



Common unofficial language, development and international tourism

Luke Emeka Okafor^{*}, Usman Khalid, Terence Then

School of Economics, University of Nottingham, Malaysia Campus, Faculty of Arts and Social Sciences, Malaysia



HIGHLIGHTS

- Common unofficial language has a significant positive impact on international tourism flows.
- Common unofficial language is a more important determinant of international tourism in Europe.
- Common official and unofficial languages promote international tourism in other regions, such as America, Asia etc.
- Regardless of the levels of development of different countries, common unofficial language promotes international tourism.
- Landlocked variable is income and region dependent.

ARTICLE INFO

Article history:

Received 21 October 2017

Received in revised form

2 January 2018

Accepted 9 January 2018

Keywords:

Tourism
Gravity model
Language
Development
Region

ABSTRACT

We employ a gravity framework to examine whether the use of common unofficial language promotes international tourist flows while considering the influence of the levels of development and regions in the underlying relationship. The empirical analysis is based on a panel data set of bilateral tourism flows among 200 countries over the period 1995 to 2015. Results show that common unofficial language is a significant determinant of international tourist flows after controlling for common official language and other classical determinants of tourist flows. This finding holds irrespective of the levels of development of different countries. Further, we show that a common unofficial language is a more significant determinant of international tourist flows than a common official language in Europe. Policies that create an enabling environment for multilingual societies to emerge in a country would help to boost international tourism.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

The tourism industry is growing at an unprecedented rate with 1.2 billion tourists crossing international borders in the year 2016. The international tourism receipts grew by approximately 15% per year from US\$ 264 billion in 1990 to US\$1, 260 billion in 2015. The World Travel and Tourism Council (WTTC) estimates that travel and tourism account for 10.2 percent of global GDP and employs 292 million people, which is 1 in 10 jobs on the planet. International tourism accounts for 7% of the world's exports in goods and services, ranking third in worldwide export after fuels and chemicals, but ahead of food and automotive products (World Tourism Organization, 2017a). This clearly underscores the importance of international tourism as a major source of income and cross-

country linkages.

Several factors have contributed to growth in international tourism such as urbanisation, increasing population, higher education level, the availability of leisure time, rising levels of income and decreasing costs of travelling (Jud, 1974). The literature is replete with studies that explore the determinants of demand for tourism and tourist flows (Crouch, 1994; Li, Song, & Witt, 2005; Lim, 1997, 1999; Song & Li, 2008). In nearly all tourism demand studies to date the focus has been on economic factors explaining differences in tourism flows; primarily income as captured by GDP per capita, changes in relative prices, transportation costs and exchange rates (Lim, 1997; Zhang & Jensen, 2007). Although income, prices and other economic variables play a vital role in determining the demand for tourism, tourism literature identifies a handful of other variables that potentially affect the demand for tourism. Amongst these variables, social variables such as cultural proximity and common language are at the top of the list.

Language plays an important role in international tourism, as it may enhance the pleasantness of a vacation, or it can act as a

^{*} Corresponding author. School of Economics, University of Nottingham, Malaysia Campus, Faculty of Arts and Social Sciences, Jalan Broga, 43500 Semenyih, Selangor, Malaysia.

E-mail address: Luke.Okafor@nottingham.edu.my (L.E. Okafor).

barrier. Similar to several aspects of consumer demand, attitudes and beliefs may also influence tourism demand (Vietze, 2012). Therefore, tourists being consumers would prefer to visit a certain destination where they believe they can easily derive satisfaction without much effort. To this end, language proximity plays a key role as sharing a common language would decrease the transaction costs of international tourism. Cross-cultural interaction is an integral part of international tourism. The tourist must integrate into a culturally distinct environment in which he or she will react with different degrees of comfort and enthusiasm (Kastenholz, 2010). Therefore, the satisfaction derived from the tourism experience is very much contingent upon the cross-cultural interaction, which is potentially facilitated by cultural and language proximity (Kastenholz, 2010).

Evidence on the link between common unofficial language and international tourism is scant. While most previous studies include common language as one of the relevant determinants of tourist flows, limited attention has been paid to the role played by common unofficial language. A quick glance at the dyadic data reveals that only 17.4% out of the 50176 country pairs share an official language. This means that typical tourist choices are limited if the decision to travel is based on a pair of countries sharing a common official language. In this scenario, a common unofficial language could influence the decision to visit a place or holiday resort thereby increasing the choice set. Against this backdrop, it is important to assess the role played by common unofficial language in determining the international tourist flows.

An isolated analysis of common unofficial language per se, however, may grossly simplify the mechanisms at work. It is likely that countries in different parts of the world and at various levels of development may have different drivers of tourist flows. Therefore, a tourist might choose a certain destination based on the level of economic development of a country or geographical location of a country. We therefore assess the role of common unofficial language in determining international tourist flows in different parts of the world and for countries at different levels of development. As pointed out earlier, the literature analyzing the impact of common unofficial language on tourist flows is sparse. There is no evidence that any previous study has explored the impact of common unofficial language on international tourism while considering the influence of the levels of development and geography in the underlying relationship.¹

The purpose of this study to examine whether the use of common unofficial language promotes international tourist flows while considering the influence of the levels of development and regions in the underlying relationship. Our contributions to the strand of literature that relates to international tourism are fourfold. First, we show that international tourists are not only strongly attracted to destinations that share the same common official language with their originating countries, but also to destinations that share the same common unofficial language.

Second, we demonstrate that this result holds irrespective of the levels of development of different countries. Third, we show that a common unofficial language is a more important determinant of international tourism compared to common official language in Europe. Several European countries do not share a common official language, and this potentially explains why common unofficial language provides a significant boost to tourism in Europe compared with common official language. Lastly, the results show that the effect of landlocked variable is income and region dependent. For instance, landlocked variable has a negative impact on

international tourism in high-income countries, especially European countries, whereas its impact on upper middle-income, lower middle-income and low-income countries or regions apart from Europe is mixed.

Several studies have analyzed the determinants of bilateral tourism. Eilat and Einav (2004) show that political risk is a crucial determinant of tourism, while the exchange rate is a crucial driver of tourism in advanced economies. Gil-Pareja, Llorca-Vivero, and Mart'inez-Serrano (2006, 2007) use gravity framework to analyse the role of embassies and common currency on tourism flows. They find that embassies and consulates have a substantial positive impact on tourism flows. Further, they show that common currency; namely, euro provides a significant boost to international tourism.

In general, several studies examine the link between tourism flows and a range of variables. For instance, Santana, Ledesma, and Pérez (2010) and Santana, Ledesma, Pérez, and Cortés (2010) examine the impact of different exchange rate regimes on tourism flows. Fourie and Santana (2011) employ a gravity equation to estimate the effect of mega-events on international tourism, whereas Fourie and Santana (2013) analyzes the effect of cultural affinity and ethnic reunion on international tourism using a gravity model. Other studies have used a gravity model to explore the implications of taxes and transport infrastructure on tourism (see, for example, Durbarry (2008); Khadaroo & Seetanah, 2008). Similarly, Neumayer (2010) uses gravity framework to assess the impact of visa restrictions on international tourism flows. In addition, Vietze (2012) analyzes the effect of the religious association on U.S. tourist arrivals, while Massidda and Etzo (2012) evaluate the impact of several variables such as price differences, expenditure, cultural activities and crime rates, among others, on domestic tourism in Italy.

The rest of this paper is structured as follows: Sections 2 and 3 explain the methodology used in this research, which encompasses the choice of econometric models, and the description of dependent and independent variables used in the empirical analysis. Section 4 discusses the findings of the paper. Section 5 concludes the paper.

2. Data and methodology

2.1. Theoretical rationale

Gravity models have been used extensively in empirical work, particularly in the field of international trade (Gil-Pareja, Llorca-Vivero, & Mart'inez-Serrano, 2007; Khadaroo & Seetanah, 2008; Tinbergen, 1962). International tourism is a form of international trade in services and thus, an augmented gravity approach is an appropriate methodology to adopt (Khadaroo & Seetanah, 2008).

The gravity model of international trade is built upon Newton's law of universal gravitation as follows:

$$F = G \frac{m_1 m_2}{r^2} \quad (1)$$

F represents the gravitational force between two masses, being directly proportional to the first mass (m_1) and second mass (m_2), and negatively proportional to the square of the distance between the masses (r^2). G is the gravitational constant.

An analogy to international trade would be that the amount of trade between two economies is directly proportional to the economic sizes and inversely proportional to the distance between them. The gravity equation below includes an error term, e_{ij} , which is log-normally-distributed as follows:

¹ The only study that uses common unofficial language as a determinant of tourist flows is Culiuc (2014).

$$F_{ijt} = G^{\beta_0} \frac{m_i^{\beta_1} m_j^{\beta_2}}{r_{ij}^{\beta_3}} e_{ij} \quad (2)$$

A gravity model of international trade is employed to model international tourism flows. The independent variables used in the analysis are the baseline gravity variables augmented with other determinants of international tourism flows. Population is used to capture the economic size. It is an appropriate variable to capture economic size for studies dealing with international tourism or migration (Gil-Pareja et al., 2007). We also include real GDP per capita at purchasing power parity as individuals from richer countries are more likely to demand tourism services or face a lower barrier to international migration (Gil-Pareja et al., 2007). It also helps to mitigate the influence of price effects in the underlying relationship.

To capture the influence of the levels of development and regions in the underlying relationship, the sample is disaggregated into different income groups and different regions. GDP per capita is used as a proxy for the level of development. Similar to World Bank's new country classifications by income level for the fiscal year 2018, we classify low-income countries as those with a GDP per capita of \$1005 or less in 2011, lower middle-income countries as those with a GDP per capita between \$1006 and \$3955 in 2011, upper middle-income countries as those with a GDP per capita between \$3956 and \$12235 in 2011, and high-income countries as those with a GDP per capita greater than \$12235 in 2011 (World Bank, 2017a,b). To ensure that the estimates are not sensitive to the year used as a benchmark for the classifications, we reclassify countries into different income groups using the fiscal year 2015 in lieu of the fiscal year 2011. The disaggregation of the sample into groups is performed separately for the origin and destination countries.

Further, the sample is disaggregated into different groups based on regions, namely Europe, America and Latin America and the Caribbean (LATCA), Asia, Middle East, North Africa and Oceania, and Sub-Saharan Africa (SSA). Similar to income groups, the disaggregation into different regions is performed separately for the originating and destination countries.

The independent variables consist of different fixed or varying characteristics of the origin and the destination countries. The variables included are relevant to the origin-destination linkage, origin specific or destination specific characteristics of different countries.

The gravity function is specified as follows:

$$TF_{odt} = f(POP_{ot}, POP_{dt}, GDPPC_{ot}, GDPPC_{dt}, DISTAN_{od}, CONTI_{od}, COMOFFL_{od}, COMNOFFL_{od}, FCOL_o, FCOL_d, LAND_{od}, ISLA_{od}) \quad (3)$$

where *odt* indexes origin, destination and time, *POP* denotes population, *GDPPC* is GDP per capital at purchasing power parity (PPP), *DISTAN* is distance, *CONTI* is contiguity, *COMOFFL* is common official language dummy, *COMNOFFL* is common unofficial language dummy, *FCOL* is former colony dummy defined in terms of the colonial link between the originating country and destination country, *LAND* is landlocked dummy and *ISLA* is Island dummy.

A particular focus of the present study relates to the role of common unofficial language in overall international tourism attractiveness. To assess the role of common unofficial language in promoting international tourism, we use a dummy variable (*COMNOFF*) that takes the value 1 if the origin country shares a common unofficial language with the destination country, spoken by at least 9% of the population and zero otherwise. The potential

benefit of common unofficial language in facilitating international tourism could be achieved if a sizable number of individuals, say, at least 9% of the populations in both the originating and destination countries can speak the same common unofficial language.

Specifically, two countries or regions share a common unofficial language if such a language is not officially approved or accepted by their respective governments, though, such a language is relatively widely used in both countries. In the context of the present study, at least 9% of the populations in both countries speak such a language before it is considered to be relatively widely used. For instance, USA and Chile share a common unofficial language as Spanish is spoken by at least 9% of the populations in both countries.

In contrast, two countries or regions share a common official language if such a language is accepted or approved by their respective governments. The official language is given a legal status in the sense that it is the language used for official government business, taught in public schools, used in the administration of justice, among others. For instance, a significant number of countries use the English language as their official language, countries such as South Africa, Nigeria, Zambia, Australia, and Bahamas etc. Similarly, Spanish is the official language in about 20 countries, such as Mexico, Paraguay, Chile, Puerto Rico, and Cuba, among others. To be more specific in terms of example, Nigeria and Zambia share a common official language as the English language is the official language in both countries.

As pointed out earlier, a significant fraction of country pairs does not share a common official language. This suggests that common unofficial language potentially plays an important in promoting international tourism. Fidrmuc and Fidrmuc (2016) show that widespread use of foreign languages is a significant driver of international trade. In the context of international tourism, the ability to speak a foreign language potentially confers benefits as it makes it easier for one to visit foreign countries, meet new people and establish new contacts. Common unofficial language variable is sourced from the Head, Mayer, and Ries (2010) gravity data set that also includes information on common official language, contiguity and colonial linkages.

2.2. Empirical strategy

Following the above theoretical rationale, the reduced-form augmented gravity model is specified as follows:

$$\begin{aligned} LnTF_{odt} = & \beta + \alpha_1 LnPOP_{ot} + \alpha_2 LnPOP_{dt} + \alpha_3 LnGDPPC_{ot} \\ & + \alpha_4 LnGDPPC_{dt} + \alpha_5 LnDISTAN_{odt} + \alpha_6 CONTI_{od} \\ & + \alpha_7 COMOFFL_{od} + \alpha_8 COMNOFFL_{od} + \alpha_9 FCOL_o \\ & + \alpha_{10} FCOL_d + \alpha_{11} LAND_{od} + \alpha_{12} ISLA_{od} + \varepsilon_{odt} \end{aligned} \quad (4)$$

where Ln denotes natural logarithm; β is a constant that captures the unobserved quality of destination in a given year to inhabitants from the originating country; ε is an error term that is independent and identically distributed across country pairs and over time. α 's are parameters to be estimated. Detailed description of the variables used in the empirical analysis is provided in the section that relates to overview of data and measures below.

Equation (4) is estimated using four different types of estimators, namely OLS, OLS with fixed effects, Random Effects estimator and Poisson Pseudo-Maximum Likelihood (PPML). OLS estimates are likely to be biased, as multilateral resistance terms (MRTs) are not directly observable. Analogous to the empirical analysis of international trade, one of the appropriate ways of circumventing this issue is to use country fixed effects as a proxy for MRTs (for more

detail, see Anderson & Van Wincoop, 2004). The country fixed effects help to capture all country-specific characteristics, and thus control for a country's overall level of tourist flows from origin to destination countries. As a result, we augment the base model with origin and destination fixed effects and rerun the regression.

To check for the robustness of the parameter estimates we employ random effects and PPML estimators. The random effect is applied as the fixed effect is not a viable option given that the variable of interest in this paper, namely common unofficial language is time-invariant. Further, to ensure that the estimates are not sensitive to 'zero tourist flows' issue, which is more likely to be the case with the use of disaggregated data, we employ PPML. PPML estimator is appropriate for dealing with zeros in the tourist matrix, in order to obtain unbiased and consistent estimates.

3. Overview of data and measures

The empirical analysis is carried out using data obtained from different sources, namely the United Nations World Tourism Organisation (UNWTO), World Development Indicators, and Gravity dataset. UNWTO database contains the flow of visitors differentiated by the country of origin and destination, for 222 countries, from the year 1995–2015 (World Tourism Organisation, 2017b). World Development Indicators compiled by World Bank contains information relating to Gross Domestic Product and GDP per capita at PPP for 217 countries used in the empirical analysis (World Bank, 2017a,b). The Gravity data set from Head et al. (2010) which is derived from CEPII. Data for 224 countries on contiguity, colonial links and language.

The three data sets were merged, creating a panel data consisting of 200 countries. After dropping observations with missing data, the final data set is an unbalanced panel with 196,398 observations. Most of the missing data occur in the earlier period of the sample and/or due to the paucity of data from developing countries.

3.1. Dependent variable

To capture the flow of tourists which we denote as TF , we use *tourist arrivals* which is the most commonly used measure of tourism demand in the existing literature (Gil-Pareja et al., 2007; Khadaroo & Seetanah, 2008). The flow of tourist is captured by including all countries worldwide, in terms of origin and destination. The data set is derived from national sources and was compiled and disseminated by UNWTO. UNWTO ensures that all countries adhere to a uniform definition regime.

To capture *tourist arrivals*, tourism is defined by the UNWTO as "the activities of persons travelling to and staying in places outside their usual environment for not more than one consecutive year for leisure, business and other purposes not related to the exercise of an activity remunerated from within the place visited." (Tourism Satellite Account: Recommended Methodological Framework, Eurostat, OECD, WTO, UNSD, 2001, paras 1.1 and 2.1). In particular, international tourists are "tourists who stay at least one night in a country where they are not residents," where a resident is "a person who has lived for most of the past year in a country" (Eilat & Einav, 2004).

Using tourist arrivals as a measure of tourist flows is more practical and precise as it is easier to count the number of people entering a country. This measure enables us to capture the demand side of tourism. Expenditures of tourists are also used in the literature as a proxy for the flow of tourists, however, they need to be estimated. In addition, the tourism receipts data published in the balance of payments suffer from inaccuracies (Sinclair, 1998). In general, the viability of using tourist arrivals as a proxy for the value of trade in international tourism is supported by its high correlation

with tourism receipts (Neumayer, 2004).

3.2. Control variables

Following the literature, the estimation model includes a set of control variables that are relevant determinants of international tourism flows (Gil-Pareja et al., 2007; Khadaroo & Seetanah, 2008; Martins, Gan, & Ferreira-Lopes, 2017). First, we include a dummy variable, which takes the value 1 if the two countries share a common language, and zero otherwise, denoted as $COMOFFL$.

Second, the levels of economic development of the origin and destination countries are controlled for by including real GDP per capita at PPP. The flow of visitors is likely to be correlated with the level of economic development. However, this may work differently for visitors from different countries and/or regions. For example, visitors from richer countries may be more attracted to tourism destinations that are developed, while visitors from poorer countries may be indifferent. $GDPPC_o$ denotes origin country's GDP per capita at PPP while $GDPPC_d$ is destination country's GDP per capita at PPP.

Third, the population of origin and destination countries are controlled for to account for the relative size of the destination and origin countries. POP_o represents the population of the originating country, whilst POP_d is the population of the destination country. Fourth, we also include a measure of distance between the two countries as a control variable. Distance does not only serve as an analogy of the distance between two masses in the gravity equation, but plays a crucial role as a proxy for the cost of travelling to the destination country for tourism. Air travel is the predominant mode of transportation for international tourism (Khadaroo & Seetanah, 2008). The practice of price discrimination complicates the task of estimating the cost of travelling to the tourism destination. We denote the distance between the origin country and the destination country as $DISTAN_{od}$.

Fifth, contiguity is captured using a border dummy variable, $CONT_{od}$, which takes the value of 1 for countries sharing borders, and 0 otherwise. The primary purpose of including this variable is to further reveal the preferences of visitors. Similarly, we control for the influence of colonial linkages between the origin country and destination country or vis-à-vis using former colony dummies. $FCOL_o$ is a dummy that takes the value 1 if the origin country is a former colony of the destination country and zero otherwise. $FCOL_d$ is a dummy that takes the value 1 if the destination country is a former colony of the origin country.

An island dummy and a landlocked dummy are included to control for the geographical characteristics of the origin and destination countries. Traditionally, island countries have been successful in the area of attracting international tourists. In view of the tourist potential of island countries, Wilkinson (1989) notes that if the trend of the growth of international tourism persists, it is a matter of time for island countries to develop the tourism industry. Island dummy, $ISLA_{od}$, takes the value 1 if the origin country and destination country are an island and zero otherwise. A tourist travelling from a country of origin that is an Island to destination that is also an Island has the option of travelling by sea in addition to travelling by air. A tourist from a country of origin that is not an Island has a limited choice set as the option of travelling by sea is limited or may be not be available.

Lastly, while island countries present extraordinary attractions for the international tourism, landlocked countries are the polar opposite due to the absence of coastline. Controlling for other factors, landlocked countries on average are less endowed with natural resources and thus less of a magnet for international tourists relative to island countries. Landlocked dummy, $LAND_{od}$, takes a value 1 if the origin and destination country are landlocked

and zero otherwise. The predominant means of transportation for tourists originating from countries that are landlocked is by air. Tourists originating from countries that are not landlocked could chose to travel by sea.

3.3. Summary statistics

The summary statistics are presented in [Table 1](#). Countries that share a common unofficial language tend to attract more tourists relative to countries that do not, judging from the unconditional means. They are also more likely to share common official languages. This suggests that common unofficial language is one of the potential drivers of international tourist flows after controlling for the impact of common official language.

Countries that share a common unofficial language tend to have slightly lower population than countries that do not. This applies whether the population is measured based on the origin or destination country. Analogous to GDP as a measure of economic size, population is one of the potential determinants of international tourism flows and is expected to have a positive impact on international tourist flows. The impact of population, however, may be income group and region dependent. For instance, individuals from wealthier countries are more likely to travel across borders for tourism compared with individuals from developing countries.

GDP per capita at PPP in the originating countries tends to be higher in countries that do not share a common unofficial language relative to countries that do share a common unofficial language. Sharing a common unofficial language potentially decreases the costs of international tourism, thus, helping to dilute the impact of GDP per capita at PPP in the originating countries. Conversely, GDP per capita at PPP in the destination countries tend to be higher in countries that share a common unofficial language than their counterparts. Countries that share a common unofficial language may be rich enough to build the infrastructural base necessary to entice international tourists.

In general, countries that share a common unofficial language

are more likely to be closer to one another, share a common official language, have colonial linkages, and be an Island than countries that do not share a common unofficial language. These commonalities are more likely to promote international tourism, especially for countries that share a common unofficial language.

Countries that share a common unofficial language are on average less likely to be landlocked compared their counterparts. Landlocked countries might attract a smaller fraction of tourists than countries that are not landlocked. However, the dampening effect of landlocked variable on tourist flows may be income group and region dependent. For instance, landlocked European countries might attract a smaller share of tourists than their counterparts that are not landlocked. Given the closeness of European countries to one another, landlocked countries are disadvantaged in terms of accessibility by sea, assuming a significant number of European tourism is by sea. In contrast, the landlocked variable may not have any significant impact on international tourist flows for the regions such as Africa and South America. The distances between Africa and South America are large. This suggests that most international tourists between the two regions travel by air.

Tests for multicollinearity are reported in [Table 2](#). As shown, there is no evidence of severe multicollinearity issue. The 'rule of thumb' in the econometric literature is that a tolerance level less than 0.1 or a VIF larger than 10 is suggestive of serious multicollinearity problem.

4. Empirical results

4.1. The link between common unofficial language and international tourism

[Table 3](#) presents parameter estimates obtained using different estimators. The first column reports OLS estimates controlling for year fixed effects, whereas the second column shows random effects estimates controlling for year fixed effects. The third column presents parameter estimates obtained using Poisson Pseudo-

Table 1
Summary statistics.

Variable	(1)	(2)	(3)
	Do not share a common unofficial language	Do Share a common unofficial language	All countries
Ln tourist flow	6.77 (3.22)	7.66 (3.27)	6.95 (3.25)
Ln Population (Origin)	16.31 (1.82)	15.82 (2.15)	16.21 (1.90)
Ln Population (Destination)	16.23 (2.02)	15.78 (2.24)	16.14 (2.07)
Ln GDP per capita, PPP (Origin)	9.45 (1.18)	9.19 (1.26)	9.40 (1.20)
Ln GDP per capita, PPP (Destination)	9.34 (1.05)	9.38 (1.11)	9.34 (1.06)
Ln Distance	8.56 (0.87)	8.31 (1.04)	8.51 (0.91)
Contiguity	0.03 (0.17)	0.08 (0.27)	0.04 (0.20)
Common Official Language	0.03 (0.18)	0.80 (0.40)	0.19 (0.39)
Former colony (Destination)	0.01 (0.09)	0.05 (0.21)	0.02 (0.12)
Former colony (Origin)	0.01 (0.07)	0.02 (0.13)	0.01 (0.09)
Landlocked	0.02 (0.14)	0.01 (0.12)	0.02 (0.13)
Island	0.03 (0.18)	0.14 (0.34)	0.05 (0.23)
Observations	155734	40664	196398

Table 2

Tests of multicollinearity: Variance inflation factors (VIF) and Tolerance. Dependent variable: Ln tourist flows.

Variable	VIF	Tolerance
Common Official Language	2.84	0.35
Common Unofficial Language	2.73	0.37
Ln Distance	1.24	0.8
Contiguity	1.24	0.81
Island	1.17	0.86
Ln Population (Destination)	1.15	0.87
Ln Population (Origin)	1.10	0.91
Ln GDP per capita, PPP (Destination)	1.07	0.94
Former colony (Destination)	1.06	0.94
Ln GDP per capita, PPP (Origin)	1.05	0.95
Former colony (Origin)	1.03	0.97
Landlocked	1.03	0.97
Mean VIF	1.39	

Note: The 'rule of thumb' in the econometric literature is that a variance inflation factors greater than 10 or a tolerance level less than 0.1 is a sign of a severe multicollinearity problem.

Maximum Likelihood Estimator (PPML), whereas the final column report OLS estimates controlling for year, origin and destination fixed effects (OLSFE).

In general, the parameter estimates obtained using different estimators are consistent in terms of direction, however, landlocked variable shows a difference in terms of confidence level. This suggests that the influence of landlocked variable on tourist

attractiveness potentially depend on the level of development or region of host or origin countries. In view of the fact that the estimates are consistent across different estimators, the discussion of the results will be focused on the estimates obtained using OLSFE.

As reported in Table 3, column 4, the GDP per capita at PPP and the population variables in the origin and destination countries have a significant positive impact on tourist flows. For instance, as reported in column 4, a one percentage change in population results in around 0.25% change in tourist flows for origin countries, whereas its impact is about 0.71% change for destination countries. This suggests that the greater the population, the larger the tourist demand of the origin countries, whereas the bigger the population, the larger the tourist services supply of the destination countries.

As shown in column 4, a one percentage change in GDP per capita at PPP results in around 0.77% change in tourist flows for origin countries, while its impact is about 1.08% change for destination countries. This indicates that richer originating countries have a larger demand for international tourism, whereas, richer destination countries are better equipped to supply a larger pool of tourist services. This finding suggests that international tourism is a normal good in consumption and for a majority of individuals tends to be a luxury good. On the one hand, it is normal good for some individuals since an increase in income is associated with an increase in demand for tourism services. On the other hand, it is a luxury good if some tourists spend a significant amount of money on non-necessity goods such as first-class airfares and five-star hotels. The demand for tourism in this regard tends to depend on

Table 3

The link between common unofficial language and international tourism.

Variable	(1) _{OLS}	(2) _{RE}	(3) _{PPML}	(4) _{OLSFE}
Ln Population (Origin)	0.80*** (0.00)	0.76*** (0.01)		0.25*** (0.04)
Ln Population (Destination)	0.66*** (0.00)	0.68*** (0.01)		0.71*** (0.05)
Ln GDP per capita, PPP (Origin)	1.18*** (0.00)	0.98*** (0.02)	0.77*** (0.10)	0.77*** (0.03)
Ln GDP per capita, PPP (Destination)	0.83*** (0.00)	0.98*** (0.02)	0.20*** (0.07)	1.08*** (0.03)
Ln Distance	-1.36*** (0.00)	-1.45*** (0.02)	-1.11*** (0.02)	-1.69*** (0.00)
Contiguity	1.50*** (0.03)	1.48*** (0.10)	1.16*** (0.04)	1.18*** (0.02)
Common Official Language	1.26*** (0.02)	1.27*** (0.07)	0.49*** (0.04)	0.97*** (0.02)
Common Unofficial Language	0.27*** (0.02)	0.30*** (0.07)	0.45*** (0.04)	0.28*** (0.02)
Former colony (Destination)	1.00*** (0.03)	1.46*** (0.14)	0.33*** (0.04)	0.62*** (0.03)
Former colony (Origin)	1.40*** (0.04)	1.33*** (0.17)	0.21*** (0.05)	1.05*** (0.05)
Landlocked	-0.26*** (0.04)	-0.25** (0.11)	-0.23*** (0.05)	0.13*** (0.03)
Island	1.16*** (0.02)	1.16*** (0.07)	0.20*** (0.04)	0.27*** (0.02)
Ln Population (Origin)_scaled			0.88*** (0.18)	
Ln Population (Destination)_scaled			0.39** (0.16)	
Year Effects	Yes	Yes	Yes	Yes
Country Effects (Origin)	No	No	Yes	Yes
Country Effects (Destination)	NO	NO	Yes	Yes
Observations	196398	196398	196398	196398
R-squared	0.69		0.87	0.84
R-squared Within		0.29		

Notes: The dependent variable in columns 1, 2 and 4 is natural log of tourist flow, whereas the dependent variable in column 3 is tourist flow. In column 3, Ln Population (Origin)_scaled and Ln Population (Destination)_scaled were scaled by dividing the variables by 1000,000 before taking logs in order to achieve convergence. Ln denotes natural logarithm. All regressions include constant but they are not reported.

discretionary income of the consumers (Crouch, 1992). In support of the notion that tourism is a luxury for some consumers, Smeral (2003) show that the income elasticity for international tourism is greater than unity.

In line with expectation, distance has a negative impact on international tourist flows. This is not surprising as the transaction costs associated with tourism tend to increase the longer the distance between the originating and the destination countries. The high transaction costs, such as traveling, and search costs associated with longer distances tend to lower the demand for international tourism.

Common official language and common unofficial language have a significant positive impact on tourist flows. More specifically, the flow of international tourists to a destination will be 163% higher for the country pairs that share a common official language compared with those that do not. Further, country pairs that share a common unofficial language will attract 32.31% larger pool of international tourists relative to those that do not.

These results show that sharing a common official language or common unofficial language helps to increase the number of tourists. Countries sharing a common official language and/or common unofficial language are found to be a magnet for international tourists. These findings demonstrate that tourists have a strong preference for destination countries that share the same official language with their origin countries. Further, tourists are strongly attracted to a destination that shares a common unofficial language that is spoken by at least 9% of the population.

As expected, sharing a common border has a substantial positive impact on tourist flows. Tourists face lower transportation and search costs with shared borders, and this helps to facilitate international tourism. Further, previous colonial ties have a strong positive impact on tourist flows. This finding potentially stems from the fact that countries with prior colonial relationships attract a larger share of tourists from each other. Similarly, island countries attract more tourists than non-island countries. In line with expectation, a larger number of tourists tend to visit island countries, perhaps due to the perceived pleasure of vacationing on island environment or the relative ease of accessibility.

We check whether the estimates are sensitive to price effects by including GDP deflator at purchasing power parity (PPP) for the originating and destination countries as additional control variables. Overall, the parameter estimates of common unofficial language and other relevant control variables are consistent with the results obtained earlier across different estimators. The parameter estimates are not reported, but they are available from the authors upon request.

4.2. *The link between common unofficial language and international tourism by income groups*

Tables 4 and 5 report parameter estimates obtained by rerunning the regressions for the sub-samples of countries by income groups. Table 4 presents the estimates for sub-sample of origin countries. Column 1 reports estimates for high-income countries, column 2 shows the estimates for upper middle-income countries, column 3 reports estimates for lower middle-income countries, whereas column 4 reports estimates for low-income countries.

As shown in columns 1 to 4, common unofficial language has significant positive impact on tourist flows regardless of income levels of the originating countries. This suggests that tourists from the originating countries are significantly attracted to destination countries that share common unofficial language, regardless of the stage of development after controlling for the influence of common official language and other relevant variables.

In general, the other parameter estimates are consistent with

the ones reported in Table 3, though there are some notable differences. Population in the originating countries has a negative impact on tourist flows in the low-income countries. Most individuals in low-income countries tend to be poor and thus potentially could not afford the costs associated with international tourism. In addition, a significant number of individuals in low income countries might have a high preference for internal tourism as opposed to international tourism. Further, tourists from the upper middle-income, lower middle-income and low-income countries are less likely to visit a country if the destination country is a former colony of the origin country.² In contrast, tourists from high-income countries are more likely to travel to a country if the destination country is a former colony of the origin country.

Besides, landlocked has a strong negative impact on tourist flows in the high-income countries, whereas its impact on upper middle-income and lower middle-income countries is positive and statistically significant whereas its impact for low-income countries is statistically insignificant. This suggests that the rich countries that are landlocked would attract fewer tourists compared to a rich country that is not landlocked. Landlocked countries tend to be more expensive for tourists, and this appears to be significant for rich countries.

Table 5 reports the parameter estimates of the link between common unofficial language and international tourism by income levels of the destination countries. Columns 1, 2, 3 and 4 show estimates for high-income, upper middle-income, lower middle-income and low-income countries respectively. The estimates reported in columns 1 to 4 confirm that common unofficial language helps to promote international tourism. These results show that the estimates are not susceptible to the way in which the countries are classified by income levels either by using the sub-sample of the originating countries as discussed earlier or the destination countries. This finding reinforces the notion that common unofficial language helps to promote international tourism across countries.

Overall, the other estimates are in line with the results presented in Table 3 with some exceptions. Population in lower middle-income destination countries has a negative impact on tourist flows, whereas its impact is positive in high-income, upper middle-income and low-income destination countries. These results suggest that larger population in lower middle-income destination countries leads to a lower supply of tourist services, whereas larger populations in high-income, middle-income and low-income destination countries lead to greater supply of tourist services. Low-income destination countries with larger populations may find it easier to supply a larger pool of tourist services by leveraging on low wages. As low-income destination countries with larger populations transform into lower middle-income countries, the pool of tourist services on offer might decrease with rising wages and increasing demand for social safety net. As lower middle-income destination countries with larger populations transform into upper middle-income and high-income countries, the pool of tourist services available tend to increase with rising productivity, sophistication in service delivery, and better technologies for service delivery.

Further, tourists from low-income originating countries are less likely to travel to a country if the originating country is a former colony of the destination country. Besides, the estimate in column 1

² Using country of destination for classifying countries into income groups, there are 23 low-income countries, resulting in 412 country pairs. Former colony (destination) is set equal to 1 if the destination country is a former colony of the originating country. With respect to the second dummy variable, there are 2 low-income countries, resulting in 32 country pairs. Former colony (origin) is set to 1 if the origin country is a former colony of the destination country.

Table 4
The link between common unofficial language and international tourism by income groups based on country of origin.

Variable	(1)High-income	(2)Upper Middle-income	(3)Lower Middle-income	(4)Low-income
Ln Population (Origin)	0.41*** (0.07)	0.18 (0.12)	-0.14 (0.15)	-0.56** (0.23)
Ln Population (Destination)	0.43*** (0.06)	0.69*** (0.10)	1.25*** (0.13)	1.07*** (0.15)
Ln GDP per capita, PPP (Origin)	0.87*** (0.06)	0.89*** (0.05)	0.61*** (0.07)	0.54*** (0.06)
Ln GDP per capita, PPP (Destination)	1.16*** (0.03)	1.11*** (0.05)	1.21*** (0.06)	0.91*** (0.09)
Ln Distance	-1.51*** (0.01)	-1.70*** (0.01)	-1.91*** (0.01)	-2.00*** (0.02)
Contiguity	0.76*** (0.04)	1.30*** (0.05)	1.01*** (0.05)	0.88*** (0.07)
Common Official Language	0.82*** (0.03)	1.17*** (0.03)	1.01*** (0.03)	0.81*** (0.03)
Common Unofficial Language	0.22*** (0.02)	0.35*** (0.03)	0.49*** (0.03)	0.30*** (0.04)
Former colony (Destination)	1.02*** (0.03)	-1.00*** (0.13)	-0.82*** (0.21)	-1.47* (0.89)
Former colony (Origin)	0.50*** (0.06)	1.22*** (0.09)	1.99*** (0.13)	1.81*** (0.25)
Landlocked	-0.39*** (0.03)	0.20** (0.08)	0.83*** (0.08)	-0.01 (0.08)
Island	0.05** (0.02)	0.35*** (0.04)	0.41*** (0.05)	0.18** (0.08)
Year Effects	Yes	Yes	Yes	Yes
Country Effects (Origin)	Yes	Yes	Yes	Yes
Country Effects (Destination)	Yes	Yes	Yes	Yes
Observations	86122	42984	34286	19572
R-squared	0.86	0.83	0.82	0.78

Table 5
The link between common unofficial language and international tourism by income groups based on country of destination.

Variable	(1)High-income	(2)Upper Middle-income	(3)Lower Middle-income	(4)Low-income
Ln Population (Origin)	0.49*** (0.07)	0.05 (0.08)	0.20** (0.09)	-0.23 (0.15)
Ln Population (Destination)	0.49*** (0.08)	1.62*** (0.10)	-0.31** (0.10)	1.39*** (0.27)
Ln GDP per capita, PPP (Origin)	0.72*** (0.04)	0.94*** (0.05)	0.71*** (0.06)	0.86*** (0.08)
Ln GDP per capita, PPP (Destination)	0.93*** (0.05)	1.01*** (0.04)	1.86*** (0.07)	1.15*** (0.08)
Ln Distance	-1.62*** (0.01)	-1.67*** (0.01)	-1.73*** (0.01)	-1.53*** (0.02)
Contiguity	0.48*** (0.04)	1.81*** (0.04)	1.51*** (0.05)	0.92*** (0.05)
Common Official Language	0.79*** (0.03)	1.01*** (0.03)	0.94*** (0.03)	0.92*** (0.04)
Common Unofficial Language	0.25*** (0.03)	0.46*** (0.03)	0.22*** (0.03)	0.27*** (0.04)
Former colony (Destination)	0.43*** (0.06)	0.68*** (0.05)	0.61*** (0.05)	1.04*** (0.08)
Former colony (Origin)	1.65*** (0.06)	0.06 (0.08)	0.61*** (0.23)	-1.42*** (0.26)
Landlocked	-1.00*** (0.05)	-0.26*** (0.06)	0.25*** (0.06)	0.39*** (0.05)
Island	0.37*** (0.03)	0.04 (0.03)	0.11** (0.04)	-0.18 (0.13)
Year Effects	Yes	Yes	Yes	Yes
Country Effects (Origin)	Yes	Yes	Yes	Yes
Country Effects (Destination)	Yes	Yes	Yes	Yes
Observations	72921	59317	33358	17368
R-Squared	0.87	0.84	0.84	0.80

of Table 5 is consistent with the finding that high-income and upper middle-income landlocked countries are less attractive destinations for tourists. In contrast, there is no strong evidence that lower middle-income and low-income landlocked countries face a similar obstacle in terms of attracting tourists.

We performed a robustness check by rerunning the regressions for the income classifications in which the fiscal year 2015 is used in lieu of the fiscal year 2011 as noted earlier. In general, the parameter estimates of common unofficial language and other relevant variables are in line with the results obtained earlier for the originating and destination countries. The parameter estimates are not reported, but they are available from the authors upon request.

4.3. The link between common unofficial language and international tourism by region of origin

Table 6 reports the parameter estimates of the link between common unofficial language and international tourism by region of origin. Column 1 presents estimates for sub-sample of European countries, column 2 shows estimates for the sub-sample of North America, Latin American and Caribbean countries, column 3 reports estimates for Asia, Oceania, Middle East, and North Africa countries, whereas column 4 displays estimates for Sub-Saharan Africa countries.

As reported in columns 1 to 4, common unofficial language has a strong positive impact on tourist flows across different regions of the world after controlling for common official language and other relevant covariates. This suggests that the demand for tourism is higher in countries where a significant share of the population speaks a common unofficial language.

In general, the other parameter estimates are in line with the results reported in Tables 4 and 5 with some exceptions. Population in the origin countries has a negative impact on tourist flows in

Europe, whereas its impact is positive in other regions of the world such as North America, Latin America and the Caribbean, Asia, among others. This indicates that the demand for international tourism in Europe is higher in smaller countries than larger countries. Previous colonial ties established because of the destination country being a former colony of the origin country boost international tourism in European countries, but they have the opposite effect in Asia, Middle East, North Africa, Oceania and Sub-Saharan Africa countries. This happens to be the case if the sample is subdivided into regions based on the region of origin. Landlocked has a negative impact on tourist flows in Europe, whereas its impact is positive in Asia, Oceania, Middle East, North Africa, and Sub-Saharan Africa. This is consistent with the notion that wealthy landlocked economies are less attractive for international tourists.

4.4. The link between common unofficial language and international tourism by region of destination

Table 7 presents parameter estimates on the link between common unofficial language and international tourism by region of destination. In line with results presented in Table 6, column 1 reports parameter estimates for the sub-sample of European countries, column 2 presents parameter estimates for the sub-sample of North America, Latin American and Caribbean countries, column 3 shows estimates for Asia, Oceania, Middle East, and North Africa countries, whereas column 4 presents parameter estimates for Sub-Saharan Africa countries.

As presented in columns 1 to 4, common unofficial language has a substantial positive impact on the tourism attractiveness of different regions of the world after controlling for the influence of common official language and other relevant covariates. Interestingly, common unofficial language is particularly important in European countries given that it has a statistically positive impact on

Table 6
The link between common unofficial language and international tourism by region of origin.

Variable	(1)Europe	(2)America & LATCA	(3)Asia, Middle East, North Africa & Oceania	(4)SSA
Ln Population (Origin)	-0.27** (0.13)	1.64*** (0.20)	0.77*** (0.07)	0.27* (0.15)
Ln Population (Destination)	0.53*** (0.06)	1.41*** (0.10)	0.49*** (0.10)	0.91*** (0.11)
Ln GDP per capita, PPP (Origin)	1.15*** (0.06)	0.74*** (0.07)	0.88*** (0.05)	0.36*** (0.05)
Ln GDP per capita, PPP (Destination)	1.16*** (0.03)	1.01*** (0.05)	1.13*** (0.05)	0.94*** (0.06)
Ln Distance	-1.48*** (0.02)	-1.75*** (0.01)	-1.71*** (0.01)	-1.71*** (0.02)
Contiguity	0.81*** (0.03)	0.96*** (0.04)	1.52*** (0.05)	1.18*** (0.05)
Common Official Language	0.23*** (0.03)	1.18*** (0.03)	1.02*** (0.03)	0.86*** (0.03)
Common Unofficial Language	0.45*** (0.03)	0.24*** (0.03)	0.11*** (0.03)	0.43*** (0.03)
Former colony (Destination)	1.18*** (0.03)	0.09 (0.18)	-0.26** (0.10)	-1.83*** (0.48)
Former colony (Origin)	1.01*** (0.08)	-0.69*** (0.09)	1.52*** (0.08)	0.48*** (0.18)
Landlocked	-0.10*** (0.03)	0.002 (0.09)	0.66*** (0.08)	0.47*** (0.05)
Island	0.11*** (0.03)	0.56*** (0.03)	0.28*** (0.03)	0.92*** (0.06)
Year Effects	Yes	Yes	Yes	Yes
Country Effects (Origin)	Yes	Yes	Yes	Yes
Country Effects (Destination)	Yes	Yes	Yes	Yes
Observations	62769	37022	63790	32021
R-squared	0.89	0.89	0.82	0.83

Table 7
The link between common unofficial language and international tourism by region of destination.

Variable	(1)Europe	(2)America & LATCA	(3)Asia, Middle East, North Africa & Oceania	(4)SSA
Ln Population (Origin)	0.76*** (0.09)	0.35*** (0.09)	0.19*** (0.06)	0.19* (0.11)
Ln Population (Destination)	-2.60*** (0.21)	1.97*** (0.16)	0.99*** (0.08)	1.47*** (0.14)
Ln GDP per capita, PPP (Origin)	0.63*** (0.06)	0.81*** (0.05)	1.00*** (0.04)	0.78*** (0.06)
Ln GDP per capita, PPP (Destination)	-0.35*** (0.07)	1.06*** (0.06)	1.29*** (0.04)	1.08*** (0.07)
Ln Distance	-1.41*** (0.02)	-1.74*** (0.01)	-1.57*** (0.01)	-1.70*** (0.02)
Contiguity	0.88*** (0.04)	1.00*** (0.04)	1.66*** (0.05)	1.33*** (0.05)
Common Official Language	0.05 (0.05)	0.83*** (0.03)	1.11*** (0.03)	0.96*** (0.04)
Common Unofficial Language	0.29*** (0.05)	0.75*** (0.03)	0.16*** (0.02)	0.13*** (0.04)
Former colony (Destination)	0.68*** (0.06)	0.39*** (0.06)	0.38*** (0.05)	1.06*** (0.06)
Former colony (Origin)	1.46*** (0.06)	-1.09*** (0.08)	0.56*** (0.07)	-1.14*** (0.31)
Landlocked	-0.45*** (0.04)	-0.09 (0.06)	0.51*** (0.05)	0.37*** (0.05)
Island	-0.03 (0.06)	0.35*** (0.03)	0.36*** (0.03)	0.18*** (0.05)
Year Effects	Yes	Yes	Yes	Yes
Country Effects (Origin)	Yes	Yes	Yes	Yes
Country Effects (Destination)	Yes	Yes	Yes	Yes
Observations	40540	43902	80964	30597
R-squared	0.90	0.91	0.82	0.80

tourist flows, whereas the impact of common official language is positive, but statistically insignificant. This suggests that European countries where common unofficial language is prevalent attract more tourists than others where they are not prevalent. Common unofficial language is potentially more important as a driver of international tourism compared to common official language in Europe. Most European countries do not share a common official language, and this might explain why common unofficial language transcends common official language in terms of promoting international tourism.

Overall, the other parameter estimates are in line with the results reported in Tables 5 and 6 with some notable differences. Population in the origin countries has a positive impact on tourist flows in the European countries, whereas, population in the destination countries has a negative impact on tourist flows. These findings suggest that the larger the population in the originating countries, the more the demand for tourism services. The larger the population in the destination countries the more likely the tourism industry focuses more on attracting domestic tourists or providing services or advertisements targeting internal tourists as opposed to international tourists.

Further, GDP per capita in the originating European countries has a positive impact on tourist flows, whereas GDP per capita in the destination European countries has a negative impact on international tourism. This suggests that GDP per capita plays an important role in terms of the demand for tourist services relative to the supply of tourist services. Previous colonial ties established as a result of the origin country being a former colony of the destination country promote international tourism in European countries, Asia, Oceania, and Middle East, whereas they have the opposite effect in North America, Latin America and the Caribbean, and Sub-Saharan Africa. This happens to be the case if the sample is subdivided into regions based on the region of destination. These

findings suggest that the impact of prior colonial linkages on international tourism is region dependent.

5. Conclusion

This study examines the role of common unofficial language in promoting international tourist flows while considering the influences of the levels of development and regions in the underlying relationship. An augmented gravity model is employed to estimate the impact of common unofficial language on international tourism for the whole sample and the sub-samples based on the income group and region classifications. This study uses a panel data set of bilateral tourism flows augmented with classical gravity variables for 200 countries over the period of 1995–2015.

We can draw three main conclusions from the analysis. First, international tourists, are not only strongly attracted to destinations that share the same common official language with their origin countries, but also to destinations that share the same common unofficial language. We argue that the positive impact of common unofficial language on tourist flows emanates from cultural proximity. Going to a destination that shares a common unofficial language not only reduces the transaction costs of tourism, but also increases the chances of cross-cultural interactions and provide tourists with a sense of comfort and enthusiasm.

Second, the positive impact of common unofficial language holds irrespective of the levels of development of different countries. Thus, the levels of income of the destination country and the origin country are less critical factors in the decision-making process of a typical tourist. Third, common unofficial language is a more important determinant of international tourism compared to common official language in Europe. A plausible explanation for this is the fact that several European countries do not share a common official language. The limited scope of common official

language in Europe potentially explains why common unofficial language provides a significant boost to international tourism compared with common official language.

Considering the findings discussed above, two important policy implications emerge. First, the importance of common unofficial language in international tourism underscores the importance of multilingual population in particular and multicultural society in general in promoting tourism. Therefore, any efforts to attract international tourists should include the promotion of the cultural heritage of the destination country. Second, policies that create an enabling environment for citizens of a country to learn at least two foreign languages would make it possible for a multilingual country to emerge over time. Multilingual country, in turn, tends to be a magnet for international tourism.

It is worthwhile to note that the main finding of this study should be interpreted as suggesting that common unofficial language is one of the plausible determinants of international tourism. The data we used for our empirical analysis allows us to capture a common unofficial language if at least 9% of the populations in the originating and destination countries can speak the same common unofficial language. A more detailed data set that includes different measures of languages, linguistic distance and different cutoffs for common unofficial languages could allow for a deeper investigation of the link between common unofficial language and international tourism. On a related development, appropriately collected primary data from respondents from a cross-section of countries could be used to confirm or better hone our results.

A second limitation of our study is that we did not distinguish between the short and the long run effects of common unofficial language on international tourism. It might be the case that the existing relationship between tourist flows and common unofficial language differ in the short and the long run. Future studies can address this by distinguishing between the short run and the long run in exploring the relationship tested in this study.

6. Contribution made by each author to the paper

Then conceived the idea of testing the potential impact of unofficial language on international tourism, while Okafor augmented the idea by incorporating the influence of development and geography in the underlying relationship. Then formulated in part the research questions, while Okafor and Khalid refined the research questions. Okafor and Khalid specified the empirical models, while Okafor estimated the regression equations. Then collected and compiled the data from three different sources, namely UNWTO, World Bank, and CEPII. He also performed preliminary quantitative analyses. Khalid drafted the introduction and review of related literature. Okafor drafted the theoretical rationale, data description and discussion of the findings. Okafor and Khalid drafted the conclusion.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.tourman.2018.01.008>.

References

- Anderson, J. E., & Van Wincoop, E. (2004). Trade costs. *Journal of Economic Literature*, 42(3), 691–751.
- Crouch, G. I. (1992). Effect of income and price on international tourism. *Annals of Tourism Research*, 19(4), 643–664.
- Crouch, G. I. (1994). The study of international tourism demand: A survey of practice. *Journal of Travel Research*, 32(4), 41–55.
- Culiuc, A. (2014). *Determinants of international tourism*. IMF working paper.
- Durberry, R. (2008). Tourism taxes: Implications for tourism demand in the UK. *Review of Development Economics*, 12, 21–36.

- Eilat, Y., & Einav, L. (2004). The determinants of international tourism: A three dimensional panel data analysis. *Applied Economics*, 36, 1315–1328.
- Fidrmuc, J., & Fidrmuc, J. (2016). Foreign languages and trade: Evidence from a natural experiment. *Empirical Economics*, 50(1), 31–49. <https://doi.org/10.1007/s00181-015-0999-7>.
- Fourie, J., & Santana, M. (2011). The impact of mega-events on tourist arrivals. *Tourism Management*, 32, 1364–1370.
- Fourie, J., & Santana, M. (2013). Cultural affinity and ethnic reunion. *Tourism Management*, 36, 411–420.
- Gil-Pareja, S., Llorca-Vivero, R., & Mart'inez-Serrano, J. A. (2006). The impact of embassies and consulates on tourism. *Tourism Management*, 28, 355–360.
- Gil-Pareja, S., Llorca-Vivero, R., & Mart'inez-Serrano, J. A. (2007). The effect of EMU on tourism. *Review of International Economics*, 15(2), 302–312.
- Head, K., Mayer, T., & Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics*, 81(1), 1–14.
- Jud, G. D. (1974). Tourism and economic growth in Mexico since 1950. *Inter-American Economic Affairs*, 28, 19–43.
- Kastenholz, E. (2010). 'Cultural proximity' as a determinant of destination image. *Journal of Vacation Marketing*, 16(4), 313–322.
- Khadaroo, J., & Seetanah, B. (2008). The role of transport infrastructure in international tourism development: A gravity model approach. *Tourism Management*, 29(5), 831–840. <https://doi.org/10.1016/j.tourman.2007.09.005>.
- Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research*, 24, 835–849.
- Lim, C. (1999). A meta-analytic review of international tourism demand. *Journal of Travel Research*, 37, 273–284.
- Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research*, 44, 82–99.
- Martins, L. F., Gan, Y., & Ferreira-Lopes, A. (2017). An empirical analysis of the influence of macroeconomic determinants on World tourism demand. *Tourism Management*, 61(Supplement C), 248–260. <https://doi.org/10.1016/j.tourman.2017.01.008>.
- Massidda, C., & Etzo, I. (2012). The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management*, 33, 603–610.
- Neumayer, E. (2004). The impact of political violence on tourism dynamic cross-national estimation. *Journal of Conflict Resolution*, 48(2), 259–281.
- Neumayer, E. (2010). Visa restrictions and bilateral travel. *The Professional Geographer*, 62(2), 171–181.
- Santana, M., Ledesma, F. J., & Pérez, J. V. (2010). Exchange rate regimes and tourism. *Tourism Economics*, 16, 25–43.
- Santana, M., Ledesma, F. J., Pérez, J. V., & Cortés, I. (2010). Does a common currency promote countries' growth via trade and tourism? *The World Economy*, 33, 1811–1835.
- Sinclair, M. T. (1998). Tourism and economic development: A survey. *Journal of Development Studies*, 34(5), 1–51.
- Smeral, E. (2003). A structural view of tourism growth. *Tourism Economics*, 9(1), 77–93.
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—a review of recent research. *Tourism Management*, 29(2), 203–220.
- Tinbergen, J. (1962). *Shaping the world Economy: Suggestions for an international economic policy*. New York: Twentieth Century Fund.
- Vietze, C. (2012). Cultural effects on inbound tourism into the USA: A gravity approach. *Tourism Economics*, 18, 121–138.
- Wilkinson, P. F. (1989). Strategies for tourism in island microstates. *Annals of Tourism Research*, 16(2), 153–177.
- World Bank. (2017a). *Country and lending Groups: Country classifications*. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>. (Accessed 13 December 2017).
- World Bank. (2017b). *World Bank development Indicators*. Washington, USA.
- World Tourism Organisation. (2017b). *UNWTO's tourism statistics*. Madrid, Spain: CD-ROM.
- World Tourism Organization. (2017a). *UNWTO annual report 2016*. Madrid: UNWTO.
- Zhang, J., & Jensen, C. (2007). Comparative advantage: Explaining tourism flows. *Annals of Tourism Research*, 34(1), 223–243.



Okafor Luke Emeka is an Assistant Professor at the University of Nottingham, Malaysia Campus (UNMC). He received his Ph.D. in Economics from Monash University, Australia in 2015. His research interests include international economics, applied economics, development economics, trade and finance.



Usman Khalid is an Assistant Professor at the University of Nottingham, Malaysia Campus (UNMC). He received a Ph.D. from Lund University, Sweden in 2016. His research interests include institutional economics, development economics, international trade, and economic growth.



Terence has graduated from University of Nottingham Malaysia Campus in 2017, majoring in Economics. He is passionate in applying statistical analyses in solving problems of different domains besides being interested in interdisciplinary research, particularly those involving economics and geography. He currently works as a consultant in the field of credit risk modelling.