Faculty of Economics and Business

The Role of Language in Influencing International Tourism Demand in ASEAN+3 Countries

Thien Fung Thai

Doctor of Philosophy
2022

The Role of Language in Influencing International Tourism Demand in ASEAN+3 Countries

Thien Fung Thai

A thesis submitted

In fulfillment of the requirements for the degree of Doctor of Philosophy

(Tourism Development)

Faculty of Economics and Business
UNIVERSITI MALAYSIA SARAWAK
2022

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Malaysia Sarawak. Except where due acknowledgements have been made, the work is that of the author alone. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

……

Signature

Name: Thien Fung Thai

Matric No.: 15010129

Faculty of Economics and Business

Universiti Malaysia Sarawak

Date :

ACKNOWLEDGEMENT

I would like to take this opportunity to thank those who have contributed directly or indirectly to the completion of this thesis.

First of all, I would like to thank my family and friends who has supported me all the way in my journey of studying degree of Doctor of Philosophy. Their continuous support has provided me the courage needed to complete my study.

Moreover, I am also grateful to my supervisory committee, Professor Dr. Puah Chin Hong and Professor Dr. Venus Liew Khim Sen for their precious time and efforts devoted to supervise me. This thesis would not have been completed if it is not for their advices, guidance and supervisions.

Also, my sincere gratitude to the Centre for Graduate Studies, for the advice and support given during my period of study in Universiti Malaysia Sarawak.

Finally, I would like to thank the management of the Universiti Malaysia Sarawak for making it possible for me to complete my study here in Sarawak. Thank you all.

ABSTRACT

This thesis empirically examined the tourism demand for the ASEAN+3 countries, namely Brunei, Cambodia, China, Indonesia, Japan, Korea, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand and Vietnam from macroeconomic perspective. The static linear panel model analysis was applied to examine the influence of tourists' income, destination countries income, distance, free trade agreement, adjacent country, tourism price and language towards tourism demand in these countries. Tourism sector is vulnerable to external shocks, therefore, ASEAN+3 governments need to identify determinants of tourism demand to fully utilise their resources and efforts in sustaining the development of tourism sector in these countries. The results obtained in this study showed that tourist income, distance, adjacent country, tourism price and language are more influential determinants while the effect of destination income and free trade agreement are relatively less significant. Identification of these determinants allow governments of ASEAN+3 countries to design individual policy and also establish collaboration among each other to strengthen tourism sector performance.

Keywords: Tourism demand, determinant, panel data analysis, ASEAN+3, language

Peranan Bahasa Dalam Menpengaruhi Permintaan Pelancongan Antarabangsa di Negara ASEAN+3

ABSTRAK

Tesis ini mengkaji secara empirikal permintaan pelancongan untuk negara-negara ASEAN+3, iaitu Brunei, Cambodia, China, Indonesia, Japan, Korea, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand dan Vietnam dari perspektif makroekonomi. Analisis model panel linear static telah digunakan untuk mengkaji pengaruh pendapatan pelancong, pendapatan negara destinasi, jarak, perjanjian perdagangan bebas, negara bersebelahan, harga pelancongan dan bahasa terhadap permintaan pelancongan di negara-negara tersebut. Sektor pelancongan adalah terdedah kepada kejutan luaran, oleh itu, kerajaan ASEAN+3 perlu mengenal pasti penentu permintaan pelancongan bagi menggunakan sepenuhnya sumber dan usaha mereka dalam mengekalkan pembangunan sektor pelancongan di negara-negara tersebut. Hasil yang diperoleh dalam kajian ini menunjukkan bahawa pendapatan pelancong, jarak, negara bersebelahan, harga dan bahasa pelancongan adalah penentu yang lebih berpengaruh manakala pengaruh pendapatan negara destinasi dan perjanjian perdagangan bebas adalah kurang signifikan secara bandingan. Pengenalpastian penentu-penentu ini boleh membantu kerajaan negara-negara ASEAN+3 dalam perancangn dasar individu dan juga merangsangi kerjasama antara satu sama lain untuk memperkukuh prestasi sektor pelancongan.

Kata kunci: *Permintaan pelancongan, penentu, analisa data panel, ASEAN+3, bahasa*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARDL	Autoregressive Distributed Lag
ASEAN	Association of Southeast Asian Nations
ASEAN+3	ASEAN Plus Three
ATMS	ASEAN Tourism Marketing Strategy
ATSP	ASEAN Tourism Strategy Plan
BPLM	Breusch-Pagan Lagrange Multiplier
CEPII	Centre d'Etudes Prospectives et d'Informations Internationales
CLMV	Cambodia, Laos, Myanmar and Vietnam
CPI	Consumer Price Index
EPI	English Proficiency Index
FEM	Fixed Effects Model
FMOLS	Fully Modified Ordinary Least Square
GDP	Gross Domestic Products
GMM	Generalised Method of Moments
OECD	Organisation for Economic Co-operation and Development
PCSE	Panel -Corrected Standard Errors
POLS	Pooled Ordinary Least Squares
PPML	Poisson Pseudo-Maximum Likelihood
REM	Random Effects Model
SDGs	Sustainable Development Goals
SH	Sargan-Hansen
UNESCO	United Nations Educational, Scientific and Cultural Organization

UNWTO	United Nations World Tourism Organization
USA	United States of America
VECM	Vector Error Correction Model
VIF	Variance inflation factor
WTTC	World Travel and Tourism Council

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter introduces the tourism industry and has been organized in the following manner. Section 1.2 provides a background on the world tourism, Section 1.3 focuses on the regional tourism and Section 1.4 discusses on country specific tourism. Section 1.5 elaborates this thesis's problem statement while Section 1.6 provides the objectives of study. The significance of the study is discussed in Section 1.7 and lastly, Section 1.8 details the structure of this thesis.

1.2 World Tourism Background

The role of tourism after World War II has gone through tremendous evolution. In modern days, tourism is identified as the tool to generate foreign exchange earnings; create job opportunity; enhance the socio-economic status; improve trade performance; and encourage infrastructure development through its multiplier effect. Within the focus of 2030 Agenda for Sustainable Development, whereby a global framework for peace and prosperity for humanity and earth has been set up, there are 17 Sustainable Development Goals (SDGs) with tourism playing a part in it.

In the United Nations World Tourism Organization (UNWTO) Annual Report 2017 (2018), it is documented that tourism sector has been explicitly included in SDGs,

particularly Goals 8, 12 and 14[1]. The report emphasised that tourism has a cross-cutting nature and impact that paves the way towards achieving all the 17 SDGs, as well as an overall economy growth for a nation. In acknowledging tourism's contribution towards the agenda, 2017 was declared as the "International Year of Sustainable Tourism for Development" by the United Nations General Assembly.

In the past few decades, the growth of tourism has benefited from a few factors such as rapid transportation development (especially air travel), affordable travel cost, increase in disposable income, visa facilitation and an increasing number of tourist destinations worldwide. Apart from that, technology enhancement has allowed borderless information sharing on destinations and even easy purchasing of air tickets and accommodation without going through a travel agency; such connectivity is unimaginable prior to the dot.com boom. The stages of revolution in the past few decades have transformed tourism from a luxury activity into a leisure activity that is affordable to many.

This is evident from the statistics reported in the UNWTO Tourism Highlights (2019), which stated that 56 percent of international tourist arrival in 2018 were for leisure, recreation and holidays. About 27 percent of international tourist travelled to visit friends and relatives, health treatment, religious reasons. Business travels contributed to 13 percent international arrival. The remaining 4 percent traveller did not specify their travelling purpose (refer to Figure 1.1). In the report of World Travel and Tourism Council's (WTTC) World Economic Impact (2019a), leisure spending contributed 78.5 percent towards world

[1] Goal 8 of SDGs promotes "decent work and economic growth"; Goal 12 focus on "responsible consumption and production" and Goal 14 mandates to "life below water".

Gross Domestic Products (GDP) in 2018. These are strong supports to the claim that tourism is more of a leisure activity nowadays.

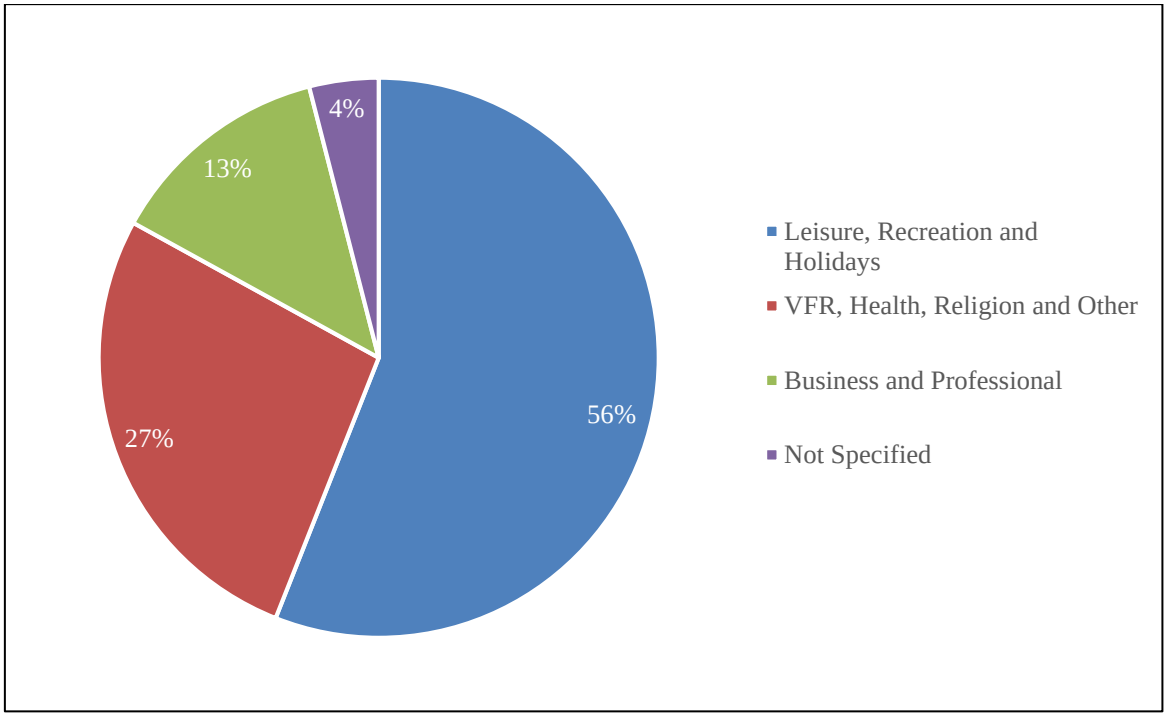

Figure 1.1: International Tourists' Purpose of Visit, 2018
(Source: UNWTO, 2019)

The transformation of tourism into a leisure activity has created opportunities for many destinations around the globe to develop and maintain the facilities to welcome international tourist. The rewards for local governments are, for most of the time, favourable for positive growth in foreign exchange earnings and employment rate. According to the WTTC (2019a), tourism has generated a total of USD 8,811.0 billion (or 10.4 percent of world GDP) in 2018 and this figure was forecasted to increase to USD 9,123.7 billion (or 10.4 percent of world GDP) in 2019. The long-term forecast has projected that the total contribution will reach USD 13,085.7 billion (or 11.5 percent of world GDP) in 2029 with an annual increment of 3.7 percent.

In terms of employment opportunities, the WTTC (2019a) reported that the tourism industry had created 318.8 million jobs worldwide in 2018 (or 10.0 percent of total employment). By the end of 2019, this number was projected to reach 328.2 million jobs (or 10.1 percent of total employment). This is anticipated to increase to 420.6 million jobs (or 11.7 percent of total employment) by 2029 with an average annual growth of 2.5 percent. These positive forecasts imply that tourism is a potential and worthwhile sector to be developed, especially for countries with abundant natural resources and cultural richness.

Despite the occasional shocks, the tourism industry has seen steady growth in international tourist arrival from 1990 to 2018 (See Figure 1.2). From 438 million travellers recorded in 1990, the number of travellers increased to 680 million in 2000, 890 million in 2005, 952 million in 2010, and 1,401 million in 2018. Notably, in 2012, the tourism sector surpassed its benchmark of 1 billion international tourist arrivals after 60 years of continuous expansion and diversification, making the sector as one of the largest and fastest growing sectors around the globe (UNWTO, 2013).

The continuous expansion of tourism sector had been trending upwards in international tourism receipt from 1990 to 2018 as shown in Figure 1.2. The international tourism receipt in 1990 was USD 271 billion, which increased to USD 416 billion in 1995, USD 495 billion in 2000, USD 703 billion in 2005, USD 966 billion in 2010 and USD 1,451 billion in 2018. International tourism receipt reached the USD 1,000 billion landmark in 2011 with a total of USD 1,081 billion received.

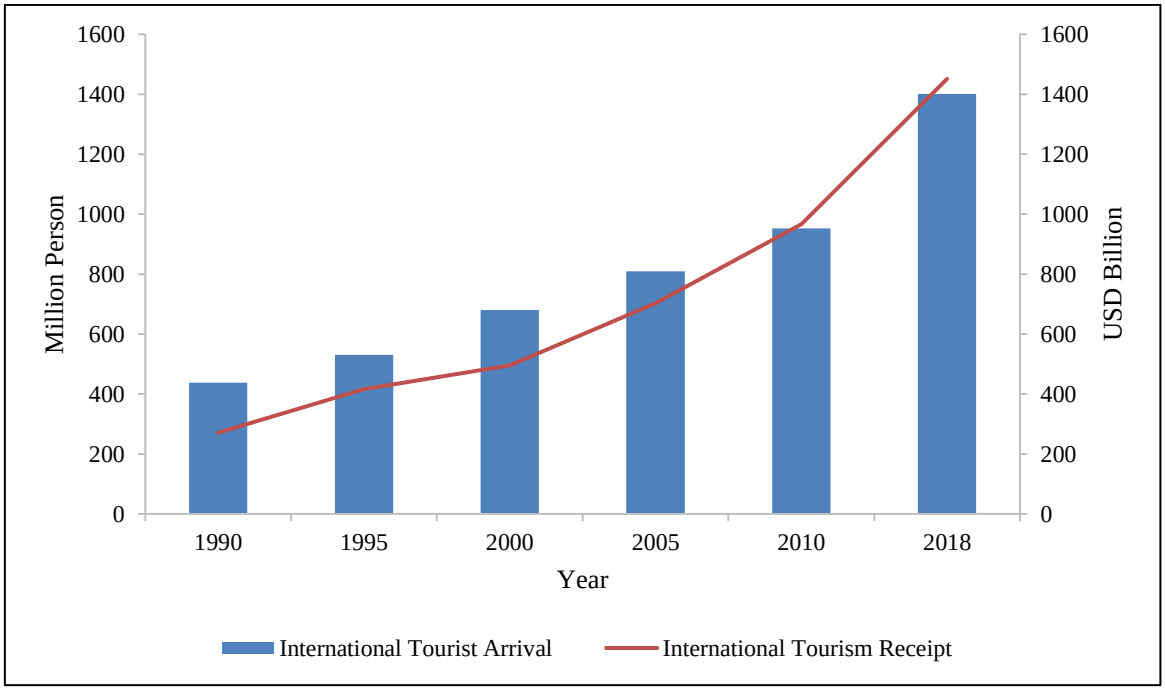

Figure 1.2: International Tourist Arrivals and International Tourism Receipts for World Tourism, 1990 to 2018
(Source: UNWTO, 2015; 2019)

From macro-economic point of view, the exports of a tourist destination country include the expenditure by international visitor, which is also an import for the tourist origin country. Tourism as a key category of international trade in service sector is able to generate income through international passenger transport services rendered to non-residents. In 2018, a total of USD 256 billion were receipted from passenger transport earning. The earnings had pushed the total value of tourism exports to USD 1.7 trillion (on average, USD 5 billion per day). Besides that, the share of international tourism in the world's exports in goods and services was 7 percent in 2018. Tourism was performing so well that it was ranked the third top performing world's export behind chemicals and fuel, and in front of food and automotive products (UNWTO, 2019).

1.3 Regional Tourism

The increase in the numbers of tourist destinations around the globe that have invested and expanded their tourism sector is apparent. Efforts have been continuously invested with the purpose of transforming the local tourism sector into a main engine for socio-economic progress through an anticipated increase in jobs, businesses, export revenues, and infrastructure. Apart from the historical expansion and diversification in the past 60 years, many new destinations have emerged away from those traditional favourites of Europe and North America, most notably in the Asia-Pacific region.

Table 1.1 shows the international tourist arrivals for all five regions, namely Europe, Asia and the Pacific, Americas, Africa and Middle East have generally increased from 1990 until 2018. Among these regions, Asia and the Pacific region has recorded a strong performance for the nine consecutive years since 2010 with an average growth of 6.16 percent - the highest among all regions. The recorded average growth for Europe was 4.8 percent), Americas was 4.6 percent, Africa was 3.6 percent, and Middle East was 0.9 percent for the same period as shown in Table 1.2.

In addition to an upward trend, it is noticeable that the growth of international tourist arrivals in the Asia-Pacific is more stable as compared with other regions. For the past two years, Asia-Pacific recorded a growth rate of 5.7 percent and 7.3 percent in 2017 and 2018, respectively. Meanwhile, reported respectively for 2017 and 2018, Europe had recorded 8.5 percent and 5.5 percent; the Americas recorded 4.7 percent and 2.3 percent; Africa recorded 8.5 percent and 7.0 percent; and lastly, the Middle East recorded 4.1 percent and 4.7 percent (refer Table 1.2). With all regions considered, only Asia-Pacific and Middle East had recorded improved growth from previous year.

Table 1.1: International Tourist Arrivals by Region, 1990-2018

Region	International Tourist Arrivals (million)						
	1990	1995	2000	2005	2010	2015	2018
Europe	216.5	304.5	386.6	453.2	489.4	607.7	710.0
Northern Europe	28.7	36.4	44.8	59.9	62.8	75.9	78.9
Western Europe	108.6	112.2	139.7	141.7	154.4	180	200.4
Central/ Eastern Europe	33.9	57.9	69.6	95.3	98.9	126.6	141.4
Southern/Mediterranean Europe	90.3	98	132.6	156.4	173.3	225.2	289.4
-of which EU-28	230.1	268	330.5	367.9	384.3	478.4	562.9
Asia and the Pacific	55.9	82.1	110.4	154	205.5	279.2	347.7
North-East Asia	26.4	41.3	58.3	85.9	111.5	142.1	169.2
South-East Asia	21.2	28.5	36.3	49	70.5	104.6	128.7
Oceania	5.2	8.1	9.6	10.9	11.4	14.2	17.0
South Asia	3.2	4.2	6.1	8.2	12.1	18.3	32.8
Americas	92.8	108.9	128.2	133.3	150.2	192.6	215.7
North America	71.8	80.5	91.5	89.9	99.5	127.6	142.2
Caribbean	11.4	14	17.1	18.8	19.5	23.9	25.7
Central America	1.9	2.6	4.3	6.3	7.9	10.3	10.8
South America	7.7	11.7	15.3	18.3	23.2	30.8	37.0
Africa	14.8	18.7	26.2	34.8	50.4	53.5	67.1
North Africa	8.4	7.3	10.2	13.9	19.7	18	23.9
Subsaharan Africa	6.4	11.5	16	20.9	30.7	35.4	43.3
Middle East	9.6	12.7	22.4	33.7	54.7	53.3	60.5

(Source: UNWTO Tourism Highlights, 2015; 2019)

Table 1.2: Market Share, Yearly Growth, Yearly Average Growth of International Tourist Arrivals by Region

Region	Market Share (%)	Growth (%)		Yearly Average (%)
	2018	2017/2016	2018/2017	2010-2018
Europe	50.7	8.6	5.5	4.8
Northern Europe	5.6	6.0	0.5	4.1
Western Europe	14.3	6.1	4.0	3.3
Central/Eastern Europe	10.1	5.6	5.0	4.7
Southern/Mediterranean Europe	20.7	12.9	8.2	6.3
-of which EU-28	40.2	7.7	4.4	5.0
Asia and the Pacific	24.8	5.7	7.3	6.6
North-East Asia	12.1	3.4	6.1	5.3
South-East Asia	9..2	8.8	6.8	7.8
Oceania	1.2	6.1	2.8	5.1
South Asia	2.3	6.2	19.4	10.5
Americas	15.4	4.7	2.3	4.6
North America	10.1	4.3	3.7	4.6
Caribbean	1.8	3.0	-1.4	3.5
Central America	0.8	4.6	-2.4	4.2
South America	2.6	7.7	1.2	5.8
Africa	4.8	8.5	7.0	3.6
North Africa	1.7	14.7	10.1	2.4
Subsaharan Africa	3.1	5.5	5.4	4.4
Middle East	4.3	4.1	4.7	0.9

(Source: UNWTO Tourism Highlights, 2019)

The corresponding international tourism receipts depicted a slightly different scenario. In 2018, Asia and the Pacific had received a total of USD 435.5 billion tourism receipts, this was behind the Europe at USD 570.5 billion, but far exceeded those of the Americas (USD 333.6 billion), Africa (USD 38.4 billion) and Middle East (USD 73.0 billion). The market share for Asia and the Pacific (30.0 percent) is also catching up with

Europe (39.3 percent) and leading the Americas (23.0 percent), Africa (2.6 percent) and Middle East (5.0 percent) (See Table 1.3).

Table 1.3: International Tourism Receipts and Market Share by Region, 2017-2018

Region	International Tourism Receipts		
	Market Share (%)	Receipts (USD billion)	
	2018	2017	2018
Europe	39.3	519.5	570.5
Northern Europe	6.5	88.6	93.9
Western Europe	12.9	170.9	187.5
Central/Eastern Europe	4.7	60.7	68.7
Southern/ Mediterranean Europe	15.2	199.3	220.4
-of which EU-28	33.1	438.6	480.7
Asia and the Pacific	30.0	396.0	435.5
North-East Asia	13.0	168.1	188.4
South-East Asia	9.8	130.6	142.3
Oceania	4.2	57.4	61.1
South Asia	3.0	39.9	43.6
Americas	23.0	325.8	333.6
North America	17.8	252.4	258.9
Caribbean	2.2	31.8	32.0
Central America	0.9	12.6	12.8
South America	2.1	29.0	29.9
Africa	2.6	36.4	38.4
North Africa	0.7	9.9	10.7
Subsaharan Africa	1.9	26.5	27.7
Middle East	5.0	68.4	73.0

(Source: UNWTO Tourism Highlights, 2019)

The strong and stable performance of the Asia and the Pacific region in recent years has proven that its tourism sector has the potential to compete with the traditional favourites like Europe and the Americas. This is particularly obvious when the comparison is made in terms of tourism receipts. In terms of international tourist arrivals, the Asia and the Pacific

(347.7 billion tourists) received far less tourists compared with Europe (710.0 billion tourists), but its outlook remains positive.

1.4 Country Background

The strong growth of the Asia and the Pacific region warrants a deep investigation and study into its regional and sub-regional tourism. South-East Asia is a sub-region that has been focusing and relying on its tourism sector as one of its major engines of growth. A distinctive advantage of South-East Asia is the rich and unique set of natural resources and cultural background in both rural and urban areas. To be precise, this sub-region has 11 natural and 17 cultural heritage sites registered with United Nations Educational, Scientific and Cultural Organization (UNESCO) World Heritage List. In terms of cultural richness, the region has one of the most interesting conglomerates of influences of the Arabian, Chinese, Indian and European culture. The appeal of its outstanding natural heritage, its rural landscapes, and it vibrant urban centres is further enriched by Buddhist, Hindu, Muslim and Christian religious traditions, and vernacular architecture, music, literature, and indigenous knowledge (ASEAN Secretariat, 2015).

The socio-economic stability and advancement of the region is achieved and maintained through an intense collaboration among the members of the Association of Southeast Asian Nations (ASEAN), including tourism. The aim is to promote the ASEAN countries as a single tourism destination, which is evident in the ASEAN Tourism Marketing Strategy (ATMS) 2012-2015 (ASEAN Secretariat, 2012) as well as the ASEAN Tourism Strategy Plan (ATSP) 2016-2025 (ASEAN Secretariat, 2015). In ATMS 2012-2015, one of its main attentions was to encourage tourists to visit more than one country within a single visit to the South-East Asian region.

In the recent ATSP 2016-2025, the vision for the next decade is to establish ASEAN as a single tourism destination with quality offerings of unique, and diverse experience that are responsible, sustainable, inclusive and balanced. This is an extension to the previous policy that is expected to further improve the socio-economic well-being of the ASEAN people. The latest promotion campaign was launched in 2017 to venerate the 50th anniversary of ASEAN, named *Visit ASEAN@50 Golden Celebration Campaign* to promote ASEAN to achieve this vision.

ASEAN is also interested in strengthening its influence as a regional grouping to have a large extra-regional trade component. Mohamed Ariff (2008) opined this can be achieved through several options available, which include expanding its membership, deepening intra ASEAN economic cooperation, strengthening extra ASEAN linkage, participating in wider regional grouping and forming a coalition with third countries. To date, strengthening extra ASEAN linkage has been carried out through establishing cooperation with individual country. The collaboration of ASEAN, China, Japan and Korea to form the ASEAN Plus Three (ASEAN+3) is also an effort to further strengthening its regional grouping status.

The tourism sector also falls under the radar of ASEAN+3. This was mentioned during the first meeting of the ASEAN+3 Tourism Ministers in Indonesia during 2002. The Ministers acknowledged that, by establishing a cooperation in tourism, this would encourage greater tourism exchange, social interaction and more integrated economic unity in East Asia. Under the ASEAN+3framework, cooperation in tourism covers the following aspects: i) facilitation of travel (includes the abolition of travel barriers); ii) tourism promotion between destinations in ASEAN and Plus Three Countries; iii) encouragement to private

sector cooperation; iv) joint programmes on human resources development, research, and information technology; and v) investment promotion.

The most recent development within the ASEAN+3 framework entails the signing of a memorandum of cooperation in January 2016. Tourism is identified as a significant instrument for promoting connection between people-to-people and this memorandum will serve its purpose as a mechanism that induces more action-oriented cooperation as well as to stimulate ASEAN+3 countries to improve facilitation of travel and tourist visits, development of quality tourism and linkages among education and training institutions.

Given the intense cooperation among the ASEAN+3 countries to develop the tourism sector, having an empirical study of the macroeconomic determinants that influence tourism demand in ASEAN+3 countries are undoubtedly beneficial to discover more insights for better policy formulation and cooperation. Beyond the ASEAN+3 countries, the study outcome is expected to also contribute towards achieving the vision of ATSP 2016-2025 for ASEAN countries.

1.4.1 International Tourist Arrival from 2000-2018[2]

Figure 1.3 shows the international tourist arrival for ASEAN countries from 2000 to 2018. International tourist arrival has been increasing steadily except during the SARS outbreak in 2003 and global financial crisis in 2009, which caused the rate to either

[2] Brunei redefined its definition of international tourist in 2004 and this has resulted in a structural break in it time series data. In 2003, the international tourist arrival was calculated at 944,130 visitors while in 2004, the international tourist arrival only amounted to 118,863 visitors.

slowdown or decline. Most ASEAN countries reached the highest recorded international tourist arrival in 2018.

The international tourist arrival for China, Japan and Korea from 2000 to 2018 are depicted in Figure 1.4 and reveals a similar trend with the ASEAN countries. Again, in 2003 and 2009, the rate declined due to the SARS outbreak and global financial crisis, respectively. In 2018, China, Japan and Korea received a total of 141.2 million; 31.1 million and 15.3 million international tourists, respectively. Figure 1.3 and 1.4 jointly showed that the international tourist arrivals at the ASEAN+3 countries were similar and increasing.

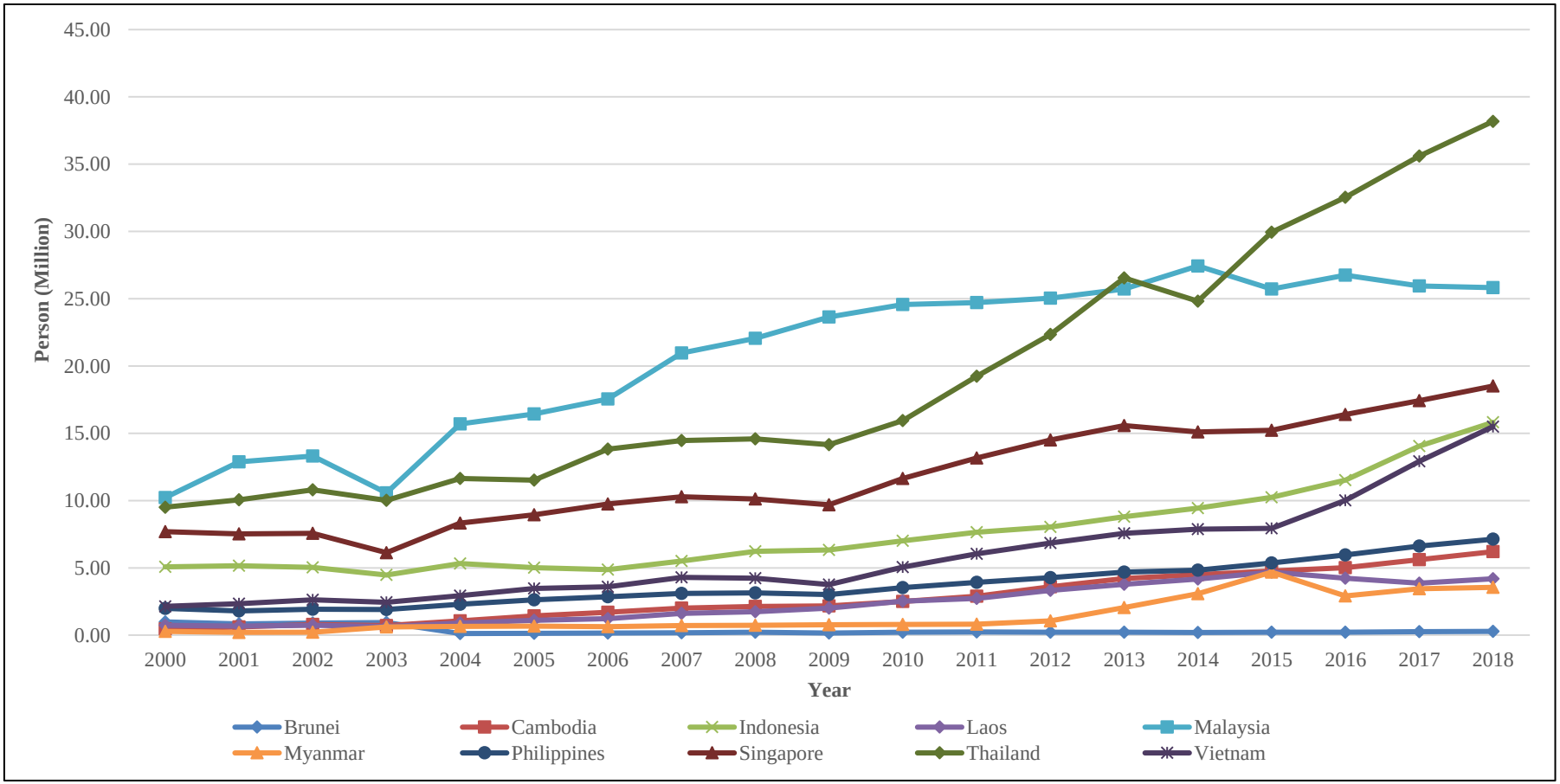

Figure 1.3: International Tourist Arrival in ASEAN Countries, 2000-2018
(Sources: Department of Statistics, Brunei, various issues; Ministry of Tourism, Cambodia, 2019; BPS-Statistics Indonesia, various issues; Ministry of Information, Culture and Tourism, Laos, various issues; Tourism Malaysia, 2019; Ministry of Hotels and Tourism, Myanmar, various issues; Philippine Statistics Authority, various issues; Singapore Tourism Analytics Network, 2019; Ministry of Sports and Tourism, Thailand, various issues; Vietnam National Tourism Administration, 2019)

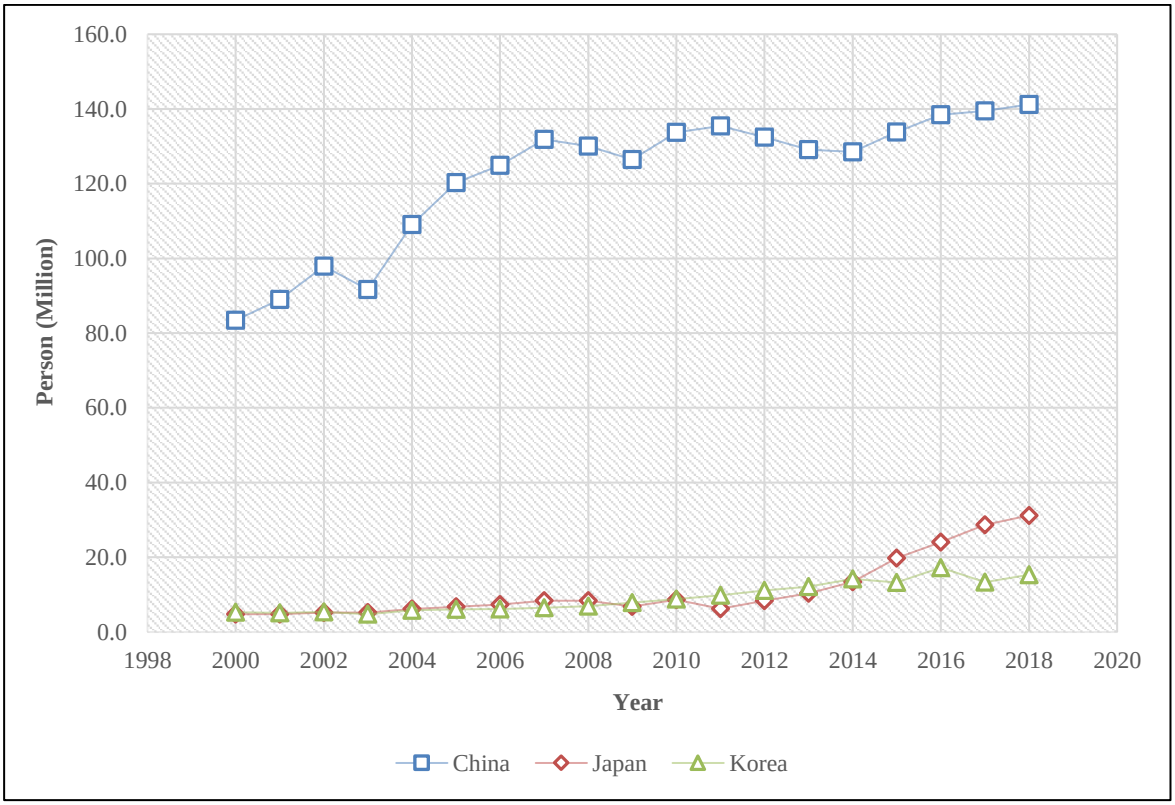

Figure 1.4: International Tourist Arrival in China, Japan and Korea, 2000-2018 (Sources: Ministry of Culture and Tourism of the People's Republic of China, various issues; Japan National Tourist Organization, 2019; Korea Tourism Organization, 2019)

1.4.2 Tourism Sector's Growth in Its Total Contribution to Gross Domestic Product (GDP)[3] from 2000-2019

Figure 1.5 shows the five-year average growth rate of the tourism sector's total contribution to ASEAN countries' GDP from 2000 to 2019. For Brunei, the tourism sector's

[3] WTTC defines tourism sector total contribution to GDP as the sum of direct, indirect and induced impacts of travel and tourism to GDP. WTTC measures direct contribution as those generated by industries that deal directly with tourist while indirect contribution is stated as tourism and travel's contribution in GDP that are supported by capital investment, government collective spending and supply-chain effects of the industry. Lastly, induced contribution is described as the spending of those who are directly or indirectly employed by travel and tourism industry.

contribution to GDP declined before going into negative in 2010 until 2014. The growth recovered in 2015 and continued to grow thereafter. Cambodia, on the other hand, showed a declining growth over the entire period of observation.

Indonesia, Laos, Malaysia and Vietnam tourism sectors' contribution to GDP dropped after the 2005-2009 period. Myanmar and Philippines tourism sectors' contribution improved after experiencing a negative growth in the first five-year period. However, the growth reduced from 2015 to 2019. Singapore's tourism sector contribution growth improved during the 2000-2014 period, but declined during the final five-year period. Thailand's tourism sector improved during the 2000 to 2009 period before experiencing a drop in 2010 to 2014. Nevertheless, the tourism sector's contribution increased again during the final five-year period.

Figure 1.6 depicts the five-year average growth rate of the tourism sector's contribution to GDP for China, Japan and Korea from 2000 to 2019. For China, its tourism sector's contribution growth increased for a prolonged period from 2000 until 2014 before it slowed down during the final five years. On the contrary, from 2000 to 2009, the tourism sector's contribution to GDP in Japan was slow, before it went into negative growth from 2010 to 2014. Nevertheless, the contribution growth increased again in the final five-year period. Korea experienced a decline in its tourism sector's contribution growth for a relatively short period from 2005 to 2009. Throughout the rest of the observed period, its contribution growth illustrated an improving trend. Albeit with different degree of fluctuation, it can be observed that the tourism sector's five-year average growth rate in its total contribution to GDP has declined during the observed period for most ASEAN+3 countries.

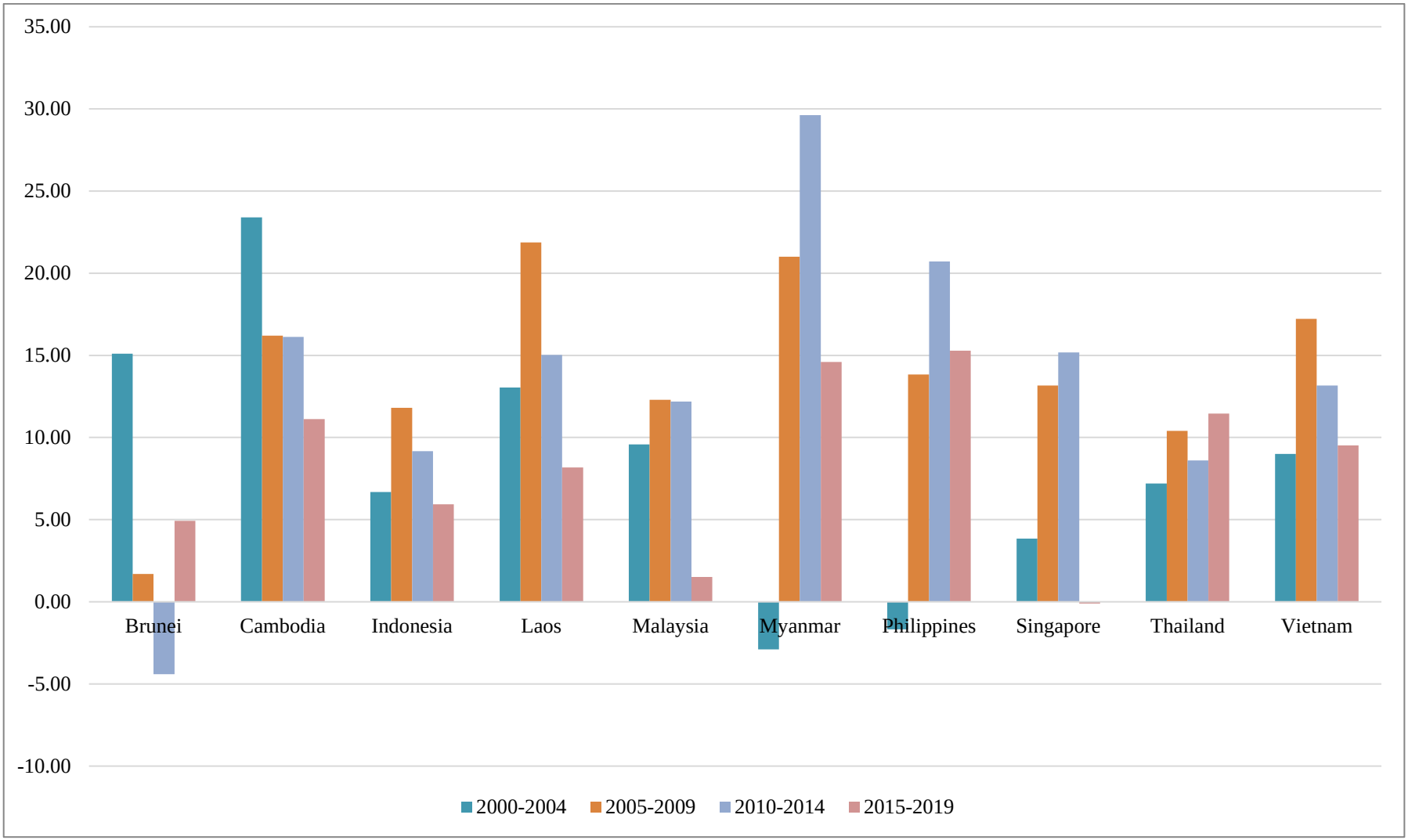

Figure 1.5: Tourism Sector's Growth in Its Total Contribution to Gross Domestic Product in ASEAN Countries, 2000-2019 (Sources: WTTC, 2019b)

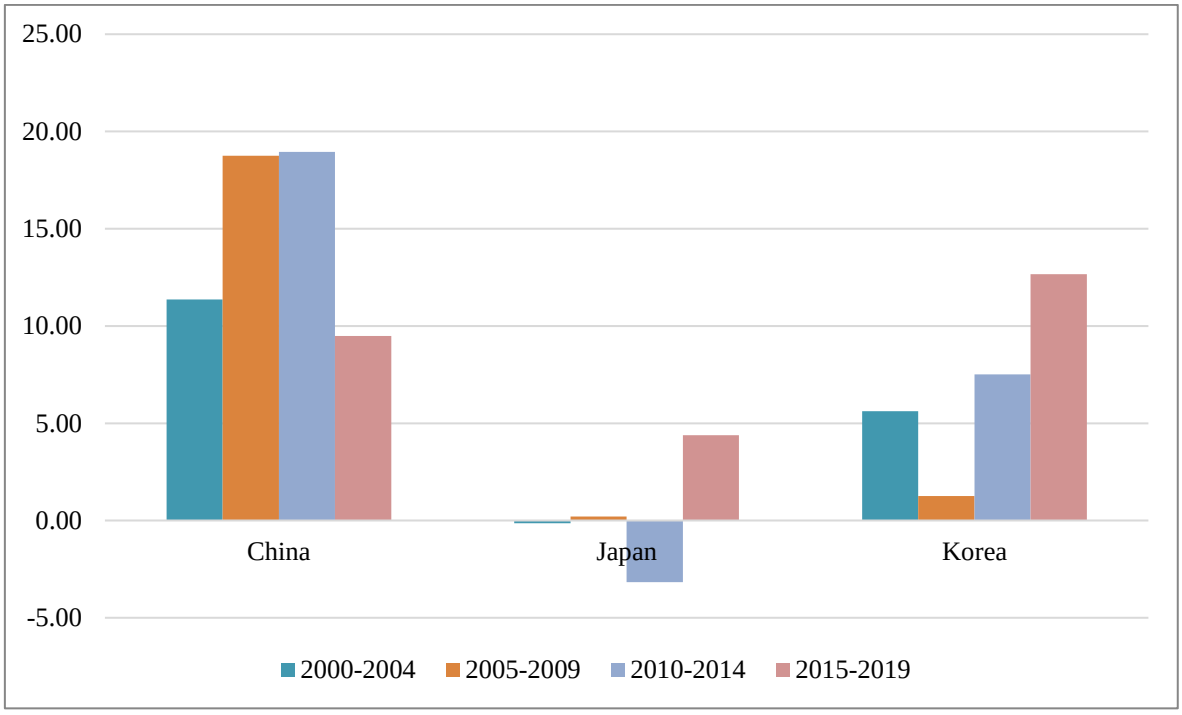

Figure 1.6: Tourism Sector's Growth in Its Total Contribution to Gross Domestic Product in China, Japan and Korea, 2000-2019
(Sources: WTTC, 2019b)

1.4.3 Tourism Sector's Growth in Its Total Contribution to Employment from 2000-2019

Figure 1.7 depicts the five-year average growth of the tourism sector in its total contribution to employment in the ASEAN countries from 2000-2019. Brunei showed a prolonged period of declining contribution from 2005 to 2014 after some growth in the first five-years. Nonetheless, the growth recovered strongly from 2015 to 2019. In Cambodia, the growth has shown a generally downward trend, which improved in 2010 until 2014. Indonesia and Myanmar's tourism sector contribution growth rate were similar. Both countries' tourism sector showed improvement after a prolonged period of negative growth from 2000 to 2009. Laos's tourism sector depicted a consistent growth in total contribution to employment throughout the observed period except for a slowdown from 2010 to 2014 only.

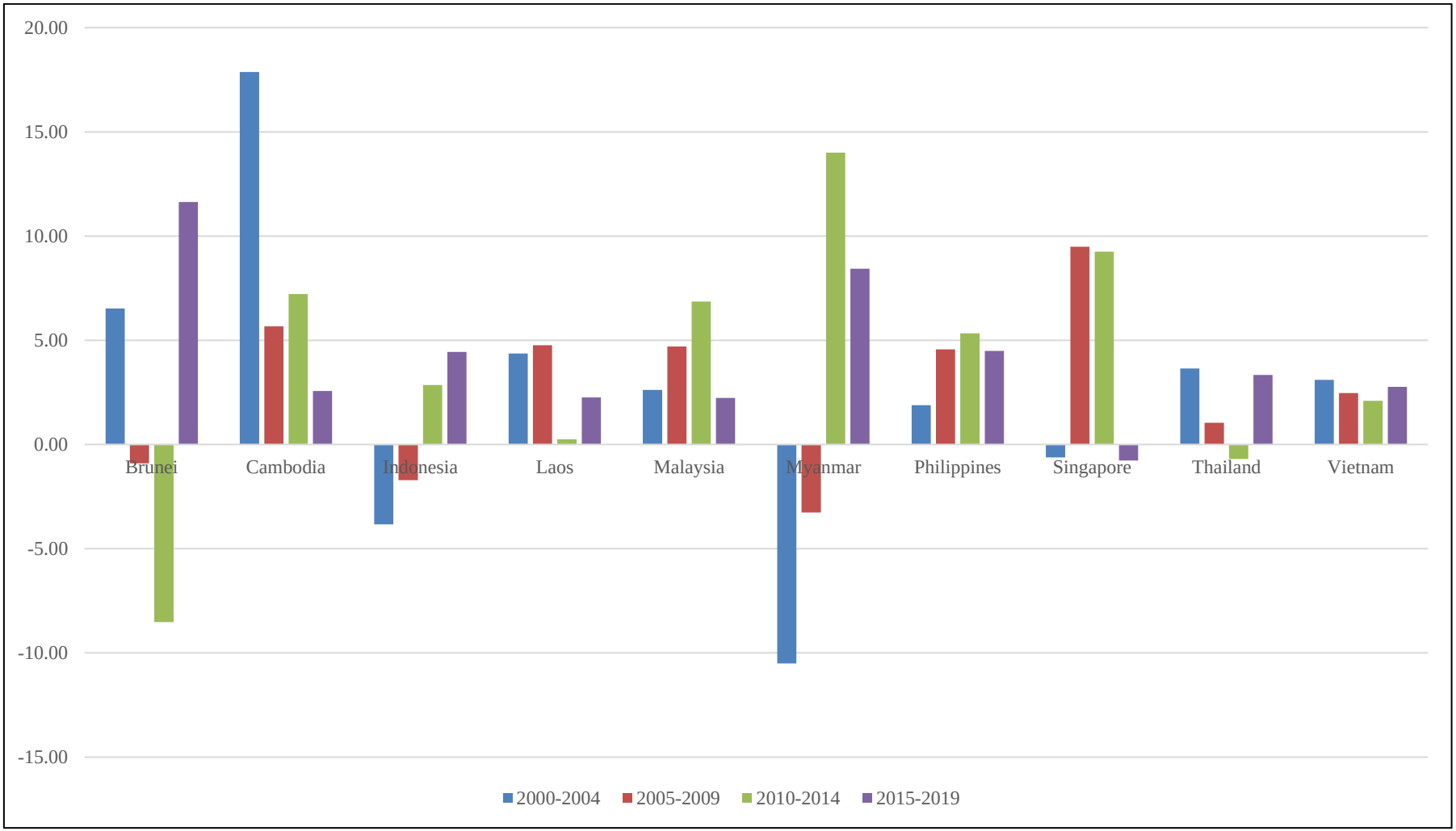

Figure 1.7: Tourism Sector's Growth in Its Total Contribution to Employment in ASEAN Countries, 2000-2019 (Sources: WTTC, 2019b)

Malaysia and Philippines, on the other hand, initially depicted a promising upward trend that reached its optimal growth rate during the 2010-2014 period. Since then, the growth had declined. Singapore experienced negative growth during the first and last five years of the observed period. Aside from that, the growth from 2005 to 2014 was strong. Thailand showed a downward trend from 2000 to 2009 before the growth went into negative from 2010 to 2014 and finally bounced back strongly from 2015 to 2019. Similar with Thailand, Vietnam also illustrated a downward trend in its tourism sector's contribution to employment albeit with less fluctuation.

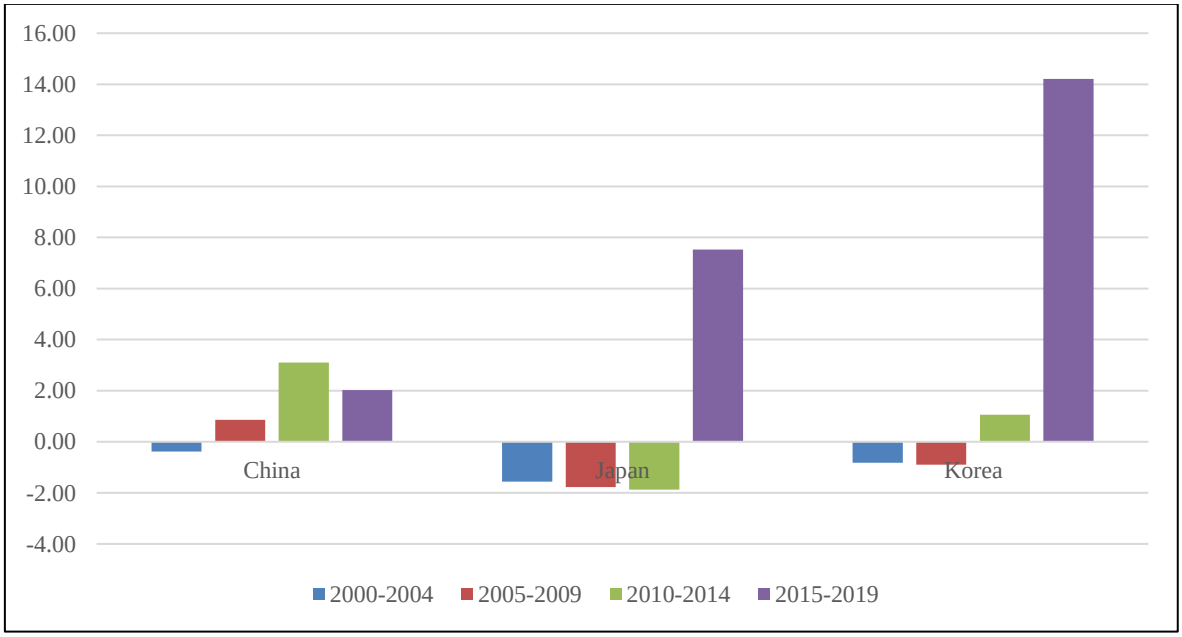

Figure 1.8: Tourism Sector's Growth in Its Total Contribution to Employment in China, Japan and Korea, 2000-2019
(Sources: WTTC, 2019b)

Figure 1.8 illustrates the tourism sector's five-year average growth in its total contribution to employment in China, Japan and Korea from 2000 to 2019. During the first five years, the growth in China was negative before it improved significantly for the next ten years and slightly slowed down during the last five years. Meanwhile, Japan has gone

through a period of negative growth from 2000 to 2014 before it recovered strongly during the last five years. Similar to Japan, Korea also experienced a period of negative growth from 2000 to 2009. The growth improved since 2010 and was strong during the last five years. It is noticeable that most ASEAN+3 countries' five-year average growth rates were volatile and fluctuated in varying degrees.

1.4.4 Top Ten Tourist Origin Countries Based on 2017 Ranking

The top ten tourist origin countries for ASEAN and Plus Three countries (China, Japan and Korea) in year 2018 are presented in Table 1.13 and Table 1.14, respectively. The figures are presented in both number of international tourist arrivals and its corresponding percentage. For Brunei, the top ten countries had contributed 84.37 percent of total visitors. This was up to 79.51 for Cambodia, 64.49 percent for Indonesia, 94.94 percent for Laos, 87.11 percent for Malaysia, 79.11 percent for Myanmar, 80.99 percent for Philippines, 74.10 percent for Singapore, 67.91 percent for Thailand, 82.99 percent for Vietnam, 89.54 percent for China, 88.18 percent for Japan and 82.54 percent for Korea. In brief, the international tourist from top ten tourist origin countries in ASEAN+3 countries have counted to at least 64 percent of total international tourist arrivals.

Table 1.4: Top Ten Tourist Origin Countries for ASEAN Countries

Year	2018	2018	2018	2018	2018	2018	2018	2018	2018	2018
Destination	Brunei	Cambodia	Indonesia	Laos	Malaysia	Myanmar	Philippines	Singapore	Thailand	Vietnam
	Person	Person	Person	Person	Person	Person	Person	Person	Person	Person
Ranking	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*
	China	China	Malaysia	Thailand	Singapore	Thailand	Korea	China	China	China
1	65,563	2,024,443	2,503,344	1,929,934	10,615,986	291,231	1,587,959	3,417,604	10,534,340	4,966,468
	23.57	*32.65*	*15.83*	*46.10*	*41.10*	*20.83*	*22.28*	*18.47*	*27.59*	*32.05*
	Malaysia	Vietnam	China	Vietnam	Indonesia	China	China	Indonesia	Malaysia	Korea
2	59,981	800,128	2,139,161	867,585	3,277,689	333,085	1,255,258	3,021,455	4,032,139	3,485,406
	21.57	*12.90*	*13.53*	*20.72*	*12.69*	*23.82*	*17.61*	*16.32*	*10.56*	*22.49*
	Philippines	Laos	Singapore	China	China	Japan	United States	India	Korea	Japan
3	22,319	426,180	1,768,744	805,833	2,944,133	104,376	1,034,396	1,442,277	1,796,401	826,674
	8.02	*6.87*	*11.19*	*19.25*	*11.40*	*7.47*	*14.51*	*7.79*	*4.71*	*5.33*
	Indonesia	Thailand	Australia	Korea	Thailand	United States	Japan	Malaysia	Laos	Taiwan
4	27,462	382,317	1,301,478	174,405	1,914,692	65,057	631,801	1,254,022	1,664,456	714,112
	9.87	*6.17*	*8.23*	*4.17*	*7.41*	*4.65*	*8.86*	*6.78*	*4.36*	*4.61*
	Singapore	Korea	Japan	United States	Brunei	Korea	Australia	Australia	Japan	United States
5	14,091	301,770	530,573	49,178	1,382,031	72,852	279,821	1,107,224	1,655,996	687,226
	5.07	*4.87*	*3.36*	*1.17*	*5.35*	*5.21*	*3.93*	*5.98*	*4.34*	*4.43*
	United Kingdom	USA	India	France	India	Singapore	Taiwan	Japan	India	Russia
6	11,966	250,813	595,636	39,315	600,311	58,657	240,842	829,676	1,595,754	606,637
	4.30	*4.04*	*3.77*	*0.94*	*2.32*	*4.20*	*3.38*	*4.48*	*4.18*	*3.91*
	Korea	Japan	Korea	Japan	Korea	Vietnam	Canada	Philippines	Russia	Malaysia
7	9,125	210,471	358,885	38,985	616,783	53,329	226,429	778,141	1,472,765	540,119
	3.28	*3.39*	*2.27*	*0.93*	*2.39*	*3.81*	*3.18*	*4.20*	*3.86*	*3.49*

Table 1.4 continued

8	India	Malaysia	United Kingdom	United Kingdom	Japan	France	United Kingdom	Korea	United States	Australia
	8,635	201,116	392,112	26,801	394,540	43,218	201,039	629,454	1,122,088	386,934
	3.10	*3.24*	*2.48*	*0.64*	*1.53*	*3.09*	*2.82*	*3.40*	*2.94*	*2.50*
9	Australia	United Kingdom	United States	Germany	Philippines	United Kingdom	Singapore	United States	Singapore	Thailand
	9,702	162,395	387,856	22,915	396,062	36,609	171,795	643,268	1,066,219	349,310
	3.49	*2.62*	*2.45*	*0.55*	*1.53*	*2.62*	*2.41*	*3.48*	*2.79*	*2.25*
10	Thailand	France	Philippines	Australia	United Kingdom	Malaysia	Malaysia	Vietnam	United Kingdom	United Kingdom
	5,828	170,844	217,874	19,607	361,335	47,632	143,566	591,614	986,666	298,114
	2.10	*2.76*	*1.38*	*0.47*	*1.40*	*3.41*	*2.01*	*3.20*	*2.58*	*1.92*
Subtotal	234,672	4,930,477	10,195,663	3,974,558	22,503,562	1,106,046	5,772,906	13,714,735	25,926,824	12,861,000
	84.37	*79.51*	*64.49*	*94.94*	*87.11*	*79.11*	*80.99*	*74.10*	*67.91*	*82.99*
Total	278,136	6,201,077	15,810,305	4,186,432	25,832,354	1,398,098	7,127,678	18,508,302	38,178,194	15,497,791
	100	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*

(Source: Department of Statistics, Brunei, 2019; Ministry of Tourism, Cambodia, 2019; BPS-Statistics Indonesia, 2019; Ministry of Information, Culture and Tourism, Laos, 2018; Tourism Malaysia, 2019; Ministry of Hotels and Tourism, Myanmar, 2019; Department of Tourism, Philippine, 2019; Singapore Tourism Analytics Network, 2019; Ministry of Sports and Tourism, Thailand, 2019; Vietnam National Tourism Administration, 2019)

Table 1.5: Top Ten Tourist Origin Countries for China, Japan and Korea

Year	2018	2018	2018
Destination	China	Japan	Korea
	Person	Person	Person
Ranking	*Percent*	*Percent*	*Percent*
	Hong Kong	China	China
1	79,370,000	8,380,034	4,789,512
	56.21	*26.87*	*31.21*
	Macau	Korea	Japan
2	25,150,000	7,538,952	2,948,527
	17.81	*24.17*	*19.21*
	Taiwan	Taiwan	Taiwan
3	6,136,081	4,757,258	1,115,333
	4.35	*15.25*	*7.27*
	Korea	Hong Kong	United States
4	4,193,500	2,207,804	967,992
	2.97	*7.08*	*6.31*
	Japan	United States	Hong Kong
5	2,691,400	1,526,407	683,818
	1.91	*4.89*	*4.46*
	Russia	Thailand	Thailand
6	2,415,500	1,132,160	558,912
	1.71	*3.63*	*3.64*
	United States	Australia	Philippines
7	2,484,600	552,440	460,168
	1.76	*1.77*	*3.00*
	Mongolia	Malaysia	Vietnam
8	1,494,300	468,360	457,818
	1.06	*1.50*	*2.98*
	Malaysia	Philippines	Malaysia
9	1,291,500	503,976	382,929
	0.91	*1.62*	*2.50*
	Philippines	Singapore	Russia
10	1,205,000	437,280	302,542
	0.85	*1.40*	*1.97*
Subtotal	126,431,881	27,504,671	12,667,551
	89.54	*88.18*	*82.54*
Total	141,198,300	31,191,856	15,346,879
	100	*100*	*100*

(Source: Ministry of Culture and Tourism of the People's Republic of China, 2019; Japan National Tourist Organization, 2019; Korea Tourism Organization, 2019)

1.5 Problem Statement

The statistics showed earlier clearly implied that the tourism sector plays an important role in the economy development of ASEAN+3 countries. Apart from contributing significantly to a country's income, it generates foreign exchange earnings, creates employment opportunities, and supports other industries through its industry multiplier effect. Yazdi and Khanalizadeh (2017) expressed that, in economy where tourism is vital, even short-term rejuvenation of the tourism sector is able to contribute more to the economy than other sectors such as agriculture and mining that require a longer period due to the substantial structural changes necessary to recuperate these sectors.

Despite receiving growing tourist arrivals in recent years, the tourism sector has generally shown a fluctuating performance accompanied by a gradual slowdown in its contribution to national income and employment in some ASEAN+3 countries. Without prior knowledge of the possible determinants of tourism demand, the governments will undoubtedly face considerable challenges to respond appropriately and effectively.

The consequences of a missed opportunity can be many folds. Firstly, the government may end up wasting the efforts and resources invested into the tourism sector. As a result, the tourism sector will fail to reach its optimal performance, bringing optimal contribution towards the economy within the nation and for the region. Secondly, this dampens the ASEAN's aim to become a single tourist destination as documented in the ATSP 2016-2025. In addition, the cooperation between ASEAN+3 will also be affected.

The ASEAN+3 countries have been welcoming millions of international visitors from numerous countries. During their visits, these tourists need to interact actively with the local communities. As such, if the language barrier is not addressed effectively, the it will become a hindrance to tourism development. This has been highlighted by Chockalingam and Ganesh (2010), in which the study has identified the local language as one of the 15 exclusive tourist problems identified in Tamil Nadu, India.

Okafor et al. (2018) concluded that language has significant influence on international tourism because it can either improve or destroy the experience. The absence of a common language to bridge the communication and interaction between the international tourists and the local community is known as the language barrier. Hara (2013) stated that negative perception of language barrier influences tourists' travelling decision. This influence is far greater when the tourist has never travelled to a destination or learnt about its language. The same conclusion has been reached by Biswas and Mamun-Or-Rashid (2019).

Language barrier presents an unfavourable condition for ASEAN+3 countries because these countries are famous for their unique culture, evident through the documentation of 17 cultural heritage sites within the South-East Asia sub-region. If a language barrier exists and cannot be adequately solved, tourists will become dispirited since they cannot understand, least appreciate, the diversity in culture. This hinders a tourist from visiting, re-visiting or recommend a country as a vacation destination.

The solution to language barrier is a common language. Within the ASEAN region, some countries share some common languages due to a historical linkage such as a colonization or international migration. The role of common language in attracting tourist arrival is often taken as a dummy variable in tourism demand studies. However, from empirical point of view, the use of a dummy variable poses a problem especially for time series analysis, since the dummy variable is time invariant. The use of dummy variable also disallows empirical studies to examine the effect of language proficiency.

In addition, as time passes, the effect of a common language may not be as influential as it used to be. This is because most languages will be localised to the home nation's national language and lost its originality over time. Is common language still influential in attracting tourist arrival under such scenario? From practical point of view, heavy reliance on common language established from past relationship is unrealistic because countries that does not cross path in the past would not share a common language. Subsequently, effect of common language cannot be determined for these countries.

Theoretically, English is the international language and a common language (lingua franca) to bridge the communication between tourists and local community. Erazo et al. (2019) stated that English is a crucial element in tourism sector because it almost guarantees the best service for foreign visitors who do not understand the local language. The function of English as a lingua franca in international tourism has also been reiterated in Zahedpihseh et al. (2017), which stated that it is so important until it resulted in a boom in teaching English.

However, English is not the flawless solution to language barrier, particularly in countries with low English literacy as emphasized by Biswas and Mamun-Or-Rashid (2019). For this reason, using English as a common language may result in different results. Hence, the role of English is still ambiguous despite its general acceptance as the international language. Is English able to function as a common language to attract international tourists? How does English literacy level affect the choice of English as a common language? This holds true for the ASEAN+3 countries whereby the English literacy level differs widely across these countries.

In short, tourism sector has been contributing significantly to ASEAN+3 countries in their economy development and regional collaboration. Language as a medium of communication plays an important role that will influence the experience of tourist during their trip. In order to attract more international tourist into these countries, the role of language needs to be properly assessed. This will allow tourism authorities to develop suitable policy to develop the tourism sector in enhancing their performance and subsequently offer a pleasure experience to the tourists.

1.6 Objectives

1.6.1 General Objective

The general objective of this study is to assess the role of language as a determinant of international tourism demand of ASEAN+3 countries.

1.6.2 Specific Objectives

The specific objectives of this study are as follow:

i. To assess the influence of common language on international tourism demand model;

ii. To determine the role of English as a lingua franca in international tourism demand model; and

iii. To identify the effect of English as a common language for tourism destination with different level of English proficiency.

1.7 Significance of the Study

The tourism sector has been one of the economy drivers for ASEAN+3 countries throughout their economic development history, making significant contribution to the countries' income and employment. However, the statistics presented in earlier section showed that the tourism sector's contributions to both national income and employment are unstable in the long term. A study is needed to understand the international tourism demand determinants for the ASEAN+3 countries to improve and sustain the tourism sector's performance.

At the same time, this study focused on the role of language as a determinant in international tourism demand model for ASEAN+3 countries. This study was intended to find out whether the attractiveness of common language has declined over time, after it has been localized with the local language? Practically, it is only rational for ASEAN+3 countries to promote common language as one of its attractiveness when they are certain that a common language does attract tourist arrival.

From an empirical perspective, the common language is usually represented by a dummy variable in a tourism demand study. However, an inherent problem lies in its invariance to time, which means it cannot be estimated in a time series analysis. Therefore, the effect of a common language cannot be investigated for a country using time series analysis. Another problem, both practically and empirically, is that the role of language cannot be determined when tourist origin and destination countries do not share a common language. Hence, this study attempted to fill up this gap by investigating the role of English as a common language.

From a theoretical perspective, English with its international status is able to function as a common language and attract international tourist arrival by removing the language barrier. English is widely taught as a second language throughout the world and this guarantee its identical use and acceptance globally. Nevertheless, since the access to education and learning varies around the world, it is necessarily to find out the suitability of English as a common language in attracting international tourist arrival. It is also crucial to determine how English literacy level affects the choice of English as a common language, particularly within the ASEAN+3 countries due to a difference in English literacy level over the region. This difference may have different implications in attracting international tourist arrival.

1.8 Structure of the Study

The structure of this thesis has been arranged in the following manner. Chapter 2 provides a review of previous literature on tourism studies. Chapter 3 discusses the methodology used in this study to examine the relationships among the variables under

study. Chapter 4 presents and discusses the empirical findings. Lastly, Chapter 5 concludes this study with some policy recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter presents the literature review conducted on a bank of related models and studies that have been formulated in relation to the topic. The structure is divided into four sections. The first section concentrates on the previous survey of literature as an introduction. Section 2.3 briefly discusses the gravity model of international tourism demand. Section 2.4 focuses on reviewing the tourism demand literature and lastly, the concluding remarks shall be presented in Section 2.5.

2.2 Introduction

The published literature on tourism demand have seen a few authors concentrating their efforts on providing an overall and updated review on the progress in the development of related studies. These have become valuable guidance to many. Among others, Li et al. (2005) had focused on 84 post-1990 empirical studies of international tourism demand applying econometric techniques while Song and Li (2008) had surveyed 121 papers published on tourism modelling and forecasting since 2000. Song et al. (2012), on the other hand, provided a more updated survey by including the studies published up to 2011. Li et al. (2005) stated that the two major types of tourism demand studies concentrated on analysing the effects of determinants in tourism demand and accurate forecasting of the future tourism demand. Their literature review showed that annual data was the most used data frequency followed by quarterly and monthly frequencies. The increasing interest in using quarterly data was associated with analysing seasonality in international tourism

demand. With regard to such pattern, Song and Li (2008) stated that most studies used annual data probably because explanatory variables at higher frequency are difficult to obtain.

Li et al. (2005) revealed that throughout the papers published in 1990-2004, the Western Europe and North American countries had received most attention in studies prior to 1990. This happened because these countries accounted for a very large proportion of global tourism flows. Nonetheless, countries in East Asia-Pacific region had shown a steadfast growth in international tourism in 1990s. Similar pattern was found in the survey of Song and Li (2008) as well. However, the measures of tourism demand have not changed much in the 1990s and 2000s. The most popular measure was still tourist arrival and followed by tourist expenditure, in aggregate and per capita form. The only noticeable new trend was that the studies in the 1990s started to focus on disaggregated tourism markets by travel purpose. In other words, tourism demand was now subcategorised into different group of tourists according to their visiting purposes.

Song and Li (2008) further confirmed this finding when their survey revealed that tourist arrivals are detailed into different types of arrival such as holiday tourist arrivals, business tourist arrivals, tourist arrivals for visiting friends and relatives, and so forth. In addition, the tourism expenditure has also been disaggregated to expenditure on certain tourism products categories such as, meal expenditure, sightseeing experience and shopping, and more. Apart from that, Song et al. (2012) found out that the length of stay has been used as an alternative measurement in recent tourism demand studies. The reason behind this particular interest was because different market segments are associated with different influencing factors and varying decision-making processes. As such, a disaggregated level study provides more precise insights into the features of a particular market segment. As a

result, more specific and accurate information can be provided to develop efficient marketing strategies.

Meanwhile, both Li et al. (2005) and Song and Li (2008) found out that the widely used explanatory variables in modelling tourism demand are income, tourism price, substitute price, travel costs, exchange rates, dummy variable and deterministic trends. Song and Li (2008) also documented that the quantitative approach is more widely applied as compared to qualitative approach in tourism demand study. Additionally, Song et al. (2012) has pointed out that methodology in tourism demand study has diverted from the use of static regression models to more advanced dynamic specifications. It is highlighted that future studies should explore more into the dynamic panel data techniques as well as other more advanced methodologies in modelling tourism demand.

2.3 Gravity Model of International Tourism Demand

The application of gravity model was popularised in the 1960s and widely applied in the explanation of international trade flows, migration and foreign direct investment. In this particular model, bilateral flows between two countries are directly proportional to both countries' economic masses and inversely proportional to the distance between them. Initially, the gravity model was criticised for lack of theoretical foundation, but researchers in different eras have shown that gravity model can be derived from the Ricardian, Heckscher-Ohlin and monopolistic competition models (Kimura & Lee, 2006).

Considering tourism is a type of services trade, gravity model had also been applied in studying the flow of international tourist when gravity model was first introduced. Nonetheless, such applications were soon being criticised for its lack of theoretical foundation. As stated by Sheldon and Var (1985, as cited in Morley et al., 2014), the gravity

model in its initial formulation predicted that tourism flows from Region A to Region B are the same as those from Region B to Region A, a condition that is not common for the case of tourism flows.

The criticism has led to gravity model applications in the area of tourism demand or tourism flows no longer favourable and soon disregarded starting from the 1980s and 1990s. This situation is evident when the literature surveys done by Li et al. (2005), Song and Li (2008), and Song et al. (2012) no longer mentioned gravity model in their reviews. Nevertheless, the application of gravity model in tourism demand study has resurfaced in the 2000s when researchers are able to provide theoretical supports and proofs. Among others, Keum (2010) tested international trade theories on international tourism and trade flows for comparison, and the author concluded that the gravity model used was robust for both flows. In a more recent study, Morley et al. (2014) have provided a theoretical foundation of the application of gravity model in tourism demand based on consumer's utility theory and further proved that the application in tourism demand studies is valid and robust[4].

On a wider scope, Kimura and Lee (2006) proved that the gravity model had performed better for the case of international services trade than goods trade in their efforts to compare the performance of gravity model in estimating the impact of factors affecting both trades. Since the 2000s, there are many researchers that have adopted the gravity model including Kimura and Lee (2006); Khadaroo and Seetanah (2008); Keum (2010); Yang et al. (2010); Massida and Etzo (2012); Ahmad Kosnan et al. (2013); Chasapopoulos et al. (2014); Kunroo and Azad (2015); Lorde et al. (2015); Santeramo and Morelli (2015); Alawin

[4] See Morley et al. (2014) for further discussion.

and Abu-Lila (2016); Karaman (2016); Rossello et al. (2017); Yazdi and Khanalizadeh (2017); Okafor et al. (2018); Puah et al. (2019); Tatoglu and Gul (2020); and Ulucak et al. (2020).

The gravity model draws its inception from the Newton's law of gravitation in the seventeenth century. This law explains that two bodies are subjected to a force of attraction that depends positively on the product of their masses and negatively on their distance. This law was applied to social phenomena by social scholars in the nineteenth century on the studies of migration, international trade and foreign direct investment. A more modern version can be attributed to the basic gravity model developed by Tinbergen (1962), which states that trade flow between two countries is determined by their economic size and distance between these two countries. To be more specific, trade flow is positively associated with national income of both countries (proxy for economic size) and negatively associated with distance between two countries.

Applying gravity model into a tourism demand study would see trade flow being specified to tourism demand. Tourism demand is determined by national incomes of origin and destination countries and distance between them. The next section shall provide discussion on the measurements of variables as well as the determinants utilised in tourism demand study.

2.4 Literature Review

2.4.1 Measurement of Tourism Demand

Consistent with the findings of Li et al. (2005), Song and Li (2008), and Song et al. (2012), the literature reviewed in this study also revealed that most tourism demand studies have utilized the tourist arrival as the dependant variable, followed by tourist expenditure (or tourism receipts) and length of stay. For example, Dritsakis (2004) used tourist arrival from Germany and Great Britain to study the tourism demand in Greece.

Halicouglu (2004); Aguilo et al. (2005); Eita et al. (2011); Ekanayake et al. (2012); Lorde et al. (2015); Perles-Ribes et al. (2017); and Wamboye et al. (2020) have used tourist arrival as the dependent variable in their studies for different countries. Halicouglu (2004) studied tourism demand for Turkey; Aguilo et al. (2005) for Balearic Islands of Spain; Eita et al. (2011) for South Africa; Ekanayake et al. (2012) for United States of America (USA); Lorde et al. (2015) for Caribbean countries; Perles-Ribes et al. (2017) for Spain; and Wamboye et al. (2020) for Tanzania.

Another popular measurement for tourism demand is the tourist expenditure. This was used by Divisekera (2003) in estimating tourism demand model for United Kingdom, Australia, New Zealand and USA; Botti et al. (2007) for France; Brida et al. (2008) for Mexico; Dritsakis (2012) for Spain, France, Italy, Greece, Turkey, Cyprus and Tunisia; and Ahmad Kosnan et al. (2013) for Malaysia, for example. The third most commonly used dependent variable is length of stay, as seen in the studies of Garin-Munoz (2007) for Spain; Athanasopoulus and Hyndman (2008) for Australia; Garin-Munoz (2009) for Galicia of Spain; Falk (2010) for Austria; and Yang et al. (2010) for China in their tourism demand model.

2.4.2 Determinants of Tourism Demand

2.4.2.1 Destination Country's Economic Size

The gravity model applied in tourism demand study disclosed that tourism demand is determined by the economic size or mass of origin and destination countries, and distance between the pairs of country. Generally, a country's economic size is measured by its national income represented by gross domestic product (GDP). In tourism study, there is a popular genre that focuses on the relationship of economy performance with tourism demand, known as the tourism-led-growth hypothesis. Katircioglu (2009), Dritsakis (2012), Eeckels et al. (2012) and Perles-Ribes et al. (2017) are among the researchers that investigated this hypothesis.

While these studies attempted to verify the tourism-led-growth hypothesis, where the tourism sector performance was expected to stimulate a country's income level, the opposite was found to be true as well. Perles-Ribes et al. (2017) found a bidirectional relationship between tourism demand and economy performance; and Pablo-Romero and Molina (2013) revealed that nine out of 87 studies they reviewed have identified unidirectional relationship from economy growth to tourism demand. These findings have highlighted the role of a nation's economy growth as a factor responsible for bringing in more international tourists.

Higher income nation provides better country image that attracts tourist arrival. To be more specific, high income country has a more stable exchange rate; lower crime rate; more frequent air flights to choose from; more secured and safer environment; and many more factors that encourage international tourists to come and enjoy their stay. Alawin and Abu-Lila (2016) indicated that higher tourist destination's national income means that the economy growth has led to infrastructure development and this also includes upgrades of

accommodation and facilities that will attract more tourists. In fact, many studies have included destination country's income in modelling their tourism demand model, and most measured this in terms of GDP.

The variants used in tourism demand studies to measure national income include GDP; real GDP; GDP per capita; and real GDP per capita. Ahmad Kosnan et al. (2013) used GDP to represent Malaysia's income level and discovered that higher destination country's income attracts international tourist. Katircioglu (2009) estimated Turkish tourism demand and utilised real GDP as proxy for income level, but the author did not find enough evidence to support the tourism-led-growth hypothesis. Keum (2010) also used real GDP and found that Korea's income attracts tourist arrival. Meanwhile, GDP per capita was adopted in the Lorde et al. (2015) study for Caribbean tourism demand. Real GDP per capita was utilised by Rossello et al. (2017) in their tourism demand study for a panel of 196 countries. Both studies found expected positive effect of destination countries' income on tourist arrival.

2.4.2.2 Origin Country's Economic Size

The second explanatory variable in the gravity model is the tourist origin country's economic size. Botti et al. (2007) stated that consumer demand function is associated with income and price. When applied to tourism studies, it can be interpreted as the tourism demand is affected by the income of tourist origin countries and the price level at the destination. Similarly, Kadir et al. (2013) also documented that income and price type factors play a significant role in affecting international tourism demand. An increase in origin country's income means improvement in tourists' spending ability, and thus an increase in demand for tourism. This is also known as the wealth effect. Fereidouni et al. (2017) have also examined the relationship between wealth effect from real estate and outbound tourism

in Malaysia and concluded that wealth effect significantly affects the outbound tourism demand. In brief, tourists' desire to travel are dependent on their wealth or income.

Increase in income directly influences a person's desire to spend more. This has been emphasized in the study of Botti et al. (2007) whereby the authors found that income positively affects the decision of a tourist to visit France. This also means that tourism demand and tourist income have a positive relationship. Garin-Munoz (2007) verified that German tourist income is an important factor that induces tourist arrival in Spain. Brida et al. (2007) arrived at similar conclusion for USA tourist in studying Mexican tourism demand. For Malaysia, Habibi and Abdul Rahim (2009) investigated the effect of tourist income on its top ten source countries and found out that this variable is significant in all countries except Brunei, Australia and United Kingdom. Similarly, Jerabek (2019) confirmed the positive relationship between tourist income and tourist arrival for the top four tourist source countries for Czech Republic.

However, there are also cases where income insignificantly affects tourist decision to travel. Mohd Salleh et al. (2007) examined Malaysian tourism demand for Singapore, Japan, Hong Kong and Australian tourist using Autoregressive Distributed Lag (ARDL) cointegration technique and their findings showed that Australian tourists' decision to travel to Malaysia are not influenced by their income. Besides that, Habibi et al. (2009) also studied tourism demand in Malaysia by selecting the top 15 tourist generating countries. The conclusion reached is in line with Mohd Salleh et al. (2007) and Habibi et al. (2009), suggesting that international tourists considered tourism in Malaysia as a non-luxury service. In other words, travelling to Malaysia is within their ability anytime and income does not affect their travelling decision.

2.4.2.3 Distance

The third variable of a basic gravity model is the distance between origin and destination countries. Tinbergen (1962) mentioned that the trade flow between a pair of countries depends on transportation cost, which is often calculated based on the geographical distance between two countries. In general, the assumption that distance negatively influences trade flow, i.e., the further the distance, the lesser the trade flow. In terms of tourism, this also means that, when a destination country is too far away from the origin country, the tourists are less likely to travel due to increased travelling cost and time. Ulucak et al. (2020) stated that, apart from travelling cost, long distance travel is also uncomfortable.

Most studies using gravity model have thus far followed the definition introduced by Tinbergen (1962) to measure distance, which is the geographical distance between two nations. For instance, Leitao (2010) investigated tourism demand in Portugal; Alawin and Abu-Lila (2016) studied tourist flow for Jordan; and Tatoglu and Gul (2020) estimated tourism demand model for the 14 most visited countries in the world. In general, all studies found a negative relationship between tourism demand and distance, implying that tourist flows reduce as the distance between origin and destination countries increases.

Despite its wide adoption, Lorde et al. (2015) suggested that using the geographical distance as a proxy for transportation cost comes with a drawback because it is time invariant. This has been specifically mentioned by Durbarry (2008) in his study on tourism demand in United Kingdom. The logic is that a time invariant variable is problematic for time series analysis. Hence, Lorde et al. (2015) proposed to modify this proxy by multiplying the geographical distance with average oil price to obtain a dynamic transportation cost proxy. Rodriguez et al. (2012); Ekanayake et al. (2012); Puah et al. (2019); and Wamboye et al. (2020) have also adopted the same derivation of dynamic transportation cost in their

studies examining the tourism demand for Spain, United States, Vietnam and Tanzania, respectively. The outcome does not deviate - distance has an inverse relationship with tourism demand. As such, this implies that the modification is feasible.

2.4.2.4 Travel Cost

Travel or transportation cost in monetary term is another widely investigated variable in tourism demand study. This provides a variation to measurement using travel distance. The variable has been used by Dritsakis (2004) to examine the role of travel cost in determining the travel decision of German and British visitors to Greece based on the average economy class airfare prices of different airport companies from the origin country to Greece. The study concluded that a negative relationship exists between travel cost and tourism demand, suggesting that an increase in travel cost does discourages tourists inflow to Greece. Other studies that included airfare in modelling tourism demand are Chaiboonsri et al. (2010) and Nelson et al. (2011) for Thailand and Hawaii. Particularly, Nelson et al. (2011) are concerned with significant rise of airfare during their study period for the case of Hawaii. Both studies also confirmed the negative influence of airfare on tourism demand.

However, Mohd Salleh et al. (2007) and Mohd Salleh et al. (2008) argued that difficulties exist in determining the exact flight that tourists boarded and hence the exact airfare. Nelson et al. (2011) also pointed out that the wide variety of airfares available in reality and the lack of information on actual consumer payment on airfare are the two hurdles to utilising airfare to measure travel cost. It is in the highlight of this situation that Mohd Salleh et al. (2007) and Mohd Salleh et al. (2008) suggested using the crude oil price as a proxy for travel cost.

The adaptation of this proxy is commonly found in tourism demand studies and some examples are Halicouglu (2004); Garin-Munoz (2006; 2007); Garin-Munoz and Montero-Martin (2007); Habibi and Abdul Rahim (2009); and Jerabek (2019). Distance is a time invariant proxy. An alternative measurement of travel distance to overcome this time invariance characteristic is the crude oil price so that time series analyses can be done. Among others, Halicouglu (2004) and Habibi and Abdul Rahim (2009) examined tourism demand for Turkey and Malaysia, respectively, using ARDL. Jerabek (2019) estimated tourism demand in Czech Republic using the Vector Error Correction Model (VECM).

Regardless of the authors' view on the accuracy of travel cost measured using either airfare or crude oil price, the findings of Halicouglu (2004); Garin-Munoz (2006; 2007); Garin-Munoz and Montero-Martin (2007); Mohd Salleh et al. (2007); Mohd Salleh et al. (2008); Habibi and Abdul Rahim (2009); and Jerabek (2019) have collectively supported the finding of Dritsakis (2004) in which a negative and significant relationship existed in the tourism demand model for Turkey, Spain, Malaysia, and Czech Republic, respectively. Nevertheless, in the study of Garin-Munoz (2009) that modelled international and domestic tourism demand for Galicia of Spain, while international tourism demand is sensitive to travel cost, the domestic tourism demand is not affected by travel cost. Garin-Munoz (2009) stated that this is possibly because local tourists travel by car, which can be significantly inexpensive and thus, does not affect their travel decision.

2.4.2.5 Tourism Price

The computation of tourism price generally has two approaches. Firstly, it can be derived by taking the price ratio (the cost of living of the tourists in the tourist destination relative to the cost of living in the tourists' origin country) adjusted by the exchange rate of these two countries. Such computation has been utilized in the studies of Garin-Munoz and

Montero-Martin (2007); Habibi and Abdul Rahim (2009); Habibi et al. (2009); Leitao (2010); Untong et al. (2015); Tang and Tan (2016); and Puah et al. (2019).

A common practice in computing tourism price is using the Consumer Price Index (CPI) ratio of tourist origin and destination countries adjusted by the exchange rate between these two countries (Chasapopoulos et al., 2014). According to demand theory, an inverse relationship is expected between tourism price and tourism demand. The adoption of this particular computation can be found in, for example, Garin-Munoz and Montero-Martin (2007) in estimating tourism demand for Balearic Islands.

Similarly, in examining the tourism demand for Malaysia, Habibi and Abdul Rahim (2009); Habibi et al. (2009); and Tang and Tan (2016) also utilised the same computation adopted by Garin-Munoz and Montero-Martin (2007) to find out the effect of tourism price on tourism demand. An important point to note is that all these studies utilised different analysis methods, namely time series analysis, dynamic panel estimation, and non-stationary panel data approach, respectively. Other researchers that adopted the same measurement include Leitao (2010) and Untong et al. (2015), that focused on tourism demand in Portugal and Thailand, respectively. All studies concluded that tourism price discouraged tourism demand.

In a study done by De Vita and Kyaw (2013), aimed at determining the role of exchange rate in tourism demand model, the authors tested two models. The first model incorporated exchange rate and price level (CPI ratio) separately in the tourism demand equation. The second model adjusted price ratio by exchange rate. Results for first model indicated that exchange rate and price level (CPI ratio) are not significant. The second model was statistically significant. Therefore, De Vita and Kyaw (2013) suggested that tourism

price should enter the tourism demand model as measured by price ratio adjusted by exchange rate instead of including price level and exchange rate as separate variables.

Another computation of tourism price is derived by taking the ratio of price level in the tourist destination and tourist origin country. This ratio is known as the relative price or cost of living. The derivation does not consider the adjustment of exchange rate between tourist destination and tourist origin country. Exchange rate in this case, is utilized as another explanatory variables in the demand model. For example, Mohd Salleh et al. (2007) and Mohd Salleh et al. (2008) analysed tourism price effect in attracting tourist arrival in Malaysia using ARDL analysis. Mohd Salleh et al. (2007) selected Brunei, China, Hong Kong, Indonesia, Japan, Singapore and Thailand which have the highest market share while Mohd Salleh et al. (2008) selected Australia, Hong Kong, Japan and Singapore based on their importance as representatives from their regions. The results were a mix depending on the selected countries.

At the same time, Massidda and Etzo (2012); Rodriguez et al. (2012); and Ulucak et al. (2020) applied panel data analysis to study the role of tourism price on tourism demand using price ratio as a proxy. Massidda and Etzo (2012) focused their attention on tourism demand in Italy; Rodriguez et al. (2012) chose Galicia of Spain as the destination; and Ulucak et al. (2020) selected Turkey. Massidda and Etzo (2012) and Ulucak et al. (2020) verified the inverse relationship between tourism price and tourism demand, but Rodriguez et al. (2012) did not find any significant effect of tourism price on tourism demand. Nonetheless, Rodriguez et al. (2012) study is different from Massidda and Etzo (2012) and Ulucak et al. (2020); it focused on the academic tourism in Galicia and thus, the targeted tourist group was mainly foreign students. The study explained that this group of tourists do

not notice any substantial differences between the cost of living in their origin countries and Galicia.

In most studies, the CPI is used as a proxy for price level to compute tourism price, regardless of whether it is adjusted with exchange rate or not. Although it is ideal to use tourism related price index to represent the price level, this data is not commonly available, hence CPI is adopted as an alternative approach. In fact, Lorde et al. (2015) credited the use of CPI as logical in the sense that it is widely available and make comparison between countries less problematic.

Nonetheless, there are few studies that managed to use tourism related price index in modelling tourism demand. For example, Garin-Munoz (2006) constructed a tourism price index that considers a basket of more representative goods of tourist consumption in Canary Islands. The author then adjusted tourism price index of Canary Islands and CPI of origin countries with exchange rate to compute the tourism price variable. Although Garin-Munoz (2006) managed to construct tourism price index for Canary Islands, but the same index could not be constructed for the origin countries, and thus CPI was used as an alternative.

Similar with Garin-Munoz (2006), Ekanayake et al. (2012) computed the tourism price variable using tourism related price index for inbound tourism in USA, which is the Tourism Price Index of United States with CPI of origin countries. However, the authors only utilised the simple price ratio between origin and destination countries without exchange rate adjustment such as that done by Garin-Munoz (2006).

Some rare cases where tourism related price indexes for both origin and destination countries were available were found in the studies of Brida et al. (2008) and Jerabek (2019). On one hand, Brida et al. (2008) studied American tourism demand for Mexico. This study has utilised the Index of Tourist Product in USA and Index of Tourisms in Hotel and Expenses of Tourist in Mexico to compute the price ratio with exchange rate adjustment as a measurement for tourism price.

On the other hand, Jerabek (2019) modelled tourism demand in South Moravian Region of Czech Republic for tourists from Germany, Poland, Austria and Slovakia using time series approach. The author used CPI for restaurants and hotels of both origin and destination countries, adjusted by exchange rate, as the measurement for tourism price. This computation, however, did not capture the price level of other tourism products and services. As such, it does not seem to be an ideal proxy.

Nonetheless, the two computation methods for tourism price have consistently proven that a negative relationship exists between tourist demand and tourism price. This has been seen in the studies of Garin-Munoz (2006); Garin-Munoz and Montero-Martin (2007); Mohd Salleh et al. (2007); Habibi and Abdul Rahim (2009); Massidda and Etzo (2012); De Vita and Kyaw (2013); Untong et al. (2015); Puah et al. (2019); and Ulucak et al. (2020). The authors have jointly stated that an increase in tourism price, *ceteris paribus*, will result in a decline in tourism demand in the tourism destination under study. This means that tourists will either opt for domestic tourism or substitute with other tourism destinations.

2.4.2.6 Exchange Rate

The inclusion of exchange rate in tourism demand model is an interesting case. It can either embedded in the computation of tourism price or taken as an independent variable by itself. The former method has been discussed in earlier section. Aguilo et al. (2005) have resorted to the latter approach and suggested that the inclusion of exchange rate and tourism price together in tourism demand model tends to cause multicollinearity problem. This happens when the exchange rate has already been used as a component in the computation of tourism price. Dogru et al. (2017) have also reached the same conclusion and the authors have shown in their study that tourism price standardised by exchange rate has almost perfect negative correlation with exchange rate. The importance of exchange rate is emphasised by Aguilo et al. (2005), who opined that tourists have a higher change to respond to exchange rate movement instead of inflation (price level) when making their travelling decision because it is easier to understand.

There are two ways of viewing exchange rate movement and depending on the point of view, the relationship between exchange rate and tourism demand can be positive or negative. Firstly, when viewing it from the tourist perspective, exchange rate is measured as destination currency over origin currency (destination currency/origin currency). Appreciation or strengthening of tourists' currency in this scenario reflects that domestic price in destination country has become relatively cheaper than the origin country and this encourages tourist arrival into tourism destinations (Wamboye et al., 2020); this indicates a positive relationship.

Adopting this measurement of exchange rate, Wamboye et al. (2020) found a positive relationship between exchange rate and tourism demand in Tanzania with its top fifteen tourist origin countries from 2000 to 2016. The authors concluded that tourist is attracted to

travel to Tanzania when exchange rate is favourable to them. Chaiboonsri et al. (2010) examined the long run relationship between exchange rate and tourist arrival in Thailand for a panel of six tourist origin countries (Malaysia, Japan, Korea, China, Singapore and Taiwan) through the use of Dynamic Ordinary Least Squares and Fully Modified Ordinary Least Square (FMOLS). The study confirmed a positive impact of exchange rate on tourist arrival.

In examining the role of world heritage sites as a determinant of tourism demand for China, Yang et al. (2010) included exchange rate as one of the explanatory variables. The study adopted the direct quotation of exchange rate as the ratio of foreign currency to Chinese Yuan. This is the same measurement used in Wamboye et al. (2020). Yang et al. (2010) verified that depreciation of Chinese Yuan or appreciation of tourist origin country currency attracts more international tourist arrivals into China.

Another study that examined the role of exchange rate as determinant of tourism demand using direct quotation method was done by Eita et al. (2011) for South Africa between 1999 to 2007. The study investigated the determinants affecting top 27 tourist origin countries for South Africa. The findings confirmed that depreciation of South Africa's currency attracts international tourist arrival. Generally, a positive relationship indicates that appreciation (depreciation) of exchange rate (tourist currency) will encourage (discourage) more tourists to travel to the destination country.

Alternatively, exchange rate movement can also be viewed from the tourist destination point of view. In this situation, exchange rate is measured by tourist currency over destination currency (tourist currency/destination currency). Strengthening (weakening) of destination currency in this case results in decline (increase) of tourist arrival into the destination due to lower purchasing power in tourist destination, which implies an

inverse relationship between exchange rate and tourist arrival. This drives the tourists to opt for domestic tourism or other tourist destinations.

For instance, Alawin and Abu-Lila (2016) applied this measurement in determining the factors affecting tourists' decision to visit Jordan. Using gravity model, the authors examined the role of exchange rate among a set of explanatory variables for a panel of 22 tourist origin countries. The findings confirmed the inverse relationship between exchange rate and tourist arrival. When Jordanian dinar appreciates, less tourists will visit Jordan.

Yazdi and Khanalizadeh (2017) adopted the same measurement in studying tourism demand for USA from 1995 to 2014. Using gravity model, the study chose 14 major tourist generating countries of USA and applied panel unit root test, panel cointegration test and panel ARDL test to find out the determinants of international tourism demand in USA. The authors found out that exchange rate is negatively related to tourism demand and this is parallel with the finding of Alawin and Abu-Lila (2016).

2.4.2.7 Substitute Price

Apart from the price level at tourism destination, some researchers are interested in finding out the role of price level at alternative tourism destination. The price level at an alternative tourism destination is termed as the "substitute price". Martin and Witt (1988) in their study has divided substitute price into two components which are; tourist cost of living in substitute destination and transportation cost to substitute destination. However, recent studies seem to have focused more on tourist living cost in substitute destination.

Martin and Witt (1988) used the weighted cost of living at substitute destinations to measure substitute price. The weighted substitute price is computed by weighing the price level of substitute destinations according to the share of international tourist arrival. This computation produces a substitute price index of selected alternative destinations. This approach is also adopted by Durbarry (2008); Habibi and Abdul Rahim (2009); and Tang and Tan (2016).

Durbarry (2008) studied tourism demand model for United Kingdom and has also adopted a weighted substitute price. The author used five substitute destinations where the author explained that the top five countries usually account for more than 60 percent of tourist outflows and is deemed sufficient. The study discovered that an increase in substitute price of these five alternative destinations will increase tourist arrival in United Kingdom.

Focusing on identifying the determinants of Malaysian tourism demand from its top ten tourist generating countries, Habibi and Abdul Rahim (2009) applied ARDL model and utilised quarterly data from 1998 to 2007. Similar with Durbarry (2008), Habibi and Abdul Rahim (2009) also selected five substitute destinations based on geographic and cultural characteristics. The weighted substitute price is also adopted. The outcome pointed out that most of the tourist generating countries have a positive relationship with price level in these five substitute destinations.

Tang and Tan (2016) also used weighted substitute price in their study of tourism demand. In modelling Malaysian tourism demand for its 12 major tourist source countries, the authors chose four substitute destinations, which are Indonesia, Singapore, Thailand and Philippines based on the geographical and culture characteristics as well. FMOLS was

estimated and the outcome revealed that substitute price has a negative effect on tourism demand in Malaysia.

At the same time, there are researchers who included substitute price that is solely of one destination to determine its role on the country of study. This type of substitute price measurement is similar to tourism price, whereby it is measured using CPI ratio or CPI ratio adjusted by exchange rate. Some of these researchers are Mohd Salleh et al. (2007), Kusni et al. (2013), and Untong et al. (2015).

Assessing a set of explanatory variables' effect on tourist arrival in Malaysia, Mohd Salleh et al. (2007) focused on important representative countries from different regions with Singapore, Indonesia and Thailand as the selected substitute countries. Their empirical outcomes revealed a mixed result.

Kusni et al. (2013) also studied tourism demand in Malaysia based on the member countries of the Organisation for Economic Co-operation and Development (OECD) as the tourist source countries together with Singapore and Thailand as the alternative tourist destinations. Their findings concluded that Singapore is a substitute destination for Malaysia due to a significantly positive relationship between Singapore substitute price and tourism demand in Malaysia.

Untong et al. (2015) focused on the Chinese tourism demand for Thailand. Five countries have been selected as the alternative tourist destinations in this study, namely Hong Kong, South Korea, Singapore, Malaysia and Vietnam. The study reported a long run positive relationship between substitute price and Chinese tourist arrival in Thailand.

Past literature revealed that substitute price could have a positive or negative relationship with tourism demand. When an alternative tourism destination has a positive relationship with the tourism demand of a nation, this indicates that the alternative tourism destination is a substitute destination. Increase in price level in alternate tourism destination will bring in more tourist arrival into the tourism destination under study (Kadir et al., 2013). Durbarry (2008), Tang and Tan (2016) and Untong et al. (2015) found proofs in support of this statement. Meanwhile, an alternative tourism destination is said to be a complementary destination when there is a negative relationship between tourism demand and substitute price (Mohd Salleh et al., 2007). Studies that have found such relationship included Habibi and Abdul Rahim (2009) and Lorde et al. (2015).

2.4.2.8 Word-of-Mouth Effect

A more recent study by Seabra et al. (2020) suggested that there is a clear indication of memory effect when new tourist arrivals are found to be somewhat affected by past tourist arrivals. This happens when potential tourists are influenced by other past tourists' experiences and reviews on a certain place of interest or even their own past experiences visiting the place. From their study, Seabra et al. (2020) revealed that tourists who enjoyed memorable experience in the past have a high tendency to revisit the same tourist destination to relive the experience again, suggesting a positive memory effect.

Apart from that, vast literatures also demonstrated that past demand of tourism can influence current demand and this influence represents habit persistent or word-of-mouth effect. Past tourism demand is proxied by a lagged dependant variable in most empirical studies and this variable has been included as an explanatory variable by researchers such as Garin-Munoz (2006; 2007); Garin-Munoz and Montero-Martin (2007); Khadaroo and

Seetanah (2008); Habibi et al. (2009); Chasapopoulos et al. (2014); Karaman (2016); Dogru et al. (2017); and Seabra et al. (2020).

According to Garin-Munoz (2006; 2007), Garin-Munoz and Montero-Martin (2007), and Habibi et al. (2009), including a lagged dependant variable in the tourism demand model is justified for two reasons. Firstly, when tourists travel to a tourism destination that they have visited before, these tourists have less uncertainty with regard to that particular tourism destination as compared to their first visit. This will encourage repeat visitors and over time, it will become a habit persistent.

Furthermore, the spread of information regarding a tourism destination happens when the tourists share their experiences with their friends and families after returning from a trip. Such spread of information reduces the uncertainty with regard to that particular tourism destination for these first-time tourists or become a push factor that attract them to also visit the same destination. This spread of information is known as word-of-mouth effect. Therefore, a lagged dependant variable is expected to have a positive relationship with tourism demand in which it represents indicate habit persistent or repeat visitors and word-of-mouth effect that encourages more visitors to visit a tourism destination.

Both Garin-Munoz (2006) and Garin-Munoz and Montero-Martin (2007) have chosen island tourist destinations for their studies. The former selected Canary Islands and the latter focused on Balearic Islands. Similarly, both studies employed dynamic panel analysis of Generalised Method of Moments (GMM) to investigate the influence of habit persistent and/or word-of-mouth on tourism demand. Both studies confirmed a significant effect in this case.

Garin-Munoz (2007) focused on the German tourism demand for Spain. The author noticed that German tourists have decreased over the years despite being the top two tourist source markets for Spain. With this background, the author aimed to find out the determinants that affected this demand with habit persistent included in the model. The author concluded that habit persistent is an important variable in explaining the demand.

Besides that, Khadaroo and Seetanah (2008) included a lagged dependant variable to represent repeat visitors in their tourism demand study for a panel of 28 countries. The authors utilised annual bilateral tourism flows from 1990 to 2000 to conduct the GMM estimation. The estimation outcomes showed that the lagged dependant variable is positive and significant for all the countries except for Africa. The study further stated that repeat visitors are present around the world.

Karaman (2016) studied the tourism demand in Turkey with the attention focused on visa policy in Turkey. The study also considered the effect of habit persistent. Using the annual observation from 2000 to 2013, the author discovered a highly persistent nature of tourism demand in Turkey. Dogru et al. (2017) also examined the determinants of tourism demand in Turkey. The study concluded that habit persistent and/or word-of-mouth effect has a positive and significant relationship for most of its top nine tourist source countries.

2.4.2.9 Trade Related Variable

Trade related variable is also a common explanatory variable in tourist demand model, as shown in literature review. Chasapopoulos et al. (2014) explained that trade reduces cultural distance between two countries and encourages tourism in both destinations. Bilateral trade effect on tourist inflow has been studied by Mohd Hanafiah et al. (2010) for

Malaysia using annual observation from 1997 to 2008 for selected Asian countries. The study reached the same conclusion.

Chasapopoulos et al. (2014) also examined the role of bilateral trade in influencing tourism demand in Greece for a panel of 31 tourist origin countries. The authors explained that trade serves as an informational knowledge platform among the countries. The empirical estimation concluded that bilateral trade has a positive influence on tourism demand in Greece.

Both Mohd Hanafiah et al. (2010) and Chasapopoulos et al. (2014) also concluded that trade activities stimulate tourism demand. This is consistent with the view of Leitao (2010), in which it stated that bilateral trade creates home-country preferences and reduces transaction cost amid home and host country. In addition, Leitao (2010) also proved that bilateral trade attracts international tourist arrival in his modelling of tourism demand for Portugal.

Apart from bilateral trade, the role of total trade has been examined by Karaman (2016) for the tourism demand in Turkey; the study noted a positive impact. Besides that, Tatoglu and Gul (2020) examined the role of export and import distinctively using the panel data for 14 most visited countries in the world with 30 different origin countries. The study concluded that both export and import bring favourable effects on tourist flow.

Meanwhile, the role of trade openness in attracting tourist arrival has been examined in the studies done by Wong and Tang (2010) and Mohamed Ali Ibrahim (2011). Mohamed Ali Ibrahim (2011) stated that the level of business activities affects the international tourist arrival, especially in a nation where the economy activities are largely generated by international business such as Egypt.

Similarly, Wong and Tang (2010) studied the relationship between trade openness and tourism demand in Singapore. The study concluded that, because of Singapore's economic openness status after the free trade regime and an open foreign investment policy implementation, its economy has bloomed as a result. The study of Wong and Tang (2010) further implied that a bidirectional causality exists between trade openness and tourism demand in Singapore. Meanwhile Mohamed Ali Ibrahim (2011) has also discovered that trade openness has attracted more tourists visiting Egypt.

The function of trade agreement or membership is also a factor that has been studied. This variable has been included as one of the explanatory variables by Tinbergen (1962) in an extended version of gravity model, whereby the author used dummy variables to represent British Commonwealth preference and Benelux preference. Tinbergen (1962) documented that trade agreement has positive influence on trade flow due to preference treatment. This approach has been adopted by Kimura and Lee (2006), which showed that a significant impact of trade agreement on both services trade and goods trade exists.

Eita et al. (2011) discovered that being a member of South African Development Community and European Union is beneficial in attracting tourist inflow into South African while Santeramo and Morelli (2015) discovered that Schengen agreement has improved Italian tourism demand. Rossello et al. (2017) concluded that regional trade agreement plays an important role in tourism demand model estimated for a panel of 196 countries.

2.4.2.10 Adjacent/Common Border

Common border is also frequently included in gravity model as an explanatory variable such as in the study of Kimura and Lee (2006), Khadaroo and Seetanah (2008); Eita et al. (2011); Ahmad Kosnan (2013); Kunroo and Azad (2015); Karaman (2016); and Viljoen

et al. (2019). Kimura and Lee (2006) identified a significant and positive impact of common border on both bilateral goods and services trade but the impact is weaker for bilateral services trade.

Ahmad Kosnan et al. (2013) discovered that common border allows frequent tourist visits to certain destination countries as traveling to such destinations would be relatively easier, cheaper and convenient; this holds true based on their empirical findings for Malaysia tourism demand. Khadaroo and Seetanah (2008) included common border as well in their tourism demand study on tourism flows among 28 countries and they discovered that the impact is significant and positive.

Meanwhile, Eita et al. (2011) stated that South Africa has focused much of its tourism growth strategy on six bordering Southern African Development Community countries. The empirical outcomes supported that common border is an important determinant of tourism demand in South Africa. Karaman (2016) investigated the visa policy in Turkey. The study has also included common border in their empirical model and the findings indicated that common border attracts tourist arrival.

Rossello et al. (2017) conducted a comprehensive tourism demand study by using a panel of 196 destination countries and 200 origin countries from 2000 to 2013. The authors focused on the relationship between infectious disease risk and international tourism demand. The model also specified common border as one the explanatory variables. The analysis outcomes confirmed that tourists prefer to travel to countries located closer to home.

Similarly, Okafor et al. (2018) included common border as one of tourism demand determinants. The inclusion of this variable is to further reveal the preferences of tourists. The outcome of their study confirmed that common border attracts tourist inflows. The study further explained that shared border to lowers the transportation and search costs.

Viljoen et al. (2019) also investigated the role of common border for Africa continent as a whole and South Africa region in specific as compared to Eita et al. (2011), which only studied South Africa. They found that having a common border is positive for the African continent but not for the South African region, which depicted a negative and significant relationship. All the afore-mentioned studies used dummy variable to proxy for countries sharing common border. The dummy takes the value of one if the countries are sharing a common border and the value of zero if otherwise.

2.4.2.11 One-off/Special Event

Tourism demand is also affected by a one-off or special event, and this is often included as a dummy variable. A one-off or special event can bring either positive or negative impact on tourism demand, depending on the nature of the event. Events such as economic crisis, disease outbreak and terrorist attack cause negative impacts while events like holiday and tourism promotion campaign have a positive impact on tourism demand.

In the studies done by Mohd Salleh et al. (2007), Habibi et al. (2009) and Kadir et al. (2013), the effect of Asian financial crisis in 1997/98 had been studied as a dummy variable in their tourism demand model. Their findings consistently unveiled that Asian financial crisis 1997/98 had a negative effect on tourism demand.

In addition, Kusni et al. (2013) and Fereidouni et al. (2017) assessed the global financial crisis for Malaysia and found that this event had a negative effect towards inbound and outbound tourism in Malaysia. At the same time, Dogru et al. (2017) studied the impact of global financial crisis on Turkey but did not find this event to be influencing tourism demand.

Moreover, Tang and Wong (2009), Yang et al. (2010), and Kusni et al. (2013) examined the impact of SARS outbreak towards tourism demand in Cambodia, China and Malaysia, respectively. The finding of these studies showed that SARS outbreak had a negative influence on tourism demand. Besides that, Tang and Tan (2016) found out that another disease related event - Avian flu significantly affected tourism demand in Malaysia.

Nonetheless, Fereidouni et al. (2017) did not find SARS outbreak to be influencing the outflow of Malaysian tourist. Similarly, Kuo et al. (2008) investigated the impact of Avian flu in China, Hong Kong, Singapore, Taiwan, Indonesia and Vietnam, but found no significant impact on tourism demand in these countries. Rossello et al. (2017) dedicated their effort in examining the impact of infectious disease risk (Malaria, Dengue, Yellow Fever and Ebola) on international tourist arrivals and found out that Malaria caused the biggest impact.

Apart from that, Garin-Munoz (2006; 2007), Tang and Tan (2016) and Yazdi and Khanalizadeh (2017) focused on the impact of the 911 terrorist attack on tourism demand in Spain, Malaysia and United States. Garin-Munoz (2006; 2007), Tang and Tan (2016) and Yazdi and Khanalizadeh (2017) confirmed that the terrorist attack had negatively impacted all countries under study. The negative impact is obvious because, unlike many other factors, it is a factor that is uncontrollable and countries which are more vulnerable to terrorism

attack are naturally less preferred by tourists (Samitas et al., 2018). Santana-Gallego et al. (2020) detailed out that the impact of terrorism is larger for tourists visiting for personal reason than tourists visiting for business purpose. Owing to the fact that many travels for personal reason, this has become unfavourable for most destination countries which rely on tourism as a source of income.

Besides using a dummy variable to capture the unfavourable event, it can also be used to capture the effect of the favourable event. For instance, the study of Garin-Munoz (2009) investigated the event of Holy Year in Galicia, Spain, and Falk (2010) studied on the effect of Easter holiday on Austria tourism demand. Kadir et al. (2013), on the other hand, has studied the "Visit Malaysia Year" effect on Malaysia. Both Garin-Munoz (2009) and Falk (2010) discovered that these events have a positive influence on tourism demand while Kadir et al. (2013) did not found evidence supporting that Visit Malaysia Year has a significant impact on tourism demand in Malaysia.

In another study conducted by Athanasopoulos and Hyndman (2008), they investigated the effect of Sydney Olympic on different type of visitors. They found out that even though the Sydney Olympic event had successfully attracted more visitors into Australia, some business travels were postponed until after the Sydney Olympic. The impact of Athens Olympic on Greece was studied by Chasapopoulos et al. (2014), but they discovered that the event had generally discouraged tourist inflow. The authors attributed this to the tourists' concerns on increase in prices, overcrowded condition and adequate security measures during the period due to massive inflow of visitors.

2.4.3 The Role of Language in Enhancing Tourism Demand

Okafor et al. (2018) stated that language has an important role in international tourism because it can either improve the travel experience or hinder it. A shared language improves tourists' experience since it bridges the communication gap between the tourists and the locals. On the contrary, language barrier discourages tourist inflow because the tourists will face difficulties in their daily interaction with the locals.

Based on a survey conducted in Chennai of Tamil Nadu, India by Chockalingam and Ganesh (2010), language barrier is among the 15 exclusive tourist problems identified. The survey, concluded from 150 respondents, highlighted that local language presents an unfavourable effect on tour experience.

A similar conclusion was reached by Hara (2013) who studied the perception of Americans on language barrier upon their desire to travel to Japan. The study utilised online questionnaire and collected information on demographic characteristics, travel behaviours and perceptions of 192 respondents. The aim was to find out the difference in perception of language barrier for Americans who has travelled to Japan before or studied Japanese language before compared to those who have not. Hara (2013) concluded that perception of language barrier badly influenced tourists' travelling decision, especially for those who have not travelled to Japan or studied Japanese language before.

Meanwhile, a study done by Prachanant (2012) surveyed 40 tourism employees from international tour companies in Thailand on the use of English as a medium of communication with international tourists. Interestingly, although majority of the respondents rated English as highly needed in their work, the fact that English is rarely used in daily communication has hindered them from communicating effectively with tourists.

This suggests that there is a need to improve the English skillsets for employees in the Thailand tourism sector.

The issue of multilingualism has been studied by Suhaimi and Abdullah (2017) in their survey on 150 shopkeepers and interview with 13 tourists in Kuala Terengganu, Malaysia. While the shopkeepers opined that multilingualism is insignificant and language barrier can be overcome through the aids of non-verbal cues, the tourists stated that multilingualism is important and also allow them to learn the local language during their visit. Nevertheless, some have stated that English is sufficient within the country.

These studies have shown that local language is indeed a barrier, especially for tourists that have no prior knowledge of local language. Nonetheless, when both locals and tourists are able to communicate in a common language such as English, it can effectively mitigate the problem.

In fact, language is a common variable of interest in international tourism demand studies. This variable has been included as a common language in numerous studies. Most studies proxied this common language using a dummy variable that takes the value of one if a common language exists between the origin and destination countries, otherwise zero. Kimura and Lee (2006) in their study on the bilateral services trade (which included tourism) for OECD countries, mentioned that the existence of a common language has improved approximately 50 percent service export between the two countries. This means that the common language has stimulated service exports, including tourism.

Meanwhile, Durbarry (2008) tested the effect of common language in determining tourism demand in United Kingdom. Since English is the native language, tourists who speak English are naturally considered as sharing a common language with the locals. The author found out that English speaking tourists are more inclined to visit and spend in United Kingdom than those non-English speaking tourists. Specifically, common language is more influential on tourist arrival than tourist spending.

Utilising the bilateral tourist flows data of 28 countries, Khadaroo and Seetanah (2008) examined the tourism demand factors that also included language. The authors segregated the sample according to origin and destination continent before conducting a dynamic panel estimation using GMM. Similar with the finding of Durbarry (2008), the empirical outcomes revealed that common language attracts tourist arrivals. The results were consistent across continent-wised origin and largely the same across continent-wised destination.

Besides that, Seetanah et al. (2010) studied the determinants for South Africa from 1985 to 2000 using FMOLS. Interestingly, the authors also segregated the sample as that done by Durbarry (2008). The sample was segregated by region to further distinguish the different sets of determinants for regional tourism demand. Common language is found to be consistently significant and positive under all the estimated models. This indicates that common language is an important determinant that is able to attract tourist arrivals into South Africa.

In a similar study set up as Durbarry (2008), Alawin and Abu-Lila (2016) aimed to identify the effect of common language on Arab speaking and non-Arab speaking tourists' inflows into Jordan. The authors reached the same conclusion as Durbarry (2008) that common language is an important factor to attract tourist inflow into Jordan. Likewise, Arab speaking tourists are more motivated to visit Jordan since they share a common language with the locals.

On a wider scope, Rossello et al. (2017) investigated the function of a set of explanatory variables, which included common language, for 208 origin countries and 196 destination countries between 2000 and 2013. The gravity model was adopted and the POLS was estimated. Again, the same conclusion was reached as all studies aforementioned.

One may expect that the existence of a common language and border sharing boost tourist visits because it makes the travel easier, cheaper and more convenient. However, Ahmad Kosnan et al. (2013) only found partial support to this claim in examining the factors affecting tourist arrival in Malaysia. The coefficient of common language is positive and significant for tourist receipts model, but not for the tourist arrival model where the result is positive, but insignificant.

Moreover, when examining the demand for academic tourism in Spain, Rodriguez et al. (2012) discovered that common language is not significant from their empirical findings. Common language is defined as Spanish or similar language (Portuguese or Italian) in this study. Rodriguez et al. (2012) suggested that speaking a similar language is not necessarily an advantage since in many cases, student mobility programme is used precisely to learn new languages.

In a more recent study conducted by Okafor et al. (2018), the authors have investigated the role of common official and unofficial languages in international tourism demand model. The dummy variable representing common unofficial language, which is defined as spoken by at least 9 percent of the populations in both origin and destination countries, takes the value of one and zero if otherwise. This study concluded that common unofficial language is able to attract tourist arrivals more than common official language, especially, in the Europe region because most European countries do not share common official language.

Takahashi (2020) studied the difference in tourism demand determinants for economies with two different structure using French Polynesia and Singapore as the targeted countries. French Polynesia has been chosen because it is highly dependent on the tourism sector while Singapore has a diversified service industry. Common language is included in the set of explanatory variables chosen. The empirical estimation revealed that common language is only significant in Singapore because it is a multilingual country that shares a common language with more than one tourist origin countries.

The role of language, as shown in past studies, seem to differ depending on the scope as well as study background. Nonetheless, whenever language barrier exists, it becomes undoubtedly unfavourable to attract tourist inflows. On the contrary, when a common language exists, tourists are more likely to visit a tourism destination.

2.5 Concluding Remarks

To sum up, the literature reviewed in this chapter has provided useful insights on the variables used in tourism demand studies. There are three fundamental variables used in gravity model which are origin and destination countries income together with distance between these countries. Apart from that, price component variables are also a common explanatory variable included in modelling tourism demand and these can be divided into tourism price and travel cost.

Tourist arrivals is the most common measurement used to capture tourism demand followed by tourism receipts and length of stay. There is no clear advantage of each measurement except that data for tourist arrivals is usually more complete and easily available than tourism receipts and length of stay. For income, this variable is usually represented by GDP in its various form (nominal, real and per capita). Similar with tourism demand, there is also no distinctive benefit of using a specified measurement.

Distance is typically measured by the geographical distance in a gravity model but such measurement is time invariant and would cause inconvenience in empirical estimation, especially, for time series analysis. Hence, a dynamic way to measure distance in tourism demand is to multiple crude oil price with geographical distance. In such computation, the fluctuation of crude oil price is captured together with the constant geographical distance to reach a tourism destination. This dynamic measurement is termed as the “travel cost” in some literature. Despite airfare being a more precise measurement of travel cost, it is difficult to obtain such data, causing most researchers to use crude oil price as an alternative. It is also worth highlighting that airfare does not take into consideration the distance to be covered to reach a tourism destination.

Tourism price has two common ways of computing. The more general practice is to take the price ratio of origin and destination countries, adjusted by exchange rate. The second measurement is the simple ratio of price level between origin and destination countries. Furthermore, exchange rate is known to affect tourism demand, but but including this variable more than once in a model will lead to multicollinearity problem. Hence, to avoid this, the exchange rate has to be dropped from the tourism demand model when using the first measurement of tourism price. Otherwise, exchange rate can be included as an explanatory variable when using the second measurement of tourism price.

Under a dynamic setting, a lagged dependant variable can be used to estimate the word-of-mouth effect and/or habit persistence. This lagged dependant variable is found to have a positive relationship with tourism demand in most studies as tourists tend to feel safer, more secured and attracted to revisit a place they have visited before and in long run, this would become a habit. Potential tourists in origin country would also come when they receive positive feedback on the tourism destination from other past tourists. Since tourism is a type of services trade, trade related variable is often included as one of the tourism demand determinants. For instance, bilateral trade, trade openness and trade agreement are variables that have been widely studied in the tourism demand model.

Language has received attention as a factor influencing tourism demand because of the convenience it would bring to tourists during their trip. The existence of a common language between origin and destination countries generally encouraged tourist arrival. Typically, common language is represented by a dummy variable but just like geographical, this makes the measurement time invariant and unfavourable for time series analysis. The

same applies to common border represented as a dummy variable. Nonetheless, sharing a common border generally encourages border tourism and is beneficial for both countries.

Apart from that, dummy variable is often used to represent one-off event that may have occurred at origin or destination countries. A one-off event that has been included in some empirical studies are like the Asian financial crisis in 1997/1998, the terrorist attack in September 2001, the SARS outbreak in 2003, the recent global financial crisis in 2009 and tourism promotion campaign, among others.

From the literature reviewed, the studies in tourism demand model have been extensively carried out since a few decades ago. However, in regard to the adaptation of gravity model, there is a period of time where this model has not been given attention because it lacks the support of theoretical foundation and empirical evidence. Nevertheless, since the 2000s, application of gravity model in international tourism demand has been revitalised after some researchers have proven the validity of this model in modelling international tourism demand using empirical proofs as well as a sound theoretical basis.

Some studies have paid much of its attention on quantifiable economic variables and less on qualitative and linguistic variable. For example, there are limited studies focusing on assessment of a linguistic variable in attracting international tourist arrival although a linguistic variable is an important factor for tourists since it is needed for daily communication with the locals.

From the scarce literature in international tourism demand that examined the influence of a linguistic variable, a single and widely adopted proxy is using a dummy variable to indicate the existence of a common language between a pair of countries.

However, because dummy variable is time invariant, the interpretation of results may become limited, not to mention that it causes estimation error for time series analysis.

At the same time, the common languages that have been defined and used in previous studies have mostly been established from the past relationship between two countries, especially through colonial and migration relationship. Most of these studies define or perceive the existence of a common language when there is a certain percentage of population in both countries are speaking the same language.

The problem of this definition or perception is that, for countries with no historical ties, a common language is as good as non-existent albeit the reality is that English is indeed the international language spoken by many. In other words, English can function as a common language for these countries with no historical ties.

Furthermore, the use of a dummy as a proxy for common language does not allow for precise determination of the effects of language proficiency level on the travelling decision of tourists. While there may be a huge population in the destination country that speak the same language as the tourists' country of origin, they may not be fluent or proficient enough to communicate effectively and enhance the tourists' travel experience.

Therefore, this study aimed to assess the role of English as a common language and the international language of choice. English can be the common language for countries that do not have any past relationship and this trend is expected to continue. Apart from that, English proficiency has been assessed through the results of formal examinations to determine the effect of language proficiency on tourists' decision to travel.

In order to achieve the objectives of this study, the gravity model was adopted with some modification. Gravity model has been chosen because of its well-known fitting for the studies on international trade, migration and foreign direct investment (Morley et al., 2014). Tourism demand studies have been done extensively using this model with satisfactory results. Among others, Keum (2010); Morley et al. (2014); Lorde et al. (2015); Yazdi and Khanalizadeh (2017); Okafor et al. (2018); and Tatoglu and Glu (2020) have used gravity model in their tourism demand studies.

Apart from the original gravity model, its augmented version has also been widely used with inclusion of other explanatory variables to segregate between origin and destination countries (Morley et al., 2014). The augmented gravity model allows researchers to include their variable of interest and assess its impact on the dependant variable. For example, Khadaroo and Seetanah (2008) studied the role of transport infrastructure on tourism development while Yang et al. (2010) researched on the role of world heritage sites in influencing tourism demand.

Therefore, the selection of explanatory variables in this study was mainly based on the gravity model. Tinbergen (1962) specified the initial gravity model as the trade flow between two countries, influenced by the income levels of both countries and the distance between them. Tinbergen (1962) then expanded the initial gravity model to include dummy variables representing neighbouring countries and trade agreements (Commonwealth and Benelux preferences). Thus, following gravity model, five determinants were included in the tourism demand model of this study, which were origin countries' income; destination countries' income; distance; neighbouring countries; and free trade agreements.

In addition, this study also considered the demand theory whereby demand is dependent on the income of consumer and price of product. Since income was included in the tourism demand model, price factor was then added to represent tourism price. Next, language was added into the tourism demand model. This variable was of main interest to assess its role through the augmented gravity model.

CHAPTER 3

METHODOLOGY

3.1 Overview

This chapter presents the methodology adopted to conduct this study, divided into three sections. The first section, Section 3.2, describes the data utilized in this study and its sources. Section 3.3 discusses the static linear panel data models. The last section, Section 3.4, describes the empirical models that were estimated in this study.

3.2 Data Description

The dependent variable of this study is the international tourist arrivals from the top ten countries for ASEAN+3 countries. The data were collected from the respective tourism authorities as shown in Table 3.1. Meanwhile, Table 3.2 listed the top ten tourist origin countries for ASEAN+3 countries.

Table 3.1: Source of International Tourist Arrivals for ASEAN+3 Countries

Country	Source
Brunei	Department of Statistics
Cambodia	Ministry of Tourism
Indonesia	BPS-Statistics Indonesia
Laos	Ministry of Information, Culture and Tourism
Malaysia	Tourism Malaysia
Myanmar	Ministry of Hotels and Tourism
Philippines	Philippine Statistics Authority
Singapore	Singapore Tourism Analytics Network
Thailand	Ministry of Sports and Tourism
Vietnam	Vietnam National Tourism Administration
China	Ministry of Culture and Tourism of the People's Republic of China
Japan	Japan National Tourism Organization
Korea	Korea Tourism Organization

Table 3.2: Top Ten Tourist Origin Countries for ASEAN+3 Countries

Ranking	Brunei	Cambodia	Indonesia	Laos	Malaysia	Myanmar	Philippines	Singapore	Thailand	Vietnam	China	Japan	Korea
1	China	China	Malaysia	Thailand	Singapore	Thailand	Korea	China	China	China	Hong Kong	China	China
2	Malaysia	Vietnam	China	Vietnam	Indonesia	China	China	Indonesia	Malaysia	Korea	Macau	Korea	Japan
3	Philippines	Laos	Singapore	China	China	Japan	United States	India	Korea	Japan	Taiwan	Taiwan	Taiwan
4	Indonesia	Thailand	Australia	Korea	Thailand	United States	Japan	Malaysia	Laos	Taiwan	Korea	Hong Kong	United States
5	Singapore	Korea	Japan	United States	Brunei	Korea	Australia	Australia	Japan	United States	Japan	United States	Hong Kong
6	United Kingdom	USA	India	France	India	Singapore	Taiwan	Japan	India	Russia	Russia	Thailand	Thailand
7	Korea	Japan	Korea	Japan	Korea	Vietnam	Canada	Philippines	Russia	Malaysia	United States	Australia	Philippines
8	India	Malaysia	United Kingdom	United Kingdom	Japan	France	United Kingdom	Korea	United States	Australia	Mongolia	Malaysia	Vietnam
9	Australia	United Kingdom	United States	Germany	Philippines	United Kingdom	Singapore	United States	Singapore	Thailand	Malaysia	Philippines	Malaysia
10	Thailand	France	Philippines	Australia	United Kingdom	Malaysia	Malaysia	Vietnam	United Kingdom	United Kingdom	Philippines	Singapore	Russia

The data for destination and origin countries' income, price levels and exchange rate were compiled from the International Financial Statistics published by International Monetary Fund. The Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) provided the data for distance between destination and origin countries and further information such as (1) whether an origin country is an adjacent or neighbouring country to a destination country of interest (sharing a common border); (2) the existence of trade agreement between the destination and origin countries; and (3) whether the destination and origin countries are sharing a common language (defined as a language that is spoken by at least 9 percent of population). The crude oil price was obtained from the Global Economic Monitor published by the World Bank while the English Proficiency Index (EPI) was collected from the EF EPI report produced by the Education First Limited[5]. The period of study covers from 2012 until 2017, utilising annual observation.

3.3 Static Linear Panel Models[6]

3.3.1 Pooled OLS, Fixed and Random Effects Models

In static panel data analysis, there are three possible ways of estimating a model depending on the assumptions made on the intercept term. First, there is the Pooled Ordinary Least Squares (POLS) model that treats the intercept and slope as constant across units and

[5] The EPI data is based on annual test data from test takers who took the EF Standard English Test or one of the Education First Limited English placement tests. In its latest 2020 edition, there are more than 2.2 million test takers in 2019. Only cities, regions and countries with minimum 400 test takers are included for the EPI, these data are highly correlated with TOEFL iBT 2018 and IELTS Academic Test 2018, thus ensuring credibility of the data.

[6] The discussion here is mainly adopted from Song and Witt (2011).

times. The second model is the Fixed Effects model (FEM) that uses the dummy variables to allow the intercepts to vary across cross-section units. This means that each cross-section unit has fixed unobserved differences. Lastly, the Random Effects model (REM) also allows for intercepts to vary between cross-section units such as the FEM, but the variation is treated as randomly determined.

All the models have its advantages and disadvantages. The POLS model gives a straightforward estimation, but it is estimated under a restrictive condition. In addition, in most cases, assuming that the cross-unit individual effects are invariant can be unrealistic. Meanwhile, the FEM is able to capture the unobserved differences between cross-section units using the dummy variables, but this comes with a price of decreased degree of freedom. This causes the FEM estimation to become less efficient than those of the REM. On the contrary, the REM is estimated with the assumption that the unobserved cross-section individual effects are not correlated with the explanatory variables given that these effects are randomly determined. Nevertheless, this assumption may be improper, and the violation of this assumption will result in bias and inconsistent estimation of the REM.

Since the FEM and REM models are more powerful that the POLS model, the following discussion shall focus on the nature of FEM and REM for a better understanding on these two models. Consider a classical linear regression model of Equation 3.1 with observations on dependent variable *y* across cross-sections $i = 1, \ldots, N$ over time periods $t = 1, \ldots, T$ and $k = 1, \ldots, K$ explanatory variables denoted by *K* x 1 vector *x*.

$$y_{it} = \alpha_i + \beta'_{it} + \varepsilon_{it} \qquad \text{Equation 3.1}$$

where the intercept and slope coefficients are allowed to vary across cross-sections. For all *i* and *t*, the error term is independently and identically distributed with mean zero and variance of σ_ε^2 or i.i.d. $(0, \sigma_\varepsilon^2)$ for all *i* and *t*.

Assume that the error term in Equation 3.1 has been re-specified as

$$\omega_{it} = \mu_i + \varepsilon_{it} \qquad \text{Equation 3.2}$$

where μ_i denotes the unobservable cross-section specific effects and ε_{it} is the usual disturbance term which varies across regions and time. With regard to μ_i, there are two assumptions that can be made. First, in the case of μ_i is correlated with x_{it}, this is the FEM. Second, when μ_i is uncorrelated with x_{it}, this is the REM.

Choosing between FEM and REM is an art by itself. This is because these two models can provide considerable differences in coefficients estimation in a typical panel data set, especially when the number of time series observations is small and the number of cross-section is large (Hausman, 1978). Given that the FEM assumes that correlation exists between μ_i and x_{it}, therefore the model must be estimated conditionally on the presence of μ_i, which is now treated as a fixed parameter to be estimated as in Equation 3.3.

$$y_{it} = \mu_i + b'x_{it} + \varepsilon_{it} \qquad \text{Equation 3.3}$$

The OLS estimation of the FEM will be bias because of the correlation between μ_i and x_{it}. In this case, unbiased OLS estimates can be obtained either through first-differencing the variables or alternatively, differencing them by cross-section-specific means. In either case, the fixed effect μ_i is eliminated and this immediately eliminates the correlation problem as well. Through such treatment, the fixed effect and correlation

problem are taken care of as well as any observed variable which is time invariant, making the FEM robust towards omission of any relevant time-invariant explanatory variable.

Theoretically, FEM is the model from which inferences can be made contingent on the observed sample. This means that FEM is suitable when the focus of estimation is on the sample itself. On the contrary, if the estimation is centred on making inferences about the population from which the sample is drawn, then it is more appropriate to consider μ_i as randomly distributed across cross-section units:

$$y_{it} = \alpha + b'x_{it} + \mu_i + \varepsilon_{it} \qquad \text{Equation 3.4}$$

where μ_i and ε_{it} are i.i.d. $(0, \sigma_\mu^2)$ and i.i.d. $(0, \sigma_\varepsilon^2)$, respectively. Assuming that μ_i is not correlated with x_{it}, the OLS is asymptotically unbiased, but less efficient than REM.

The existence of random effects can be tested using a straightforward Lagrange multiplier test. This test is based on the OLS residuals from the model in which both slope and intercept terms are assumed constant (Breusch & Pagan, 1980). The test is based on:

$$H_0 = \sigma_\mu^2 = 0 \quad (\text{or } Corr[w_{it}, w_{is}] = 0$$

$$H_1 = \sigma_\mu^2 \neq 0 \quad (\text{or } Corr[w_{it}, w_{is}] \neq 0$$

with the test statistic defined as:

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}[\sum_{t=1}^{T} e_{it}]^2}{\sum_{i=0}^{n}\sum_{i=1}^{T} e_{it}^2} - 1\right]^2 \qquad \text{Equation 3.5}$$

and e_{it} taken as the residual from the POLS regression. The Lagrange multiplier test statistic is distributed as chi-squared with one degree of freedom under the null hypothesis. Acceptance of null hypothesis means that random effects do not exist, which means the

POLS is favoured. Lagrange multiplier test enables us to choose the better model between POLS and REM.

Equivalently, the REM can be tested directly against the FEM using the Hausman test developed by Hausman (1978). This test is devised from the idea that if there is no correlation between μ_i and x_{it}, systematic differences do not exist between the FEM and REM estimation due to the consistency of both OLS in FEM model and Feasible Generalised Least Squares in REM. The Hausman test states that:

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}[\sum_{t=1}^{T} e_{it}]^2}{\sum_{i=0}^{n}\sum_{i=1}^{T} e_{it}^2} - 1\right]^2 \qquad \text{Equation 3.6}$$

where $\widehat{\sum}_{fe}$and $\widehat{\sum}_{re}$ are the estimated slope covariance matrices for the FEM and REM, respectively. The Hausman test statistic is distributed asymptotically as chi-squared with *K* degrees of freedom under the null hypothesis, which states that $\mu_i$is uncorrelated with x_{it}. This test suggests that REM is relevant when the null hypothesis is accepted with the following hypotheses:

$$H_0 = REM \qquad\qquad H_1 = FEM$$

The better model between POLS and REM is decided using the Lagrange multiplier test and the preferred model between REM and FEM is confirmed using the Hausman test.

As an alternative, test of overidentifying using Sargan-Hansen statistic is conducted using Stata command - xtoverid (Schaffer & Stillman, 2010) when Hausman test outcome provides a negative test statistic or when the difference of covariance matrix is not definite positive. Both conditions do not fulfil the requirement of Hausman test. Under Sargan-Hansen statistic, FEM uses the orthogonality conditions that the regressors are uncorrelated

with the idiosyncratic error while REM uses the additional orthogonality conditions that the regressors are uncorrelated with the group-specific error. These additional orthogonality conditions are overidentifying restriction.

Stata command xtoverid implements the test using the artificial regression approach described by Arellano (1993) and Woolridge (2002) where REM is re-estimated augmented with deviations-from-mean form. The test statistics is a Wald test of the significance of these additional regressors. The interpretation of the test outcome is same with Hausman test where rejection of null hypothesis favours FEM. After that, the Chow test is carried out to decide between POLS and FEM to choose the better model. Chow test is a F-test conducted to detect whether the cross-sections share a common intercept or not. Rejection of F-test means that each cross-section has different intercept, which is in favour of FEM.

Past literature has suggested the importance of estimating tourism demand in a dynamic set up to capture the effects of habit persistence and/or word-of-mouth in affecting tourist decision to travel. The dynamic panel estimation technique of Generalised Method of Moments (GMM) is often used in tourism demand studies to examine this effect, which is represented by a lagged dependant variable. However, Pablo-Romero and Molina (2013) and Viljoen et al. (2019) stated that GMM is more desirable for panel data with small *T* and large *N*. Roodman (2009) stated that, although there is no exact definition for large *N*, but applying GMM to panels with *N* at 20 would already cause concern. The *N* in this study is small and less than 20, and thus, the static linear panel models have been chosen. To be specific, this study used the top ten tourist generating countries for each ASEAN+3 country ($N = 10$) from 2012 to 2017 ($T = 6$).

Apart from the dynamic panel estimator of GMM, Shepherd (2016) introduced an alternative estimator for gravity model of international trade - the Poisson Pseudo-Maximum Likelihood (PPML) estimator. The estimator is powerful enough to even consider null observed trade value. Shepherd (2016) stated that observations with zero trade value are always omitted from the Ordinary Least Squares model because the logarithm of zero cannot be defined and will potentially lead to sample selection bias. Nonetheless, since there were no observations with zero value, the PPML estimator was no adopted in this study.

3.4 Empirical Model

Chasapopoulos et al. (2014) mentioned that, in empirical examination of tourism demand, most studies start off with the consumer demand theory and tend to model the demand for tourism based on income and price variables. In this study, consumer demand theory was augmented into the gravity model and also added with a linguistic variable to examine the tourism demand:

$$TA = F(OY, DY, DIST, ADJ, FTA, TP, LANG) \quad \text{Equation 3.7}$$

where TA represents tourist arrival from top ten origin countries to the ASEAN+3 countries and OY and DY proxy the real income level of origin and destination countries, respectively. DIST is the distance between destination and origin countries, ADJ is a dummy variable representing adjacent or neighbouring countries, FTA is a dummy variable representing free trade agreement between destination and origin countries, TP is the tourism price at tourism destination and lastly, LANG is a linguistic variable.

As stated aforehand, the dependent variable, TA, is measured by the number of tourist arrival from the top ten origin countries into the ASEAN+3 countries. Meanwhile, the real income of origin and destination countries are computed by deflating the gross domestic product of these countries with their respective consumer price index. The formula to calculate the real income level of origin country is as follows:

$$OY = \frac{GDP_{origin}}{CPI_{origin}}$$

The real income level of destination country is calculated by replacing the origin country with destination country such as follows:

$$DY = \frac{GDP_{destination}}{CPI_{destination}}$$

The distance between destination and origin countries in this study were obtained by multiplying the geographical distance with travel cost (represented by real crude oil price) to overcome the disadvantage of using only geographical distance that is time invariant. The following formula illustrates the calculation of distance variable used in this study:

$$DIST = geographical\ distance\ (km) x\ real\ crude\ oil\ price$$

The crude oil price was chosen to represent travel cost because this study encountered a similar practical issue raised by Mohd Salleh et al. (2007) and Mohd Salleh et al. (2008) which was the difficulties to obtain the exact airfare as a measurement for travel cost. Therefore, as a solution, this study utilised crude oil price which is also widely adopted by studies such as Garin-Munoz (2007); Habibi and Abdul Rahim (2009); Ekanayake et al. (2012); Lorde et al. (2015); and Puah et al. (2019). The geographical distance between origin

and destination countries has been measured using the great circle formula that uses the latitudes and longitudes of the most important cities/agglomerations (in terms of population) while real crude oil price has been computed by deflating the West Texas Intermediate spot price per barrel with USA CPI.

Adjacent or neighbouring country is a dummy variable that takes the value of 1 if the destination and origin countries are sharing a common border, otherwise it takes the value of 0. FTA is a dummy variable that takes the value of 1 if a free trade agreement exists between the destination and origin countries, otherwise it takes the value of 0. The tourism price is derived by dividing the consumer price index of the countries under study with the consumer price index of their respective top ten origin countries, and adjusted with the exchange rate between the destination and origin countries. The formula to calculate tourism price is as follows:

$$TP = \frac{CPI_{destination}}{CPI_{origin} * ER_{destination/origin}}$$

The linguistic variable has two measurements. The first measurement captures the existence of common language between destination and origin countries using a dummy that takes the value of 1 if there is a common language, otherwise it takes the value of 0. Common language in this study follows the definition in CEPII, which is defined as a language that is spoken by at least 9 percent of population between origin and destination countries. The second measurement measures the level of English proficiency in the destination countries or known as EPI and this data is compiled from EF EPI report produced by the Education First Limited. All variables except the dummy variables have been transformed into natural logarithm form before any estimation is conducted. Finally, the gravity model of tourism demand can be expressed as follows:

$$LTA_{ijt} = \beta_0 + \beta_1 LOY_{jt} + \beta_2 LDY_{it} + \beta_3 LDIST_{ijt} + \beta_4 ADJ_{ijt} + \beta_5 FTA_{ijt} + \beta_6 LTP_{ijt} + \beta_7 LANG_{ijt} + \mu_{ijt} \quad \text{Equation 3.8}$$

The dependant variable LTA_{ijt} is the tourist arrivals from origin *j* to destination *i* at time *t*. For explanatory variables, LOY_{jt} and LDY_{it} represent the destination and origin countries real income level, and $LDIST_{ijt}$ is the distance between origin and destination countries. Dummy variables that proxy adjacent country and free trade agreement are ADJ and FTA, respectively. LTP_{ijt} measures tourism price and lastly, LANG is the linguistic factor.

To begin the estimation, a traditional gravity model of tourism demand was firstly established using the following model represented by Equation 3.9, whereby the demand for tourism was determined based on the destination and origin countries income level, geographical distance, adjacent country and free trade agreement. Modified from Equation 3.9, the tourism price and linguistic variable were included in Equation 3.10 to test the influence of these variables. Equation 3.10 was estimated twice as linguistic variable in this study used two measurements.

$$LTA_{ijt} = \beta_0 + \beta_1 LOY_{jt} + \beta_2 LDY_{it} + \beta_3 LDIST_{ijt} + \beta_4 ADJ_{ijt} + \beta_5 FTA_{ijt} + \mu_{ijt} \quad \text{Equation 3.9}$$

$$LTA_{ijt} = \beta_0 + \beta_1 LOY_{jt} + \beta_2 LDY_{it} + \beta_3 LDIST_{ijt} + \beta_4 ADJ_{ijt} + \beta_5 FTA_{ijt} + \beta_6 LTP_{ijt} + \beta_7 LANG_{ijt} + \mu_{ijt} \quad \text{Equation 3.10}$$

The first measurement used common language between destination and origin countries as a proxy while the second measurement utilised EPI to investigate the role of English as a medium of communication when the international tourists do not share a

common language with the residents of the tourism destination. Equation 3.10 can be rewritten as Equation 3.10.1 and Equation 3.10.2 in which the COML proxies the common language and EPI means English Proficiency Index. The afore-mentioned estimations were repeated for individual countries, ASEAN as a region and ASEAN+3 countries as a group to find out the effect of these determinants at different level.

$$LTA_{ijt} = \beta_0 + \beta_1 LOY_{jt} + \beta_2 LDY_{it} + \beta_3 LDIST_{ijt} + \beta_4 ADJ_{ijt} + \beta_5 FTA_{ijt} + \beta_6 LTP_{ijt} + \beta_7 COML_{ijt} + \mu_{ijt} \quad \text{Equation 3.10.1}$$

$$LTA_{ijt} = \beta_0 + \beta_1 LOY_{jt} + \beta_2 LDY_{it} + \beta_3 LDIST_{ijt} + \beta_4 ADJ_{ijt} + \beta_5 FTA_{ijt} + \beta_6 LTP_{ijt} + \beta_7 LEPI_{ijt} + \mu_{ijt} \quad \text{Equation 3.10.2}$$

Lastly, in order to find out if different level of English proficiency at destination countries would affect the tourism demand, the ASEAN+3 countries were grouped into high, medium and low English proficiency level. According to these levels, Equation 3.10.2 was estimated again.

The score for EPI proficiency is shown in Table 3.3[7]. For this study, countries with "high" and "very high" proficiency bands were grouped together to become "high English proficiency" while countries with "low" and "very low" proficiency band were grouped as "low English proficiency".

[7] Please refer to https://www.ef.com/wwen/epi/about-epi/#proficiency-bands/ for details.

The purpose of doing such grouping was to enrich the data since there were only one country with very high (Singapore) and very low proficiency (Cambodia) level as shown in Table 3.4. Moreover, the estimation outcome for Singapore and Cambodia under this circumstance would be the same with the estimation for individual country and would not provide extra information. It is important to note that EPI data are not available for Brunei and Myanmar while Laos and Philippines only have 2 years observations, which are insufficient for estimation.

Table 3.3: Definition and Score of EF EPI Proficiency Band

Proficiency Band	Score
Very High	600 and above
High	550-599
Moderate	500-549
Low	450-499
Very Low	Below 450

(Source: EF Education First Ltd, 2017)

Table 3.4: ASEAN+3 Countries EPI Level

Destination Countries	EPI Level
Cambodia	Very low
Indonesia	Low
Malaysia	High
Singapore	Very high
Thailand	Low
Vietnam	Moderate
China	Low
Japan	Low
Korea	Moderate

(Source: EF Education First Ltd, 2017)

3.4.1 The Expected Sign of International Tourism Demand Determinants

The estimated coefficient signs allow us to understand the relationship between tourism demand and its determinants. Firstly, it is hypothesised that LOY has a positive relationship with LTA whereby increase in origin countries' real income is expected to improve tourist expenditure power. Improvement in income will subsequently encourage tourist to travel to ASEAN+3 countries since it has now become more affordable. This positive relationship has been found in the studies of Brida et al. (2008); Ahmad Kosnan et al. (2013); and Jerabak (2019), among others. It is also notable that tourists visiting ASEAN+3 countries are mostly from emerging and developed countries. Therefore, it is expected that improvement in spending power will induce their travelling intention.

Meanwhile, for destination countries' real income level, this study expected the coefficient sign of LDY would be positive. This means that increase in ASEAN+3 countries' real income level was expected to attract more tourist inflows such as that identified in Eeckels et al. (2012); Lorde et al. (2015); Alawin and Abu-Lila (2016); and Seabra et al. (2020). Since tourism sector is one of the engines of growth in ASEAN+3 countries, economic improvement in these countries was expected to be re-invested back into the tourism sector for infrastructure development and it was expected that such development would attract more tourist.

ASEAN+3 countries that are far from tourist origin countries are likely to be an unfavourable choice for tourist. This is because of the increased travelling time and cost associated with the distance, especially for ASEAN+3 countries. Anticipation on this negative effect of LDIST on tourist arrival is parallel with the findings of Rodriguez et al. (2012), Ekanayake et al. (2012), Santana-Gallego et al. (2015) and Wamboye et al. (2020).

On the contrary, adjacent or neighbouring countries was hypothesised to have favourable influence on tourists' travelling decision because of the associated travelling time and cost savings resultant from a shorter distance. This favourable effect may be further enhanced by the similarity and familiarity in culture, climate and language that exist between ASEAN+3 countries and its adjacent countries. Thus, a positive coefficient sign was expected for ADJ. This corresponded to studies by Ahmad Kosnan et al. (2013); Khadaroo and Seetanah (2008); Eita et al. (2011); Karaman (2016); Rosello et al. (2017); and Okafor et al. (2018).

FTA was hypothesised to have a positive influence on tourism demand in ASEAN+3 countries. Such expectation was based on the establishment of free trade agreements between the ASEAN+3 countries with their trading partners. Frequent trading activities were expected to enhance the familiarity and sense of curiosity towards ASEAN+3 countries and thus, attracting tourists from these trading partners to visit. Eita et al. (2011); Morelli (2015); and Rossello et al. (2017) verified that trade agreement induces tourism demand.

Tourism price was expected to have a negative effect on tourism demand in this study. Increase in ASEAN+3 countries' tourism price would discourage tourist arrivals into these countries, causing them to choose cheaper destinations. The expectation of an adverse relationship between tourism price with tourism demand has been found in the studies of Garin-Munoz and Montero-Martin (2007); Fabibi and Abdul Rahim (2009); De Vita and Kyaw (2013); Untong et al. (2015); and Ulucak et al. (2019).

Lastly, language was hypothesised to be positively related to tourism demand in this study. The existence of a common language and English was expected to bridge the communication gap between international tourists and the locals and attracts more visits.

This is especially true in ASEAN+3 countries where most societies are multilinguistic and the study of English as a second language is generally implemented. Durbarry (2008); Khadaroo and Seetanah (2008); Ahmad Kosnan et al. (2013); Alawin and Abu-Lila (2016); Rosello et al. (2017); and Okafor et al. (2018) consistently verified the positive influence of language on tourism demand.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 Overview

This chapter is structured as follows. Section 4.2 provides a brief introduction mainly to reiterate the purpose of the study. Section 4.3 reports the descriptive statistics and correlation analysis that have been conducted. The estimation results of static linear panel models are presented in Section 4.4.

4.2 Introduction

This study has been conducted with the aim to find out the factors affecting tourism demand for ASEAN+3 countries under three model specifications. The factors included in this study are the tourist origin country's income level; tourism destination country's income level; distance; free trade agreement; common border; tourism price; common language; and English Proficiency Index (EPI). The period of study spanned from 2012 to 2017 using annual observations.

This chapter presents the empirical results of the estimation techniques that have been discussed in Chapter 3. To identify the relationship between tourism demand and its determinants for the countries under study, static linear panel models were estimated and the best fit models were selected using the empirical testing procedure as explained in Chapter 3. Nevertheless, it is important to present and discuss on the descriptive statistics and correlation analysis results beforehand.

4.2.1 Descriptive Statistics and Correlation Analysis

Table 4.1 summarizes the descriptive statistics for all countries under study. The tabulated statistics included the overall, between and within standard deviations for the variables used in this study which are tourist arrivals; origin countries' real income; destination countries' real income; distance; free trade agreement; adjacent countries; tourism price; common language; and EPI.

For Brunei, the source of variation came from between countries except for Brunei's national income. The standard deviation was not calculated for EPI because this data was unavailable for Brunei. For Cambodia, the source of variation for the variables under study came from between countries as well, except for Cambodia's national income and EPI. It is worthwhile to mention that Cambodia does not share a common language with its top ten origin countries and hence, the standard deviation is zero. Similarly, Indonesia's national income and EPI had the source of variation coming from within countries while the rest of the variables' variation came from between countries.

For Laos, the variations came from between countries apart from its national income. The standard deviation of EPI for Laos was not computed since the data only spanned two years. For Malaysia, the national income and EPI variations were from within countries. The variations for the remaining variables were from between countries.

Table 4.1: Summary of Descriptive Statistics

Brunei										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.840	1.118	0.208	0.810	0.303	0.303	4.465	0.462	-
	Between	0.848	1.166	0.000	0.753	0.316	0.316	4.667	0.483	-
	Within	0.217	0.074	0.208	0.369	0.000	0.000	0.052	0.000	-
Cambodia										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.651	2.116	0.109	1.192	0.462	0.462	3.606	0.000	0.024
	Between	0.648	2.211	0.000	1.185	0.483	0.483	3.768	0.000	0.000
	Within	0.197	0.082	0.109	0.369	0.000	0.000	0.094	0.000	0.024
Indonesia										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.745	1.418	0.078	0.943	0.403	0.303	2.346	0.403	0.008
	Between	0.751	1.480	0.000	0.907	0.422	0.316	2.451	0.422	0.000
	Within	0.196	0.076	0.078	0.369	0.000	0.000	0.074	0.000	0.008
Laos										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.488	1.377	0.124	1.195	0.494	0.462	3.432	0.303	-
	Between	1.535	1.437	0.000	1.188	0.516	0.483	3.586	0.316	-
	Within	0.241	0.078	0.124	0.369	0.000	0.000	0.100	0.000	-
Malaysia										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.071	1.749	0.094	1.035	0.303	0.494	2.955	0.494	0.018
	Between	1.111	1.826	0.000	1.010	0.316	0.516	3.088	0.516	0.000
	Within	0.135	0.094	0.094	0.369	0.000	0.000	0.087	0.000	0.018

Table 4.1 continued

Myanmar										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.694	1.579	0.110	1.035	0.462	0.403	3.289	0.000	-
	Between	0.581	1.649	0.000	1.010	0.483	0.422	3.436	0.000	-
	Within	0.415	0.074	0.110	0.369	0.000	0.000	0.075	0.000	-
Philippines										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.788	1.336	0.037	0.892	0.494	0.000	2.357	0.494	-
	Between	0.798	1.394	0.000	0.849	0.516	0.000	2.463	0.516	-
	Within	0.195	0.081	0.037	0.369	0.000	0.000	0.063	0.000	-
Singapore										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.616	1.524	0.043	1.108	0.303	0.303	3.382	0.403	0.043
	Between	0.635	1.591	0.000	1.092	0.316	0.316	3.535	0.422	0.000
	Within	0.100	0.078	0.043	0.369	0.000	0.000	0.063	0.000	0.043
Thailand										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.610	2.017	0.073	1.030	0.462	0.403	2.980	0.303	0.042
	Between	0.605	2.103	0.000	1.005	0.483	0.422	3.114	0.316	0.000
	Within	0.193	0.140	0.073	0.369	0.000	0.000	0.085	0.000	0.042
Vietnam										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.717	1.305	0.058	0.889	0.469	0.303	2.183	0.000	0.018
	Between	0.707	1.356	0.000	0.845	0.434	0.316	2.279	0.000	0.000
	Within	0.240	0.140	0.058	0.369	0.219	0.000	0.113	0.000	0.018

Table 4.1 continued

China										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.343	2.125	0.083	0.806	0.475	0.494	2.255	0.494	0.022
	Between	1.400	2.216	0.000	0.748	0.408	0.516	2.355	0.516	0.000
	Within	0.102	0.144	0.083	0.369	0.271	0.000	0.101	0.000	0.022
Japan										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.170	1.473	0.107	0.749	0.502	0.000	2.082	0.000	0.020
	Between	1.122	1.538	0.000	0.681	0.497	0.000	2.173	0.000	0.000
	Within	0.467	0.068	0.107	0.369	0.159	0.000	0.098	0.000	0.020
Korea										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.053	1.636	0.058	0.830	0.503	0.000	2.521	0.000	0.014
	Between	1.074	1.705	0.000	0.776	0.502	0.000	2.634	0.000	0.000
	Within	0.232	0.135	0.058	0.369	0.150	0.000	0.088	0.000	0.014
ASEAN										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.641	1.600	1.454	1.029	0.434	0.385	4.985	0.433	0.143
	Between	1.633	1.604	1.457	0.965	0.431	0.386	5.005	0.435	0.152
	Within	0.225	0.094	0.104	0.367	0.069	0.000	0.082	0.000	0.029
ASEAN+3										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.714	1.663	2.042	0.982	0.468	0.375	4.652	0.417	0.117
	Between	1.701	1.665	2.046	0.914	0.455	0.376	4.666	0.418	0.125
	Within	0.246	0.101	0.100	0.367	0.113	0.000	0.085	0.000	0.026

Table 4.1 continued

High EPI										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.874	1.655	0.073	1.074	0.301	0.435	3.287	0.492	0.035
	Between	0.885	1.688	0.003	1.031	0.308	0.444	3.358	0.503	0.012
	Within	0.119	0.086	0.073	0.368	0.000	0.000	0.076	0.000	0.033
Moderate EPI										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	0.902	1.506	1.142	0.867	0.490	0.219	2.923	0.000	0.022
	Between	0.889	1.532	1.165	0.803	0.463	0.224	2.985	0.000	0.016
	Within	0.235	0.137	0.058	0.368	0.187	0.000	0.101	0.000	0.016
Low EPI										
Standard Deviation	Variable	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
	Overall	1.297	1.890	2.256	0.969	0.491	0.401	4.076	0.348	0.104
	Between	1.281	1.903	2.273	0.905	0.475	0.404	4.109	0.351	0.112
	Within	0.260	0.107	0.091	0.367	0.140	0.000	0.090	0.000	0.026

For Myanmar and Vietnam, the source of variation was from between countries, with the exception for both nations' national income. In addition, Myanmar does not share a common language with its tourist origin countries while Philippines has no adjacent countries since it is an island country, and thus the standard deviations were zero. Moreover, since Myanmar does not have EPI data available and Philippines has insufficient EPI data, the standard deviations have been omitted.

There were some similarities noted for Singapore, Thailand and Vietnam. The source of variation mostly came from between countries except for national income and EPI, where the variations were from within countries. The same pattern has been identified for China, Japan and Korea. It is noteworthy to mention that neither Japan nor Korea share a common language with its tourist origin countries and both countries do not have adjacent countries.

Meanwhile, for ASEAN and ASEAN+3 countries, the source of variation for all variables under study came from between countries. For high EPI countries, the variations in national income and EPI were from within countries while the rest were between countries. Moderate and low EPI countries both had the source of variation for all variables coming from between countries. Moderate EPI countries do not share a common language with its tourist origin countries and therefore, the standard deviation had been zero. In summary, the source of variation for most variables came from between countries and this justifies the use of panel data estimation techniques.

Table 4.2 shows the summary of correlation analysis for all variables and countries under study. The purpose of a correlation analysis is to identify the correlation relationship among independent variables. For Brunei, the correlations among the independent variables were satisfactory except for the correlation between LTP and FTA which was relatively high

at 0.785. Meanwhile, for Cambodia, a few relatively high correlations were detected among LDIST and LROY; FTA and LDIST; ADJ and LROY; and ADJ and LDIST; whereby the values were 0.772; -0.776; -0.771; and -0.744, respectively. A high correlation of 0.972 was detected for LEPI and LRDY. These values suggested a possible multicollinearity problem that would require attention in the static linear panel estimation later.

There were two relatively high correlations for Indonesia, which were 0.764 for LDIST and LROY as well as -0.775 for COML and LDIST. The remaining correlations were satisfactory. At the same time, there were three relatively high correlations found for Laos, which were 0.733 (LDIST and LROY); -0.810 (ADJ and LDIST); and -0.740 (LTP and LDIST). The remaining correlations were acceptable. Correlations among explanatory variables for Malaysia were reasonable except for ADJ and LDIST at -0.789. Myanmar also produced similar outcome as Malaysia whereby there was only one relatively high correlation between FTA and LDIST.

Table 4.2: Summary of Correlation Analysis

Brunei									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.060	1.000							
LRDY	-0.053	0.015	1.000						
LDIST	-0.350	0.678	0.443	1.000					
FTA	0.071	-0.316	0.000	-0.575	1.000				
ADJ	0.586	-0.348	0.000	0.266	0.111	1.000			
LTP	-0.148	-0.227	0.005	-0.493	0.785	-0.074	1.000		
COML	0.521	-0.501	0.000	-0.551	0.218	0.509	0.196	1.000	
LEPI	-	-	-	-	-	-	-	-	-
Cambodia									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.258	1.000							
LRDY	0.097	0.014	1.000						
LDIST	-0.457	0.772	-0.201	1.000					
FTA	0.572	-0.486	0.000	-0.776	1.000				
ADJ	0.541	-0.771	0.000	-0.744	0.429	1.000			
LTP	0.649	-0.678	0.006	-0.596	0.693	0.675	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	0.100	0.017	0.972	-0.177	0.000	0.000	0.003	-	1.000
Indonesia									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.281	1.000							
LRDY	-0.085	0.017	1.000						
LDIST	-0.665	0.764	0.177	1.000					
FTA	0.492	-0.507	0.000	-0.620	1.000				
ADJ	0.440	-0.410	0.000	-0.467	0.167	1.000			
LTP	-0.355	-0.066	0.006	0.010	0.476	-0.139	1.000		
COML	0.658	-0.615	0.000	-0.775	0.250	0.667	-0.342	1.000	
LEPI	-0.182	-0.009	0.169	0.223	0.000	0.000	-0.005	0.000	1.000

Table 4.2 continued

Laos									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.576	1.000							
LRDY	0.037	-0.001	1.000						
LDIST	-0.883	0.733	-0.271	1.000					
FTA	0.560	-0.455	0.000	-0.656	1.000				
ADJ	0.925	-0.465	0.000	-0.810	0.535	1.000			
LTP	0.585	-0.650	0.011	-0.740	0.695	0.467	1.000		
COML	0.651	-0.412	0.000	0.560	0.272	0.509	0.091	1.000	
LEPI	-	-	-	-	-	-	-	-	1.000
Malaysia									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.269	1.000							
LRDY	-0.021	0.021	1.000						
LDIST	-0.813	0.557	0.304	1.000					
FTA	0.318	-0.232	0.000	-0.478	1.000				
ADJ	0.726	-0.668	0.000	-0.789	0.272	1.000			
LTP	-0.179	0.204	0.023	0.031	0.435	-0.021	1.000		
COML	0.774	-0.291	0.000	-0.531	0.272	0.583	-0.138	1.000	
LEPI	0.032	-0.010	-0.682	-0.304	0.000	0.000	-0.020	0.000	1.000
Myanmar									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.189	1.000							
LRDY	-0.390	0.008	1.000						
LDIST	-0.397	0.746	0.342	1.000					
FTA	0.306	-0.519	0.000	-0.738	1.000				
ADJ	0.711	0.101	0.000	-0.443	0.327	1.000			
LTP	-0.007	-0.413	0.012	-0.446	0.595	-0.019	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	-	-	-	-	-	-	-	-	-

Table 4.2 continued

Philippines									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.578	1.000							
LRDY	0.170	0.003	1.000						
LDIST	-0.089	0.474	-0.278	1.000					
FTA	0.228	-0.225	0.000	-0.449	1.000				
ADJ	-	-	-	-	-	-			
LTP	0.598	-0.033	0.000	-0.543	0.364	-	1.000		
COML	0.068	0.035	0.000	0.595	-0.250	-	-0.342	1.000	
LEPI	-	-	-	-	-	-	-	-	-
Singapore									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.136	1.000							
LRDY	0.070	0.023	1.000						
LDIST	-0.398	0.666	-0.067	1.000					
FTA	0.341	-0.567	0.000	-0.496	1.000				
ADJ	0.134	-0.337	0.000	-0.687	0.111	1.000			
LTP	-0.046	-0.495	-0.002	-0.285	0.414	-0.286	1.000		
COML	0.393	0.157	0.000	-0.055	-0.167	0.167	-0.457	1.000	
LEPI	0.092	0.008	0.645	-0.280	0.000	0.000	-0.006	0.000	1.000
Thailand									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.144	1.000							
LRDY	0.102	0.016	1.000						
LDIST	-0.263	0.806	-0.062	1.000					
FTA	0.388	-0.385	0.000	-0.706	1.000				
ADJ	0.219	-0.754	0.000	-0.696	0.327	1.000			
LTP	-0.033	-0.524	0.002	-0.440	0.450	0.334	1.000		
COML	-0.115	-0.771	0.000	-0.597	0.218	0.667	0.651	1.000	
LEPI	0.132	0.000	0.638	-0.181	0.000	0.000	0.005	0.000	1.000

Table 4.2 continued

Vietnam									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.450	1.000							
LRDY	0.268	-0.020	1.000						
LDIST	-0.262	0.551	-0.343	1.000					
FTA	0.362	-0.039	-0.006	-0.265	1.000				
ADJ	0.697	0.440	0.000	-0.177	0.227	1.000			
LTP	0.388	-0.218	0.018	-0.456	0.253	-0.114	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	0.171	-0.029	0.668	-0.396	-0.036	0.000	0.021	-	1.000
China									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.112	1.000							
LRDY	0.015	-0.015	1.000						
LDIST	-0.253	0.498	-0.282	1.000					
FTA	-0.021	-0.219	-0.079	0.072	1.000				
ADJ	0.456	-0.593	0.000	-0.127	-0.144	1.000			
LTP	-0.259	-0.331	0.016	-0.674	-0.216	0.067	1.000		
COML	0.602	-0.323	0.000	-0.103	0.289	0.167	-0.492	1.000	
LEPI	0.027	-0.003	0.776	-0.188	-0.012	0.000	0.006	0.000	1.000
Japan									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.454	1.000							
LRDY	-0.236	0.005	1.000						
LDIST	-0.679	0.136	0.206	1.000					
FTA	-0.724	-0.581	-0.040	0.249	1.000				
ADJ	-	-	-	-	-	-			
LTP	0.420	-0.212	0.039	-0.716	-0.149	-	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	-0.305	-0.006	0.685	0.302	-0.057	-	0.028	-	1.000

Table 4.2 continued

Korea									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.719	1.000							
LRDY	0.080	0.015	1.000						
LDIST	-0.585	0.028	-0.127	1.000					
FTA	-0.331	-0.197	0.057	0.382	1.000				
ADJ	-	-	-	-	-	-			
LTP	-0.330	-0.488	0.009	-0.235	0.064	-	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	0.038	0.008	0.286	-0.222	0.051	-	-0.004	-	1.000
ASEAN									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.071	1.000							
LRDY	0.357	0.033	1.000						
LDIST	-0.419	0.694	0.121	1.000					
FTA	0.358	-0.358	0.089	-0.534	1.000				
ADJ	0.431	-0.524	-0.090	-0.630	0.257	1.000			
LTP	0.349	-0.260	0.134	-0.257	0.365	0.146	1.000		
COML	0.455	-0.250	0.215	-0.323	0.216	0.312	0.382	1.000	
LEPI	0.247	-0.003	0.580	-0.021	0.176	-0.032	0.390	0.473	1.000
ASEAN+3									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.015	1.000							
LRDY	0.383	-0.129	1.000						
LDIST	-0.404	0.556	0.030	1.000					
FTA	-0.080	-0.249	-0.279	-0.244	1.000				
ADJ	0.420	-0.476	-0.037	-0.497	0.131	1.000			
LTP	0.323	-0.285	0.182	-0.292	0.169	0.195	1.000		
COML	0.428	-0.230	0.032	-0.267	0.199	0.347	0.335	1.000	
LEPI	0.087	0.011	0.271	-0.014	0.129	-0.073	0.305	0.356	1.000

Table 4.2 continued

High EPI									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.144	1.000							
LRDY	0.002	0.013	1.000						
LDIST	-0.629	0.612	0.163	1.000					
FTA	0.314	-0.382	0.000	-0.484	1.000				
ADJ	0.566	-0.560	0.016	-0.709	0.193	1.000			
LTP	-0.140	-0.093	0.000	-0.108	0.407	-0.206	1.000		
COML	0.547	-0.029	-0.019	-0.235	0.068	0.236	-0.141	1.000	
LEPI	0.018	0.057	-0.008	-0.214	0.000	-0.119	0.085	0.140	1.000
Moderate EPI									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	0.585	1.000							
LRDY	0.109	-0.205	1.000						
LDIST	-0.446	0.290	-0.174	1.000					
FTA	-0.076	-0.094	-0.152	0.082	1.000				
ADJ	0.357	0.308	-0.229	-0.087	0.184	1.000			
LTP	0.010	-0.419	0.595	-0.365	0.026	-0.195	1.000		
COML	-	-	-	-	-	-	-	-	
LEPI	0.138	-0.150	0.713	-0.340	-0.104	-0.160	0.420	-	1.000
Low EPI									
	LTA	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML	LEPI
LTA	1.000								
LROY	-0.102	1.000							
LRDY	0.468	-0.128	1.000						
LDIST	-0.344	0.624	0.052	1.000					
FTA	-0.186	-0.293	-0.270	-0.292	1.000				
ADJ	0.362	-0.598	-0.028	-0.487	0.077	1.000			
LTP	0.462	-0.393	-0.375	-0.392	-0.056	0.301	1.000		
COML	0.475	-0.419	0.224	-0.343	0.107	0.403	0.094	1.000	
LEPI	0.204	-0.009	0.817	0.142	-0.105	-0.170	-0.017	0.135	1.000

The correlations among the explanatory variables were satisfactory for both Philippines and Singapore - none of the correlation was larger than 0.7. On one hand, Thailand had four relatively strong correlations among LDIST and LROY; ADJ and LROY; COML and LROY; and FTA with LDIST. The correlation values were 0.806; -0.754; -0.771;

and -0.706, respectively. Vietnam, on the other hand, had reasonable correlations among the independent variables. China yielded the same correlation outcomes as Vietnam.

Japan had one comparatively high correlation at -0.716 between LTP and LDIST. The rest were satisfactory. Korea had an acceptable result in which no high correlation was found among the explanatory variables. The same outcome was observed for both ASEAN and ASEAN+3 countries.

For high EPI countries, there was one relatively high correlation detected for ADJ and FTA at -0.709. Otherwise, the correlations were satisfactory. Finally, for moderate and low EPI countries, there was no high correlation identified among the independent variables. In general, most correlation coefficients were within reasonable ranges and provided good assurance to the choice of variables used in this study. The results also suggested that multicollinearity needed to be checked to ensure the robustness of the static linear panel estimation.

4.3 Static Linear Panel Models Selection Procedure

This section focuses on the discussion of the static linear panel models results. The variables included in the static linear panel models were tourist arrival (TA); real income level of tourist origin country (ROY); real income level of tourism destination country (RDY); distance measurement (DIST); tourism price (TP); and English Proficiency Index (EPI) and dummy variables representing free trade agreement (FTA); adjacent/neighbouring country (ADJ); and common language (COML).

The final model selection procedures have been detailed out in Table 4.3 to Table 4.17 for augmented gravity model using common language; in Table 4.18 to Table 4.28 for augmented gravity model using EPI; and in Table 4.29 to Table 4.31 for tourism demand model grouped by English proficiency level. The summarised results of the final static linear panel models are presented in Table 4.32, Table 4.34 and Table 4.37. Discussion in this section shall be based on the final model for each individual country, ASEAN countries, ASEAN+3 countries and countries categorised by levels of EPI (high, moderate and low).

4.3.1 Model Selection Procedure for Augmented Gravity Model (Common Language)

The empirical results for Brunei are documented in Table 4.3. Comparison between POLS and REM using Breusch-Pagan Lagrange multiplier (BPLM) test revealed that REM was preferred over POLS with test statistic significant at 1 percent level. Next, Chow test statistics was significant at 1 percent level and thus, reject the null hypothesis of common intercept. Sargan-Hansen (SH) test statistic was found to be significant and this led to the conclusion that FEM was the preferred final model.

A series of diagnostic test was carried out to make sure the model was robust. The model was inspected for multicollinearity, heteroskedasticity and serial correlation. The mean value of variance inflation factor (VIF) was 3.12, hence there was no multicollinearity problem for Brunei but there are problems of heteroscedasticity and serial correlation. The model was estimated again using panel-corrected standard errors (PCSE) to these problems.

Table 4.3: Results of Static Linear Panel Data Models for Brunei

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-10.30*** (-6.56)	-10.28*** (-4.66)	-12.07** (-2.16)	-9.80*** (-7.62)
LROY	0.64*** (7.90)	0.62*** (3.18)	1.05* (1.92)	0.59*** (9.25)
LRDY	1.87*** (4.66)	0.86* (1.78)	0.30 (0.51)	1.56*** (4.13)
LDIST	-1.24*** (-6.99)	-0.64** (-2.39)	-0.32 (-1.01)	-1.05*** (-5.73)
FTA	-0.02 (-0.06)	0.77 (0.67)	-	0.05 (0.16)
ADJ	1.21*** (5.29)	1.15 (1.39)	-	1.21*** (7.70)
LTP	-0.10*** (-4.71)	-0.10 (-1.38)	-0.47 (-0.58)	-0.09*** (-3.58)
COML	0.32* (1.81)	0.78 (1.34)	-	0.38* (1.77)
Chow Test (*p*-value)		18.73*** (0.00)		
BPLM Test (*p*-value)		70.61*** (0.00)		-
SH Test (*p*-value)		8.25* (0.08)		-
Multicollinearity	-	-	3.12	-
Heteroskedasticity	-	-	788.31*** (0.00)	-
Serial Correlation	-	-	17.82*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Table 4.4: Results of Static Linear Panel Data Models for Cambodia

Dependent Variable: LTA				
Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	-10.89* (-1.83)	-10.08*** (-3.19)	-10.23*** (-3.06)	-8.80*** (-3.00)
LROY	0.18*** (2.72)	0.21* (1.71)	0.40 (1.10)	0.22** (2.05)
LRDY	1.41 (1.64)	1.27*** (2.93)	1.19** (2.52)	1.10*** (2.82)
LDIST	0.09 (0.42)	0.05 (0.45)	0.05 (0.42)	-
FTA	0.50 (1.11)	0.42 (0.83)	-	0.33 (0.89)
ADJ	0.81** (2.54)	0.82 (1.46)	-	0.77* (1.87)
LTP	0.10** (2.27)	0.11 (1.43)	0.25 (0.68)	0.11** (2.47)
COML	-	-	-	-
Chow Test (*p*-value)		26.56*** (0.00)		27.12*** (0.00)
BPLM Test (*p*-value)		97.07*** (0.00)		97.13*** (0.00)
Hausman Test (*p*-value)		0.31 (0.98)		0.32 (0.95)
Multicollinearity	-	13.09	-	2.34
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity. The variable with highest variance inflation factors, LDIST was removed and the model was re-estimated with the same procedure. The final model was REM after re-estimation.

Table 4.4 shows the empirical results for Cambodia. The REM was the preferred model as suggested by BPLM test statistic, which was significant at 1 percent level. The Chow test statistic was statistically significant at 1 percent level, favouring FEM. The Hausman test was insignificant and this led to the conclusion that the preferred final model was REM. The mean value of VIF was 13.09, which suggested that there was multicollinearity problem. The same testing procedure was repeated after dropping LDIST and the final model preferred is REM.

Petersen (2009) stated that Feasible Generalized Least Square of REM produces unbiased standard errors only when the firm effect[8] is permanent. For this study, "firm effect" refers to cross-section correlation of residuals from different countries. Hence, it is more suitable to be termed as "country effect". Under the default REM estimation, the standard errors are biased when country effect is present because it violates the assumption that residuals are independent and individually distributed. This can be overcome by using REM clustered by country effect (cross-section unit) to produce unbiased standard errors because the standard errors can be adjusted for country effect to get a robust result.

For Indonesia, the empirical results are tabulated in Table 4.5. The BPLM test revealed that REM was preferred since the test statistic was 75.26 and significant at 1 percent level. The Chow test statistic was 18.40 and significant, meaning that the cross-sections did not share a common intercept. The Hausman test statistic was 1.85, but insignificant. This, the REM was final model. The calculated mean value of VIF at 3.34 showed that there was no multicollinearity problem. After that, the REM was estimate again using standard errors clustered by country (cross-section unit) to get a robust result.

Table 4.6 illustrates the empirical results for Laos. The BPLM test statistic was statistically significant, indicating REM was the preferred model. Next, the SH test statistic was 29.82 and significant at 1 percent level, which and favoured FEM over REM. Lastly, the Chow test statistic was statistically significant at 1 percent level. Thus, the final model was FEM. There was no problem of multicollinearity (VIF=9.74), but heteroskedasticity and

[8] Firm effect is defined as the residuals of a cross-section unit that are correlated across years.

serial correlation were detected. These problems were rectified by estimating the model again with the PCSE.

Table 4.5: Results of Static Linear Panel Data Models for Indonesia

Dependent Variable: LTA				
Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	-1.07 (-0.24)	-2.86 (-1.13)	-3.67 (-1.04)	-2.86** (-2.55)
LROY	0.27*** (6.44)	0.36*** (3.13)	0.68** (2.24)	0.36*** (3.48)
LRDY	-0.26 (-0.49)	-0.07 (-0.26)	-0.17 (-0.54)	-0.07 (-0.62)
LDIST	-0.28** (-2.60)	-0.38*** (-6.71)	-0.38*** (-6.02)	-0.38*** (-4.20)
FTA	1.32*** (8.22)	1.34*** (3.19)	-	1.34*** (4.11)
ADJ	0.30* (1.74)	0.33 (0.59)	-	0.33*** (3.44)
LTP	-0.17*** (-7.82)	-0.18** (-2.49)	-0.15 (-0.48)	-0.18** (-2.56)
COML	0.46** (2.07)	0.44 (0.78)	-	0.44 (1.31)
Chow Test (*p*-value)		18.40*** (0.00)		
BPLM Test (*p*-value)		75.26*** (0.00)		-
Hausman Test (*p*-value)		1.85 (0.76)		-
Multicollinearity	-	3.34	-	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

The empirical results for Malaysia are documented in Table 4.7. POLS was the preferred model as suggested by the BPLM test statistic that was not significant. The SH test statistic was significant at 1 percent level and showed that FEM was the preferred model. Meanwhile, the Chow test statistic was significant at 1 percent and therefore, it was concluded that the final model was FEM. The mean value of VIF is 2.93, implying that there

was no multicollinearity problem. Heteroskedasticity and serial correlation were detected, and thus the model was estimated again with the PCSE to rectify these problems.

Table 4.6: Results of Static Linear Panel Data Models for Laos

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-9.07** (-2.07)	-5.83 (-1.42)	-3.15 (-0.61)	-2.76 (-0.95)
LROY	0.00 (0.13)	0.03 (0.58)	1.06** (1.98)	0.06 (0.94)
LRDY	1.09 (1.61)	0.62 (0.99)	-0.61 (-1.01)	0.15 (0.33)
LDIST	0.27 (1.21)	0.09 (0.43)	-0.32* (-1.69)	-0.06 (-0.43)
FTA	-0.31** (-2.37)	-0.32** (-2.04)	-	-0.32* (-1.96)
ADJ	2.48*** (9.03)	2.28*** (8.29)	-	2.12*** (8.51)
LTP	0.18*** (4.81)	0.16*** (4.05)	0.65 (1.49)	0.13*** (3.65)
COML	1.83*** (6.45)	1.67*** (5.54)	-	1.53*** (6.85)
Chow Test (*p*-value)		5.33*** (0.00)		
BPLM Test (*p*-value)		11.55*** (0.00)		-
SH Test (*p*-value)		29.82*** (0.00)		-
Multicollinearity	-	-	9.74	-
Heteroskedasticity	-	-	33.85*** (0.00)	-
Serial Correlation	-	-	39.90*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Table 4.7: Results of Static Linear Panel Data Models for Malaysia

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-15.56*** (-5.24)	-9.78*** (-3.45)	1.09 (0.41)	-10.55*** (-4.36)
LROY	0.25*** (9.40)	0.25*** (5.22)	-0.07 (-0.25)	0.25*** (7.53)
LRDY	2.41*** (5.81)	1.30*** (3.00)	0.05 (0.11)	1.58*** (4.33)
LDIST	-0.81*** (-11.29)	-0.47*** (-4.76)	-0.00 (-0.08)	-0.64*** (-7.16)
FTA	-0.09 (-0.61)	0.30 (1.12)	-	0.11 (0.64)
ADJ	0.39** (2.59)	0.87*** (3.51)	-	0.65*** (3.56)
LTP	-0.06*** (-4.51)	-0.08*** (-3.23)	-0.32 (-0.63)	-0.07*** (-5.36)
COML	0.76*** (8.80)	0.78*** (4.93)	-	0.78*** (9,91)
Chow Test (*p*-value)		10.60*** (0.00)		
BPLM Test (*p*-value)		0.34 (0.28)		-
SH Test (*p*-value)		64.87*** (0.00)		-
Multicollinearity	-	-	2.93	-
Heteroskedasticity	-	-	457.79*** (0.00)	-
Serial Correlation	-	-	162.82*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Referring to Table 4.8 for Myanmar, REM was the preferred model as suggested by the BPLM test that was significant at 5 percent level. Then, SH test was estimated and the test statistic was 10.24 and significant at 5 percent level, which favoured FEM. The Chow test statistic was 3.29 and statistically significant, concluding that the final model was FEM. The mean value of VIF was 9.90, showing that multicollinearity was not a problem.

However, there were problems of heteroskedasticity and serial correlation. The model was estimated again using the PCSE to overcome these problems.

Table 4.8: Results of Static Linear Panel Data Models for Myanmar

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	4.40 (0.88)	6.07 (1.06)	-1.58 (-0.22)	4.44 (0.86)
LROY	0.28** (2.43)	0.25 (1.62)	0.49 (0.40)	0.23 (1.52)
LRDY	-1.14 (-1.19)	-1.53 (-1.36)	-2.01 (-0.99)	-1.22 (-1.23)
LDIST	-0.41 (-1.58)	-0.29*** (-0.91)	0.00 (0.01)	-0.30 (-0.99)
FTA	0.17 (0.80)	0.29 (0.77)	-	0.26 (0.95)
ADJ	0.58* (1.88)	0.68 (1.53)	-	0.70* (1.88)
LTP	-0.01 (-0.82)	-0.01 (-0.48)	-1.55 (-1.14)	-0.01 (-0.79)
COML	-	-	-	-
Chow Test (*p*-value)		3.29*** (0.00)		
BPLM Test (*p*-value)		2.83** (0.04)		-
SH Test (*p*-value)		10.24** (0.03)		-
Multicollinearity	-	-	9.90	-
Heteroskedasticity	-	-	131.11*** (0.00)	-
Serial Correlation	-	-	231.70*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

The results for Philippines are presented in Table 4.9. BPLM test indicated that REM was preferred at 1 percent significance level. The Chow test statistic was significant at 1 percent level, indicating that the cross-sections had different intercept. Hausman test statistic was insignificant. Thus, REM was the preferred final model. The mean VIF value was 2.11,

hence there was no multicollinearity problem. Finally, REM was estimated again using standard errors clustered by country (cross-section unit) to get a robust result.

Table 4.9: Results of Static Linear Panel Data Models for Philippines

Dependent Variable: LTA				
Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	-13.86 (-1.54)	-17.32*** (-3.26)	-15.59** (-2.55)	-17.32*** (-4.45)
LROY	0.46*** (12.68)	0.45*** (5.33)	0.14 (0.23)	0.45*** (5.40)
LRDY	1.33 (1.18)	1.73*** (2.65)	1.85** (2.57)	1.73*** (3.77)
LDIST	-0.33*** (-3.87)	-0.27*** (-4.06)	-0.25*** (-3.47)	-0.27*** (-5.93)
FTA	0.26*** (3.18)	0.28 (1.14)	-	0.28** (2.26)
ADJ	-	-	-	-
LTP	0.17*** (8.60)	0.17*** (3.33)	-0.27 (-0.35)	0.17*** (6.51)
COML	0.77*** (7.45)	0.72*** (2.87)	-	0.72*** (5.20)
Chow Test (*p*-value)		18.08*** (0.00)		
BPLM Test (*p*-value)		79.59*** (0.00)		-
Hausman Test (*p*-value)		0.40 (0.98)		-
Multicollinearity	-	2.11	-	-
Observations	60	60	60	60

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Table 4.10 illustrates the results for Singapore. REM was the preferred model as suggested by the statistically significant BPLM test. SH test was significant at 1 percent level, favouring FEM. Chow test statistic was also statistically significant, leading to the same conclusion. The mean VIF value was 2.77, showing no multicollinearity problem.

However, there are heteroskedasticity and serial correlation issue and thus, the model was re-estimated using the PCSE to rectify these problems.

Table 4.10: Results of Static Linear Panel Data Models for Singapore

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	2.18 (0.27)	-8.42*** (-3.22)	-8.01** (-2.57)	-7.07* (-1.78)
LROY	0.34*** (7.73)	0.33*** (3.68)	0.55 (1.65)	0.31*** (7.19)
LRDY	-0.27 (-0.27)	0.53* (1.66)	0.37 (0.95)	0.61 (1.23)
LDIST	-0.57*** (-6.47)	-0.12*** (-3.42)	-0.12** (-2.54)	-0.30*** (-3.61)
FTA	0.89*** (4.77)	1.27*** (3.30)	-	1.07*** (7.10)
ADJ	-0.85*** (-2.90)	0.39 (0.99)	-	-0.16 (-0.63)
LTP	-0.01 (-0.67)	0.05 (1.27)	0.18 (0.46)	0.02 (0.67)
COML	0.46*** (3.63)	0.70*** (2.60)	-	0.61*** (3.69)
Chow Test (*p*-value)		74.69*** (0.00)		
BPLM Test (*p*-value)		46.34*** (0.00)		-
SH Test (*p*-value)		20.44*** (0.00)		-
Multicollinearity	-	-	2.77	-
Heteroskedasticity	-	-	423.61*** (0.00)	-
Serial Correlation	-	-	29.55*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Thailand's empirical results are shown in Table 4.11. The BPLM test statistic was statistically significant and favoured REM. SH test statistic was significant at 5 percent level, favouring FEM over REM. The Chow test statistic was significant at 1 percent level, which concluded FEM as the final model. The mean VIF value was 3.90, showing no

multicollinearity problem. However, heteroskedasticity and serial correlation were detected and rectified by estimating the model again using the PCSE.

Table 4.11: Results of Static Linear Panel Data Models for Thailand

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-1.46 (-0.22)	-5.91** (-2.36)	-7.00*** (-2.96)	-4.56 (-1.48)
LROY	0.38*** (5.95)	0.51*** (4.70)	1.06*** (4.24)	0.34*** (4.37)
LRDY	0.23 (0.31)	0.37 (1.29)	0.02 (0.09)	0.47 (1.28)
LDIST	-0.50*** (-3.47)	-0.28*** (-5.08)	-0.32*** (-5.78)	-0.32*** (-3.33)
FTA	0.05 (0.24)	0.44 (1.09)	-	0.25 (0.95)
ADJ	0.95*** (4.25)	1.37** (2.25)	-	1.00*** (6.72)
LTP	0.02 (0.92)	0.01 (0.14)	0.66* (1.69)	0.01 (0.56)
COML	-0.33 (-0.92)	0.39 (0.42)	-	-0.19 (-0.44)
Chow Test (*p*-value)		44.23*** (0.00)		
BPLM Test (*p*-value)		94.42*** (0.00)		-
SH Test (*p*-value)		12.55** (0.01)		-
Multicollinearity	-	-	3.90	-
Heteroskedasticity	-	-	1519.02*** (0.00)	-
Serial Correlation	-	-	10.80*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA, ADJ and COML are omitted by Stata because of collinearity.

Referring to Table 4.12 for Vietnam, REM was the preferred model as suggested by BPLM test statistic at 1 percent significance level. The Chow test statistic was statistically significant at 1 percent level. The SH test was also found to be significant. Therefore, the final model was FEM. The mean VIF value was 2.31, showing no multicollinearity problem

but there were problems of heteroscedasticity and serial correlation. A re-estimation of the model using PCSE provided a robust result.

Table 4.12: Results of Static Linear Panel Data Models for Vietnam

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-27.45*** (-4.85)	-27.86*** (-6.21)	-27.37*** (-5.25)	-26.24*** (-6.63)
LROY	0.13** (2.69)	0.12 (1.28)	0.12 (0.50)	0.14*** (2.61)
LRDY	3.60*** (4.67)	3.67*** (6.29)	3.66*** (5.72)	3.45*** (6.41)
LDIST	0.06 (0.74)	0.07 (0.81)	0.08 (0.76)	0.04 (0.60)
FTA	0.18** (2.15)	0.17* (1.95)	0.17 (1.62)	0.12 (1.28)
ADJ	1.49*** (7.77)	1.54*** (3.53)	-	1.53*** (5.80)
LTP	0.16*** (7.73)	0.17*** (3.19)	0.22 (0.72)	0.17*** (5.38)
COML	-	-	-	-
Chow Test (*p*-value)		14.67*** (0.00)		
BPLM Test (*p*-value)		72.29*** (0.00)		-
SH Test (*p*-value)		36.69*** (0.00)		-
Multicollinearity	-	-	2.31	-
Heteroskedasticity	-	-	2845.20*** (0.00)	
Serial Correlation	-	-	29.143*** (0.00)	
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

Table 4.13 shows the results for China. The BPLM test was statistically significant and in favour of REM. The SH test statistic was 65.12 and significant at 1 percent level, thus FEM was favoured over REM. Lastly, Chow test statistic was significant at 1 percent level. The final model was FEM. The mean VIF value was 3.22 and showed no multicollinearity

problem. However, heteroskedasticity and serial correlation were detected and rectified by re-estimating the model using the PCSE.

Table 4.13: Results of Static Linear Panel Data Models for China

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	37.97*** (2.89)	-1.97 (-0.50)	-1.38 (-0.50)	5.59 (0.74)
LROY	0.52*** (7.78)	0.41*** (4.51)	0.04 (0.26)	0.50*** (7.24)
LRDY	-2.85** (-2.67)	-0.08 (-0.26)	0.27 (1.11)	-0.46 (-0.75)
LDIST	-1.29*** (-6.37)	-0.14** (-2.05)	-0.04 (-0.73)	-0.66*** (-4.12)
FTA	0.32* (1.68)	-0.02 (-0.37)	-0.00 (-0.01)	0.25* (1.74)
ADJ	2.18*** (9.75)	1.84*** (5.55)	-	2.11*** (7.13)
LTP	-0.18** (-2.09)	0.14* (1.75)	-0.30 (-1.16)	0.02 (0.36)
COML	1.28*** (4.36)	2.21*** (5.83)	-	1.85*** (7.71)
Chow Test (*p*-value)		170.80*** (0.00)		
BPLM Test (*p*-value)		39.18*** (0.00)		-
SH Test (*p*-value)		65.12*** (0.00)		-
Multicollinearity	-	-	3.22	-
Heteroskedasticity	-	-	3046.80*** (0.00)	-
Serial Correlation	-	-	6.88** (0.02)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

The empirical results presented in Table 4.14 are for Japan. REM was the preferred model as suggested by BPLM test at 1 percent significance level. The SH test was significant at 1 percent level, making FEM the favoured model over REM. The Chow test was significant at 1 percent level. The final model was FEM. The mean VIF value was 1.91 and

showed no multicollinearity. The model was re-estimated with PCSE to rectify heteroskedasticity and serial correlation.

Table 4.14: Results of Static Linear Panel Data Models for Japan

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	20.85*** (3.41)	17.12*** (5.34)	12.07* (1.90)	20.66*** (4.95)
LROY	0.23*** (4.44)	0.45*** (3.64)	2.50*** (3.10)	0.27*** (5.71)
LRDY	-1.46** (-2.49)	-1.28*** (-4.19)	-2.41*** (-2.75)	-1.49*** (-3.79)
LDIST	-0.91*** (-7.34)	-0.97*** (-11.25)	-0.81*** (7.96)	-0.88*** (-8.18)
FTA	-0.95*** (-5.97)	-0.25 (-1.45)	0.11 (0.61)	-0.84*** (-5.03)
ADJ	-	-	-	-
LTP	0.00 (0.13)	0.02 (0.32)	1.13 (1.21)	0.02 (0.51)
COML	-	-	-	-
Chow Test (*p*-value)		26.36*** (0.00)		
BPLM Test (*p*-value)		64.07*** (0.00)		-
SH Test (*p*-value)		35.78*** (0.00)		-
Multicollinearity	-	-	1.91	-
Heteroskedasticity	-	-	67.99*** (0.00)	-
Serial Correlation	-	-	84.91*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

Korea's results are tabulated in Table 4.15. The BPLM test statistic at 5 percent significance level favoured REM. SH test was significant at 1 percent level with a test statistic of 41.70 and favoured FEM. The Chow test statistic was 10.45 and significant at 1 percent level. Thus, the final model was FEM. The mean VIF value was 1.27 and meant no

multicollinearity problem. However, heteroskedasticity and serial correlation were spotted and thus, the model was estimated again using the PCSE to overcome these problems.

Table 4.15: Results of Static Linear Panel Data Models for Korea

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	4.65 (0.63)	-1.54 (0.25)	-12.74** (-2.21)	-1.07 (-0.18)
LROY	0.43*** (13.72)	0.43*** (8.60)	-0.22 (-0.61)	0.43*** (11.44)
LRDY	-0.32 (-0.42)	0.16 (0.26)	1.38** (2.18)	0.21 (0.35)
LDIST	-0.85*** (-13.97)	-0.62*** (-8.53)	-0.26*** (-3.13)	-0.75*** (-9.71)
FTA	0.14 (1.46)	-0.04 (-0.34)	0.01 (0.06)	0.08 (0.66)
ADJ	-	-	-	-
LTP	-0.06*** (-3.29)	-0.04 (-1.48)	-0.87 (-1.55)	-0.05** (-2.59)
COML	-	-	-	-
Chow Test (*p*-value)		10.45*** (0.00)		
BPLM Test (*p*-value)		3.38** (0.03)		-
SH Test (*p*-value)		41.70*** (0.00)		-
Multicollinearity	-	-	1.27	-
Heteroskedasticity	-	-	144.40*** (0.00)	-
Serial Correlation	-	-	10.43** (0.01)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

As illustrated in Table 4.16 for ASEAN, the BPLM test statistic at 1 percent significance level suggested REM over POLS. The Hausman test statistic was 43.97 and significant at 1 percent level, which favoured FEM. The Chow test statistic was significant and implied FEM as the final model. The mean VIF value was 1.72, showing no

multicollinearity problem. Despite so, there were heteroskedasticity and serial correlation issues, so the model was re-estimated with the PCSE.

Table 4.16: Results of Static Linear Panel Data Models for ASEAN

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-6.32*** (-11.69)	-5.98*** (-8.44)	-4.63*** (-4.61)	-7.41*** (-14.90)
LROY	0.32*** (9.09)	0.36*** (5.81)	0.62*** (4.97)	0.26*** (6.68)
LRDY	0.76*** (25.64)	0.45*** (7.92)	-0.01 (-0.13)	0.73*** (19.47)
LDIST	-0.48*** (-6.92)	-0.32*** (-13.29)	-0.31*** (-13.14)	-0.26*** (-6.49)
FTA	0.09 (0.74)	0.28** (2.39)	0.31** (2.47)	0.23** (2.18)
ADJ	1.41*** (10.46)	1.47*** (4.93)	-	1.57*** (9.84)
LTP	-0.01 (-1.57)	-0.01 (-0.60)	0.13 (0.92)	-0.01 (-0.99)
COML	0.24** (2.35)	0.56** (2.14)	-	0.29** (2.57)
Chow Test (*p*-value)		131.59*** (0.00)		
BPLM Test (*p*-value)		1340.52*** (0.00)		-
Hausman Test (*p*-value)		43.97*** (0.00)		-
Multicollinearity	-	-	1.72	-
Heteroskedasticity	-	-	14981.83*** (0.00)	-
Serial Correlation	-	-	180.15*** (0.00)	-
Observations	600	600	600	600

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

Table 4.17 tabulates the empirical results for ASEAN+3. The BPLM test statistic was statistically significant, which indicates REM was the preferred model. Hausman test statistic was also statistically significant and FEM was favoured over REM. Lastly, Chow test statistic was significant at 1 percent level and it can be concluded that the final model

was FEM. The mean value of VIF (1.43) implied that there was no multicollinearity problem. The model was re-estimated with PCSE to rectify the problems of heteroskedasticity and serial correlation that were detected.

Table 4.17: Results of Static Linear Panel Data Models for ASEAN+3

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-5.28*** (-12.03)	-5.58*** (-9.09)	-2.41** (-2.59)	-6.66*** (-15.62)
LROY	0.38*** (13.99)	0.37*** (7.22)	0.52*** (4.69)	0.33*** (10.59)
LRDY	0.59*** (31.78)	0.44*** (10.84)	-0.12 (-1.26)	0.58*** (23.30)
LDIST	-0.57*** (-11.30)	-0.37*** (-16.05)	-0.35*** (15.80)	-0.32*** (-8.95)
FTA	0.08 (0.96)	0.14** (2.05)	0.15** (2.21)	0.14* (1.77)
ADJ	1.42*** (12.45)	1.57*** (5.92)	-	1.62*** (11.38)
LTP	-0.01** (-2.06)	-0.01 (-0.68)	-0.09 (-0.67)	-0.00 (-0.75)
COML	0.64*** (7.02)	0.73*** (3.14)	-	0.65*** (6.42)
Chow Test (*p*-value)		112.03*** (0.00)		
BPLM Test (*p*-value)		1704.20*** (0.00)		-
Hausman Test (*p*-value)		67.11*** (0.00)		-
Multicollinearity	-	-	1.43	-
Heteroskedasticity	-	-	16455.89*** (0.00)	-
Serial Correlation	-	-	188.40*** (0.00)	-
Observations	780	780	780	780

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ and COML are omitted by Stata because of collinearity.

4.3.2 Model Selection for Augmented Gravity Model (EPI)

The empirical results for Cambodia are documented in Table 4.18. REM was the preferred model as suggested by BPLM test statistic at 1 percent significance level. Next, Chow test statistics was statistically significant at 1 percent level, but Hausman test was found to be insignificant. Therefore, the final model was REM. The mean value of VIF was 17.29, indicating the presence of multicollinearity problem. After dropping LDIST from the model, the same procedure was repeated and the preferred REM was re-estimated using standard errors clustered by country (cross-section unit) for a robust result.

Table 4.18: Results of Static Linear Panel Data Models for Cambodia

Dependent Variable: LTA

Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	-10.13 (-0.37)	-11.64 (-1.34)	-12.14 (-1.33)	-11.85** (-2.00)
LROY	0.19** (2.16)	0.17 (1.25)	-0.24 (-0.50)	0.18* (1.73)
LRDY	0.89 (0.22)	0.10 (0.08)	-0.00 (-0.00)	-0.00 (-0.01)
LDIST	0.11 (0.37)	0.01 (0.14)	0.07 (0.49)	-
FTA	0.63 (1.06)	0.40 (0.73)	-	0.38 (0.91)
ADJ	1.00** (2.33)	0.74 (1.20)	-	0.75** (2.20)
LTP	0.07 (1.22)	0.09 (1.12)	0.14 (0.40)	0.09** (2.11)
LEPI	0.37 (0.03)	2.20 (0.52)	3.64 (0.80)	2.43 (1.03)
Chow Test (*p*-value)		31.66*** (0.00)		32.74*** (0.00)
BPLM Test (*p*-value)		46.93*** (0.00)		47.32*** (0.00)
Hausman Test (*p*-value)		1.03 (0.95)		0.84 (0.93)
Multicollinearity	-	-	17.29	7.85
Observations	40	40	40	40

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity. The variable with highest variance inflation factors, LDIST was removed and the model was re-estimated with same procedure. The final model was RE after re-estimation.

Meanwhile for Indonesia, the empirical results are shown in Table 4.19. BPLM test revealed that REM was preferred with test statistic of 92.06 at 1 percent significance level. The Chow test statistic was statistically significant and confirmed that the cross-sections did not share common intercept. The Hausman test statistic was 0.68 and insignificant, showing that REM was the preferred final model. The calculated mean value of VIF at 2.30 showed no multicollinearity problem. Re-estimating REM was done with standard errors clustered by country (cross-section unit) to produce a robust result.

Table 4.19: Results of Static Linear Panel Data Models for Indonesia

Dependent Variable: LTA

Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	24.35 (1.12)	36.09*** (3.39)	35.04*** (3.03)	36.09*** (4.38)
LROY	0.25*** (5.60)	0.25** (2.47)	0.37 (1.31)	0.25*** (2.70)
LRDY	0.07 (0.14)	-0.12 (-0.52)	-0.14 (-0.50)	-0.12 (-1.41)
LDIST	-0.38*** (-4.11)	-0.26*** (-4.32)	-0.26*** (-4.01)	-0.26*** (-4.15)
FTA	1.29*** (7.83)	1.51*** (3.78)	-	1.51*** (6.03)
ADJ	0.52*** (3.44)	0.61 (1.35)	-	0.61*** (3.51)
LTP	-0.19*** (-9.09)	-0.22*** (-3.67)	-0.28 (-0.98)	-0.22*** (-3.63)
LEPI	-6.93 (-1.27)	-9.77*** (-3.77)	-9.54*** (-3.48)	-9.77** (-4.46)
Chow Test (*p*-value)		25.64*** (0.00)		
BPLM Test (*p*-value)		92.06*** (0.00)		-
Hausman Test (*p*-value)		0.68 (0.98)		-
Multicollinearity	-	2.30	-	-
Observations	60	60	60	60

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity.

As shown in Table 4.20 for Malaysia, the BPLM test favoured REM and was statistically significant. The SH test statistic was 10.66 and significant at 5 percent level, which favoured FEM. The Chow test statistic was 32.17 and significant at 1 percent level, and thus the final model was FEM. The mean value of VIF was 2.98 and indicated no multicollinearity problem. Heteroskedasticity and serial correlation problems were solved by re-estimating the model using the PCSE.

Table 4.20: Results of Static Linear Panel Data Models for Malaysia

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	21.71 (1.09)	-9.45 (-0.97)	-11.91 (-1.21)	-1.89 (-0.18)
LROY	0.32*** (8.08)	0.26** (2.38)	-0.02 (-0.07)	0.32*** (6.71)
LRDY	1.64** (2.20)	-0.04 (-0.12)	-0.09 (-0.18)	1.07*** (2.60)
LDIST	-0.88*** (-7.65)	-0.00 (-0.02)	0.12 (0.85)	-0.49*** (-3.98)
FTA	0.17 (0.73)	1.18* (1.84)	-	0.66*** (2.66)
ADJ	0.84*** (3.82)	1.98*** (4.20)	-	1.42*** (5.02)
LTP	-0.10*** (-4.83)	-0.13** (-2.22)	-0.23 (-0.45)	-0.12*** (-5.25)
LEPI	-7.65* (-1.96)	1.47 (0.69)	2.90 (1.39)	-1.63 (-0.69)
Chow Test (*p*-value)		32.17*** (0.00)		
BPLM Test (*p*-value)		60.57*** (0.00)		-
SH Test (*p*-value)		10.66** (0.03)		-
Multicollinearity	-	-	2.98	-
Heteroskedasticity	-	-	485.93*** (0.00)	-
Serial Correlation	-	-	75.01*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity.

Table 4.21 presents the static linear panel models result for Singapore. The BPLM test statistic was 24.56, significant and favoured REM. The SH test was significant, and therefore FEM was preferred. The Chow test statistic was 40.76 at 1 percent significance level and concluded that the final model was FEM. The mean VIF value was 3.78 and indicated no multicollinearity. However, heteroskedasticity and serial correlation were detected and then, rectified by re-estimating the model with the PCSE.

Table 4.21: Results of Static Linear Panel Data Models for Singapore

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	4.38 (0.73)	-7.05** (-2.35)	-10.30*** (-3.19)	4.48 (0.82)
LROY	0.38*** (11.78)	0.31** (3.37)	0.63* (1.95)	0.38*** (10.67)
LRDY	5.63*** (5.38)	0.75 (1.07)	-0.76 (-1.12)	5.43*** (5.55)
LDIST	-1.14*** (-13.94)	-0.16 (-1.49)	0.05 (0.53)	-1.12*** (-13.02)
FTA	0.47*** (3.38)	1.22*** (3.02)	-	0.49*** (3.92)
ADJ	-2.34*** (-9.31)	0.27 (0.54)	-	-2.29*** (8.75)
LTP	-0.10*** (-6.75)	0.00 (0.17)	0.28 (0.74)	-0.10*** (-6.23)
LEPI	-10.62*** (-8.31)	-0.43 (-0.36)	2.13* (2.00)	-10.31*** (-8.11)
Chow Test (*p*-value)		40.76*** (0.00)		
BPLM Test (*p*-value)		24.56*** (0.00)		-
SH Test (*p*-value)		38.23*** (0.00)		-
Multicollinearity	-	-	3.78	-
Heteroskedasticity	-	-	233.46*** (0.00)	-
Serial Correlation	-	-	14.60*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity.

The results for Thailand are tabulated in Table 4.22. REM was the preferred model by BPLM test statistic at 1 percent significance level. FEM was preferred by statistically significant SH test. The Chow test statistic was statistically significant and hence, FEM was the final model. The mean value of VIF was 3.89, suggesting that there was no multicollinearity. The detection of heteroskedasticity and serial correlation was solved by re-estimating the model with PCSE.

Table 4.22: Results of Static Linear Panel Data Models for Thailand

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	-0.56 (-0.08)	-6.69** (-2.49)	-7.84*** (-3.03)	-4.79 (-1.64)
LROY	0.40*** (6.12)	0.51*** (4.73)	1.04*** (4.12)	0.35*** (5.65)
LRDY	0.60 (0.60)	0.18 (0.51)	-0.14 (-0.38)	0.43 (0.94)
LDIST	-0.52*** (-3.20)	-0.26*** (-4.08)	-0.29*** (-4.49)	-0.30*** (-3.25)
FTA	0.08 (0.38)	0.42 (1.04)	-	0.31 (1.37)
ADJ	0.87*** (4.02)	1.58*** (2.87)	-	0.97*** (5.30)
LTP	0.01 (0.53)	0.03 (0.44)	0.59 (1.50)	0.00 (0.46)
LEPI	-1.05 (-0.55)	0.54 (077)	0.56 (0.82)	0.06 (0.09)
Chow Test (*p*-value)		44.51*** (0.00)		
BPLM Test (*p*-value)		95.88*** (0.00)		-
SH Test (*p*-value)		11.17** (0.04)		-
Multicollinearity	-	-	3.89	-
Heteroskedasticity	-	-	950.32*** (0.00)	-
Serial Correlation	-	-	16.81*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity.

Table 4.23 illustrates the static linear panel models result for Vietnam. BPLM test indicated that REM was preferred at 1 percent significance level. Chow test statistic (significant at 1 percent level) indicated that the cross-sections did not share common intercept. Insignificant Hausman test concluded that REM was the final model. The mean value of VIF was 2.57, hence there was no multicollinearity problem. REM was re-estimated using standard errors clustered by country (cross-section unit) to get a robust result.

Table 4.23: Results of Static Linear Panel Data Models for Vietnam

Dependent Variable: LTA

Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	-29.51** (-2.69)	-42.61*** (-2.80)	-69.62** (-2.51)	-42.61*** (-4.19)
LROY	0.13** (2.46)	0.06 (0.58)	0.10 (0.41)	0.06 (0.62)
LRDY	3.50*** (3.89)	3.89*** (6.24)	4.61*** (5.25)	3.89*** (5.74)
LDIST	0.07 (0.75)	0.25 (1.28)	0.58 (1.71*)	0.25 (1.61)
FTA	0.18** (2.14)	0.16* (1.85)	0.16 (1.54)	0.16*** (2.61)
ADJ	1.50*** (7.48)	1.76*** (3.57)	-	1.76*** (5.52)
LTP	0.17*** (7.52)	0.19*** (3.30)	0.28 (0.93)	0.19*** (5.26)
LEPI	0.68 (0.22)	3.15 (1.01)	8.11 (1.55)	3.15* (1.68)
Chow Test (*p*-value)		15.37*** (0.00)		
BPLM Test (*p*-value)		72.52*** (0.00)		-
Hausman Test (*p*-value)		1.45 (0.96)		-
Multicollinearity	-	2.57	-	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

Referring to Table 4.24 for China, REM was the preferred model as suggested by BPLM test statistic that was statistically significant. The SH test was estimated and the test statistic was 162.52 and significant at 1 percent level, which favoured FEM. Next, Chow test statistic was also statistically significant at 1 percent level, concluding that the final model was FEM. The mean value of VIF was 2.37, indicating that multicollinearity was not a problem. However, heteroskedasticity and serial correlation were spotted. Hence, the model was re-estimated using the PCSE to solve these problems.

Table 4.24: Results of Static Linear Panel Data Models for China

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	53.33*** (3.17)	-0.24 (-0.04)	-3.88 (-1.33)	20.55* (1.67)
LROY	0.40*** (5.65)	0.12 (1.24)	0.08 (0.52)	0.33*** (4.16)
LRDY	-5.37*** (-2.97)	-0.64 (-1.11)	-0.29 (-0.82)	-2.32* (-1.78)
LDIST	-1.86*** (-10.17)	-0.19** (-2.09)	-0.05 (-1.10)	-1.38*** (-7.27)
FTA	0.30 (1.34)	0.00 (0.09)	-0.03 (-0.51)	0.31 (1.36)
ADJ	2.08*** (8.02)	1.56*** (3.77)	-	1.99*** (5.91)
LTP	-0.48*** (-8.42)	-0.18** (-2.29)	-0.14 (-0.55)	-0.39*** (-6.25)
LEPI	5.01 (0.78)	2.32 (1.21)	2.18** (2.13)	3.69 (0.83)
Chow Test (*p*-value)		251.02*** (0.00)		
BPLM Test (*p*-value)		9.15*** (0.00)		-
SH Test (*p*-value)		162.52*** (0.00)		-
Multicollinearity	-	-	2.37	-
Heteroskedasticity	-	-	2976.54*** (0.00)	-
Serial Correlation	-	-	4.84* (0.05)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

Table 4.25 documents the empirical results for Japan. The BPLM test was significant at 1 percent level and favoured REM. SH test was statistically significant, suggesting FEM was favoured over REM. Lastly, Chow test statistic was 28.36 and significant at 1 percent level, and thus the final model was FEM. The mean VIF value was 2.19, indicating no multicollinearity problem but heteroskedasticity and serial correlation were detected. These problems were solved by re-estimating the model using the PCSE.

Table 4.25: Results of Static Linear Panel Data Models for Japan

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	40.80*** (3.17)	33.74*** (5.19)	28.11*** (3.23)	33.61*** (3.86)
LROY	0.22*** (4.29)	0.44*** (3.63)	2.27*** (2.97)	0.25*** (6.14)
LRDY	-0.65 (-0.88)	-0.72** (-2.10)	-1.94** (-2.29)	-0.88 (-1.58)
LDIST	-0.82*** (-6.30)	-0.84*** (-9.22)	-0.71*** (-6.88)	-0.86*** (-7.73)
FTA	-1.00*** (-6.29)	-0.25 (-1.55)	0.07 (0.42)	-0.92*** (-5.79)
ADJ	-	-	-	-
LTP	0.02 (0.57)	0.06 (0.74)	1.19 (1.34)	0.02 (0.57)
LEPI	-7.34* (-1.75)	-5.89*** (-2.90)	-4.94** (-2.55)	-4.89* (-1.73)
Chow Test (*p*-value)		28.36*** (0.00)		
BPLM Test (*p*-value)		68.29*** (0.00)		-
SH Test (*p*-value)		45.97*** (0.00)		-
Multicollinearity	-	-	2.19	-
Heteroskedasticity	-	-	281.46*** (0.00)	-
Serial Correlation	-	-	40.74*** (0.00)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

The results illustrated in Table 4.26 are for Korea. BPLM favoured REM with the test statistics significant at 5 percent level. SH test statistic was also statistically significant and favoured FEM over REM. The Chow test statistic was significant at 1 percent level and thus, the final model was FEM. The mean value of VIF at 1.29 indicated no multicollinearity problem. Nevertheless, heteroskedasticity and serial correlation were detected and re-estimating the model with the PCSE rectified these problems.

Table 4.26: Results of Static Linear Panel Data Models for Korea

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	39.61*** (3.07)	24.68** (2.18)	0.91 (0.09)	29.08** (2.54)
LROY	0.43*** (14.85)	0.43*** (8.99)	-0.24 (-0.67)	0.43*** (12.89)
LRDY	0.26 (0.36)	0.53 (0.86)	1.59** (2.51)	0.36 (0.62)
LDIST	-0.89*** (-15.44)	-0.68*** (-9.28)	-0.32*** (-3.59)	-0.82*** (-11.82)
FTA	0.18* (2.00)	-0.00 (-0.03)	0.02 (0.13)	0.13 (1.21)
ADJ	-	-	-	-
LTP	-0.07*** (-3.77)	-0.05* (-1.74)	-1.01* (-1.81)	-0.06*** (-3.03)
LEPI	-10.06*** (-3.19)	-7.31*** (-2.70)	-3.86 (-1.66)	-7.78*** (-2.96)
Chow Test (*p*-value)		8.58*** (0.00)		
BPLM Test (*p*-value)		2.72** (0.04)		-
SH Test (*p*-value)		45.16*** (0.00)		-
Multicollinearity	-	-	1.29	-
Heteroskedasticity	-	-	194.06*** (0.00)	-
Serial Correlation	-	-	7.95** (0.02)	-
Observations	60	60	60	60

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

As illustrated in Table 4.27 for ASEAN, REM was the preferred model as suggested by the significant BPLM test. The Hausman test was insignificant, which again favoured REM. Chow test statistic was 71.17 at 1 percent significance level, which rejected the null hypothesis of a common intercept. The mean value of VIF was 1.88 and indicated no multicollinearity. The REM was re-estimated with the standard errors clustered by country (cross-section unit) to provide a robust result.

Table 4.27: Results of Static Linear Panel Data Models for ASEAN

Dependent Variable: LTA

Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	0.86 (0.78)	-1.37 (-1.01)	-2.70 (-1.58)	-1.37 (-1.00)
LROY	0.30*** (11.69)	0.29*** (6.23)	0.42*** (3.53)	0.29*** (4.33)
LRDY	0.40*** (10.98)	0.40*** (6.20)	0.38** (2.31)	0.40*** (6.85)
LDIST	-0.47*** (-8.94)	-0.29*** (-10.18)	-0.28*** (-9.12)	-0.29*** (-6.62)
FTA	0.14 (1.47)	0.24*** (2.88)	0.26*** (2.75)	0.24* (1.87)
ADJ	0.86*** (8.25)	1.07*** (4.97)	-	1.07*** (4.48)
LTP	0.04*** (6.78)	0.04*** (3.15)	0.17 (1.14)	0.04*** (3.23)
LEPI	-0.95*** (-3.22)	-0,73** (-2.09)	-0.56 (-1.33)	-0.73** (-2.34)
Chow Test (*p*-value)		71.17*** (0.00)		
BPLM Test (*p*-value)		686.10*** (0.00)		-
Hausman Test (*p*-value)		4.56 (0.60)		-
Multicollinearity	-	-	1.88	-
Observations	340	340	340	340

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

Referring to Table 4.28 for ASEAN+3 results, the BPLM test statistic was statistically significant and preferred REM. Hausman test statistic was 30.26 at 1 percent significance level, implying that FEM was favoured. Lastly, Chow test statistic was also statistically significant. Hence, the FEM was the final model. The calculated mean VIF value was 1.43 and indicated no multicollinearity. The problem of heteroskedasticity and serial correlation were solved by estimating the model again with the PCSE.

Table 4.28: Results of Static Linear Panel Data Models for ASEAN+3

Dependent Variable: LTA				
Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	0.95 (0.80)	4.58*** (3.21)	6.48*** (3.75)	-0.15 (-0.14)
LROY	0.37*** (15.34)	0.39*** (8.66)	0.45*** (3.89)	0.33*** (11.62)
LRDY	0.31*** (14.60)	0.32*** (7.32)	0.06 (0.38)	0.31*** (10.94)
LDIST	-0.54*** (-12.61)	-0.41*** (-15.66)	-0.41*** (-14.75)	-0.30*** (-8.64)
FTA	-0.00 (-0.01)	0.13** (2.27)	0.14** (2.35)	0.07 (0.78)
ADJ	1.30*** (12.76)	1.40*** (6.15)	-	1.49*** (9.67)
LTP	0.05*** (6.54)	0.06*** (3.43)	-0.08 (-0.56)	0.05*** (7.19)
LEPI	-0.82*** (-2.71)	-2.08*** (-5.97)	-2.12*** (-4.89)	-0.90*** (-3.21)
Chow Test (*p*-value)		72.03*** (0.00)		
BPLM Test (*p*-value)		1055.32*** (0.00)		-
Hausman Test (*p*-value)		30.26*** (0.00)		-
Multicollinearity	-	-	1.43	-
Heteroskedasticity	-	-	27933.36*** (0.00)	-
Serial Correlation	-	-	122.90*** (0.00)	-
Observations	520	520	520	520

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

4.3.3 Model Selection for Gravity Model Grouped by English Proficiency Level

For high EPI countries, the results are presented in Table 4.29. Comparison between POLS and REM using BPLM test revealed that REM was preferred with a test statistic of 246.90 at 1 percent significance level. The Chow test statistic at 1 percent significance level confirmed that the cross-sections did not share a common intercept. The SH test statistic was 7.57 and insignificant. Thus, REM was the final model. The calculated mean value of VIF at 2.15 showed that there was no multicollinearity. REM was re-estimated using standard errors clustered by country (cross-section unit) to provide a robust result.

Table 4.29: Results of Static Linear Panel Data Models for High EPI Countries

Dependent Variable: LTA

Variable	POLS	REM	FEM	REM (Clustered by country)
Constant	4.67 (0.54)	-8.72*** (-3.21)	-8.00*** (-2.73)	-8.72*** (-2.95)
LROY	0.26*** (6.31)	0.18** (2.06)	0.11 (0.55)	0.18** (2.41)
LRDY	1.40* (1.88)	-0.21 (-0.86)	-0.24 (-0.71)	-0.21 (-1.13)
LDIST	-0.63*** (-6.33)	0.04 (0.58)	0.10 (1.30)	0.04 (0.50)
FTA	0.48** (2.15)	1.13** (2.00)	-	1.13*** (4.01)
ADJ	0.41* (1.90)	1.42*** (3.34)	-	1.42*** (3.86)
LTP	-0.05*** (-2.76)	-0.03 (-0.67)	-0.17 (-0.57)	-0.03 (-0.55)
LEPI	-3.37* (-1.96)	1.73** (2.34)	2.10*** (2.74)	1.73** (2.46)
Chow Test (*p*-value)		125.16*** (0.00)		
BPLM Test (*p*-value)		246.90*** (0.00)		-
SH Test (*p*-value)		7.57 (0.10)		-
Multicollinearity	-	2.15	-	-
Observations	120	120	120	120

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. FTA and ADJ are omitted by Stata because of collinearity.

The results for moderate EPI countries are presented in Table 4.30. REM was the preferred model by the statistically significant BPLM test. SH test statistic was at 1 percent significance level and favoured FEM. The Chow test statistic was significant at 1 percent and confirmed FEM as the final model. The mean VIF value was 1.79 and indicated no multicollinearity. The detected heteroskedasticity and serial correlation were solved by re-estimating the model using the PCSE.

Table 4.30: Results of Static Linear Panel Data Models for Moderate EPI Countries

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	25.96** (2.57)	12.27* (1.77)	-1.94 (-0.26)	23.08*** (3.41)
LROY	0.48*** (16.19)	0.40*** (5.65)	-0.09 (-0.44)	0.46*** (13.36)
LRDY	0.24*** (4.32)	0.27** (2.46)	1.83*** (4.44)	0.23*** (3.66)
LDIST	-0.70*** (-13.53)	-0.51*** (-7.48)	-0.32*** (-4.12)	-0.61*** (-9.01)
FTA	0.12 (1.56)	0.14 (1.57)	0.10 (1.05)	0.11 (1.33)
ADJ	0.33* (1.71)	0.63*** (1.27)	-	0.45** (1.97)
LTP	0.00 (0.01)	-0.00 (-0.06)	-0.37 (-1.29)	0.01 (0.48)
LEPI	-6.98** (-2.71)	-3.81** (-2.25)	-3.16* (-1.87)	-6.36*** (-3.72)
Chow Test (*p*-value)		21.55*** (0.00)		
BPLM Test (*p*-value)		147.70*** (0.00)		-
SH Test (*p*-value)		27.27*** (0.00)		-
Multicollinearity	-	-	1.79	-
Heteroskedasticity	-	-	2597.78*** (0.00)	-
Serial Correlation	-	-	10.28*** (0.00)	-
Observations	120	120	120	120

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

As shown in Table 4.31 for low EPI countries, the significant BPLM test favoured REM over POLS. The Hausman test was 47.06 and significant at 1 percent level, which favoured FEM. The Chow test statistic was 87.72 at 1 percent significance level, and thus the final model was FEM. The mean VIF value was 3.08 and indicated no multicollinearity. However, heteroskedasticity and serial correlation were spotted and the model was re-estimated using PCSE to obtain robust results.

Table 4.31: Results of Static Linear Panel Data Models for Low EPI Countries

Dependent Variable: LTA

Variable	POLS	REM	FEM	Panel-Corrected Standard Errors
Constant	10.86*** (2.90)	5.91*** (2.82)	5.66*** (2.63)	1.16 (0.57)
LROY	0.38*** (9.72)	0.48*** (6.38)	0.73*** (4.70)	0.35*** (7.64)
LRDY	0.49*** (8.07)	0.35*** (4.97)	-0.69*** (-3.25)	0.34*** (7.29)
LDIST	-0.61*** (-8.65)	-0.47*** (-12.82)	-0.47*** (-13.56)	-0.35*** (-6.47)
FTA	0.00 (0.02)	0.11 (1.37)	0.10 (1.17)	0.02 (0.16)
ADJ	1.37*** (8.45)	1.69*** (4.34)	-	1.54*** (6.26)
LTP	0.02 (1.26)	0.04 (1.17)	-0.14 (-0.68)	0.06*** (4.13)
LEPI	-3.70*** (-3.44)	-2.64*** (-4.67)	-0.73 (-1.21)	-1.28** (-2.11)
Chow Test (*p*-value)		87.72*** (0.00)		
BPLM Test (*p*-value)		574.03*** (0.00)		-
Hausman Test (*p*-value)		47.06*** (0.00)		-
Multicollinearity	-	-	3.08	-
Heteroskedasticity	-	-	2195.37*** (0.00)	-
Serial Correlation	-	-	95.53*** (0.00)	-
Observations	280	280	280	280

Notes: Asterisk *** and ** represent 1 and 5 percent level of significance, respectively. The figures in brackets are t-statistics. ADJ are omitted by Stata because of collinearity.

4.4 Static Linear Panel Models Results

4.4.1 Augmented Gravity Model (Common Language)

The results of augmented gravity model (common language) are reported in Table 4.32. Firstly, the main variable of interest in this study which is the language is discussed. For Cambodia, Myanmar, Vietnam Japan, and Korea; these countries and their top ten tourist origin countries do not share a common language and hence, COML was omitted from the model. COML was found out to be significant for Brunei, Laos, Malaysia, Philippines, Singapore, China, ASEAN and ASEAN+3 countries, but not for Indonesia and Thailand. The results showed that COML is able to attract tourist arrivals due to its ability to lower language barrier which is similar to the findings of Durbarry (2008), Seetanah (2010), Ahmad Kosnan et al. (2013), and Alawin and Abu-Lila (2016), among others.

Next, for tourist origin country's income level (LROY), all the countries resulted in the expected positive relationship with tourist arrival into tourist. This shows that wealth effect exists, which means tourist are keen to travel when their income level improve. *Ceteris paribus*, more tourist will visit these countries when there is an increase in their income level (Botti et al., 2007; Habibi & Abdul Rahim, 2009; Jerabek, 2019). Nonetheless, this variable was found to be insignificant for Laos and Myanmar.

Tourist destination's income (LRDY) seems to play a minor role in explaining the tourist inflow for all countries. This variable is significant and has the correct positive sign for Brunei, Cambodia, Malaysia, Philippines, Vietnam, ASEAN and ASEAN+3 countries. The positive coefficient sign suggests that economy development in tourism destination countries attracts more tourist inflows (Pablo-Romero & Molina, 2013; Alawin & Abu-Lila,

2016). A good economic performance allows for better tourism sector development and maintenance, hence attracting more tourists to visit these nations.

Only Japan has a significant and negative relationship between tourism demand and income (Durbarry, 2008). This is perhaps relatable to the Dutch Disease, a condition where the initial rapid inflow of tourist stimulates faster development of tourism sector and resulted in currency appreciation of the destination country. The currency appreciation then reduces the competitiveness of other sectors (especially international trade), adversely affecting the national income. A decline in national competitiveness and income will eventually lead to the currency depreciation and thus, encourage more tourists to visit Japan.

Distance (LDIST) seems to be performing decently as this variable has the anticipated negative and significant relationship with tourist inflow for most countries. The results were on a par with previous literatures such as Leitao (2010), Ekanayake et al. (2012) and Wamboye et al. (2020). In general, it can be interpreted as most tourists are less likely to travel to destinations far from their origin countries because of increased travelling cost and time as well as the discomfort that comes with long travelling time (Ulucak et al., 2020). Nevertheless, the tourism demand for Laos, Myanmar and Vietnam is not influenced by distance while for the case of Cambodia, the LDIST has the highest VIF. As such, it was omitted from the empirical model in order to rectify the multicollinearity problem.

Meanwhile, free trade agreement (FTA) has the correct positive coefficient sign and has been significant for Indonesia, Philippines, Singapore, China, ASEAN, and ASEAN+3. In line with past studies such as Tinbergen (1962), Santeramo and Morelli (2015), and Rossello et al. (2017), FTA attracts more tourist inflows when there are more business trips; better spreading of information through the businessperson to their contacts after their own

visit; removal or easing of travel restrictions such as visa applications; elimination of border control; and so forth.

However, the opposite was found for Japan and Laos. FTA has discouraged tourist arrival into both countries instead of attracting more tourists. This has already been reported by Saayman et al. (2016), whereby mixed effects of the formal trade agreement and informal economic cooperation on international tourism demand were detected. They explained that, in early years of agreements, the effect towards international tourism demand used to be more apparent, but the effect had diminished and turned stagnant or negative over time. Meanwhile, FTA was found to be insignificant for Brunei, Cambodia, Malaysia, Myanmar, Thailand, Vietnam, and Korea.

Besides that, visitors from adjacent/neighbouring countries has been a crucial source of tourist inflow for most countries. ADJ was a significant factor under this model, except for Singapore. This variable was also omitted from Japan, Korea and Philippines's models because Japan and Philippines do not share a common border with any country while for Korea, North Korea was the only adjacent/neighbouring country that shares a common border with Korea, but North Korea is not in the top ten tourist generating countries for Korea. This outcome is parallel with the findings of Eita et al. (2011), Ahmad Kosnan et al. (2013), and Karaman (2016), for example.

Table 4.32: Summary of Augmented Gravity Model (Common Language)

Country	Final Model	Constant	LROY	LRDY	LDIST	FTA	ADJ	LTP	COML
Brunei	PCSE	-9.80***	0.59***	1.56***	-1.05***	0.05	1.21***	-0.09***	0.38*
		(-7.62)	(9.25)	(4.13)	(-5.73)	(0.16)	(7.70)	(-3.58)	(1.77)
Cambodia	REM	-8.80**	0.22*	1.10**	-	0.33	0.77*	0.11**	-
		(-3.00)	(2.05)	(2.82)		(0.89)	(1.87)	(2.47)	
Indonesia	REM	-2.86**	0.36***	-0.07	-0.38***	1.34***	0.33***	-0.18**	0.44
		(-2.55)	(3.48)	(-0.62)	(-4.20)	(4.11)	(3.44)	(-2.56)	(1.31)
Laos	PCSE	-2.76	0.06	0.15	-0.06	-0.32*	2.12***	0.13***	1.53***
		(-0.95)	(0.94)	(0.33)	(-0.43)	(-1.96)	(8.51)	(3.65)	(6.85)
Malaysia	PCSE	-10.55***	0.25***	1.58***	-0.64***	0.11	0.65***	-0.07***	0.78***
		(-4.36)	(7.53)	(4.33)	(-7.16)	(0.64)	(3.56)	(-5.36)	(9.91)
Myanmar	PCSE	4.44	0.23	-1.22	-0.30	0.26	0.70*	-0.01	-
		(0.86)	(1.52)	(-1.23)	(-0.99)	(0.95)	(1.88)	(-0.79)	
Philippines	REM	-17.32***	0.45***	1.73***	-0.27***	0.28**	-	0.17***	0.72***
		(-4.45)	(5.40)	(3.77)	(-5.93)	(2.26)		(6.51)	(5.20)
Singapore	PCSE	-7.07*	0.31***	0.61	-0.30***	1.07***	-0.16	0.02	0.61***
		(-1.78)	(7.19)	(1.23)	(-3.61)	(7.10)	(-0.63)	(0.67)	(3.69)
Thailand	PCSE	-4.56	0.34***	0.47	-0.32***	0.25	1.00***	0.01	-0.19
		(-1.48)	(4.37)	(1.28)	(-3.33)	(0.95)	(6.72)	(0.56)	(-0.44)
Vietnam	PCSE	-26.24***	0.14***	3.45***	0.04	0.12	1.53***	0.17***	-
		(-6.63)	(2.61)	(6.41)	(0.60)	(1.28)	(5.80)	(5.38)	
China	PCSE	5.59	0.50***	-0.46	-0.66***	0.25*	2.11***	0.02	1.85***
		(0.74)	(7.24)	(-0.75)	(-4.12)	(1.74)	(7.13)	(0.36)	(7.71)
Japan	PCSE	20.66***	0.27***	-1.49***	-0.88***	-0.84***	-	0.02	-
		(4.95)	(5.71)	(-3.79)	(-8.18)	(-5.03)		(0.51)	
Korea	PCSE	-1.07	0.43***	0.21	-0.75***	0.08	-	-0.05**	-
		(-0.18)	(11.44)	(0.35)	(-9.71)	(0.66)		(-2.59)	
ASEAN	PCSE	-7.41***	0.26***	0.73***	-0.26***	0.23**	1.57***	-0.01	0.29**
		(-14.90)	(6.68)	(19.47)	(-6.49)	(2.18)	(9.84)	(-0.99)	(2.57)
ASEAN+3	PCSE	-6.66***	0.33***	0.58***	-0.32***	0.14*	1.62***	-0.00	0.65***
		(-15.62)	(10.59)	(23.30)	(-8.95)	(1.77)	(11.38)	(-0.75)	(6.42)

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. Because of multicollinearity problem, LDIST for Cambodia was removed as it had the highest variance inflation factor (VIF).

Finally, for tourism price (LTP), review from past literature revealed that this variable is expected to have an inverse relationship with tourism demand. Higher tourism price will result in decline for tourism demand of a particular tourism destination. In this study, LTP was significant for Brunei, Cambodia, Indonesia, Laos, Malaysia, Philippines,

Vietnam, and Korea. For the case of Brunei, Indonesia, Malaysia, and Korea, the coefficient followed past literature (Untong et al., 2015; Tang & Tan, 2016; Puah et al., 2019). However, this variable has a positive relationship with Cambodia, Laos, Philippines and Vietnam. In other words, an increase in tourism price will induce more tourist arrival into these countries.

Puah et al. (2019) found similar result and recognised such situation as the "demand push inflation". It is a consequence of strong demand from tourists that has eventually pushed up tourism price level. Therefore, despite increase in price level, tourists still visit these countries. Another study done by Tkales and Vizek (2016) also supports the view of demand push inflation, whereby they verified that strong demand for various tourism related goods and services not only improves the region where tourism activities are dominant, the effects are also influential nationwide.

4.4.2 Augmented Gravity Model (EPI)

This section focuses on the estimation outcome for augmented gravity model with one modification done to the previous empirical model. The common language (COML) was replaced with English Proficiency Index (EPI) to test the functionality of EPI in reducing language barrier. In other words, EPI was tested for its suitability as a common language for international tourist and local community. Estimation results are not available for Brunei, Laos, Myanmar and Philippines in Table 4.34 because the EPI data published by Education First Limited is not available for Brunei and Myanmar while Laos and Philippines only have 2 years observations each. Table 4.33 details the EPI level for the remaining countries.

Table 4.33: Destination Country EPI Level

Destination Country	EPI Level
Cambodia	Low
China	Low
Indonesia	Low
Japan	Low
Thailand	Low
Vietnam	Moderate
Korea	Moderate
Malaysia	High
Singapore	High

(Source: EF Education First Ltd, 2017)

Table 4.34: Summary of Augmented Gravity Model (EPI)

Country	Final Model	Constant	LROY	LRDY	LDIST	FTA	ADJ	LTP	LEPI
Cambodia	REM	-11.85** (-2.00)	0.18* (1.73)	-0.00 (-0.01)	-	0.38 (0.91)	0.75** (2.20)	0.09** (2.11)	2.43 (1.03)
Indonesia	REM	36.09*** (4.38)	0.25*** (2.70)	-0.12 (-1.41)	-0.26*** (-4.15)	1.51*** (6.03)	0.61*** (3.51)	-0.22*** (-3.63)	-9.77** (-4.46)
Malaysia	PCSE	-1.89 (-0.18)	0.32*** (6.71)	1.07*** (2.60)	-0.49*** (-3.98)	0.66*** (2.66)	1.42*** (5.02)	-0.12*** (-5.25)	-1.63 (-0.69)
Singapore	PCSE	4.48 (0.82)	0.38*** (10.67)	5.43*** (5.55)	-1.12*** (-13.02)	0.49*** (3.92)	-2.29*** (-8.75)	-0.10*** (-6.23)	-10.31*** (-8.11)
Thailand	PCSE	-4.79 (-1.64)	0.35*** (5.65)	0.43 (0.94)	-0.30*** (-3.25)	0.31 (1.37)	0.97*** (5.30)	0.00 (0.46)	0.06 (0.09)
Vietnam	REM	-45.42*** (-4.33)	0.05 (0.49)	4.02*** (5.36)	0.28* (1.79)	0.15** (2.33)	1.75*** (5.23)	0.18*** (4.24)	3.56* (1.81)
China	PCSE	20.55* (1.67)	0.33*** (4.16)	-2.32* (-1.78)	-1.38*** (-7.27)	0.31 (1.36)	1.99*** (5.91)	-0.39*** (-6.25)	3.69 (0.83)
Japan	PCSE	33.61*** (3.86)	0.25*** (6.14)	-0.88 (-1.58)	-0.86*** (-7.73)	-0.92*** (-5.79)	-	0.02 (0.57)	-4.89* (-1.73)
Korea	PCSE	29.08** (2.54)	0.43*** (12.89)	0.36 (0.62)	-0.82*** (-11.82)	0.13 (1.21)	-	-0.06*** (-3.03)	-7.78*** (-2.96)
ASEAN	REM	-1.46 (-1.06)	0.29*** (4.31)	0.40*** (6.74)	-0.28*** (-6.41)	0.24* (1.83)	1.08*** (4.48)	0.04*** (3.07)	-0.70** (-2.22)
ASEAN+3	PCSE	-0.15 (-0.14)	0.33*** (11.62)	0.31*** (10.94)	-0.30*** (-8.64)	0.07 (0.78)	1.49*** (9.67)	0.05*** (7.19)	-0.90*** (-3.21)

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics. Because of multicollinearity problem, LDIST for Cambodia was removed as this was the variable with highest variance inflation factor (VIF).

The main explanatory variable of interest under this model specification is LEPI, which determines the role of English as a tool for communication between two parties that may have different mother tongue. LEPI was found to be significant for Indonesia, Singapore, Vietnam, Japan, Korea, ASEAN and ASEAN+3 countries, but insignificant for the remaining countries. Surprisingly, LEPI does not help to attract more tourists into the destination countries apart from Vietnam. LEPI as a common language is able attracts tourist to visit Vietnam and this is comparable with the findings of Durbarry (2008). On the contrary, LEPI seems to discourage visitors from visiting these destination countries. Looking into the detailed distribution of origin countries in Indonesia, Singapore, Japan and Korea, there are some interesting revelations that are noteworthy (see Table 4.35).

In Indonesia, 42.48 percent of the 68.32 percent tourists (top ten tourist origin countries contribution) were from countries with high level of EPI, but Indonesia is a country with low EPI level itself. Hence, a big gap of English proficiency between the destination and origin countries has resulted in English being a barrier instead of a common language that attracts tourist inflows. On the contrary, Singapore has such a high EPI that it has discouraged tourists from countries with moderate and low EPI to visit. Tourists from moderate and low EPI countries made up 54 percent of total tourist arrivals into Singapore.

Furthermore, majority of the tourists visiting Japan have low EPI (44.99 percent) and with Japan having low EPI as well, the use of English as a language for communication does not seem to be suitable and this discourages tourist inflows. Korea is a country with moderate EPI, but a large portion of its visitors have low EPI (61.30 percent) and this again represents a gap in English proficiency between origin and destination countries, which hinders the use of English as a proper common language to attract tourist.

Table 4.35: EPI Level Distribution by Country

Indonesia		Singapore		Japan		Korea	
Person *Percent*	EPI Level	Person *Percent*	EPI Level	Person *Percent*	EPI Level	Person *Percent*	EPI Level
China 2,093,171 *14.91*	Low	China 3,228,134 *18.53*	Low	China 7,355,818 *25.64*	Low	China 4,169,353 *31.26*	Low
Japan 573,310 *4.08*	Low	Indonesia 2,954,400 *16.96*	Low	Taiwan 4,564,053 *15.91*	Low	Japan 2,311,447 *17.33*	Low
India 536,902 *3.82*	Moderate	Japan 792,873 *4.55*	Low	Thailand 987,211 *3.44*	Low	Taiwan 925,616 *6.94*	Low
Korea 423,191 *3.01*	Moderate	India 1,272,077 *7.3*	Moderate	Hong Kong 2,231,568 *7.78*	Moderate	Thailand 498,511 *3.74*	Low
Malaysia 2,121,888 *15.11*	High	Korea 631,363 *3.62*	Moderate	Korea 7,140,438 *24.89*	Moderate	Russia 270,427 *2.03*	Low
Singapore 1,554,119 *11.07*	High	Vietnam 531,359 *3.05*	Moderate	United States 1,374,964 *4.79*	High	Hong Kong 658,031 *4.93*	Moderate
Australia 1,256,927 *8.95*	High	Malaysia 1,168,384 *6.71*	High	Australia 495,054 *1.73*	High	Vietnam 324,740 *2.44*	Moderate
United Kingdom 378,131 *2.69*	High	Australia 1,082,001 *6.21*	High	Malaysia 439,548 *1.53*	High	United States 868,881 *6.52*	High
United States 344,766 *2.46*	High	Philippines 736,500 *4.23*	High	Philippines 424,121 *1.48*	High	Philippines 448,702 *3.36*	High
Philippines 308,977 *2.2*	High	United States 565,430 *3.25*	High	Singapore 404,132 *1.41*	High	Malaysia 307,641 *2.31*	High
EPI Level	*Percent*		*Percent*		*Percent*		*Percent*
Low	*18.99*		*40.03*		*44.99*		*61.3*
Moderate	*6.84*		*13.97*		*32.67*		*7.37*
High	*42.48*		*20.39*		*10.94*		*12.19*
Subtotal	*68.32*		*74.39*		*88.59*		*80.86*

(Sources: EF Education First Ltd, 2017; BPS-Statistics Indonesia, 2018; Singapore Tourism Analytics Network, 2018; Japan National Tourist Organization, 2019; Korea Tourism Organization, 2019)

Table 4.36: Tourist Arrival to ASEAN and ASEAN+3 Countries by EPI Level

	ASEAN		ASEAN+3	
EPI Level	Tourist Arrival	*Percentage*	Tourist Arrival	*Percentage*
Low	39,865,015	*47.28*	98,372,451	*39.94*
Moderate	12,975,714	*15.39*	106,994,491	*43.44*
High	31,477,952	*37.33*	40,955,995	*16.63*
Total	84,318,681	*100.00*	246,322,937	*100.00*

(Source: EF Education First Ltd, 2017)

Similarly, for ASEAN and ASEAN+3 countries, estimation outcomes showed that English discourage tourist inflow. The top ten tourist generating countries statistics for ASEAN+3 countries in earlier chapter shows that a proportion of tourists are from these ASEAN+3 countries itself. This suggests two indications. Firstly, between ASEAN+3, majority of these countries share a common language and this is their preferred communication language instead of English. Secondly, most of these countries have low and moderate EPI level, hence it limits the function and preference of using English as a common language for communication (Prachanant, 2012).

Among the ASEAN+3 countries included in this model specification that used EPI as proxy for common language, Malaysia and Singapore have high EPI level; Vietnam and Korea have moderate EPI level; and Cambodia, Indonesia, Thailand, China, and Japan have low EPI level (See Table 4.33). The number of tourists with moderate and low EPI levels visiting ASEAN were 47.28 percent and 15.39 percent, respectively and those for ASEAN+3, were 39.94 percent and 43.44 percent, respectively (see Table 4.36). This means that the use of English as a common language is limited or not preferred and perhaps have discouraged tourist movement (Prachanant, 2012).

Tourists' real income (LROY) was found to be significant for all countries except Vietnam. The positive coefficient sign suggests that wealth effect exists, which means that improvement in tourist income will induce more arrivals to tourism destinations (Fereidouni et al., 2017). Meanwhile, tourism destination income (LRDY) was less influential, but significant in Malaysia, Singapore, Vietnam, ASEAN and ASEAN+3 countries. This is most probably because a rapid economic growth has encouraged more visitors to these countries (Pablo-Romero & Molina, 2013; Perles-Ribes et al., 2017). For China, tourism destination income was significant but have a negative sign which suggests the existence of Dutch Disease, a condition where the initial rapid inflow of tourist stimulates faster development of tourism sector and resulted in currency appreciation of the destination country. The currency appreciation then reduces the competitiveness of other sectors (especially international trade), adversely affecting the national income. A decline in national competitiveness and income will eventually lead to the currency depreciation and thus, encourage more tourists to visit China.

LDIST was significant for all countries and also has the expected negative coefficient, indicating that tourists are less likely to travel to tourism destination far away from home (Ekanayake et al., 2012; Rodriguez et al., 2012; Wamboye et al., 2020). This variable was omitted for Cambodia due to high VIF value to rectify multicollinearity problem. Vietnam's tourism demand has a positive relationship with LDIST, which differs from prior expectation. This is possibly due to Vietnam opening up its tourism sector and many visitors become intrigued to visit.

The World Economic Forum's Travel and Tourism Competitiveness Report 2019 highlighted that Vietnam has been the most improved country in Asia Pacific in terms of International Openness and Air Transport Infrastructure pillars. Such improvements have had a positive effect on tourist inflows. Moreover, Vietnam has been improving its competitiveness in attracting tourists as documented in Travel and Tourism Competitiveness Report 2017 (World Economic Forum, 2017). The report stated that Vietnam was among the 15 most improved countries in Travel and Tourism Competitiveness Index in 2017 and was the only selected country from Southeast Asia.

FTA has the expected significant positive influence on all the countries except for Cambodia, Thailand, China, Korea, and ASEAN+3 countries (Tinbergen, 1962; Santeramo & Morelli, 2015; Rossello et al., 2017). For Japan, FTA has a negative coefficient sign, similar with the findings of Saayman et al. (2016), and consistent with findings from previous tourism demand model specification. From the estimated results, ADJ proves to be an important factor and has been significant for all countries. A neighbouring country has been proven as a more favourable choice rather than countries located farther away from home (Ahmad Kosnan et al., 2013; Karaman, 2016; Okafor et al., 2018). Hence, the positive connection is within rational expectation.

However, Singapore has a negative sign, indicating that visitors from neighbouring country do not prefer to travel to Singapore. Identical with the case of Singapore, a significant and negative sign for adjacent/neighbouring country has also been identified in Viljoen et al. (2019), whereby the authors revealed that tourists from adjacent countries of South Africa do not necessarily travel more to the destination under study. In Singapore case, this is possibly because Malaysian citizens prefer to travel to other countries rather than

Singapore that has too much in common with Malaysia. In addition, Singapore's strong currency has also made it an expensive option for many Malaysians.

Finally, LTP is a dominant determinant and has been significant for all countries except for Japan and Thailand. Increase in LTP has resulted in decline of LTA (Leitao, 2010; Untong et al., 2015; Tang & Tan, 2016). This means tourists tend to seek alternative tourist destination when price increases. Nonetheless, the increase in tourism price has unexpectedly attracted more tourist inflows in Cambodia, Vietnam, ASEAN and ASEAN+3 countries. In this study, the computation of LTP involved two components – inflation (price level) and exchange rate, hence the movement of these two components have been crucial in determining the relationship between LTP and LTA.

The adjustment of exchange rate, when become more influential than the effect of inflation, does not deter tourists to visit the destination. This means that, if the origin country's currency strengthens against the destination's currency more than the increase in inflation, this will still encourage the tourist to visit the destination. This is true for Cambodia, Vietnam, ASEAN and ASEAN+3 countries. This positive relationship may also be an outcome of demand push inflation (Tkales & Vizek, 2016; Puah et al., 2019).

4.4.3 Augmented Gravity Model (Grouped by English Proficiency Level)

In the next estimation procedure, the countries were grouped into three categories according to the EPI level to gain further insight on the function of English as a *lingua franca*. This is to discover the role of English in removing communication barrier for destination countries at different level of English proficiency as shown in Table 4.37.

Table 4.37: Summary of Augmented Gravity Model According to EPI Level

Country	Final Model	Constant	LROY	LRDY	LDIST	FTA	ADJ	LTP	LEPI
High EPI	REM	-8.72***	0.18**	-0.21	0.04	1.13***	1.42***	-0.03	1.73**
		(-2.95)	(2.41)	(-1.13)	(0.50)	(4.01)	(3.86)	(-0.55)	(2.46)
Moderate EPI	PCSE	23.08***	0.46***	0.23***	-0.61***	0.11	0.45**	0.01	-6.36***
		(3.41)	(13.36)	(3.66)	(-9.01)	(1.33)	(1.97)	(0.48)	(-3.72)
Low EPI	PCSE	1.16	0.35***	0.34***	-0.35***	0.02	1.54***	0.06***	-1.28**
		(0.57)	(7.64)	(7.29)	(-6.47)	(0.16)	(6.26)	(4.13)	(-2.11)

Notes: Asterisk ***, ** and * represent 1, 5 and 10 percent level of significance, respectively. The figures in brackets are t-statistics.

LEPI has been significant for all three categories of English proficiency level. For high English proficient countries, it is beneficial and has encouraged more visitors to visit these countries. Nevertheless, for moderate and low EPI countries, it has had adverse impact on tourist flow. Looking into the distribution of tourist origin countries for each country grouping, it is obvious that, for low EPI countries, almost half of its tourists originated from countries that have moderate English proficiency (refer to Table 4.38). As the empirical result suggests, this can cause tourists to become reluctant to travel to low EPI countries due to communication barrier. In their comprehensive discussion on the role of English for tourism and hospitality, Zahedpisheh et al. (2017) stated that poor English proficiency and competency can result in difficulties to attract and entertain tourists.

RocketNews24 (2015) also reported that, in countries with low English level such as Japan, tourists often find themselves caught in difficult situations such as a lack of English signboards to guide them. Another study by Wiriyachittra (2002) revealed that Thai graduates in tourism industry often do not have adequate English proficiency level to work in the hospitality sector. This has caused some foreign tourists to have an unfavourable image on Thailand's tourism sector.

Similar condition can be said for destination countries with moderate EPI that are receiving tourists who can hardly speak English. In these countries, English becomes a hindrance rather than a pull factor for tourist inflow. Even with English as a second language (ESL), tourists with low English literacy will most probably struggle to communicate and obtain essential information regarding their tour (Biswas & Mamun-Or-Rashid, 2019).

Lastly, for destination with high English proficiency, nearly half of the visitors were also from countries with high EPI. This is highly beneficial for seamless communication (Suhaimi & Abdullah, 2017). This has been supported by the empirical result that depicted a positive relationship between LTA and LEPI. In short, the finding suggests that, when English proficiency level is the same between origin and destination countries, the tourists are more likely to visit because it is easy to communicate and obtain information.

The findings for the rest of the independent variables are summarised as follow; LROY and ADJ were found to be significant and had a positive relationship with LTA for all three categories, while LRDY and LDIST were important determinants for moderate and low EPI countries and also had the expected impact. FTA has been significant for countries with high EPI and have contributed towards higher tourist inflow. Moreover, LTP only explains the tourism demand in low EPI countries and has shown that higher price level attracts more tourists. Similar to previous findings, this relationship is possibly due to a more influential currency movement than price level movement and/or the effect of demand push inflation (Puah et al., 2019; Viljoen et al., 2019).

Table 4.38: Cross Tabulation of Origin Countries EPI with Destination Countries EPI

Origin Countries		Destination Countries					
		Low EPI		Moderate EPI		High EPI	
		No of Person	*Percent*	No of Person	*Percent*	No of Person	*Percent*
	Low EPI	332,720,315	*33.42*	78,613,477	*74.53*	75,519,694	*37.27*
	Moderate EPI	556,071,424	*55.85*	11,624,669	*11.02*	26,189,948	*12.93*
	High EPI	106,846,338	*10.73*	15,234,152	*14.44*	100,895,747	*49.80*
	Total	995,638,077	*100*	105,472,298	*100*	202,605,389	*100*

(Source: EF Education First Ltd, 2017)

A few crucial information can be summarised based on the tourism demand under different model specifications. Firstly, the linguistic variable has been represented with two measurements. The first measurement is a dummy variable used to proxy common language between tourist origin and destination countries. A common language usually exists between these countries due to their historical linkage. The second measurement is the EPI due to English's status as a global language.

The key shortfall of the first measurement is that the role of a linguistic variable cannot be determined when there is no common language. Hence, English as a *lingua franca* is expected to fill this shortfall. In the estimation for common language, Cambodia, Japan, Korea, Myanmar and Vietnam have been omitted because there is no common language between tourist origin and destination countries. For the remaining countries, the overall results have been satisfactory. In general, it has been shown that common language is an important determinant of tourism demand. A common language eases the movement of a tourist because it makes daily communication easier.

As for the EPI estimation, Brunei, Laos, Myanmar and Philippines have been omitted from the estimation because of insufficient data published by the Education First Limited. Unlike the existence of a common language, which can be determined via a historical linkage, English language proficiency can only be known when it is measured and published by interested parties (in this case, Education First Limited).

The outcome of the second augmented gravity model showed that EPI has performed moderately with seven out of 11 estimation results being significant. In order to get more insights, the destination countries have been categorised into countries with low, moderate and high English proficiency. The findings suggested that EPI can be a motivator and barrier for tourism demand, depending on the composition of the tourists' English proficiency level. When there is a gap between tourist origin and destination countries', EPI discourages tourist arrival. For tourists and destinations with the same English proficiency, EPI is undoubtedly a motivator.

Next, the tourist's income is an important factor that influences the tourist's decision to travel while tourism destinations' income is relatively less important. This implies that tourist's travel intention depends more on their income level rather than the tourism development at the destination countries. This is because increase in tourist's income indicates increase in the spending power and this is more influential on their travel intention. On the other hand, an increase in a country's income does not necessarily mean that the money has been channelled to its tourism to attract tourists. Apart from that, not all tourists are aware of the tourism development of these countries as well.

Distance performed well in this study, whereby this variable has been mostly significant under different model specifications. Tourists are found more unlikely to travel to destinations that are located far from their origin countries because it increases their travel expenses, travel time and is more uncomfortable. Free trade agreement performed decently and has the expected influence in most cases. Countries with trade preferential generally have the advantage of attracting more tourists from their trade partners because of a sense of familiarity.

Adjacent/neighbouring destination proves to be a factor under most tourists' consideration as well. Majority of the empirical outcomes indicated that this variable is statistically significant and imposes the expected influence. Tourists prefer to travel to neighbouring countries since it is more time and cost saving. Tourism price is also a significant variable that affects tourist's decision to visit a country. A mixed result (both positive and negative coefficient signs) has been obtained, but it is still justifiable and safe to say that tourism price has performed moderately under both model specification for augmented gravity models. Depending on the destinations, tourism price will have different influence on a tourist's traveling decision.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter summarises and concludes this study. This begins with a summary and is followed by some policy recommendations. The limitation of the current study is also discussed and this leads to recommendation for future studies. The organisation of this chapter is as follows - Section 5.2 presents the conclusion of the study, Section 5.3 discusses the policy recommendation, Section 5.4 provides limitation of the study and Section 5.5 provides the recommendations for future studies.

5.2 Conclusion of the Study

This study has been conducted to find out the determinants of tourism demand in ASEAN+3 countries, at national and regional level. The ASEAN+3 countries are Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam, China, Japan and Korea. The augmented gravity model has been selected to estimate tourism demand for these countries. Static linear panel models of Pooled Ordinary Least Square (POLS), Random Effects Model (REM) and Fixed Effects Model (FEM) have been adopted to estimate the empirical models. The final model was decided through empirical testing procedures and then rectified following a series of diagnostic checks. The time period under study was from 2012 to 2017. The specified determinants were income level of origin and destination countries, distance, adjacent country, free trade agreement, tourism price and language.

The key findings are summarised as follows. The role of language, as represented by common language is a key determinant that affects tourist's decision to travel. When tourist share a common language with the local at the tourist destination, they are more likely to visit. In this study, dummy variable used to proxy common language had performed well while English proficiency level had performed moderately. Despite that, English proficiency proves to be an alternative measurement for common language. In the future, the role of English as an international language or *lingua franca* will continue to grow and certainly continue to play an important function. The English proficiency level of origin and destination countries pairing have different implication and it is imperative to its influence on tourism demand.

For the income related determinants, tourist's income is found to be more influential than destination's income. Distance proved to be equally important for modelling tourism demand as this variable is an important determinant as indicated by most of the results. Apart from that, trade preferential, represented by free trade agreement, performed decently and the effect of this variable is mostly parallel with published studies. The adjacent/neighbouring destination is also a component that affect tourist's decision to travel and complements the findings of distance variable. Besides that, tourism price has different impact on tourism models, depending on tourist's choice of destination.

5.3 Policy Recommendations

This section provides some policy recommendations based on the study outcome to sustain and improve the tourism sector of ASEAN+3 countries. The linguistic variable used in this study consists of two measurements, which are common language and English Proficiency Index (EPI). Empirical results showed that common language is significant in

most cases. Therefore, for countries that share a common language with ASEAN+3 countries, the common language should be a promotional item that is emphasised, especially for origin and destination countries with different English proficiency levels. The use of English as a common language is inadequate at the moment because only tourists with high English proficiency are found to be attracted by destination with high EPI.

Nonetheless, tourism authorities in ASEAN+3 countries should conduct proper training for workers in the tourism sector to improve their English level as it becomes more and more common place. This is because, unlike most common languages that has existed due to the historical linkage between two countries, English is the global *lingua franca*. Through having sufficient English proficiency only can the tourism sector in ASEAN+3 countries be prepared to serve tourists possessing different level of English proficiency. This is particularly crucial for countries with no common language such as Cambodia, Japan, Korea, Myanmar and Vietnam.

Apart from that, multilingual countries are more likely to attract international tourists and hence, policies that encourage residents of a nation to study at least two foreign languages should be developed. This policy can start off with targeting workers in tourism sector before extending to the rest of the population. Through this, the tourism sector will gain immediate advantage since the sector is directly interacting with international tourists.

In addition, collaboration between ASEAN+3 tourism authorities and industry players in organising for exchange and training programmes will also benefit each other. For instance, ASEAN+3 countries with higher English proficiency level can provide training to ASEAN+3 countries with moderate and low English proficiency level. Apart from English, the ASEAN+3 countries should also learn other foreign languages such as Chinese and

Malay from member countries in order to achieve a multilingual region status and attracts more international tourists to the region as a whole.

Next, tourism authorities need to recognize that tourist's income level affects their visit to and spending in the destination. Strategies are required to differentiate tourism services and marketing if the destination countries intend to benefit from rising income of the tourists. Therefore, the promotion ought to be specific to target tourist groups in different countries or at least, different regions, with close reference to the origin country's economy. When the economy of the tourist origin countries is performing strong, more promotional efforts can be rolled out to attract more potential tourists.

Although the destination country income level is not as influential as the origin country income level as shown by empirical results, it is nonetheless a factor that affects tourist inflow. In fact, this should be taken in the manner of facilities provided within the destination country that are catered for tourists. A direct indication of improvement in destination country's income can be linked to higher standards of accommodation and better facilities. A portion of the national income should be allocated to maintain and improve the tourism sector according to the individual ASEAN+3 countries' need. While some countries may devote their efforts to upgrade the quality of accommodation and facilities, other countries may focus their attention on increasing the accommodation capacity. Regardless of which area that a country chooses to improve, it should be to improve the tourists' experience.

The empirical findings have revealed that tourists are less interested to travel to destinations that are located far away from their origin country, hence it is important that ASEAN+3 countries should cooperate together to attract tourists from another region or

continent. In order to overcome the challenge of long distance, a trip can be breakdown into two parts with a stopover in between. Adapting this idea to ASEAN+3 countries, a stopover in China can be made available for tourists travelling from United States to Malaysia, for example. This stopover concept and related routes can be part of the information provided at tourist centres or even tour agencies to promote ASEAN+3 countries as part of the stopover options. It is even possible to work together with different airlines to promote ASEAN+3 countries through magazine or brochure provided to tourists during their flight.

Furthermore, a partnership between government, private tourism industry planners and tourism destinations would help in developing a total tourism package with stopovers that is of great help and advantage to tourism destinations. It is suggested that partnership between the ASEAN+3 governments and tourism sector industry players to collaborate and design the related tourism package. In a more advanced form, this package can be a cross-country tour package with a few members of ASEAN+3 countries as destinations, if not all. This requires massive commitment, close cooperation and rapid interaction between ASEAN+3 countries. However, it is in line with the framework of ASEAN+3 countries.

The findings of distance and adjacent country or common border indicated that tourists are more likely to travel to their neighbouring countries which are nearer to save time and cost. The computation of dynamic distance variable in this study has taken into consideration the travelling cost represented by crude oil price to reflect the travel cost based on distance covered. Therefore, monitoring the movement of crude oil price can give a signal to ASEAN+3 governments on the timing of their promotion in their neighbouring countries. When the travelling cost increases, tourists from origin countries located farther away from

ASEAN+3 countries are unlikely to travel, hence promotional effort at this time should be more focused on adjacent countries.

Besides that, ASEAN+3 countries can consider establishing more free trade agreement (FTA) with other countries since FTA does attract tourist inflow in most ASEAN+3 countries. FTA reduces trade cost as well as increase trade flows and business trip between origin and destination countries. This will create foreign product preferences and generate interest for potential tourists to visit trading partner country. The choice of countries to establish FTA with is preferably those neighbouring countries that shares similar culture and economic environment. Partnerships and regional agreements over global partnerships are also recommended because the positive effects are stronger in such forms. For example, ASEAN+3 countries could consider establishing FTA with Hong Kong, Macau and Taiwan since these countries satisfy the conditions aforementioned.

Tourism price is found to have a mixed relationship with tourism demand in this study. Demand push inflation explains the positive relationship between tourism price and tourist arrivals, but such condition, if leave unattended would eventually drive away tourists because of rapid inflation that make spending too expensive. This will lead to the tourists looking for cheaper tourism destination elsewhere. This is in line with the consumer demand theory, which generally states that increased price will reduce demand. Few recommendations can be the possible solution to overcome this situation. In the case of the tourism price itself, it is recommended that governments and industry players to monitor the price of hotel, restaurant, tourist service providers and transportation companies closely so that reasonable price is charged at all time.

Beyond merely monitoring prices, more effective competitive strategies is also recommended as this will help to achieve a fair price. This will naturally drive the tourism service providers to continuously offer tourists with products differentiation and innovation that are value for money and meeting expectations. The tourism price in this study has considered the influence of exchange rate and in the situation where the reduction in tourist arrivals or tourism receipts is caused by destination currency appreciation, substantial investment to improve tourism service quality will be able to the overcome disadvantages of strengthened currency value. A more proactive approach is suggested through the use of supply side policies such as tax reduction as incentives for industry players to invest and expand their business in spite of rising price levels. These recommendations are suitable for the ASEAN+3 countries, as it can be customized according to their unique condition of their tourism sector needs.

5.4 Limitations of the Study

In this study, the dataset used to estimate tourism demand models for the countries of interest are short in time series observation. The time period only covered annual data from 2012 to 2017. This is mainly because the EPI data only has annual observations and in order to homogenise the dataset for model estimations, all the remaining variables have been compiled in the same frequency. The standardisation of the dataset has been done at the expense of seasonal pattern of tourism demand.

Besides that, although EPI is the measurement to represent the level of English proficiency, this variable is not flawless because the test data are collected from those test takers with access to internet; interested to learn the language; have taken the test on self-initiative; and are biased towards young adults. There is an upward bias because the sample

has excluded those who have no internet access, poorer and less educated people. Nevertheless, the dataset is still credible because it has been compiled periodically and the sample necessitates a minimum of 400 test takers for any city, region and country to be included in the index calculation. So, its usage at the moment still provides valuable information on English proficiency around the world. Another limitation of this is the lack of EPI for Brunei, Laos, Myanmar and Philippines, which makes the ASEAN+3 countries set incomplete.

This study has focused on estimating the tourism demand model for aggregate international tourist arrival instead of detailing the tourist according to their visiting purpose, expenditure pattern, expenditure components, and so on. This is not feasible because of the inconsistent data among the countries under study, such as for Cambodia, Laos, Myanmar and Vietnam. Tourism data in these countries are not as comprehensive as compared to the rest of the countries. At the same time, it has limited the option to conduct a dynamic assessment of tourism model. As a consequence, a better understanding on the dynamic adjustments of tourism demand model such as the word-of-mouth effect or habit persistence is not available to give a wholesome picture to the dynamic adjustment of tourism demand within the ASEAN+3 countries.

5.5 Recommendations for Future Studies

Despite this study has achieved its objectives in finding the determinants of tourism demand for ASEAN+3 countries, there are still many areas to be covered for future studies. Firstly, the role of English as a *lingua franca* can be further examined when there are longer time series data available. A time series analysis or a long panel analysis will better suit the purpose of examining the role of English as a common language among countries around the

globe over time. A study that includes Brunei, Laos, Myanmar and Philippines when their EPI data are available to complete the ASEAN+3 countries analysis as a group is also beneficial. A long panel analysis of ASEAN+3 countries tourism demand should be conducted when more comprehensive dataset for English proficiency is available.

Next, future research should also take interest in examining the determinants of tourism demand under disaggregated tourism segments such as tourist arrival according to purpose of visit, expenditure pattern, expenditure components, and more. Different segments of tourism are expected to have different set of determinants which will provide useful insights to policy makers, tourism authorities and tourism service providers. It would be interesting as well to investigate the reaction of different tourism demand segment towards EPI. In addition, a major improvement that can be implemented is the establishment of a dynamic model to capture dynamic factors such as word-of-effect and repeated visitors.

REFERENCES

Aguilo, E., Riera, A., & Rosello, J. (2005). The short-term price effect of a tourist through a dynamic demand model. *Tourism Management, 26,* 359-365. https://doi.org /10.1016/j.tourman.2003.07.005

Ahmad Kosnan, S.S., Ismail, W.N., & Kaliappan, S.R. (2013). Determinants of international tourism in Malaysia: Evidence from gravity model. *Jurnal Ekonomi Malaysia*, *47*(1), 131–138. http://journalarticle.ukm.my/8096/1/5579-13425-1-SM.pdf

Alawin, M., & Abu-Lila, Z. (2016). Uncertainty and gravity model for international tourism demand in Jordan: Evidence from gravity model. *Applied Econometrics and International Development, 16*(1), 131-146. https://www.usc.gal/econo met/ reviews/aeid16111.pdf

Arellano, M. (1993). On the testing of correlated effects with panel data. *Journal of Econometrics, 59*(12), 87-97.

ASEAN Secretariat. (2012). *ASEAN Tourism Marketing Strategy (ATMS) 2012-2015*. https://www.asean.org/storage/2012/05/ATSP-2016-2025.pdf

ASEAN Secretariat. (2015). *ASEAN Tourism Strategy Plan 2016-2025*. https://www.asean.org/storage/2012/05/ATSP-2016-2025.pdf

Athanasopoulos, G., & Hyndman, R.J. (2008). Modelling and forecasting Australian domestic tourism. *Tourism Management, 29*, 19-31. https://doi.org/ 10.1016/j.tourman.2007.04.009

Biswas, C., & Mamun-Or-Rashid, M. (2019). Investigating barriers in tourism development: The case of Bangladesh. *International Journal of Economics, Commerce and Management, 7*(2), 654-663. http://ijecm.co.uk/wp-content/uploads/2019/02/7243.pdf

Botti, L., Peypoch, N., Randriamboarison, R., & Solonandrasana, B. (2007). An econometric model of tourism demand in France. *Tourismos: An International Multidisciplinary Journal of Tourism, 2*(1), 115-126.

BPS-Statistics Indonesia, *Statistical Yearbook of Indonesia* (various issues) [Data set].

BPS-Statistics Indonesia. (2019). *Statistical Yearbook of Indonesia 2019*.

Breusch, T.S., & Pagan, A.R. (1980). The langrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies*, *47*, 239-253. https://doi.org/10.2307/2297111

Brida, J.G., Risso, W.A., & Carrera, E. (2008). A long-run equilibrium demand function: Tourism in Mexico. *Tourismos: An International Multidisciplinary Journal of Tourism, 3*(1), 66-82.

Chaiboonsri, C., Sriboonjit, J., Sriwichailamphan, T., Chaitip, P., & Sriboonchitta, S. (2010). A panel cointegration analysis: An application to international tourism demand of Thailand. *Annals of the University of Petrosani Economics*, *10*(3), 69-86. https://www.upet.ro/annals/economics/pdf/2010/20100308.pdf

Chasapopoulos, P., den Butter, F.A.G., & Mihaylov, E. (2014). Demand for tourism in Greece: A panel data analysis using the gravity model. *International Journal of Tourism Policy, 5*(3), 173-191.

Chockalingam, M., & Ganesh, A.A. (2010). Problems encountered by tourists. *Business and Economic Horizons, 3*(3), 68-72.

De Vita, G., & Kyaw, S.K. (2013). Role of the exchange rate in tourism demand. *Annals of Tourism Research, 43*, 624-627. https://doi.org/10.1016/j.annals.2013.07.011

Department of Statistics Brunei, *Brunei Darussalam Statistical Yearbook* (various issues) [Data set]. Ministry of Finance and Economy.

Department of Statistics Brunei. (2019). *Brunei Darussalam Statistical Yearbook 2019*. Ministry of Finance and Economy.

Department of Tourism, Philippines. (2019). *Tourism Demand Statistics*. http://www.tourism.gov.ph/tourism_dem_sup_pub.aspx

Divisekera, S. (2003). A model of demand for international tourism. *Annals of Tourism Research, 30*(1), 31-49. https://doi.org/10.1016/S0160-7383(02)00029-4

Dogru, T., Sirakaya-Turk, E., & Crouch, G.I. (2017). Remodeling international tourism demand: old theory and new evidence. *Tourism Management, 60,* 47-55. https://doi.org/10.1016/j.tourman.2016.11.010

Dritsakis, N. (2004). Cointegration analysis of German and British tourism demand for Greece. *Tourism Management, 25,* 111-119.

Dritsakis, N. (2012). Tourism development and economic growth in seven Mediterranean countries: A panel data approach. *Tourism Economics*, *18*(4), 801–816. https://doi.org/10.1016/S0261-5177(03)00061-X

Durbarry, R. (2008). Tourism taxes: Implications of tourism demand in the UK. *Review of Development Economics, 12*(1), 21-36. https://doi.org/10.1111/j.1467-9361.2008.00432.x

Eeckels, B., Filis, G., & León, C. (2012). Tourism income and economic growth in Greece: Empirical evidence from their cyclical components. *Tourism Economics, 18*(4), 817–834. https://doi.org/10.5367/te.2012.0148

EF Education First Ltd. (2017). *EF English Proficiency Index*. EF Education First Ltd.

Eita, J.H., Jordaan, A.C., & Jordaan, Y. (2011). An econometric analysis of the determinants impacting on businesses in the tourism industry. *African Journal of Business Management, 5*(3), 666-675.

Ekanayake, E.M., Halkides, M., & Ledgerwood, J.R. (2012). Inbound international tourism to the United States: A panel data analysis. *International Journal of Management and Marketing Research*, *5*(3), 15-27. https://ssrn.com/abstract=2162573

Erazo, M.A.C., Ramirez, S.I.M., Encalada, M.A.R., Holguin, J.V., & Zou, J.H. (2019). English language skills required by the hospitality and tourism sector in El Oro, Ecuador. *Theory and Practice in Language Studies*, *9*(2), 156-167. http://dx.doi.org/10.17507/tpls.0902.05

Falk, M. (2010). A dynamic panel data analysis of snow depth and winter tourism. *Tourism Management, 31,* 915-924. https://doi.org/10.1016/j.tourman.2009.11.010

Fereidouni, H.G., Al-Mulali, U., & Mohammed, M.A.H. (2017). Wealth effect from real estate and outbound travel demand: The Malaysian case. *Current Issues in Tourism, 20*(1), 68-79. https://doi.org/10.1080/13683500.2014.882886

Garin-Munoz, T. (2006). Inbound international tourism to Canary Islands: A dynamic panel data model. *Tourism Management, 27,* 281-291. https://doi.org/10.1016/j.tourman.2004.10.002

Garin-Munoz, T. (2007). German demand for tourism in Spain. *Tourism Management*, *28,* 12-22. https://doi.org/10.1016/j.tourman.2005.07.020

Garin-Munoz, T. (2009). Tourism in Galicia: Foreign and domestic demand. *Tourism Economics*, *15*(4), 753-769. https://doi.org/10.5367/000000009789955107

Garin-Munoz, T., & Montero-Martin, L.F. (2007). Tourism in the Balearic Islands: A dynamic model for international demand using panel data. *Tourism Management*, *28,* 1224-1235. https://doi.org/10.1016/j.tourman.2006.09.024

Habibi, F., & Abdul Rahim, K. (2009). A bound test approach to cointegration of tourism demand. *American Journal of Economics and Business Management Administration*, *1*(2), 165-172.

Habibi, F., Abdul Rahim, K., Ramachandran, S., & Chin, L. (2009). Dynamic model for international tourism demand for Malaysia: Panel data evidence. *International Research Journal of Finance and Economics*, *33*, 207-217.

Halicioglu, F. (2004). An ARDL model of international tourist flows to Turkey. *Global Business and Economics Review 2004 Anthology*, 614-624.

Hara, T. (2013). Perception of language barriers or prior visit experience affect likelihood of visiting a country? empirical analysis of residents in the United States of America about Japan as a tourist destination. *Waseda Global Forum, 10*, 391-407.

Hausman, J.A. (1978). Specification tests in econometrics. *Econometrica, 46*, 1251-1271. https://doi.org/10.2307/1913827

Japan National Tourism Organization. (2019). *Visitor Arrivals*. https://statistics.jnto.go.jp/en/graph/#graph--inbound--travelers--transition

Jerabek, T. (2019). The effects of selected macroeconomic variables on tourism demand for the South Moravian Region of the Czech Republic from Germany, Poland, Austria and Slovakia. *Comparative Economic Research, 22*(3), 25-43. http://doi.org/10.2478/cer-2019-0021

Kadir, N., Nayan, S., & Abdullah, M.S. (2013). A panel data analysis of international tourist arrivals from ASEAN countries to Malaysia. *Procedia Economics and Finance*, *7*, 80-85. https://doi.org/10.1016/S2212-5671(13)00221-9

Karaman, A.S. (2016). The pernicious impact of visa restrictions on inbound tourism: The case of Turkey. *Turkish Studies*, *17*(3), 502-524. https://doi.org/10.1080/14683849.2016.1170602

Katircioglu, S.T. (2009). Revisiting the tourism-led-growth hypothesis for Turkey using the bound tests and johansen approach for cointegration. *Tourism Management*, *30*, 17-20. https://doi.org/10.1016/j.tourman.2008.04.004

Keum, K. (2010). Tourism flows and trade theory: A panel data analysis with the gravity model. *The Annals of Regional Science*, *44*, 541-557. https://doi.org/10.1007/s00168-008-0275-2

Khadaroo, J., & Seetanah, B. (2008). The role of transport infrastructure in international tourism development: A gravity model approach. *Tourism Management*, *29*, 831-840. https://doi.org/10.1016/j.tourman.2007.09.005

Kimura, F., & Lee, H.H. (2006). The gravity equation in international trade in services. *Review of World Economics*, *142*, 92-121. https://doi.org/10.1007/s10290-006-0058-8

Korea Tourism Organization. (2019). *Statistics of Arrivals and Departures by Item*. https://kto.visitkorea.or.kr/eng/tourismStatics/keyFacts/KoreaMonthlyStatistics.kto

Kunroo, M.H., & Azad, N.A. (2015). Theory-Based specifications of the gravity equation: An analysis using European Union as an example. *Journal of International Economics*, *2*(6), 101-117.

Kuo, H.I., Chen, C.C., Tseng, W.C., Ju, L.F., & Huang, B.W. (2008). Assessing impacts of SARS and Avian flu on international tourism demand to Asia. *Tourism Management*, *29*, 917-928. https://doi.org/10.1016/j.tourman.2007.10.006

Kusni, A., Kadir, N., & Nayan, S. (2013). International tourism demand in Malaysia by tourists from OECD countries: A panel data econometric analysis. *Procedia Economics and Finance, 7*, 28-34. https://doi.org/10.1016/S2212-5671(13)00214-1

Leitao, N.C. (2010). Does trade help to explain tourism demand? The case of Portugal. *Theoretical and Applied Economics, 17*(3), 63-74.

Li, G., Song, H., & Witt, S.F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 43*, 141-150. https://doi.org/10.1177/0047287505276594

Lorde, T., Li, G., & Airey, D. (2015). Modelling Caribbean tourism demand: An augmented gravity approach. *Journal of Travel Research, 55*(7), 945-956. https://doi.org/10.1177/0047287515592852

Martin, C.A., & Witt, S.F. (1988). Substitute prices in models of tourism demand. *Annals of Tourism Research, 15*(2), 255-268. https://doi.org/10.1016/0160-7383(88)90086-2

Massidda, C., & Etzo, I. (2012). The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management, 33*, 603-610. https://doi.org/10.1016/j.tourman.2011.06.017

Ministry of Culture and Tourism of the People's Republic of China, *The Yearbook of China Tourism Statistics* (various issues) [Data set]. China Tourism Press.

Ministry of Culture and Tourism of the People's Republic of China. (2019). *The Yearbook of China Tourism Statistics 2019*. China Tourism Press.

Ministry of Hotels and Tourism, Myanmar, *Myanmar Tourism Statistics* (various issues) [Data set].

Ministry of Hotels and Tourism, Myanmar. (2019). *Myanmar Tourism Statistics*.

Ministry of Information, Culture and Tourism, Laos, *Statistical Report on Tourism in Laos* (various issues) [Data set].

Ministry of Information, Culture and Tourism, Laos. (2018). *Statistical Report on Tourism in Laos*.

Ministry of Sports and Tourism Thailand, *Summary of International Travel Statistics* (various issues) [Data set].

Ministry of Sports and Tourism Thailand. (2019). *Summary of International Travel Statistics*.

Ministry of Tourism Cambodia. (2019). *Tourism Statistics Annual Report*.

Mohamed Ali Ibrahim, M.A. (2011). The determinants of international tourism demand for Egypt: Panel data evidence. *European Journal of Economics, Finance and Administrative Sciences, 30,* 50-58. http://dx.doi.org/10.2139/ssrn.2359121

Mohamed Ariff, A.K. (2008). *Economic Regionalism a la ASEAN, East Asia & APEC*. Kuala Lumpur: Malaysian Institute of Economic Research.

Mohd Hanafiah, M.H., Mohd Harun, M.F., & Jamaluddin, M.R. (2010). Bilateral trade and tourism demand. *World Applied Sciences Journal, 10,* 110-114.

Mohd Salleh, N.M., Law, S.H., Ramachandran, S., Shuib, A., & Mohd Noor, Z. (2008). Asian tourism demand for Malaysia: A bound test approach. *Contemporary Management Research, 4*(4), 351-368. https://doi.org/10.7903/cmr.1178

Mohd Salleh, N.M., Othman, R., & Ramachandran, S. (2007). Malaysia's tourism demand from selected countries: The ARDL approach to cointegration. *International Journal of Economics and Management, 1*(3), 345-363. http://psasir.upm.edu.my/id/eprint/672

Morley, C., Rosello, J., & Santana-Gallego, M. (2014). Gravity models for tourism demand: Theory and use. *Annals of Tourism Research, 48,* 1-10. https://doi.org/10.1016/j.annals.2014.05.008

Nelson, L.A., Dickey, D.A., & Smith, J.M. (2011). Estimating time series and cross section tourism demand models: Mainland United States to Hawaii. *Tourism Management, 32,* 28-38. https://doi.org/10.1016/j.tourman.2009.10.005

Okafor, L.E., Khalid, U., & Then, T. (2018). Common unofficial language, development and international tourism. *Tourism Management, 67,* 127-138. https://doi.org/10.1016/j.tourman.2018.01.008

Pablo-Romero, M.D.P., & Molina, J.A. (2013). Tourism and economic growth: A review of empirical literature. *Tourism Management Perspectives, 8,* 28-41. https://doi.org/10.1016/j.tmp.2013.05.006

Perles-Ribes, J.F., Ramon-Rodriguez, A.B., Rubia, A., & Moreno-Izquierdo, L. (2017). Is the tourism-led-growth hypothesis valid after the global economic and financial crisis? The case of Spain 1957-2014. *Tourism Management, 61,* 96-109. https://doi.org/10.1016/j.tourman.2017.01.003

Petersen, M.A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies, 22*(1), 435-480. https://doi.org/10.1093/rfs/hhn053

Philippine Statistics Authority, *Philippine Statistical Yearbook* (various issues) [Data set].

Prachanant, N. (2012). Needs analysis on English language use in tourism industry. *Procedia - Social and Behavioral Sciences, 66,* 117-125. https://doi.org/10.1016/j.sbspro.2012.11.253

Puah, C.H., Thien, F.T., Arip, M.A., & Chin, M.Y. (2019). Modelling a tourism demand in Vietnam. *International Journal of Economics and Management*, *13*(2), 319-329. http://www.ijem.upm.edu.my/vol13no2/4)%20Modelling%20a%20Tourism%20Demand%20in%20Vietnam.pdf

RocketNews24 (2015, January 19). *Why is Japan Such An Unpopular Tourist Destination?*. Japantoday. https://japantoday.com/category/features/travel/why-is-japan-such-an-unpopular-tourist-destination

Rodriguez, X.A., Martinez-Roget, F., & Pawlowska, E. (2012). Academic tourism demand in Galicia, Spain. *Tourism Management*, *33*, 1583-1590. https://doi.org/10.1016/j.tourman.2012.01.010

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, *9*(1), 86-136. https://doi.org/10.1177/1536867X0900900106

Rossello, J., Santana-Gallego, M., & Awan, W. (2017). Infectious disease risk and international tourism demand. *Health Policy and Planning*, *32*, 538-548. https://doi.org/10.1093/heapol/czw177

Saayman, A., Figini, P., & Cassella, S. (2016). The influence of formal trade agreements and informal economic cooperation on international tourist flows. *Tourism Economics*, *22*(6), 1274-1300. https://doi.org/10.1177/1354816616672600

Samitas, A., Asteriou, D., Polyzos, S., & Kenourgios, D. (2018). Terrorist incidents and tourism demand: Evidence from Greece. *Tourism Management Perspectives*, *25*, 23-28. https://doi.org/10.1016/j.tmp.2017.10.005

Santana-Gallego, M., Fourie, J., & Rossello, J. (2020). The effect of safety and security issues on international tourism. *Annals of Tourism Research, 80,* 1-4. https://doi.org/10.1016/j.annals.2019.02.004

Santeramo, F.G., & Morelli, M. (2015). Modelling tourism flows through gravity models: A quantile regression approach. *Current Issues in Tourism*, *19*(11), 1077-1083. https://doi.org/10.1080/13683500.2015.1051518

Schaffer, M.E., & Stillman, S. (2010). *xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor*. IDEAS. https://ideas.repec.org/c/boc/bocode/s456779.html

Seabra, C., Reis, P., & Abrantes, J.L. (2020). The influence of terrorism in tourism arrivals: A Longitudinal approach in a Mediterranean country. *Annals of Tourism Research*, *80*. https://doi.org/10.1016/j.annals.2019.102811

Seetanah, B., Durbarry, R., & Ragodoo, J.F.N. (2010). Using the panel cointegration approach to analyse the determinants of tourism demand in South Africa. *Tourism Economics*, *16*(3), 715-729. https://doi.org/10.5367/000000010792278437

Sheldon, P., & Var. T. (1985). Tourism forecasting: A review of empirical research. *Journal of Forecasting, 4*, 183-195. https://doi.org/10.1002/for.3980040207

Shepherd, B. (2016). *The gravity model of international trade: A user guide (An updated version)*. United Nations.

Singapore Tourism Analytics Network. (2019). *Tourism Statistics*. https://stan.stb.gov.sg/portal/tourism-statistics.html

Song, H., & Li, G. (2008). Tourism demand modeling and forecasting – A review of recent research. *Tourism Management, 29,* 203-220. https://doi.org/10.1016 /j.tourman.2007.07.016

Song, H., & Witt, S.F. (2011). *Tourism Demand Modelling and Forecasting: Modern Econometric Approaches.* Routledge.

Song, H., Dwyer, L., Li. G., & Cao, Z. (2012). Tourism economics research: A review and assessment. *Annals of Tourism Research, 39*(3), 1653-1682. https://doi.org /10.1016/j.annals.2012.05.023

Suhaimi, N.I., & Abdullah, A.T.H. (2017). The role of multilingualism in enhancing tourism sector in Malaysia. *International Journal of Academic Research in Business and Social Sciences, 7*(7), 816-832. http://dx.doi.org/10.6007/IJARBSS/v7-i7/3154

Takahashi, K. (2020). Comparing the determinants of tourism demand in Singapore and French Polynesia. *Tourism Analysis, 25*(1), 175-181. https://doi.org/10.3727 /108354220X15758301241585

Tang, C.F., & Tan, E.C. (2016). The determinants of inbound tourism demand in Malaysia: Another visit with non-stationary panel data approach. *Anatolia, 27*(2), 189-200. https://doi.org/10.1080/13032917.2015.1084345

Tang, T.C., & Wong, K.N. (2009). The SARS epidemic and international visitors to Cambodia: Is the impact permanent or transitory? *Tourism Economics, 15*(4), 883-890. https://doi.org/10.5367/000000009789955206

Tatoglu, F.Y., & Gul, H. (2020). Analysis of tourism demand using a multi-dimensional panel gravity model. *Tourism Review, 75*(2), 433-447. https://doi.org/10.1108/TR-05-2019-0147

Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for An International Economic Policy*. The Twentieth Century Fund.

Tkales, M., & Vizek, M. (2016). The price tag of tourism: Does activity increase the prices of goods and services? *Tourism Economics, 22*(1), 93-109. https://doi.org/10.5367/te.2014.0415

Tourism Malaysia. (2019). *MY Tourism Data.* http://mytourismdata.tourism.gov.my/

Ulucak, R., Yucel, A.G., & Ilkay, S.C. (2020). Dynamics of tourism demand in Turkey: panel data analysis using gravity model. *Tourism Economics,26*(8), 1394-1414. https://doi.org/10.1177/1354816620901956

Untong, A., Ramos, V., Kaosa-Ard, M., & Rey-Maquieira, J. (2015). Tourism Demand analysis of Chinese arrivals in Thailand. *Tourism Economics, 21*(6), 1221-1234. https://doi.org/10.5367/te.2015.0520

UNWTO. (2013). *UNWTO Tourism Highlights 2013 Edition*. https://www.e-unwto.org/doi/pdf/10.18111/9789284415427

UNWTO. (2015). *UNWTO Tourism Highlights 2015 Edition*. https://www.e-unwto.org/doi/pdf/10.18111/9789284416899

UNWTO. (2018). *UNWTO Annual Report 2017*. https://tourlib.net/wto/WTO_annual_report_2017.pdf

UNWTO. (2019). *UNWTO International Tourism Highlights 2019 Edition*. https://www.e-unwto.org/doi/pdf/10.18111/9789284421152

Vietnam National Tourism Administration. (2019). *International Visitors*. http://vietnamtourism.gov.vn/english/index.php/cat/1501

Viljoen, A., Saayman, A., & Saayman, M. (2019). Determinants influencing inbound arrivals to Africa. *Tourism Economics, 25*(6), 856-883. https://doi.org/10.1177/1354816618809840

Wamboye, E.F., Nyaronga, P.J., & Sergi, B.S. (2020). What are the determinants of international tourism in Tanzania? *World Development Perspectives, 17*. https://doi.org/10.1016/j.wdp.2020.100175

Wiriyachittra, A. (2002). English language teaching and learning in Thailand in this decade. *Thai TESOL Focus, 15*(1), 4-9.

Wong, K.N., & Tang, T.C. (2010). Tourism and openness to trade in Singapore: Evidence using aggregate and country level data. *Tourism Economics*, *16*(4), 965-980. https://doi.org/10.5367/te.2010.0017

Woolridge, J.M. (2002). *Econometric analysis of cross section and panel data.* MIT Press.

World Economic Forum. (2017). *The Travel and Tourism Competitiveness Report 2017*. http://www3.weforum.org/docs/WEF_TTCR_2017_web_0401.pdf

World Economic Forum. (2019). *The Travel and Tourism Competitiveness Report 2019. http://www3.weforum.org/docs/WEF_TTCR_2019.pdf*

WTTC. (2019a). *Travel and Tourism Economic Impact 2019: World*.

WTTC. (2019b). *Data Gateway*. https://tool.wttc.org/

Yang, C.H., Lin, H.L., & Han, C.C. (2010). Analysis of international tourist arrivals in China: The role of world heritage sites. *Tourism Management, 31*, 827-837. https://doi.org/10.1016/j.tourman.2009.08.008

Yazdi, S.K., & Khanalizadeh, B. (2017). Tourism demand: A panel data approach. *Current Issues in Tourism*, *20*(8), 787-800. https://doi.org/10.1080/13683500.2016.1170772

Zahedpisheh, N., Abu Bakar, Z., & Saffari, N. (2017). English for tourism and hospitality purposes (ETP). *English Language Teaching*, *10*(9), 86-94. http://doi.org/10.5539/elt.v10n9p86

APPENDICES

Table A: Summary of Literature Review

Author(s) (Year)	Country (Data Set)	Methodology	Variables	Finding(s)
Martin and Witt (1988)	France, Germany, United Kingdom and United States of America (Annual data 1965-1980).	• OLS. • Cochrane-Orcutt interative procedure.	• Tourist arrival. • Origin country population. • Income. • Exchange rate. • Tourism price (CPI adjusted by exchange rate). • Weighted substitute price. • Travel cost and substitute destination travel cost (airfare, gasoline cost based on distance, cost of ferry crossing). • Dummy variables represent 1974 and 1979 oil crises, and 1967-69 United Kingdom currency restrictions (for UK).	• The empirical findings suggest that both substitute destination price and substitute destinations travel cost is important in determining tourism demand but the significance varies for each country under study.

Table A continued

Divisekera (2003)	United Kingdom, Australia, New Zealand and United States of America.	• Almost Ideal Demand System. • Maximum Likelihood Method. • Log Likelihood Ratio Test.	• Tourism receipts. • Tourism price. • Tourism expenditure. • Dummies to capture country specific special event.	• The changes in expenditure and price in USA does not significantly affect UK tourism demand. • Australia, New Zealand and USA are substitute destinations for each other for UK tourist. • New Zealand tourism demand seem like expenditure inelastic but price elastic for Australia, USA and UK. • These three destinations are substitute destinations for New Zealand tourist. • USA tourist view Australia and New Zealand as up-market destinations while UK is viewed as low-market destination. • USA tourist has strong preferences for UK tourism over Australia and New Zealand. • Japanese tourist view USA and UK as substitute destinations for Australia while New Zealand is viewed as complementary destination with Australia.

Table A continued

Dritsakis (2004)	Greece (Annual data 1960-2000).	• ADF unit root test. • Johansen and Juselius cointegration test. • Error correction model.	• Tourist arrival. • Income per capita. • Tourism price (CPI ratio). • Travel cost (Oil price). • Exchange rate.	• There is a long-run relationship between tourism demand and the explanatory variables. • Income is found to have a negative relationship with tourist arrival. • German and English tourists may prefer other tourism destination when they have higher income.
Halicouglu (2004)	Turkey (Annual data 1960-2002).	• ARDL model. • CUSUM and CUSUM of square stability tests.	• Tourist arrival. • World income. • Tourism price (CPI ratio adjusted by exchange rate). • Travel cost (Oil price).	• A stable long-run relationship exists among the variables of study. • Income is the most significant variable that affects the tourism demand.
Aguilo et al. (2005)	Balearic Islands of Spain (Annual data 1960-2000).	• Difussion model. • Seemingly Unrelated Regression.	• Tourist arrival. • Income per capita. • Tourism price (CPI ratio adjusted by exchange rate). • Exchange rate. • Dummy variables representing various one-off events in origin countries.	• Exchange rate is highly elastic for France. • Tourism price only slightly affects the travel decision of German, English and Dutch tourists. • Tax of tourism will lead to decrease in tourist arrival.

Table A continued

Garin-Munoz (2006)	Canary Islands of Spain (Annual data 1992-2002).	• Difference GMM.	• Tourist arrival per capita. • Lagged dependent variable. • Tourism price (TPI and CPI ratio adjusted by exchange rate). • Travel cost (Oil price). • Income level. • Dummy variable to capture the effects of terrorist attack on September 11.	• There is a high degree of consumer loyalty or important word-of-mouth effect as indicated by significant lagged dependent variable. • Tourism activity is considered as a luxury good. • Tourism demand in Canary Islands is affected by tourism price and travel cost.
Kimura and Lee (2006)	OECD member countries (Annual data 1999 and 2000).	• Gravity model. • Fixed effect model. • Ordinary least squares.	• Bilateral service and good exports and imports. • GDP. • Distance and relative distance. • Dummy variables representing adjacency, regional trade agreement and common language, economic freedom index.	• Gravity model predicts services trade better than goods trade. • A complementary relationship exists between good exports and service imports. • Home market effect is not evident for services trade. • Economic freedom impacts more on service exports than service imports.
Botti et al. (2007)	France (Annual data 1975-2003).	• ADF unit root test. • OLS regression.	• Tourist expenditure. • Income level. • Tourism price (CPI ratio adjusted by exchange rate).	• Income level is found to have a positive effect towards tourist expenditure. • A negative relationship exists between tourist expenditure and tourism price.

Table A continued

Garin-Munoz (2007)	Spain (Annual data 1991-2003).	• Difference GMM.	• Length of stay. • Lagged dependent variable. • Income per capita. • Tourism price (CPI ratio adjusted by exchange rate). • Travel cost (Oil price). • Dummy variables representing terrorist attack of September 11 and ecotax.	• Word-of-mouth effect is important in bringing more tourists into Spain. • German tourist perceived tourism activity in Spain as a luxury good. • Price component variables negatively affect tourism demand in Spain. • Dummy variable for 2002 is not significant, indicating that terrorist attack in 2001 only impacts tourism demand in short run.
Garin-Munoz and Montero-Martin (2007)	Balearic Islands (Annual data 1991-2003).	• Difference GMM.	• Tourist arrival per capita by air. • Tourism price (CPI ratio adjusted by exchange rate). • Travel cost (real crude oil price). • Income (GDP per capita). • Lagged dependant variable. • Dummy variables representing effect of September 11 event and impact of ecotax.	• Habit persistence, income, cost of living, travel cost and event of September 11 are significant determinants and have the expected relationship. • Ecotax is found to have no impact on tourism demand.

Table A continued

Mohd Salleh et al. (2007)	Malaysia (Annual data 1970-2004).	• ADF and PP unit root tests. • ARDL model.	• Tourist arrival. • Lagged dependent variable. • Tourism price (CPI ratio). • Income. • Substitute price. • Travel cost (Oil price). • Exchange rate. • Dummy variables representing Asian financial crisis and SARS outbreak.	• There is a long-run relationship between tourist arrival and the regressors. • Tourism price affects the travel decision of Hong Kong and Singapore tourist. • Travel cost is significant but inelastic for tourist from Australia, Hong Kong and Singapore. • There is a mixture of substitute and complementary destinations for tourist from different origin countries. • Income is an important factor for all countries of origin except Australia. • Exchange rate is only significant for Hong Kong tourist.
Athanasopoulus and Hyndman (2008)	Australia (Quarterly data 1998Q1-2005Q2).	• ADF and Modified Philips-Perron unit root tests. • KPSS stationary test. • OLS regression. • Innovation state space model.	• Length of stay disaggregated based on purpose of visit. • Consumer confidence. • Price movement of domestic travel. • Income per capita. • Dummy variables representing Bali bombing and Sydney Olympic games. • Seasonal dummy variables.	• A mixture of results was found between the explanatory variables and the disaggregated dependent variable. • Business travel had increased after the Olympic Games in 2000. • Travel for visiting friends and relatives increased after the 2002 Bali bombing.

Table A continued

Brida et al. (2008)	Mexico (Annual data 1980-2006).	• ADF unit root test. • KPSS stationary test. • Johansen and Juselius cointegration test. • Error correction model. • Pairwise Granger causality test. • Impulse response function.	• Tourist expenditure. • Tourism price (Tourism price index adjusted by average nominal change). • Income per capita. • Public investment.	• Long-run relationship exists between tourism demand and the explanatory variables. • Tourism demand is positively affected by income and public investment. • Tourism demand granger cause tourism price.
Durbarry (2008)	United Kingdom (Annual data 1968-1998).	• Random effects model. • Fixed effects model.	• Tourist arrival. • Origin and destination countries' real total expenditure and real expenditure per capita. • Real GDP per capita. • Distance. • Tourism price (relative price adjusted by exchange rate). • Weighted substitute price. • Dummies representing Europe country status and common language.	• Generally, tourism demand in United Kingdom is sensitive towards price changes. • Tourists from neighbouring countries have increased tourist arrival relative to tourist from distant countries.

Table A continued

Khadaroo and Seetanah (2008)	28 countries (Annual data 1990-2000)	• Gravity model. • IPS panel unit root test. • GMM.	• Tourist arrival. • Lagged dependant variable. • Tourist income (average real income per capita). • Relative price (CPI of destination country adjusted by exchange rate). • Geographical distance. • Tourism infrastructure (number of hotel rooms). • Population. • Dummy variables representing common language, common border, and proximity. • Road infrastructure (length of paved road divided by the size of the country). • Airport infrastructure (total number of terminals in international airports), and number of ports.	• Transport capital (represented by road infrastructure, air infrastructure and number of ports) is significant to attract more tourist inflows. • All other explanatory variables are important determinants and parallel with the past studies.

Table A continued

Kuo et al. (2008)	China, Hong Kong, Singapore, Taiwan (Monthly data 2001M1-2004M12), Indonesia and Vietnam (Monthly data 2002M10-2006M9).	• ARMAX model. • Difference GMM.	• Tourist arrival. • Number of probable SARS-infected patients. • Number of confirmed cases of Avian flu. • Lagged dependent variable. • Dummy variable to capture the effect of terrorist attack in Indonesia in October 2005.	• SARS outbreak resulted temporary adverse shock. • Taiwan and China are less affected by SARS outbreak as compared with Hong Kong and Singapore. • Terrorist attack in Indonesia has a negative impact towards tourist arrival. • Avian flu does not significantly influence tourist arrival.
Mohd Salleh et al. (2008)	Malaysia (Annual data 1970-2004).	• ARDL model.	• Tourist arrival from origin countries. • Tourism price (CPI ratio). • Substitute price. • Travel cost (Oil price). • Income per capita. • Exchange rate. • Lagged dependent variable. • Dummy variables representing Asian financial crisis and SARS outbreak.	• Tourism demand in Malaysia is determined by tourism price, substitute price, travel cost, income and exchange rate both in the short-run and long-run. • Lagged dependent variable, Asian financial crisis and SARS outbreak only affect tourist arrival in the short-run.

Table A continued

Garin-Munoz (2009)	Galicia of Spain (Annual data 1999-2006).	• Difference GMM. • Random effects model. • Feasible generalized least squares.	• Length of stay per capita. • Accommodation price. • Income per capita. • Dummy variable representing Holy Year. • Travel cost (Distance multiplied by average real crude oil price). • Public expenditure on tourism promotion.	• Domestic tourism is sensitive to accommodation price while international tourism is sensitive to travel cost. • Habit persistence significantly and positively affects domestic tourism demand. • Dummy variable representing Holy Year attracts both domestic and international tourism. • Travel cost is insignificant as domestic tourist travel by car. • Promotional effort is not significant indicates the possibilities of ineffectiveness in marketing campaign.
Habibi and Abdul Rahim (2009)	Malaysia (Quarterly data 1998Q1-2007Q3).	• ADF unit root test. • ARDL model.	• Tourist arrival. • Lagged dependent variable. • Income per capita. • Tourism price (CPI ratio adjusted by exchange rate). • Travel cost (Oil price). • Weighted substitute price. • Trade value.	• Income is significant in all countries except Brunei, Australia and United Kingdom. • Except for Singapore and India, tourism price is significant for the remaining countries. • Travel cost significantly explained tourist arrival in all countries except Indonesia, China, Japan and United Kingdom. • Trade value positively affects tourist arrival from all countries except Indonesia, Japan and Philippines.

Table A continued

				• Lagged dependent variable and SARS outbreak significantly explained tourism demand in the short-run. • China, Indonesia, Singapore, Thailand and Hong Kong are complementary destination for Malaysia tourism except for Singapore, Thailand and United Kingdom.
Habibi et al. (2009)	Malaysia (Annual data 1995-2005).	• Difference GMM.	• Tourist arrival. • Lagged dependent variable. • Income per capita. • Tourism price (CPI ratio adjusted by exchange rate). • Trade openness. • Dummy variables representing Asian financial crisis and SARS outbreak.	• Lagged dependent variable is important determinant for tourism demand in Malaysia. • Tourist arrival to Malaysia is sensitive to changes in price. • Trade openness and income is not significant in explaining tourism demand in Malaysia. • The Asian financial crisis and SARS outbreak have negatively affected tourism demand.
Katircioglu (2009)	Turkey (Annual data 1960-2006).	• ADF and PP unit root tests. • ARDL bound test. • Johansen and Juselius cointegration test.	• Real GDP. • International tourist arrival. • Real exchange rate.	• Tourism-led-growth hypothesis is rejected. • No cointegration is found among real GDP, international tourist arrival and real exchange rate.

Table A continued

Tang and Wong (2009)	Cambodia (Monthly data 2003M1-2007M12).	• ADF unit root test. • Unit root with structural break tests. • Panel unit root tests.	• Tourist arrival. • Dummy variable representing SARS outbreak.	• SARS outbreak only had transitory effect on Cambodia tourism sector.
Chaiboonsri et al. (2010)	Thailand (Annual data 1986-2007).	• Panel unit root tests of Levin-Lin-Chu, Breitung, Im-Pesaran-Shin, Maddala and Wu, Choi, and Hadri. • Pedroni residual cointegration test. • Kao residual cointegration test. • Johansen Fisher panel cointegration test. • OLS, DOLS and FMOLS estimators.	• Tourist arrival. • Income. • Transportation cost (Airfare). • Exchange rate.	• Exchange rate and income level of Thailand's Asia major tourist generating countries positively affect tourist arrival in Thailand. • Transportation cost has an adverse effect toward tourism demand in Thailand.
Chockalingam and Ganesh (2010)	India	• Questionnaire. • Cluster analysis. • Chi-square test. • Multiple regression.	• Demographic characteristics. • Tour problems. • Tour satisfactions.	• Tour problem intensity experienced by tourist differs according to numerous factors that influenced tour environment as well as depending on tourist own condition.

Table A continued

Falk (2010)	Austria (Monthly data for winter from 1986/87 to 2005/2006).	• Panel unit root tests of Levin-Lin-Chu and Im-Pesaran-Shin. • Kao residual-based panel cointegration test. • Mean group estimator. • Pooled mean group estimator.	• Length of stay. • Capital utilization. • Snow accumulation. • Income per capita. • Accommodation price. • Dummy variable representing early Easter holidays.	• There is a long-run relationship between the length of stay and the explanatory variables. • The length of stay increases with the increase of snow depth although the magnitude of impact is low. • Early Easter holidays is positively related to winter tourism demand.
Keum (2010)	Korea (Annual data 1990 to 2002)	• Gravity model. • Random effects model.	• Tourism flows. • Trade flows. • Origin and destination countries' national income. • Distance.	• Distance, origin and destination countries' income have expected influence on tourism flows and trade flows. • Linder's theory or income similarity effect is not supported both in trade flows and tourism flows.
Leitao (2010)	Portugal (Annual data 1995-2006).	• Fixed effects model. • Tobit model. • System GMM.	• Tourist arrival. • Income per capita. • Bilateral trade. • Population. • Tourism price (CPI ratio adjusted by exchange rate). • Geographical distance.	• Fixed effects model disclosed that all explanatory variables have significant relationship with tourist arrival except tourism price. • However, Tobit model revealed that all explanatory variables are significant and affect tourist arrival. • System GMM results found out that lagged dependent variable significantly affects tourism demand.

Table A continued

Mohd Hanafiah et al. (2010)	Malaysia (Annual data 1997-2008).	• Gravity model.	• Tourist arrival. • Income per capita (Industrial output value). • Tourism price (CPI ratio). • Bilateral trade. • Population. • Geographical distance.	• Tourism demand in Malaysia is significantly determined by all the independent variables. • Trade activities can stimulate tourism demand in Malaysia.
Seetanah et al. (2010)	South Africa (Annual data 1985-2000)	• IPS panel unit root test. • Pedroni panel cointegration tests. • FMOLS.	• Tourist arrivals. • Tourism price (CPI adjusted by exchange rate). • Substitute price (CPI adjusted by weighted exchange rate). • Origin and destination countries' income (real GDP per capita). • Geographical distance. • Tourism infrastructure (hotel room) • Dummy variables representing common border, common language, and political stability.	• Most of the variables under study are significant and have the expected relationship with tourist arrivals, even the sample are segregated by region.

Table A continued

Wong and Tang (2010)	Singapore (Quarterly data 1986Q1-2008Q2 for openness by trade component data and 1995Q1-2008Q2 for disaggregated country level data).	• Modified Wald test.	• Tourist arrival. • Total trade openness. • Openness to merchandise trade. • Openness to services trade.	• Bidirectional causality exists between total tourist arrival and openness to merchandise trade. • Unidirectional causality runs from openness to services trade towards openness to merchandise trade. • A mixture of causality patterns exist between tourism demand and explanatory variables for different trading partners.
Yang et al. (2010)	China (Annual data 2000-2005).	• Gravity model. • POLS model. • Random effects model. • Fixed effects model.	• Tourist arrival. • Relative income (relative GDP per capita). • Origin country population. • Exchange rate. • Geographical distance. • Number of international hotels. • Infrastructure (sum of operated railways and roads in terms of kilometres). • Number of criminal cases defended per year, sanitary condition (number of hospital beds).	• Most determinants have the expected relationship with tourist arrival. • An additional World Heritage Site inscribed will attract more tourist than a new 4A- and 3A-class rated tourism spots. • Historical and cultural tourism spots possess more attractiveness than natural landscape tourism spots especial for heritage sites.

Table A continued

			• Dummy variables representing SARS outbreak, foreign direct investment, number of World Heritage Sites, 4A- and 3A-class rated tourism spots (classified into historical and cultural, natural landscape and modern facilities spots).	
Eita et al. (2011)	South Africa (Annual data 1999-2007).	• Panel unit root tests of Levin-Lin-Chu and Im-Pesaran-Shin. • POLS model. • Random effects model. • Fixed effects model.	• Tourist arrival. • Income. • Tourism price. • Exchange rate. • Geographical distance. • Infrastructure. • Dummy variables representing countries which are member of Southern Africa Development Community (SADC) and European Union (EU).	• All explanatory variables significantly affect tourism demand in South Africa. • Income, exchange rate and infrastructure positively affect tourism demand in South Africa. • Tourism price and geographical distance have a negative relationship with tourism demand in South Africa. • SADC and EU members are generally associated with increase in tourist arrivals in South Africa.

Table A continued

Mohamed Ali Ibrahim (2011)	Egypt (Annual data 1990-2008).	• Seemingly Unrelated Regression.	• Tourist arrival. • Income per capita. • Tourism price (CPI ratio adjusted by exchange rate). • Real effective exchange rate. • Trade openness. • Substitute price. • Population.	• All of explanatory variables are significant and have the correct expected sign with the exception of population which have the negative sign which is out of expectation. • Tourism in Egypt is sensitive to price changes. • Tunisia is a substitute and competitive tourism destination.
Nelson et al. (2011)	Hawaii (Annual data 1993-2007).	• Time series mixed method. • Cross section analysis.	• Tourist arrival. • Income (gross state product). • Airfare. • Linear-plateau vectors to capture two recessions and September 11 terrorist attack. • Distance to competitive destinations (Orlando and Florida). • Crude oil price. • Consumer price index. • Cold index.	• The findings revealed that airfare has been a major factor geographically, but not temporarily in travel to Hawaii. • Income elasticities revealed that higher income groups should be targeted for potential tourist. • Marketing of winter vacation to Hawaii in the Northwest and North Central regions will be beneficial in long run.

Table A continued

Dritsakis (2012)	Spain, France, Italy, Greece, Turkey, Cyprus and Tunisia (Annual data 1980-2007).	• Breitung; LL; IPS; Maddala and Wu; and Hadri panel unit root test • Pedroni; Kao; and Johansen Fisher panel cointegration test. • Fully Modified Ordinary Least Squares.	• Tourism real receipts per capita. • International tourist arrival per capita. • Real effective exchange rate. • Real GDP per capita.	• Tourism development, exchange rate and economic growth are cointegrated in the long run. • Both tourism receipts and real exchange rate affects GDP in the countries of study.
Eeckels et al. (2012)	Greece (Annual data 1976-2004).	• Spectral analysis. • Vector Autoregressive Regression. • Impulse Response Function.	• GDP. • Tourism income.	• The findings support the tourism-led-growth hypothesis in the short run. • Long term memory effect exists in Greece's GDP cycle. • Tourism income is response more to its own shock than GDP to its own shock.
Ekanayake et al. (2012)	United States (Annual data 1986-2011).	• Im-Pesaran-Shin panel unit root test. • Pedroni panel cointegration tests. • Panel OLS estimator.	• Tourist arrival. • Income per capita. • Tourism price (Tourism price index and CPI). • Exchange rate. • Travel cost (Distance multiplied by oil price). • Dummy variable representing visa requirements.	• Income per capita, tourism price, exchange rate, travel is found to have a long-run relationship with tourism demand. • The dummy variable representing visa requirements does not affect tourism demand. • Tourism activity in United States is considered as a luxury good by international tourist.

Table A continued

Massidda and Etzo (2012)	Italy (Annual data 2004-2007).	• System GMM.	• Bilateral domestic tourism flows. • Lagged dependent variable. • Population density. • Income per capita. • Outbound tourism flows. • Culture. • Degree of regional tourism vocation. • Transport infrastructure. • Public safety. • Distance. • Tourism price (CPI ratio). • Relative pollution.	• Overall, the Italian tourists are sensitive to difference in tourism price between their region and the possible tourism destinations. • The past choices and environment quality affect the domestic tourism demand in Italy. • Italian tourists perceived that domestic and international tourism act as substitute goods. • Culture activities have positive impact towards tourist arrival. • At a disaggregated level, Southern Italian tourist is more sensitive with changes in income while the Northern Italian tourist is more sensitive toward tourism price differential. • Southern Italian tourist is more concerned with environmental issues while Northern Italian is more influenced by culture activities.

Table A continued

Prachanant (2012)	Thailand	• Questionnaire. • Descriptive analysis. • Needs analysis.	• Needs of English use. • Functions of English use. • Problems of English use.	• Findings revealed that English speaking skill is most important, followed by listening, reading and writing. • The three most relevant functions of English are giving information; providing services; and offering help.
Rodriguez et al. (2012)	Galicia, Spain (Annual data 2001-2009).	• GMM	• Number of foreign student enrolment. • Tourism price (CPI ratio). • Tourist income level (GDP per capita). • Travel cost (distance multiplied by average price per barrel of oil). • Differential attractiveness between three Galician university. • Dummy variables representing Bologna system implementation period and common language, Erasmus programme effect, higher education effect in origin country. • Lagged dependant variable.	• Lagged dependant variable, Erasmus programme effect and differential attractiveness is highly significant. • Geographical proximity is important and trend shows that students with lower income are well received by the universities. • Bologna system full effect is not detected under the period under study and higher education effect in origin is not significant indicating that student despite enrolled in higher education does not necessarily go to study abroad. • Language and cost of living are not significant determinants.

Table A continued

Ahmad Kosnan et al. (2013)	Malaysia (Annual data 1998-2009).	• Gravity model. • Random effects model. • Fixed effects model.	• Tourism receipts. • Income of origin and destination countries. • Tourism price (CPI ratio). • Geographical distance. • Dummy variables representing common border and common language. • Number of hotel rooms, tourism infrastructure (road and air transport infrastructure).	• Larger market size (income) of both origin and destination countries, common border and language, tourism price, number of hotel rooms, road and air transport infrastructure have positive and significant relationship with tourism receipts while distance has a negative and significant relationship with tourism receipts.
De Vita and Kyaw (2013)	Turkey (Quarterly 1996Q1-2009Q4).	• GARCH.	• Tourist arrival. • Bilateral exchange rate. • Exchange rate volatility. • Income (real GDP per capita). • Tourism price (CPI ratio). • Alternative tourism price (CPI ratio adjusted by exchange rate).	• Exchange rate is important variable explaining tourism demand and should be included as in the model after adjusted with price level. • Exchange rate volatility should be included to explain uncertainty avoidance in travel destination decision rather than as a proxy for living costs.

Table A continued

Hara (2013)	United States of America	• Online questionnaire. • Stepwise regressions. • Contour plot.	• Demographic characteristics. • Travel behaviour. • Travel perception.	• Perception of language barrier is a serious issue to tourist who never travelled to the destination or studied the local language. • Destination marketing campaign using non-English phrase can has its advantages and disadvantages.
Kadir et al. (2013)	Malaysia (Annual data 1994-2009).	• POLS model. • Random effects model. • Fixed effects model.	• Tourist arrival. • Lagged dependent variable. • Income. • Tourism price. • Lagged tourism price. • Substitute price. • Dummy variables representing Asian financial crisis, SARS outbreak and Visit Malaysia Year 2007 campaign.	• Income, Asian financial crisis and SARS outbreak are important factors to explain tourist arrival into Malaysia. • Tourism demand in Malaysia is price inelastic and affected by current price. • Indonesia and Philippines are substitute tourism destinations while Thailand is complementary tourism destination for Malaysia.

Table A continued

Kusni et al. (2013)	Malaysia (Annual data 1995-2009).	• POLS model. • Random effects model. • Fixed effects model.	• Tourist arrival. • Lagged dependent variable. • Tourism price. • Income. • Substitute price. • Dummy variables for SARS outbreak and global financial crisis.	• Tourism price is an important factor affecting tourism demand in Malaysia. • Singapore is the substitute tourism destination for Malaysia. • SARS outbreak and global financial crisis negatively affected tourist arrival into Malaysia.
Chasapopoulos et al. (2014)	Greece (Annual data 2001-2010).	• System GMM.	• Tourist arrival. • Lagged dependant variable. • Income (GDP per capita). • Tourism price (CPI ratio adjusted by exchange rate/comparative price level). • Competitive price (destination CPI over competitive destination CPI). • Bilateral trade. • Geographical distance. • Political stability. • Transport infrastructure (gross investment spending in infrastructure). • Dummy variable representing Olympic 2004.	• Income, tourism price and transportation infrastructure are not important determinants for tourism demand in Greece while distance, bilateral trade, political stability and Olympic 2004 are significant factors affecting tourism demand. • Competitive price showed mixed results.

Table A continued

Kunroo and Azad (2015)	29 European countries (Annual data 1994-2001).	• Gravity model. • Least square dummy variable technique.	• Trade volume. • Reporting and partner countries GDP • Exporter and importer countries per capita GDP. • Product of arable land per capita in reporting country and partner country. • Dummy variables representing currency union membership, Euro currency use, number of years when reporting country (partner country) is in currency union when partner country (reporting country) is not, minimum of years in currency union together for both trading countries, common border, common official language, language when it is spoken by 9 percent of population in both countries, and landlockness.	• The authors concluded that bilateral effects and time effects are an important part of gravity model. • The two-way effects specification is preferable in the situation where variables representing country specific characteristics are explicitly included in the model.

Table A continued

			• Per capita GDP difference (proxy for economic distance). • Weighted distance. • Bilateral nominal exchange rate volatility.	
Lorde et al. (2015)	Caribbean countries (Annual data 1980-2008).	• Gravity model. • LLC panel unit root test. • GMM.	• Tourist arrivals. • Lagged dependant variable. • Origin and destination countries per capita income. • Origin and destination countries population. • Transportation cost (geographical distance multiplied with average oil price). • Tourism own- and weighted substitute prices. • Income similarity index. • Climate distance.	• Tourist arrival exhibits high degree of habit persistence. • Income of both destination and origin countries have positive influence on tourist arrival. • Destination country population is significant and has a negative relationship with tourism demand while origin country population is not significant. • Transport cost, own- and substitute price have negative impact on tourism demand. • Linder's hypothesis indicates that when income converge, tourists are inclined to travel to destination countries. • Climate distance has a positive influence on tourist arrivals.

Table A continued

Santana-Gallego et al. (2015)	195 countries (year 2012).	• Gravity model. • Helpman, Melitz and Rubistein (HMR) model. • Probit equation. • Maximum likelihood estimation.	• Export. • Tourist arrival. • Geographical distance. • Dummy variables representing common border, colonial relationships, common coloniser, common language, common first religion, common member of regional free trade agreement, same currency union. • Number of landlocked in countries pairing, number of islands in countries pairing. • Lagged tourist arrival. • Number of World Heritage Sites per destination in countries pairing. • Annual average temperatures in origin and destination countries.	• Closer countries are more likely to trade with each other. • Having a common coloniser, colonial relationship, common language, belongs to same free trade agreement, share common currency increase probability to trade while number of landlocked countries and islands reduce the probability to trade. • Entry regulations and entry cost are significantly negative. • Tourism increase the chance of trading between countries.

Table A continued

			• Number of legal procedures needed to operate a new business in origin and destination countries. • Entry cost in percentage of GDP in origin and destination countries.	
Santeramo and Morelli (2015)	Italy (Annual data 1998-2010).	• Gravity model. • Quantile regression. • Poisson Pseudo-Maximum Likelihood estimator.	• Tourist arrival and duration of stay in general touristic sector and agritourism. • Tourism supply (number of touristic structures and number of agritouristic structures). • Origin countries income (GDP per capita). • Origin countries population (total population, rural population and population working in agricultural sector). • Geographical distance. • Dummy variables representing common currency and Schengen agreement.	• Tourism supplies increase agritourism but not for overall tourism sector. • All of the demand side determinants are significant except for income in overall tourism demand specification. • The quantile regression revealed that income level and travel cost (distance) are not main determinants for explaining small and large tourism flows.

Table A continued

Untong et al. (2015)	Thailand (Annual and monthly data 1988-2013).	• Seasonality analysis. • Bootstrapping approach. • KPSS stationary test. • Grey model.	• Tourist arrivals. • Income (real GDP). • Tourism price (CPI ratio adjusted by exchange rate). • Substitute price (CPI ratio adjusted by exchange rate).	• Chinese market showed similar pattern despite became more seasonal. • All the explanatory variables are significant factors for Chinese tourist arrivals.
Alawin and Abu-Lila (2016)	Jordan (Annual data 2000-2014).	• Gravity model. • POLS model. • Random effects model. • Fixed effects model. • Panel-GARCH model.	• Tourist arrival. • Destination economic development. • Origin countries' income. • Familiarity with location (lagged dependant variable). • Tourism price. • Hotel capacity (number of hotel room). • Exchange rate. • Geographical distance. • Dummy variable representing common language. • Uncertainty.	• Destination economic development, familiarity with destination and language have a positive effect on tourism demand while cost of living, exchange rate and distance have a negative impact. • Bidirectional causality exists between tourist arrival and uncertainty.

Table A continued

Karaman (2016)	Turkey (Annual data 2000-2013).	• Gravity model. • IPS unit root test. • POLS model. • Panel OLS with time effect. • First difference GMM.	• Tourist arrival. • Lagged dependant variable. • Income and population of origin country. • Total trade between origin and destination countries. • Geographical distance. • Dummy variables representing visa restrictions, regional neighbourhood and common border.	• All the variables are significant under both static and dynamic panel analysis. • Visa restrictions impact is on the higher side for countries with almost visa-free travel.
Tang and Tan (2016)	Malaysia (Annual data 1989-2010).	• IPS, and Maddala and Wu panel unit root tests. • Pedroni cointegration test. • Group mean FMOLS.	• Tourist arrival per capita. • Income (GDP per capita). • Tourism price (CPI ratio adjusted by exchange rate). • Substitute price (weighted average of price adjusted by exchange rate). • Air pollution (per capita carbon dioxide emissions). • Crime rate.	• All the determinants have significant and expected relationship with tourism demand.

Table A continued

			• Dummy variables representing Malaysia Truly Asia promotion campaign, September 11 terrorist attack, SARS outbreak and Avian flu.	
Dogru et al. (2017)	Turkey (Quarterly data 2003Q1-2012Q4, monthly data 2003M2-2012M12).	• IPS and ADF-Fisher panel unit root tests. • Pedroni and Kao cointegrations tests. • Fully Modified Ordinary Least Squares.	• Tourist arrival. • Lagged dependant variable. • Income (GDP per capita, IPI). • tourism price (CPI ratio, CPI ratio adjusted by exchange rate). • Exchange rate. • Substitute price (weighted CPI ratio, weighted CPI ratio adjusted by exchange rate). • Seasonal dummy. • Dummy variable representing global financial crisis.	• Including tourism price standardised by exchange rate and exchange rate in the same model is likely to face multicollinearity and modelling bias problem. • Price and exchange rate should be used to compute tourism price. • Combined use of tourism and substitute prices is not practical in policy-making. • IPI is not a good proxy for income in modelling tourism demand. • Country-specific heterogeneity must be considered in panel data.

Table A continued

Fereidouni et al. (2017)	Malaysia (Quarterly data 2000Q1-2011Q4).	• ADF and PP unit root tests. • Fully Modified Ordinary Least Squares.	• Outbound Malaysian travellers to Singapore. • Income (GDP per capita). • Real effective exchange rate. • Travel cost (jet fuel price). • Terrace house price index. • Dummy variables representing quarter2-4, September 11 terrorist event, SARS outbreak, South Asian tsunami and global financial crisis.	• Income, real effective exchange rate, Wealth Effect from Real Estate (WERE) and quarter 4 (vacation period) has a positive and significant impact on Malaysian outbound travel demand.
Perles-Ribes et al. (2017)	Spain (Annual data 1957-2014).	• ADF, DF-GLS, PP, KPSS, Breitung, Ng Perron, Lee and Strazicich and Carrion-i-Silvestre unit root tests. • ARDL bound test. • Toda-Yamamoto procedure.	• Tourist arrival. • Real international tourism receipts. • Real GDP. • Real gross value added. • Number of employees. • Number of jobs. • Real effective exchange rate. • Dummy variable representing global economic crisis.	• The occurrence of Global Economic crisis resulted the cointegration relationship between tourism demand and economic output become ambiguous as compared to before crisis. • Causality analysis though tend to support bidirectional relationship between tourism demand and economic growth but it is highly responsive on variables measurement and transformation together with model specification.

Table A continued

Rossello et al (2017)	196 countries (Annual data 2000-2013).	• Gravity model. • POLS model.	• Tourist arrival. • Destination country GDP per capita and population. • Geographical distance. • Religious similarity index. • Institution quality. • Terrorism. • Number of World Heritage Sites. • Temperature. • Life expectancy at birth. • Dummy variables representing common border, common language, colonial background, common colonizer, been part of a same country and regional trade agreement, disease risk (Malaria, Dengue, Yellow Fever and Ebola).	• Most of the explanatory variables are significant and have the expected coefficient sign. • The impact of disease risk differs for each disease under study with Malaria having the biggest impact. • In general, tourist from developed countries are more sensitive with the present of disease risk.
Suhaimi and Abdullah (2017)	Malaysia	• Questionnaire. • Semi structured interviews. • Mixed methods. • Sequential explanatory.	• Demographic characteristics. • Language practices. • Challenges faced by shopkeepers and tourists.	• Findings showed that shopkeepers perceived multilingualism is insignificant. • Language barrier is overcome through the use of non-verbal cues to deliver the message.

Table A continued

Yazdi and Khanalizadeh (2017)	United States of America (Annual data 1995-2014).	• Gravity model. • LLC, IPS and Maddala and Wu panel unit root tests. • Kao and Pedroni cointegration tests. • Panel ARDL.	• Tourist arrival. • GDP (GDP per capita ratio). • Price (CPI ratio). • Total number of flights. • Real exchange rate. • Dummy variable to capture September 11 terrorist attack.	• Income elasticity suggest that tourism is a non-luxury good. • Terrorist attack, prices and real exchange rate have negative relation to tourist arrivals. • Tourism transport infrastructure is a significant determinant of tourist arrivals into USA.
Okafor et al. (2018)	200 countries (Annual data 1995-2015).	• Gravity model. • OLS. • OLS with fixed effects. • Random effects model. • Poisson Pseudo-Maximum Likelihood.	• Bilateral tourist arrival. • Population of origin and destination countries. • Income (real GDP per capita) of origin and destination countries. • Geographical distance. • Dummy variables representing common official and unofficial languages, contiguity (common border) former colonial linkage, landlocked and island.	• Common unofficial language is more important determinant than common official language regardless of level of developments in origin and destination countries. • The same findings are found for regions of origin and destination especially Europe region where common official language is not significant.

Table A continued

Samitas et al. (2018)	Greece (Monthly data 1977M1-2012M12).	• Principal Component Analysis (PCA). • ADF and PP unit root tests. • Johansen and Juselius cointegration test. • Granger causality test.	• Tourist arrival. • Terrorist incidents (a common factor extracted from three proxies – terrorist incidents with casualties, no casualties, and total terrorist incidents).	• Terrorism has a significant negative impact on tourist arrival to Greece and a unidirectional causality is detected from terrorism to tourism. • The impact persisted in the long run.
Jerabek (2019)	Czech Republic (Monthly data 2002M1-2018M5).	• ADF and PP unit root tests. • Johansen and Juselius cointegration test. • VECM. • Granger causality. • Impulse Response Function. • Variance Decomposition.	• Tourist arrival. • Income (IPI). • Tourism price (restaurant and hotel CPI ratio adjusted by exchange rate). • Travel cost (crude oil price).	• Long run relations exist for all four origin countries (Germany, Poland, Austria and Slovakia). • Income is significant for all countries while tourism price is only significant for Austria and Slovakia. • Poland is the only country that is not influenced by travel cost.
Puah et al. (2019)	Vietnam (Annual data 2011 to 2017).	• Gravity model. • POLS model. • Random effects model. • Fixed effects model.	• Tourist arrival. • Origin and destination countries income (real GDP). • Tourism price (CPI adjusted by exchange rate). • Travel cost (distance multiplied by real crude oil price).	• All independent variables are found to be significant determinants of tourism demand in Vietnam. • Destination and origin countries income and tourism price encourage tourism demand while travel cost discourage tourist arrival into Vietnam.

Table A continued

Viljoen et al. (2019)	25 African countries (Annual data 2001-2010).	• POLS model. • Random effects model. • Fixed effects models. • Least square dummy variable. • Bias-corrected LSDV. • System GMM.	• Tourist arrival. • Income (GDP). • Tourism price (CPI ratio). • Health risks. • Safety. • Telecommunication infrastructure. • Urbanisation rate. • Death rate. • Dummy variables representing terrestrial and marine protected areas, landlocked country, Northern African country, and common border with South Africa.	• Telecommunication infrastructure consistently influences tourism demand in the whole Africa and its regions followed by tourist income. • The remaining determinants have different effects on Africa and its regions.
Santana-Gallego et al. (2020)	171 destination countries (Annual data 1995-2013).	• Fixed effects model.	• Tourist arrival (personal reason and business and professional reason). • Income (real GDP per capita). • Tourism price (ratio of PPP conversion factor to market exchange rate). • Voice and accountability.	• Income and voice and accountability are significant and positively influence tourist arrival while tourism price is not significant. • Terrorism and crime are significant and reduce tourist arrival while corruption is not significant for whole sample. • The impact of terrorism and crime are larger for personal trip as compared to business trip while corruption is only significant for business trip.

Table A continued

			• Tourism threats (terrorism, corruption and crime).	• Destinations with World Heritage Sites attractiveness seem to moderate the effect of crime and corruption. • Tourist are found to tolerate a higher level of crime when they are visiting developing countries.
Seabra et al. (2020)	Portugal (Annual data 2002-2016).	• Unrestricted Vector Autoregressive model.	• Tourist arrival by region of origin. • Lagged dependant variable. • Number of terrorist attack occurred worldwide	• Terrorist attacks have a strong impact on tourist arrivals in Portugal and terrorism spill over effect exists. • The substitution and generalisation effects both exist and the effect depends on the location of the terrorist attack and the origin of tourist. • However, the overall findings indicate generalisation effect is more dominant in Portugal. • Short memory effect exists where arrival of tourist in a certain year affects the tourist arrival in following year.
Takahashi (2020)	Singapore and French Polynesia (Annual data 2008-2013)	• Random effects model.	• Tourism flow. • Origin country income (GDP per capita). • Origin country population. • Tourism price (CPI adjusted by exchange rate). • Geographical distance.	• Demand in French Polynesia tends to result in a relatively high-income market but the economy is affected by global phenomena. • Singapore with more diversified industries is likely to have good accessibility and the global economic impact is lower in the tourism market.

Table A continued

		•	• Dummy variables representing colonial relationships and common language. • Time dummy representing Global Financial Crisis 2008.	•
Tatoglu and Gul (2020)	14 most visited countries in the world (Annual data 2008-2016).	• Multi-dimensional panel gravity model. • Maximum likelihood estimation.	• Tourist arrival. • Income of origin and destination countries (GDP per capita). • Geographical distance. • Export. • Import. • Purchasing Power Parity. • Dummy variable representing Mediterranean coast.	• Income of origin and destination countries, export, import and Mediterranean coast have positive impact on tourist flow while PPP and geographical distance have adverse effect on tourist flow.

Table A continued

Ulucak et al. (2020)	Turkey (Annual data 1998-2017).	• Cross-sectional dependency tests. • Cross-sectionally augmented IPS panel unit root test. • Durbin-Hausman panel cointegration test. • Continuously updated fully modified (CupFm) and continuously updated bias-corrected (CupBC) estimators. • Gravity model.	• Tourist arrival. • Destination and origin countries income (GDP per capita). • Geographical distance. • Exchange rate. • Tourism price (CPI ratio). • Number of violence/terrorism incidents. • KOF globalisation index. • Household debt per capita (percentage of GDP). • Money supply (percentage of GDP).	• All explanatory variables are significant and have expected relationship except money supply which is not significant in CupFM estimator. • Household debt is more effective indicator of disposable income.

Table A continued

Wamboye et al. (2020)	Tanzania (Annual data 2000-2016).	• Descriptive analysis. • LLC panel unit root test. • Modified Wald test for groupwise heteroskedasticity. • Fixed effect model. • Fixed effect model with instrumental variable. • Generalized Estimating Equation (GEE) population averaged. • Linear dynamic panel estimation. • Stepwise regressions with Ordinary Least Square estimation (country level analyses).	• Tourist arrival. • Measure of prices (exchange rate, cost of living). • Transportation cost (distance multiplied with fuel price). • Infrastructure development. • Political stability.	• Income and infrastructure development are two main determinants which are robust throughout different model and sample specifications. • Transportation cost, cost of living and exchange rate have the correct relationship whenever these variables are significant. • Lastly, political stability is not significant determinant and does not improve model estimation.