calculations using World Bank data. Horizontal axis contains the value of the linear prediction of the hedonic model divided by the actual value of LN_EA in 2016. Vertical axis is the value of the unobservable country-specific component divided by the actual value of LN_EA in 2016.

Table 1: Tourist arrivals and contribution to GDP in selected countries, latest year available

Country	Tourist Arrivals	Tourist Arrivals (% of population)	Direct Contribution to GDP (% of GDP)
Maldives	1,286,000	300.6%	39.6
Seychelles	303,000	320.0%	26.4
The Bahamas	1,482,000	378.8%	19.0
Vanuatu	95,100	35.2%	18.2
Cabo Verde	598,000	110.8%	17.8
St. Lucia	348,000	195.5%	15.0
Belize	386,000	105.2%	15.0
Fiji	792,000	88.1%	14.4
Antigua & Barbuda	265,000	262.5%	13.0
Barbados	632,000	221.8%	13.0
Dominica	78,000	106.1%	12.4
São Tomé and Príncipe	**29,000**	**14.5%**	**10.8**
Jamaica	2,182,000	75.7%	10.3
Iceland	1,792,000	534.2%	8.5
Mauritius	1,275,000	100.9%	7.4

Source: UNWTO and WTTC

Table 2: Flights and Seats, airlift connectivity: 2000-2016

	2000		2005		2010		2016	
	Flights	Seats	Flights	Seats	Flights	Seats	Flights	Seats
Angola	136	17408	122	14884	238	34082	241	28920
Cameroon			11	1705				
Cape Verde			29	3538	105	14070	1	120
Gabon	96	3552	70	2870	20	940	176	4817
Ghana			74	9028			165	26730
Côte d'Ivoire	48	9600						
Nigeria			11	1705				
Portugal	48	9600	104	17614	105	20728	238	39067
São Tomé and Príncipe					210	3780	938	18996

Source: Author´s calculations.

Table 3: **Baseline Econometric analysis for Arrivals Determinants:** 2000-2016

VARIABLES	(1)	(2)
Dep. Var.	**LN_ARR_t**	**LN_EA_t**
LN_GDP_{t-1}	0.550***	-0.731*
	(0.158)	(0.377)
LN_POP_{t-1}	0.139	0.463
	(0.206)	(0.331)
LN_X_{t-1}	0.242**	0.618***
	(0.0969)	(0.181)
LN_SEATS_{t-1}	0.129**	-0.166
	(0.0649)	(0.150)
RER_EURO_{t-1}	0.604***	-0.598**
	(0.178)	(0.261)
$LN\_GDPEURO_{t-1}$	-0.507***	0.655***
	(0.158)	(0.199)
LN_LAND	-0.226*	-0.260
	(0.132)	(0.174)
Observations	542	534
Number of id	36	35

Source and notes: Author´s calculations. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include yearly dummies, not reported. Random-effects model.

Table 4: Econometric Analysis for Arrivals Determinants: 2000-2016

VARIABLES	FE	Sys-GMM	FE	Sys-GMM	FE	Sys-GMM
Dep.var.: LN_ARR_{t-1}						
LN_ARR_{t-1}	0.777***	0.768***	0.778***	0.772***	0.777***	0.740***
	(0.0321)	(0.0934)	(0.0321)	(0.0919)	(0.0321)	(0.0853)
LN_GDP_{t-1}	0.00889	0.149***	-0.00592	0.144***	0.00416	0.168***
	(0.0528)	(0.0562)	(0.0522)	(0.0552)	(0.0524)	(0.0631)
LN_POP_{t-1}	0.0940	-0.0611**	0.101	-0.0595**	0.0947	-0.0694**
	(0.169)	(0.0302)	(0.170)	(0.0297)	(0.169)	(0.0308)
LN_X_{t-1}	0.0782**	0.0584	0.0779**	0.0597	0.0780**	0.0677
	(0.0310)	(0.0495)	(0.0311)	(0.0488)	(0.0310)	(0.0470)
LN_SEATS_{t-1}	-0.00387	0.0484	-0.00161	0.0458	-0.00265	0.0509
	(0.0246)	(0.0483)	(0.0245)	(0.0483)	(0.0245)	(0.0420)
RER_EURO_t	0.116*	0.110*				
	(0.0662)	(0.0644)				
RER_USA_t			0.0737	0.0920*		
			(0.0555)	(0.0520)		
RER_POUND_t					0.0877	0.0803
					(0.0536)	(0.0506)
Observations	537	537	537	537	537	537
R-squared	0.811		0.811		0.811	
Number of id	36	36	36	36	36	36
Long-run effect	0.520	0.475	0.332	0.404	0.394	0.309
p-value =0	0.0871	0.0214	0.190	0.0199	0.109	0.138
p-value =1	0.114	0.0109	0.00826	0.000596	0.0137	0.000893
Arellano-Bond AR(2) stat		-0.309		-0.324		-0.322
AR(2) p-value		0.757		0.746		0.747
Hansen stat		16.59		16.47		16.24
Hansen p-value		0.219		0.225		0.237
# IV		37		37		37

Source and notes: Author´s calculations. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include yearly dummies, not reported. FE: fixed-effects model. Sys-GMM: Blundell and Bond (1998) System GMM model using Roodman (2009) collapsing method.

Table 5: Econometric Analysis for expenditures per Arrival Determinants: 2000-2016

VARIABLES	FE	Sys-GMM	FE	Sys-GMM	FE	Sys-GMM
Dep.var.: LN_EA_{t-1}						
LN_EA_{t-1}	0.712***	0.807***	0.715***	0.807***	0.714***	0.807***
	(0.0334)	(0.0196)	(0.0334)	(0.0198)	(0.0334)	(0.0198)
LN_GDP_{t-1}	-0.298***	0.00196	-0.274***	0.00217	-0.281***	0.00206
	(0.0871)	(0.0259)	(0.0859)	(0.0259)	(0.0864)	(0.0259)
LN_POP_{t-1}	0.282	-0.0254	0.275	-0.0255	0.276	-0.0253
	(0.278)	(0.0172)	(0.279)	(0.0171)	(0.279)	(0.0171)
LN_X_{t-1}	0.0560	-0.0238	0.0530	-0.0238	0.0549	-0.0237
	(0.0547)	(0.0219)	(0.0549)	(0.0218)	(0.0548)	(0.0218)
LN_SEATS_{t-1}	-0.0626	0.0489***	-0.0658	0.0488***	-0.0659	0.0485***
	(0.0407)	(0.0184)	(0.0407)	(0.0186)	(0.0406)	(0.0185)
RER_EURO_t	-0.309***	-0.131**				
	(0.108)	(0.0617)				
RER_USA_t			-0.226**	-0.109*		
			(0.0906)	(0.0592)		
RER_POUND_t					-0.227***	-0.103*
					(0.0875)	(0.0540)
Observations	530	530	530	530	530	530
R-squared	0.664		0.662		0.663	
Number of id	35	35	35	35	35	35
Long-run effect	-1.074	-0.678	-0.793	-0.567	-0.796	-0.533
p-value =0	0.00555	0.0296	0.0152	0.0549	0.0116	0.0489
p-value =1	0.849	0.301	0.527	0.143	0.517	0.0846
Arellano-Bond AR(2) stat		1.407		1.406		1.409
AR(2) p-value		0.160		0.160		0.159
Hansen stat		11.29		11.32		11.30
Hansen p-value		0.587		0.584		0.586
# IV		37		37		37

Source and notes: Author´s calculations. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include yearly dummies, not reported. FE: fixed-effects model. Sys-GMM: Blundell and Bond (1998) System GMM model using Roodman (2009) collapsing method.

Table 6: Difference-in-differences estimator of LGBT no criminalization: 2000-2016

VARIABLES	(1)	(2)	(3)	(4)
Dep. Var.	LN_ARR_t	LN_ARR_t	LN_EA_t	LN_EA_t
No criminalization LGBT	0.431**	0.273**	0.167	0.00475
	(0.156)	(0.0974)	(0.192)	(0.120)
Observations	251	228	246	225
R-squared	0.534	0.616	0.476	0.607
Number of id	15	15	15	15
Year dummies	YES	YES	YES	YES
Country dummies	YES	YES	YES	YES
Controls	NO	YES	NO	YES

Source and notes: Author´s calculations. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include yearly dummies, not reported.

References

Acevedo, S., Han, L., Kim, H.S. and Laframboise, N. (2016). Flying to Paradise: The Role of Airlift in the Caribbean Tourism Industry. IMF Working Papers 16/33.

Anderson, T.W. and Hsiao, C. (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association* 76, 598-606.

Anderson, T.W. and Cheng Hsiao (1982). Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics* 18, 47-82.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies, 58(2)*, 277-297.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variable Estimation of Error-Component Models. *Journal of Econometrics* 68, 29-51.

Balaguer, J. and Cantavella-Jordá, M. (2002). Tourism as a long-run economic growth factor: the Spanish case. *Applied Economics* 34, 877–884.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87(1), 115-143.

Nickel, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica* 49, 1417-1426.

Brau, R., Lanza, A. and Pigliaru, F. (2003). How fast are the tourism countries growing? The cross-country evidence. *CRENoS Centro Ricerche Economiche Nord Sud, Working Paper N0O3-09*. http://www.crenos.it/working/pdf/03-09.pdf.

Baldi, G., Texeira, M., Martin, O.A., Grau, H.R. and Jobbágy, E. (2017) Opportunities drive the global distribution of protected areas. PeerJ 5:e2989; DOI 10.7717/peerj.2989.

Chan, N. L. and Denizci Guillet, B. (2011). Investigation of social media marketing: How does the hotel industry in Hong Kong perform in marketing on social media websites? *Journal of Travel & Tourism Marketing*, *28*(4), 345–368.

Coscia, M., Hausmann, R., and Neffke, F. (2016). Exploring the Uncharted Export: An Analysis of Tourism-Related Foreign Expenditure with International Spend Data. *CID Faculty Working Paper No. 328.*

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

Culiuc, A. (2014). Determinants of International Tourism. *IMF Working Paper No. 14/82*.

Harvey, A.C. and Todd, P. H. J. (1983). Forecasting economic time series with structural and Box-Jenkins models: A case study. *Journal of Business & Economic Statistics,* 1(4), 299-307.

Gustavsson, P. and Nordström, J. (2001). The impact of seasonal unit roots and vector ARMA modelling on forecasting monthly tourism flows. *Tourism Economics,* 7(2), 117-133.

Kim, J. H. and Moosa, I. A. (2005). Forecasting international tourist flows to Australia: A comparison between the direct and indirect methods. *Tourism Management,* 26(1), 69-78.

Laframboise, N., Mwase, N., Park, J. and Zhou, Y. (2014). Revisiting Tourism Flows to the Caribbean: What is Driving Arrivals? *IMF Working Paper* 14/229.

Dritsakis, N. (2004). Tourism as a long-run economic growth factor: an empirical investigation for Greece using causality analysis. *Tourism Economics* 10, 305–316.

Durbarry, R. (2004). Tourism and economic growth: the case of Mauritius. *Tourism Economics* 10, 389–401.

Gunduz, L. and Hatemi-J, A. (2005). Is the tourism-led growth hypothesis valid for Turkey? *Applied Economics Letters* 12, 499–504.

Leung, D., Law, R., van Hoof, H. and Buhalis, D. (2013). Social Media in Tourism and Hospitality: A Literature Review. Journal of Travel & Tourism Marketing, 30(1-2), 3-22.

Lederman, D. and Maloney, W.F. (2007). *Natural Resources: Neither Curse nor Destiny*. Washington, DC: World Bank; Palo Alto, CA: Stanford University Press.

Li, J., Xu, L., Tang, L., Wang, S. and Li, L. (2018). Big data in tourism research: A literature review. Tourism Management, 68, 301-323.

McElroy, J.L. and Parry, C. (2010). The characteristics of small island tourist economies. Tourism and Hospitality Research, 10(4), 315-328.

Maloney, W. and Montes Rojas, G. (2005). How elastic are sea, sand and sun? Dynamic panel estimates of the demand for tourism. *Applied Economics Letters*, 12, 277–280.

Noone, B. N., McGuire, K. A. and Rohlfs, K. V. (2011). Social media meets hotel revenue management: Opportunities, issues and unanswered questions. *Journal of Revenue and Pricing Management*, *10*(4), 293–305.

Oh, C. (2005). The contribution of tourism development to economic growth in the Korean economy. *Tourism Management*, 26, 39–44.

Rosensweig, J. (1988). Elasticities of substitution in Caribbean tourism. *Journal of Development Economics,* 29, 89–100.

Sequeira, T.N., Campos, C. (2006). International tourism and economic growth: A panel data approach. In: Matias, A., Nijkamp, P., Neto, P. (Eds.), *Advances in Modern Tourism Research*, 153–163.

Sinclair, M.T. (1998). Tourism and economic development: a survey. *Journal of Development Studies,* 34, 1–51.

Tosun, C. (1999). An analysis of contributions international inbound tourism to the Turkish economy. *Tourism Economics,* 5, 217–250.

Wang, Y. and Fesenmaier, D. R. (2004). Modelling participation in an online travel community. *Journal of Travel Research*, *42*(3), 261–270.

Wolfe, A. and Romeu, R. (2011). Recession and Policy Transmission to Latin American Tourism: Does Expanded Travel to Cuba Offset Crisis Spillovers? *IMF Working Paper No*.11/32.

Eagles, P. F. J., et al. (2002). Sustainable tourism in protected areas: Guidelines for planning and management. *Best Practice Protected Area Guidelines Series,* 8. A.

Garín Muñoz, T. and Pérez Amaral, T. (2000). An econometric model for international tourism flows to Spain. *Applied Economics*, 7, 525–9.

Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research*, 24, 835–49.

Papatheodorou, A. (1999). The demand for international tourism in the Mediterranean region. *Applied Economics*, 31, 619–30.

McKercher, B. (1996). Differences between Tourism and Recreation in Parks. *Annals of Tourism Research*, 23(3), 563-575.

Mulholland, G. and Eagles P.F.J. (2002). African parks: combining fiscal and ecological sustainability. *Parks*, 12, 42-49.

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9(1), 86-136.

Rosensweig, J. (1988). Elasticities of substitution in Caribbean tourism. *Journal of Development Economics*, 29, 89–100.

Standing C., Tang Taye, J.P and Boyer, M. (2014). The Impact of the Internet in Travel and Tourism: A Research Review 2001–2010. *Journal of Travel & Tourism Marketing,* 31(1), 82-113.

Song, H., Romilly, P. and Liu, X. (2000). An empirical study of outbound tourism demand in the UK. *Applied Economics*, 32, 611–24.

Siikamäki, P., Kangas, K., Paasivaara, A. and Schroderus, S. (2015). Biodiversity attracts visitors to national parks. *Biodiversity and Conservation,* 24(10), 2521-2534.

Tolvanen, A. and Kangas, K. (2016). Tourism, biodiversity and protected areas - Review from northern Fennoscandia*. Journal of Environmental Management*, 169, 58-66.

UNEP-WCMC (2018). Protected Area Profile for São Tomé and Príncipe from the World Database of Protected Areas. Available at: www.protectedplanet.net

Wilkie, D. S. and Carpenter, J. (1999). Can nature tourism help finance protected areas in the Congo Basin? *Oryx*, 33(4), 333-339.

World Economic Forum (2017). *Travel & Tourism Competitiveness Report 2017*, Geneva.

Zeng, B. (2013). Social Media in Tourism. *Journal of Tourism Hospitality,* 2(1).

Zeng, B. and Gerritsen, R. (2014). What do we know about social media in tourism? A review. *Tourism Management Perspectives*, 10, 27–36.