DETERMINANTS OF INBOUND TOURISM DEMAND IN TANZANIA

Benedicto Kazuzuru

PhD (Economics) Dissertation
University of Dar es Salaam
May, 2013

DETERMINANTS OF INBOUND TOURISM DEMAND IN TANZANIA

By
Benedicto Kazuzuru

A Dissertation Submitted in (Partial) Fulfillment of the Requirements for the Degree of Doctor of Philosophy (Economics) of the University of Dar es salaam

University of Dar es Salaam
May, 2013

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance by the University of Dar es Salaam a dissertation entitled: ***Determinants of Inbound Tourism Demand in Tanzania***, in fulfillment of the requirements for the PhD degree (Economics) of the University of Dar es Salaam.

.....

Prof. Asmerom Kidane

(Supervisor)

Date: 29-4-2013

.....

Prof Nehemiah Osoro

(Supervisor)

Date: 29 April 2013

.....

Dr. Razack Lokina

(Supervisor)

Date: 29 April 2013

DECLARATION

AND

COPYRIGHT

I, **Benedicto Kazuzuru**, declare that this dissertation is my own original work and that it has not been presented and will not be presented to any other University for a similar or any other degree award

Signature----------

AKNOWLEDGEMENTS

I express my gratitude to God, the father of Jesus Christ for giving me power and strength to accomplish this research.

I thank my supervisors, Professor Asmerone Kidane, Professor Nehemiah Osoro and Dr.Razack Lokina for their advice and help which has greatly shaped my work.

I register my thanks to Dr.Rutasitara, the former Head of the Department of Economics for his efforts to ensure that PhD students in my cohort are well trained and supported accordingly. In particular I thank him for endorsing my letters requesting data from different institutions.

I express my gratitude to Professor Adolf Mkenda who is also the current Head of the Department Economics for his efforts in my training of micro econometrics and most particularly in duration analysis. His lecturing and comments made my way easy in the chapter pertaining to duration analysis.

I am also grateful to Dr. Godius Kayharara and Dr. Vicent Leyaro for the review of my work. Alongside them is Dr. Elizabeth Robinson-a visiting fellow at the University of Dar es Salaam, who also reviewed my work. Her comments have really shaped my work.

I am also very grateful to the office of the Progamme for Agricultural and Natural resources Transformation for Improved livelihoods (PANTIL) which offered me the

scholarship to pursue this study. I also register my thanks to the Postgraduate Directorate of Sokoine University of Agriculture which managed the project.

The scholarship offer could have not been realized without the support of the Department of Biometrics and Mathematics of Sokoine University of Agriculture which in collaboration with the office of the Deputy Vice Chancellor (Academics) allowed me to pursue this degree, while leaving behind a huge number of tasks to be executed. I therefore very sincerely register my thanks to the department, and in particular to the Head, Dr. Geoffrey Karugila.

My final appreciation should go to all institutions and individuals who enabled me to obtain the data for my research. In particular the following institutions are greatly appreciated: Bank of Tanzania (BOT), Ministry of Tourism and Natural Resources (MNRT), Tanzania National Parks Authority (TANAPA), Ngorongoro Conservation Area Authority (NCAA), Tanzania Tourism Board (TTB) and Zanzibar Commission for Tourism (ZCT).

In each of these institutions, there were particular individuals who played a distinctive role for me to smoothly obtain the needed information. These are madam Vilela Kejo of BOT, Pascal Mwiru of MNRT, Diocless Emmanuel of TANAPA, Deogratius Malogo and Edward Mbwiga of TTB, Amour Mtumwa Ali and Hija Makame of ZCT. Others are Judica Mlayi and Simon Kalembo of NCAA. I appreciate all of them for their help.

The information supplied to me from the BOT was jointly collected by a committee known as Tanzania Tourism Sector Survey (TTSS), which comprises five member

institutions: MNRT, BOT, National Bureau of Statistics (NBS), Immigration Department and ZCT. I would therefore like to appreciate all the individual institutions as well as the TTSS itself and in particular the Steering and Technical committees of TTSS. My work would have been very difficult without the information from TTSS.

Lastly I cannot dare to claim that this work is perfect. Therefore any error in this work is my responsibility entirely and not to any of the acknowledged individuals or institutions.

DEDICATION

To my late parents, Mr. Jacob Midaba Kazuzuru and Mrs. Karolina Rukurugu Kazuzuru, my sisters and brothers who I grew up with, and my lovely son Yalaita Kazuzuru.

ABSTRACT

The study investigates the determinants of inbound tourism demand both at macro and micro levels. At macro level the study investigates the determinants of the number of tourist arrivals, while at micro level the study investigates the determinants of the per capita expenditure of tourists, their length of stay and their choice of a package tour. In achieving the first objective, panel data regression analysis was employed using the number of arrivals obtained from the Ministry of Tourism and Natural Resources (1995-2007) as well as a number of covariates from different sources. Regarding the determinants of tourists' per capita expenditure, an OLS was applied on a cross-section of tourists surveyed in the years 2001, 2007 and 2008. As for the length of stay, a survival analysis was employed, whereas for the choice of a package tour a binary logistic regression was used, in each case using the cross-section data of the years 2001, 2007 and 2008. Among the key findings is that non price factors (such as the country's economic development) are more influential than price factors in attracting tourists to Tanzania. Trip-related characteristics of the tourists such as purpose of visit and the number in travel party were found to be the most influential variables in explaining tourists' daily spending, length of stay and choice of package tour. To promote the number of arrivals it is implied that the government should invest more in the non-price factors such as infrastructure whereas to enhance micro demand, promotion organs and other stakeholders should attune their marketing strategies more to trip-related characteristics of the tourists than to other factors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFT	Accelerated Failure Time
BOT	Bank of Tanzania
EAPA	East African Publicity Association
EATTA	East African Tourist Travel Association
GMM	Generalized Method of Moments
IMF	International Monetary Fund
MFEA	Ministry of Finance and Economic Affairs
MIT	Ministry of Information and Tourism
MNRT	Ministry of Natural Resources and Tourism,
NTB	National Tourist Board
NBS	National Bureau of Statistics
NCAA	Ngorogoro Conservation Area Authority
TANAPA	Tanzanian National Parks Authority
TRA	Theory of Reasoned Action
TTB	Tanzania Tourism Board
TTC	Tanzania Tourist Cooperation
TTSS	Tanzanian Tourism Sector Survey
VIF	Variance Inflation Factor
WTO	World Tourism Organization
WTTC	World Travel and Tourism Council
ZCT	Zanzibar Commission for Tourism

CHAPTER ONE

INTRODUCTION

1.1 Background to the problem

One of the major problems in the macroeconomic performance of most developing countries including Tanzania is the underperformance of the external sector. In Tanzania, efforts have been made to promote exports in a number of ways so as to address the problem (Mjema, 2004). Some of these efforts have led to numerous studies on exports, such as Nyoni (1996), and Rutasitara (1999), but few of them such as Kweka (2003) have addressed the trade in services and tourism in particular. Despite such a deficiency, the importance of tourism[1] cannot be overemphasized, both globally and nationally.

According to WTTC (2009) tourism contributed 10 percent to the global gross domestic product (GDP) in 2006, and 2008. WTTC (2009) states that total tourism revenue in those two years were valued at US$ 908 billions and US$1159 billions respectively. As regards to employment, the sector accounted for 9 percent of global employment in 2006, and 8 percent in 2008 (WTTC, 2009). As regards to Sub-Saharan Africa the sector accounted for 7 percent of GDP in 2006 and 8 percent in 2008, whereas the sector contributed to 5 percent of employment in the region in 2006

---

[1]Tourism is the act of people traveling to and staying in places outside their usual environment for not more than one consecutive year for leisure, business and other purposes not related to the exercise of an activity remunerated from within the place visited (WTO1995). It includes private travel for holiday and recreation purposes but also business travel (WTO, 1995). This study adopts the definition as given by World Tourism Organization (WTO). In particular the study considers travelers who visit Tanzania(inbound tourism) and not otherwise.

and 6 percent in 2008 (WTTC,2009). [2]In the case of Tanzania, WTTC estimates that in 2006 tourism receipts were valued at US$986 million, which is 11 percent of the country's GDP whereas in 2008 tourism receipts were valued at US$ 1358 million which is 10% of GDP. On average the contribution of tourism to GDP was 10% since 2002 to 2008 (WTTC,2009).With regard to exports, the contribution of tourism revenue generally rose during the years 1986 to 2008, amounting to more than 40 percent in the years 1995 and 1999 (MFEA, 2009)

Tourism demand is a broadly defined subject with some scholars using different indicator variables, such as number of arrivals, length of stay, daily spending and number of occupied rooms (Meniz and Munoz, 2006). An examination of tourism performance in Tanzania using any of the said indicators shows that the sector has been doing well during the last two decades, except in some few years, such as in 2009 when the sector was slightly affected by the 2008 global financial crisis. Revenue has been increasing as well as the number of arrivals (MFEA, 2007, 2008). This implies that even tourist per capita spending is also increasing. Survey statistics from TTSS (2001, 2004, 2005, and 2006) also indicate that tourist length of stay has remained stable since 2001, at around 12 days. Despite such good performance, this study is set out to examine determinants of inbound tourism demand. Its justification lies on the following grounds:

[2]The statistics given by WTTC on tourism's contribution to GDP appears to be different from those provided by the Ministry of Natural Resources and Tourism in Tanzania (MNRT). According to MNRT, tourism contributed on average 6% to GDP between 1993-1998, 14% in 2004 and 17% between 2005 and 2007. Unfortunately, MNRT does not have a long-term **annual** statistics on the contribution of tourism to GDP. Therefore this study solely uses the estimates as given by WTTC.

First, the country still faces a persistent balance of payments problem which needs to be addressed by improving, among other things, tourism exports. As stated, very few studies on this sector have been conducted in Tanzania especially on inbound tourism demand, and those that have been conducted have not explicitly addressed the issue of tourism demand. With the exceptional of Bashagi and Muchapondwa (2009), none of tourism studies in Tanzania directly address tourism demand (See for example, Mgani (1997), Chenja (1998), Kweka (2003), Nzuki (2006), and Anderson (2011). This scenario overlooks the fact that tourism can be a major means of addressing the balance of payment's deficit. Even though Bashagi and Muchapondwa (2009) studied determinants of tourist arrivals in Tanzania, the study did not address quite a number of issues which can affect the coming of tourists. These issues include marketing expenditures, the growth of Tanzanian economy, trade liberalization, infrastructure, and presence of neigbour countries. Besides the study did use relatively fewer number of observations much as it was a time series based study of which its data, for most African countries dates no longer than 1960.Further to this, the study had used proxies for a tourist country's income as well as for a tourist country's exchange rate against Tanzania and against Kenya. All these facts could make the study's findings unreliable.

Second, when a comparison is made between Tanzania and some neighbor countries, such as Kenya and Botswana, Tanzanian tourism is not doing well. Until recently the country's tourism revenue and arrivals had been lagging behind Kenya, Botswana, and South Africa (Fletcher and Morakabati, 2008), despite the fact that Tanzania is the only country in East Africa which has allocated 28 percent of its land for wildlife protection and nature conservation, with extraordinary variety of high quality wildlife

and resorts as well as archeological, cultural and historical assets that are in demand in the international tourism markets (MNRT, 2002).One therefore is tempted to find out why Tanzania lags behind such countries. One of the ways to address such a problem is to strongly promote the sector (MNRT,2002).To be able to do that effectively, one needs to understand the determinants of tourism demand by examining one or more of its components.

Third even though the sector is doing well this does not guarantee that in the future nothing will go wrong, and so to be prepared for any eventuality means that the determinants of tourism demand need to be understood. For example following the 2008 global financial crisis, the statistics show that the sector has been globally retarded (WTO, 2009). According the WTO (2009), major tourist destinations all around the world have suffered from weakened demand, except North Africa, Sub-Saharan Africa, Central America and Southern America. In Tanzania the number of tourist arrivals dropped from 770,376 in 2008 to 714,367 in 2009 (MFEA, 2009).Likewise tourism revenue dropped slightly from US$1199 million in 2008 to US$ 1163 million in 2009 (MFEA, 2009). This being not the end of the crisis, one is tempted to think that the tourism sector might be further affected in the future. But the question is through which aspects can the sector be affected? The answer to such a question demands clear knowledge of what determines tourism demand in Tanzania.

Fourth, even though the sector is doing well, policy makers need to understand the determinants of tourism demand, not just to promote the sector but also to safeguard the sector in the face of various policy formulations. For example, how will various tax policies on tourism affect the coming of tourists to Tanzania? Or how will

sustainable environmental policies such as restricting tourist length of stay and/ or large party visits, affect tourism revenue? The answer to the first question requires knowing what determines the number of tourist arrivals as one of the elements of tourism demand. Similarly the answer to the second question requires knowing what determines a tourist per capita spending in Tanzania as another of the elements of tourism demand.

Globally and continental-wise, tourism demand has been extensively studied. Ekanayake et al. (2012), Kordbacheh et al. (2012) provide cases of recent studies in developed countries. Mohamed (2011) and Eja et al. (2012) provide cases of recent studies in African countries. Each of these studies came up with unique results, particularly on the relative importance of the factors determining tourism demand. For example, Ekanayake et al. (2012), Kordbacheh (2012) as well as Mohamed (2011) found tourism prices to be more influential in determining tourist number of arrivals than other factors while other scholars found that prices and tourist income did not affect tourists number of arrivals (Muchapondwa and Pimhidzai 2008, Naude and Saayman 2005).This discrepancy across destinations warrants a study on tourism demand in Tanzania.

Even though tourism demand is a broadly defined subject with scholars using different indicator variables, most studies use number of tourist arrivals and/or tourists' aggregate expenditure to assess tourism demand at a particular destination (Lim, 1997; Durbarry, 2001, Zhang et al.2009). This practice however, does not provide a comprehensive picture to the destination policy makers because the models deal only with macro variables. Tourists' expenditure at a destination depends, among other

things, on the demographic characteristics and trip-related characteristics of the travellers (Wang et al. 2006), which can hardly be captured at the macro level. For example, even though travelling costs matter to tourists when visiting a particular destination, they cannot explain tourist spending patterns after reaching a destination. Similarly they cannot explain what makes a tourist stay for a short or long time at a destination.

One of the distinctive features of this study has been the attempt to assess the determinants of tourism demand both at the macro level (using the number of tourist arrivals as a response variable) and at the micro level (using tourism expenditure per person per day) so as to show a wide-ranging policy implications for Tanzanian policy makers. In addition, the study investigates in detail the factors affecting tourist per capita expenditure in Tanzania. These factors include length of stay and travel arrangements (TTSS, 2001).

The use of number of tourist arrivals instead of tourist aggregate expenditure at the macro level is motivated by the fact that the latter depends, among other things on the former. Moreover tourist per capita spending is examined at the micro level and therefore studying tourist's aggregate expenditure would be a repetition. In general, the study is based on the WTO (1995) model for determining a country's total tourism revenue. The model links total tourism revenue in a country to the number of arrivals, per capita spending and length of stay. The WTO (1995), model states that a country's total tourism revenue at any period say a year is a by product of total number of arrivals in the said period to tourist per capita expenditure per night and to average length of stay (nights) of a tourist in the period.

The logic behind this model is the more tourists a country receives, the more likely the revenue is going to be higher. However, having more tourists is not sufficient as it also depends on tourist consumption patterns, and so one needs to consider the expenditure per tourist. The higher the tourist's per capita expenditure the more revenue a country would receive. Lastly, even if a country had many tourists who are good at spending, length of stay also matters as regards revenue generation. The longer tourists stay the more revenue they typically generate.

This study addresses ways to boost tourism revenue by examining factors determining number of tourist arrivals, tourist per capita expenditure, and average length of stay .In addition tourist choice of travel arrangements (package versus non-package tour), which is thought to influence tourist spending (TTSS, 2001), is also examined (See appendix 1.1, for a diagrammatic link of these concepts).

1.2 Statement of the Problem

Tourism is an important sector both worldwide (WTTC, 2009) and countrywide (MNRT, 2002, TTB, 2006). Despite such an importance the factors determining inbound tourism demand have not been rigorously established. To the best of my knowledge, with the exceptional of Bashagi and Muchapondwa (2009) no scholarly study in Tanzania which has directly addressed the determinants of tourism demand (see for example, Mgani (1997), Chenja (1998), Kweka (2003), Nzuki (2006) and Anderson (2011)). As pointed earlier the study by Bashagi and Muchapondwa (2009) has a number of defects which can be grouped into four parts. One is the study has assessed relatively fewer explanatory variables, second the study used relatively fewer data points (121), third the study used proxy for a tourist country's income and fourth

the study could not capture individual tourist country exchange rates against Tanzania as well as against Kenya. As a consequence the study failed to prove that Kenya is among the countries which compete with Tanzania.

Lack of knowledge on the key determinants of tourism demand by both the government and other stakeholders could be detrimental to country's tourism promotion efforts especially bearing in mind that the country still faces a persistent balance of payment problem which could be addressed by among other things, through the promotion of tourism exports. Further to this, in order for Tanzania to compete effectively with neighbouring countries in the tourism sector from the demand side as well as to be prepared with the necessary tools for supporting the sector in the face of an economic crisis or new government policies, the government and private stakeholders need to understand the factors that determine tourism demand in Tanzania.

As said earlier in the background information, globally and continentally tourism demand studies have been widely covered both at macro and micro levels. But the studies' findings vary from one destination to another especially in terms of the factors which are more influential in determining tourism demand. This scenario implies that to understand the most influential determinants of tourism demand in Tanzania a study is needed. Besides the said fact, some tourism demand studies such as Durbarry (2001) have treated all neigbouring countries as competitors something which may not necessarily be the case. Further to this, most studies as according to Munoz (2004), have not assessed the impact of marketing expenditure on tourism demand.

At micro level most studies done on determinants of tourist per capita expenditure, length of stay and tourist choice of a package or non-package tour have taken for granted that all regressors are exogenous something which is not true. For example most studies modeling tourist length of stay such as Govakali et al. (2007) and Yang et al. (2011) have treated tourist mode of travel arrangement (package or non-package tour) as an exogenous regressor and similarly studies modeling mode of travel arrangement (package or non-package tour) such as Io and Hallo (2009) have treated length of stay as an exogenous regressor. On the contrary these two variables causes each other and therefore they are not exogenous to each other making the coefficient estimates from such studies to be quite unreliable. For example a tourist who plans to stay longer is likely to choose a non-package tour for many good reasons such as the high costs of longtime packages. Similarly a tourist travelling on non-package tour is likely to extend his stay at a destination if a need arise than for a tourist on package tour.

This study aims to address these gaps in the course of studying determinants of tourism demand in Tanzania. The promotion of any sector requires the use of the government's meager resources by focusing on the most influential factors. Likewise in promoting tourism demand the government needs to focus on the most influential factors in the demand for tourism. This requires a study on tourism demand, from which the most influential factors can be identified.

The study is set out to examine the determinants of tourism demand in Tanzania by focusing on four interrelated elements. These are number of tourist arrivals, tourist per capita spending, tourist length of stay and tourist choice of a package tour. In each of

these elements the study intends to identify their determinants with a view of prioritizing the factors that promote tourism demand.

Regarding the number of tourist arrivals, some determinant factors could be the relative cost of living in Tanzania, tourist income, exchange rates, infrastructure and distance (Lim, 1997; Naude and Saayman, 2005; Ekanayake et al., 2012). Based on the consumer theory of demand, some of these factors such as relative cost of living, tourist income and exchange rate, can be regarded as price factors, whereas infrastructure and marketing expenditure can be regarded as non-price factors. So in setting strategies for promotion, which factors should be given first priority? Is it price factors or non-price factors? In other words, which group between these two is more influential than the other?

Regarding tourist per capita spending, tourist length of stay and tourist choice of a package tour, some influential factors could be age, gender, travel party size, purpose of visit, familiarity with the destination, frequency of visits and destination attributes (Wang et al. 2006; Io and Hallo, 2009; Yang et al., 2011; Chaiboonsri and Chaitip 2012). Accordingly these factors may be classified into demographic, trip-related characteristics and destination attributes (Mathieson and Wall, 1984; Wang et al., 2006; Barros and Correia, 2007). So among these three groups which one should be given first priority in setting strategies for tourism promotion? Is it destination attributes as opposed to the other two or the vice-versa? In other words which one among these is more influential than others in relation to tourist spending, length of stay and choice of a package tour?

This study aims to address the questions raised above as well as the gaps identified in the past studies and set forward priority factors for promotion of inbound tourism demand in Tanzania.

1.3 Objective of the Study

The overall objective of the study is to establish the determinants of tourism demand in Tanzania, both at the macro and micro level. Specifically the study focuses on the following aspects:

(i) To examine the determinants of tourist arrivals

(ii) To examine the determinants of tourist per capita expenditure

(iii) To examine the determinants of tourist length of stay

(iv) To examine the determinants of tourist choice of a package tour

1.4 Hypotheses

The study intends to test the following hypotheses:

(i) Price factors are more influential than non-price factors in attracting number of tourist arrivals

(ii) Destination attributes are more influential than the demographic and trip-related characteristics in explaining tourist per capita spending

(iii) Destination attributes are more influential than the demographic and trip-related characteristics in explaining tourist length of stay

(iv) Destination attributes are more influential than the demographic and tri-related characteristics in explaining tourist choice of a package tour

1.5 Significance of the study

First, the significance of this study is the fact that it addresses itself on how to generate more foreign currency for Tanzania. By revealing the determinants of tourism demand, the study will contribute to tourism promotion efforts, which will in turn lead to more foreign currency generation.

Second significance of this study can be seen in the fact that it provides policy makers with the necessary promotion strategies for competing with neighbouring countries. For example, by knowing which factors matters more than the others in attracting tourist arrivals, the government will be able to use effectively its meagre resources for promotion, and thereby being able to compete effectively.

Third, the significance of the study emanates from the fact that it will contribute to policy makers' ability to support the tourism sector in the case of an economic crisis. For example, if it is found that tourists' income matters for their coming, then one of the possible measures in the face of the 2008 global financial crisis is to reduce the fee charges to tourists. These include charges by the tour operators, Hotels owners and Tanzanian National Parks Authority.

Fourth, the significance of this study lies in the fact that it provides grounds for assessing the economic impact of taxation on the tourism industry. By establishing price elasticity of tourism demand in Tanzania, based on both prices in Tanzania and in the neighbouring countries, it will help policy makers in assessing the sensitivity of taxing the tourism sector.

Fifth, the significance of this study is that it addresses the issue of sustainable tourism, which focuses on a concern for environmental protection by using promotional strategies that produce more per capita expenditure than mass tourism (WTO, 1995, Neto, 2002). This can be achieved by studying the determinants of tourist per capita expenditure and tourist length of stay as posited by this study.

1.6 Scope and Limitations of the Study.

The study examines the determinants of tourism demand at macro level using the data on tourist number of arrivals as provided by the Ministry of Natural Resources and Tourism (MNRT) and at micro level using surveys on tourists as conducted by the Tanzania Tourism Sector Survey (TTSS). All these are secondary data and are meant to cover the entire country. As it is usually the case secondary data can sometimes not fit exactly into ones' study objectives. For example in tourists' surveys, some of the needed variables were missing and ways were found to derive them from other variables. The derived variables are merely proxies to the actual variables.

Another limitation stems from the fact that even though MNRT as well as emigration department might have tried to capture the correct numbers of tourist arrivals, there some visitors who may have entered Tanzania informally. The same applies to the surveys made on tourists. These surveys were conducted in only six centers: Dar-es-salaam International Airport, Kilimanjaro International Airport, Zanzibar Airport, Namanga in Arusha Region, Kasumulo in Mbeya Region, and Holili in Kilimanjaro. There some other regions such as Kagera bordering Uganda and Rwanda, Kigoma bordering Burundi as well as Mtwara bordering Mozambique which were not included implying that the conducted surveys might have missed the tourists who entered via

the Kagera, Kigoma and Mtwara regions.

The use of consumer prices indices to calculate relative cost of living between a tourist country and the destination country which is then used as a proxy for tourism price presumes that all the goods and services consumed in a destination country can also be consumed by a tourist. But sometimes tourists consume only particular types of goods implying that tourism prices can either be overestimated or underestimated. Nevertheless this is a common weakness in most studies on tourism demand. Very few countries have been able to separate a basket of goods consumable by tourists against the one consumed by the indigenous.

Despite what has been said, these weaknesses can not invalidate the findings from this study, especially because both the number of arrivals and the total number of surveyed tourists were large enough to yield meaningful statistical inferences.

1.7 Thesis Organization.

This chapter has addressed the motivation for the study and key objectives of the study and how they are interlinked. The chapter has also given a picture of the importance of tourism to Tanzania and the relevance of the entire study. The remaining chapters are organized as follows: Chapter two gives an overview of the tourism sector in Tanzania, the challenges and opportunities it faces, and how they relate to the study. Chapter three addresses the determinants of the number of tourist arrivals, chapter four examines the determinants of tourist per capita expenditure, chapter five investigates determinants of tourist length of stay and chapter six looks at the determinants of tourist choice of package or non-package tour. Chapter 7 concludes the whole study and highlights the main findings.

CHAPTER TWO

AN OVERVIEW OF TOURISM SECTOR IN TANZANIA

2.1 Introduction

In order for one to understand thoroughly the tourism sector in Tanzania it could be important first to understand the country itself, its people and culture as well as the historical background of tourism in Tanzania based on both its management and performance. It is also important for one to understand the key challenges of tourism industries as well the national tourism policy. This chapter is geared towards these objectives.

By understanding the history of tourism one could be able to answer questions like why is it that European market over the years has been the most dominant one compared to say Asia or America. But one could also be able to know the historical challenges of tourism and therefore answer questions like why is it that tourists are overcrowded in the northern circuit of Tanzania and not in other parts of the country? Is it just because of the presence of tourism assets only? These and other questions can be answered implicitly or explicitly by going through the mentioned issues.

The remaining sections in this chapter are organized as follows: Section 2.2 gives the country's profile which includes its history, the economy, the growing importance of tourism, and the people of Tanzania. Section 2.3 addresses the history of tourism, section 2. 4 provide the performance of tourism, section 2.5 provides the challenges of tourism industry, and section 2.6 provides the national tourism policy while section 2.7 concludes the chapter.

2.2 Tanzania Country Profile, Economy and the Growing Importance of Tourism

The United Republic of Tanzania is the largest country in East Africa and the second largest in the SADC region (MNRT, 1999). It is located in East Africa between latitudes 1^0 and 11^0 south of the equator and longitudes 30^0 and 40^0 of Greenwich, covering an area of 945, 234 sq kms (MNRT, 1999).The country was formed in 1964 as a union of the two independent states of Tanganyika and Zanzibar (MNRT,2002). Tanganyika attained its independence in 1961 from Britain while Zanzibar which was being ruled by the British alongside the Oman sultanate got her independence in 1963((Ward and White, 1971).

Tanzania has passed through stages of economic reforms since it stopped socialism policies in the mid 1980's having adopted them in 1967. The government started implementing institutional reforms in the early 1990, marking a major shift from the government led economy to a private sector-led economy (MNRT, 2002).Since then there have been significant improvements in the performance of the economy. In 2004 the real GDP growth rate had reached 7.8%, the highest since the start of the economic reforms (MFEA, 2008). In the next three years it declined to 7.1% in 2007 and 7.4% in 2008(MFE, 2008). In 2009 and 2010 the growth rate was significantly reduced to 5% and 6% respectively. These reductions might have been caused by the 2008 global financial crisis.

According to the National Bureau of Statistics (NBS) the country's economy is classified in four main sectors. The first constitutes agriculture, hunting and forestry, the second constitutes fishing, the third constitutes industry and constructions and the

fourth constitutes services. These sectors have on average been contributing 28%, 1.7%, 19% and 46% respectively to total GDP since 1998 up 2007[3].

Recently, tourism has emerged as an important sector. Even though not explicitly shown as an independent sector in the national accounts, its contribution is inherently captured in other sectors (TTSS, 2001). For example, its contribution can be found in the agriculture sector through earnings from hunting, and in the services sector through earnings by hotels and restaurants, transport and communication and financial services (TTSS, 2001)

According to the Tanzania Tourism Board (TTB) 2006 report, tourism contributed 17% to GDP from 2005 to 2007. The contribution has well been noted by the government. The government views tourism as a significant industry in terms of job creation, foreign currency generation and poverty alleviation (MNRT,1999).The sector is now receiving greater attention than ever before from the government and international agencies (MNRT, 1999).

Tanzanian tourism is predominantly of wildlife nature (MNRT, 2002). The country has 15 national parks which contain various species of wildlife, ranging from mammals, birds, fish, reptiles and amphibians. In addition to the parks, there are 31 game reserves, including the famous Selous game reserve as well as 38 game controlled areas, including the popular Ngorongoro Conservation Area. These sites make up 28 percent of the entire land area of Tanzania (MNRT, 2002).

[3]The figures are based on my own calculations using statistics from the MFEA(2007)

Apart from wildlife assets, the country also has a variety of historical and archaeological assets which form part of cultural tourism. These include Stone Town in Zanzibar, Bagamoyo, Kilwa and the island of Kilwa Kisiwani, Olduvai Gorge, Isimila (near Iringa) and Tarangire (MNRT, 2002).

Among all the mentioned assets, six of them have received world recognition as being world tourist heritages. These are Ngorongoro Conservation Area, Serengeti National Park, Selous game reserve, Kilimanjaro National Park, Zanzibar, the ruins of Kilwa Kisiwani and the ruins of Songo Mnara (MNRT, 2002).

Apart from the tourist assets, the Tanzanian people themselves are a big source of inspiration to visitors. Tanzanians are a warm, open and friendly people, long known for their generosity, hospitality and wealth of folklore (MNRT, 1999). The country has 126 major ethnic groups comprising of Bantu, Nilotic and Hamitic vernacular languages, traditions and customs which greatly appeal to tourists (MNRT, 1999).Another reason for tourists' admiration of Tanzania is the fact that since independence the country has been politically and socially stable unlike neighbouring countries, such as Rwanda, Burundi and Uganda. In fact Tanzania is a very peaceful country. This view is also supported by most interviewed tourists, who assert that Tanzania is a very safe destination with friendly people (MNRT, 2002).

2.3 The 19th Century History of Modern Inbound Tourism in Tanzania

The 19th century history of modern inbound tourism activities in Tanzania dates back to the period preceding the effective establishment of colonial rule in Tanzania and Africa in general. Prior to the effective establishment of the colonial rule in Tanzania

and Africa in general there came explorers from Europe who came to find out more about the geophysical characteristics of the continent (Coupland, 1968). These explorers, such as Richard Burton, John Speke, Henry Morton Stanley and Johannes Rebmann who were in essence the agents of colonialism, were later on credited by Europeans as the discoverers of key natural tourist attractions, such as, Lake Tanganyika, Lake Victoria and Mount Kilimanjaro (Swayne, 1868). These people could be compared to modern tourists[4] in East Africa. At that time the only way to get to East Africa was by ship and Zanzibar was the only reliable port alongside East Africa (Swayne, 1868). Basic necessitates for tourists were bought in Zanzibar and brought to the mainland (Swayne, 1868; Hore, 1892).

The only means of transport in the mainland was on foot and/or ox back (Swayne, 1868; Hore, 1892). However, there were indigenous African porters, known popularly as *wapagazi* in Swahili who were specialized in carrying tourist's luggage for money as well as the tourists themselves whenever necessary (Muhammedi, 1971). The *wapagazi* were under an organized group which was accountable to the local chiefs/rulers of a particular area (Muhammedi, 1971). These people could be viewed as the ancient tour operators and guides in Tanzania. Alongside the wapagazi, the explorers were accompanied by solders hired from the Zanzibar/Coastal areas, where the Zanzibar sultan had influence (Hore 1892, Swayne, 1868, Coupland 1968). Key problems at that time were transport, diseases such as malaria and smallpox, insecurity and illiteracy among the locals (Coupland, 1968). To date some of these problems

[4] The study does not assert that these are tourists, rather comparable to modern tourists based on their purpose of visits and their practices. The only thing disqualifying them from being tourists as per WTO definition is that most of them stayed for more than one consecutive year. In either way, modern inbound tourism activities in Tanzania such as mountain climbing date back to the period of their coming.

such as malaria for coast regions and transport problems especially in western regions still persist and can discourage tourists from coming.

These explorers who featured in the 19th century were not the earliest 19th century foreigners to visit East Africa, but were comparatively similar to the modern tourists. The earliest 19th century visitors to East Africa were the Arabs (Jaffe, 1988).Of-course the Arabs came much earlier before the 19th century and between the 15th and the 17th centuries had struggled against the Portuguese over control of the East African coast (Oliver and Gervase, 1963). But the Arabs, who preceded the explorers, came not to see the natural wonders of the land and the people but rather to plunder elephant tusks, slaves and other precious resources. Their engagement in conflicts may not qualify them to be compared to modern tourists[5].

After the formal and effective colonization of Africa, tourism activities began to be handled and formally organized by the colonial governments. In Tanzania (at that time Tanganyika) the Germans (1885-1918) and later the British (1919-1961) took over the administration, including tourism issues. In Zanzibar (an independent state before the 1964 union with Tanganyika), the Oman Sultanate controlled the Island from the 1650s up to 1890, when Zanzibar was put under British protectorate (Coupland, 1968). From 1913 the island came under the British governing system whereby Governors were appointed to rule the Island until 1963, when independence was granted by the British (Ward and White, 1971). Later in 1964, indigenous

[5] The comparability here is in terms of purpose of visit and practices. Unlike the Arabs and the Portuguese the explorers never engaged themselves in wars with a view of dominating the indigenous or of being remunerated within the visited places. Their war engagement was only for defensive purpose.

Africans took over by force by overthrowing the existing leadership, which was predominantly Arabic and serving Arab interests (Ward and White, 1971).

The formal tourist activities could be credited to the German colonialists, who were the first to formally establish game reserves in Tanganyika; they were followed by the British who took over after World War1. Chachage (1998) as quoted by Kulindwa et al. (2001) gives an account of the formal establishment of tourism under the two colonial masters as follows:

"In 1890s German colonial rule established game reserves and sanctuaries beginning with those in Moshi and Kilimanjaro districts in 1891; later on in 1896 Rufiji, which is now part of Selous game reserve, and west Kilimanjaro were established. By 1908 there were 8 game reserves in Tanganyika. After the British took over, the game reserves were extended to 13 in the early 1920's. The Game Preservation Ordinance of 1921 confirmed game reserves such as Selous, Ngorongoro and Serengeti. Lake Rukwa and Usambara were later included in 1933, while Serengeti and Ngorongoro were further extended in 1936. By 1939 game reserves in the Southern highlands and Tabora were also established."

Tourism activities under both colonial governments were established to basically serve the rich Europeans who came to visit Africa (Chenjah, 1998; Kulindwa et al., 2001). This traditional market continued to be dominant during and after independence, and continues to be dominant. Africans had neither the resources nor the interest to tour the parks or sites; this indifference continues to this day. Africans' lack of interest in visiting tourist sites could have been partly attributed to the colonialists themselves, as

they restricted them from using tourists' assets. For example, under German rule no African was allowed to hunt in the game reserves (Kulindwa et al., 2001).

During British rule an attempt was made in 1938 to form the East African Publicity Association (EAPA) for promoting tourism in East Africa (Ouma, 1969). However, due to the World War II, the organ's activities were not successful (Ouma, 1969).

In 1947, two years after the end of the world war, an inter-territorial conference was held in Nairobi to discuss ways to improve tourism in East Africa; which led to the establishment of the East African Tourist Travel Association known as EATTA (Ouma, 1969). EATTA was given a ten-year review and in 1958, its activities were reviewed; and deemed as successful. At this time movements for national independence began to flourish. Besides some member countries (Uganda and Tanganyika) felt that EATTA was over-promoting Kenya at the expense of the other two (Ouma, 1969). This feeling led Uganda to establish her own promotion organ in 1956, known as the Tourism Advisory Board, followed by Tanganyika in 1962, which established the National Tourist Board of Tanganyika (Ouma, 1969). The individual countries' organs operated alongside EATTA until 1965 when the opinion of the majority was that each country should carry out its own promotion activities (Ouma, 1969).

In the period after independence, the general public could not perceive the tourism industry as an important sector (Chenjah, 1998). Its promotion was viewed as an act of embracing colonial interests in the country (Chenjah, 1998). This argument might be supported by the fact that, in the early days of independence, nationalistic feelings

were very high (Nyangwine and Maluka,2008) and the roaming back of Europeans could have been perceived as an attempt to recolonize the country. It was also felt that tourism was an economically less viable industry because it was associated with import leakages. This view may be verified by examining the reaction of the Ministry of Information and Tourism (MIT), now Ministry of Natural Resources and Tourism (MNRT) to the debate among UDSM students back in 1972 regarding the viability of the tourism industry. According to MIT (1972), some students were arguing that the industry was economically unviable. Giving a statistical account, MIT (1972) argued that the students' arguments were more theoretical and unfounded in the Tanzanian case; because the industry was more profitable even after accounting for import leakages.

Despite such views by some academicians and the general public, it was during independence when most of the game reserves were transformed into national parks. As a matter of fact, the first president of Tanzania (at that time Tanganyika) had long ago recognized the importance of tourism for Tanzania and Africa in general (TANAPA, 2002). A speech delivered by the president in September 1961, at a symposium on the conservation of nature and natural resources, came to be known as the Arusha Declaration; in which he stressed the importance of protecting wildlife for the future of our wellbeing (TANAPA, 2002). As a result the government enhanced efforts to protect wildlife and other tourist attractions.

According to TANAPA (2002), between 1960 and 1980, the following national parks attained that status of being national parks having been game reserves: Lake Manyara National Park (1960), Mikumi National Park (1964), Ruaha National Park (1964),

Arusha National Park (1967), Gombe National Park (1968), Tarangire National Park (1970), Kilimanjaro National Park (1973), Katavi National Park (1974) and Rubondo National Park (1977). Two more were established in 1980 and 1992. These are Mahale National Park (1980) and Udzungwa National Park (1992). Between 1992 and 2008 three more parks were established. These are Mkomazi National Park, Kitulo Plateau National Park and Saadani National Park. The total number of National Parks is now 15. These National Parks form the core of the tourism industry in Tanzania.

During the first decade after independence Tanzanian adopted socialist policies which led to the nationalization of all major means of production, including tourism. The National Tourist Cooperation was formed to monitor all the tourist activities. However, from the 1970's, few years after nationalization, up to the mid-1980s the tourism sector did not perform well. There was stagnant growth in both the number of arrivals and revenue.

After trade liberalization in 1986, tourism was viewed as an important sector. This led to the establishment of the National Policy on Tourism in 1992 and the enactment of the Tanzania Tourist Board in 1993 (Chenjah, 1998, MNRT, 1999).Since then, TTB, has been responsible for the promotion of tourism in the country. TTB is an organ, among many others, within MNRT, which is in charge of all the tourism activities, including the maintenance and development of tourism assets (MNRT, 2002).

2.4 The Performance of Tourism in Tanzania

2.4.1 The Evolution of Tourist Number of Arrivals and Tourism Revenue

The performance of tourism in Tanzania can be assessed in three main periods based

on both the number of arrivals and tourism revenue. First is the period after independence (1961) up to few years after the start of the socialist era(early 1970's).Second is the period between the socialist era (early 1970's) up to the adoption of the free market economy/trade liberalization(1984-86)[6]. Third is the period after the adoption of trade liberalization until now (see figures 2.1, 2.2)

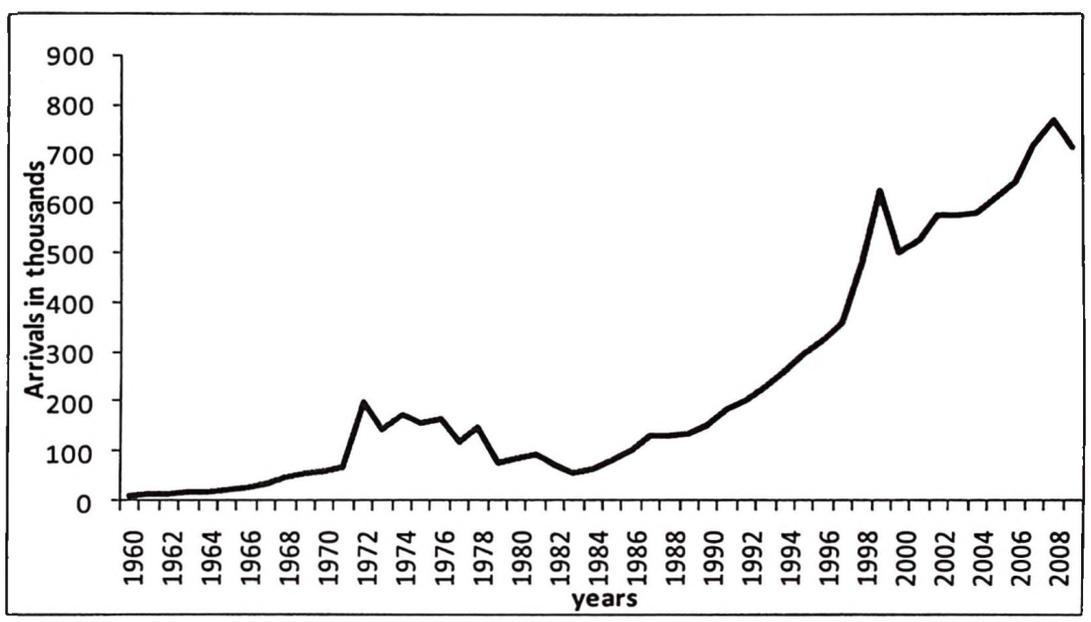


Figure 2.1: Trend in the total number of arrivals in Tanzania: 1960-2009

Source: Own drawing based on statistics from MNRT

Figure 2.1 indicates the trend in the number of tourist arrivals from 1960 to 2009. The figure shows that the number of arrivals kept on increasing from 1960 to the early 1970's, before the implementation of socialism policies and later on from 1986, when the country introduced trade liberalization policies until now. A similar situation as regards tourism revenue is depicted in figures 2.2

[6] Trade liberalization policies in Tanzania were partly adopted in 1984 by the first phase government and fully adopted in 1985/86 during the second phase government.

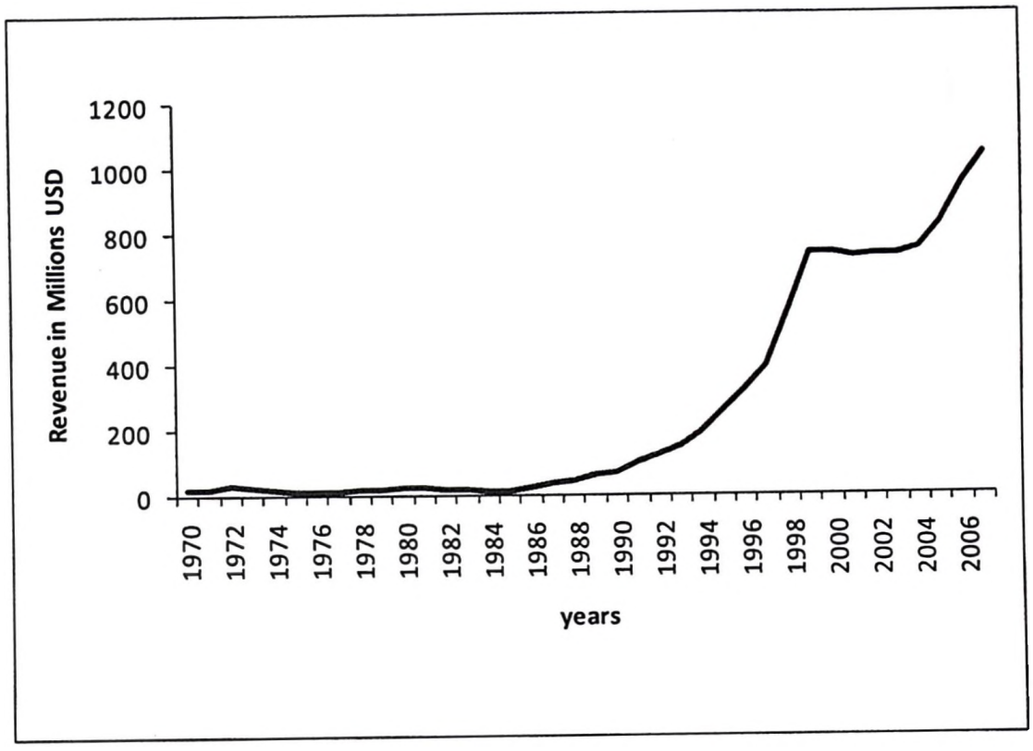


Figure 2.2: Tanzania's Tourism Revenue in millions USD: 1970-2007

Source: Own drawing based on the statistics from the MNRT.

Figure 2.2 indicates the trend in tourism revenue from 1970 to 2007. The figure indicates that tourism revenue did not do well from 1970's up to 1984. The revenue started rising in 1986 until now. The trend is similar to that of the number of tourist arrivals depicted in figure 2.1.

The evolution of tourism revenue and the number of tourist arrivals suggests that the growth in tourism came along with trade liberalization. However, formal analysis of the increase in tourists' arrivals does not yield conclusive evidence of the influence of trade liberalization on the coming of tourists.[7]

[7].A time series regression from 1970 up 2007 involving a dummy for 1984/1985/1986 and trade openness as a proxy for trade liberalization does not show a significant influence of trade liberalization on tourist arrivals. The details of these analyses are given in appendix 2.1. These preliminary analyses

The country's tourism performance has over the years been relying on the traditional markets of Europe and America. Most of the countries sending tourists to Tanzania have been in Europe. Others are the United States of America, Canada and some African countries, particularly South Africa (Table 2.1).

Table 2.1: The Top 20 source markets for Tanzania: 2003-2007

Country	2003	2004	2005	2006	2007
kenya	1	1	1	1	1
United States	3	4	4	2	2
United Kingdom	2	2	2	3	3
Italy	6	3	3	4	4
Zambia	14	6	5	6	5
Uganda	5	7	7	5	6
South Africa	4	5	6	7	7
Germany	9	9	10	8	8
France	8	8	8	9	9
Malawi	11	10	9	10	10
Netherlands	10	12	13	11	11
Canada	15	14	15	14	12
Rwanda	12	20	12	12	13
India	7	11	11	13	14
Spain	17	13	14	17	15
Burundi	13	30	19	18	16
Australia	16	16	17	15	17
Sweden	18	18	18	19	18
Switzerland	19	19	22	16	19

Source: **Own drawing based on the arrivals statistics from Ministry of Natural Resources and tourism (MNRT)**

Table 2.1 shows top 20 source markets for Tanzania from 2003 up to 2007. It is worth noting that a good number of visitors from some African countries, such as Kenya, which appear to be in the lead are not citizens of those countries (TTSS, 2001). According to the survey by TTSS (2001), 45 percent of the interviewees from Kenya

---

were done because in the final analysis provided in chapter three it is was not possible to include these dummies. much as the analysis involved panel data from 1995 up to 2007.

were not Kenyan nationals. In fact some of them could be foreigners who were passing through Kenya on their way to Tanzania, or who are living there as foreign citizens. When a comparison of regions is made, Africa leads in sending tourists to Tanzania, followed by Europe, Asia, North America, Middle East and South America (Figure 2.3)

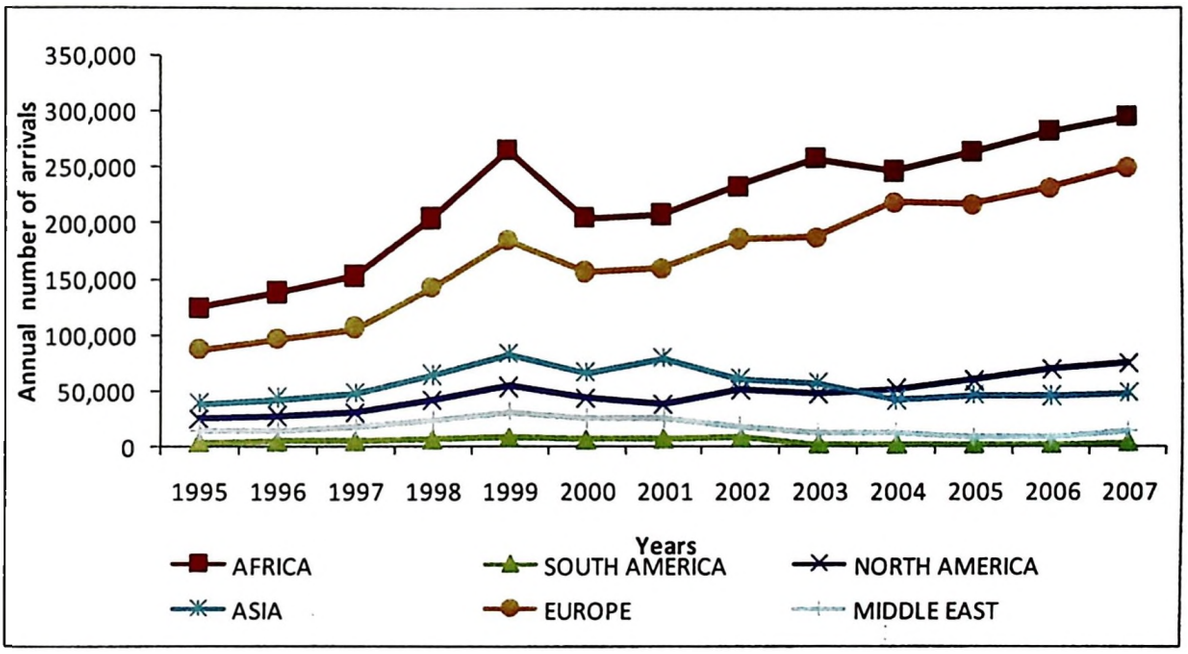


Figure 2.3. Arrivals in Tanzania by region: 1995-2007

Source: Own drawing based on statistics from MNRT

2.4.2 The Evolution of Tourist Per Capita Spending

As far as tourist per capita spending is concerned there are no long time series data, unlike with tourism revenue and number of tourist arrivals. The most recent and more detailed statistics start from 2001 and are provided by TTSS. Figure 2.4 provides the trend of tourist spending from 2001 up to 2008 based on TTSS statistics.

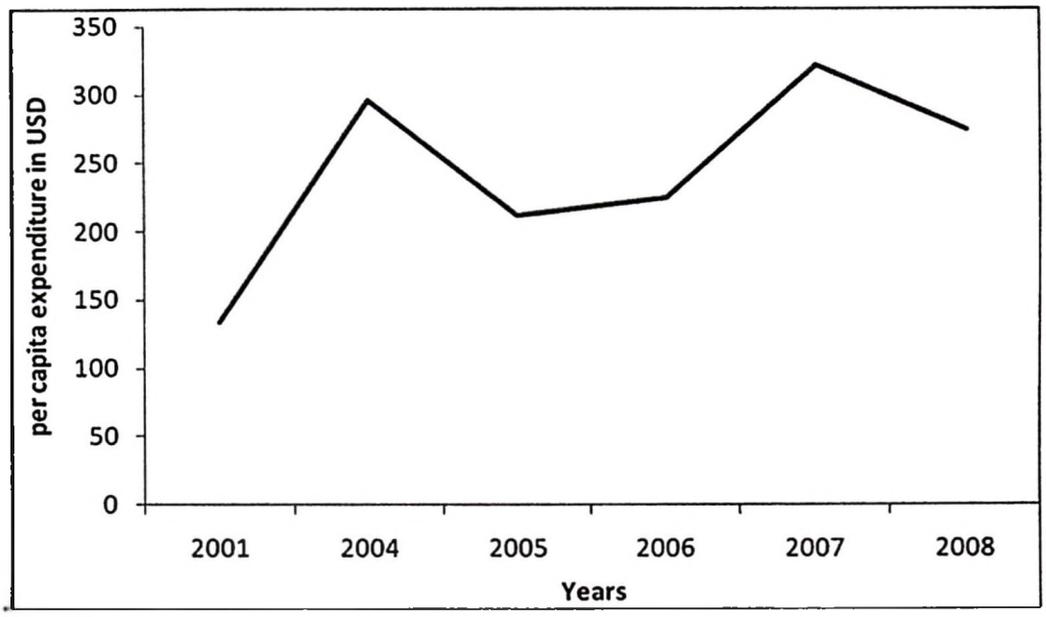


Figure 2.4: Tourist per capita spending in Tanzania: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.4 indicates tourist per capita spending as surveyed by TTSS in the years 2001 to 2008. The figure shows that with the exception of the year 2004 and the year 2008 tourist spending has been increasing over time. While for the year 2008 global financial crisis could have been a reason for a decline in tourist spending, the reason for the sudden rise in tourist spending for the year 2004 is not clear. Owing to the lack of observations in the years 2002 and 2003 there is nothing substantial that can be inferred for the scenario in the year 2004.The message from figure 2.4 is that tourist spending is doing well and the challenges ahead lie in the sustainability and improvement of tourist spending in Tanzania. When various regions are examined, the trend appears to have been stable over time. Figure 2.5 provides the trend of tourist spending across the regions

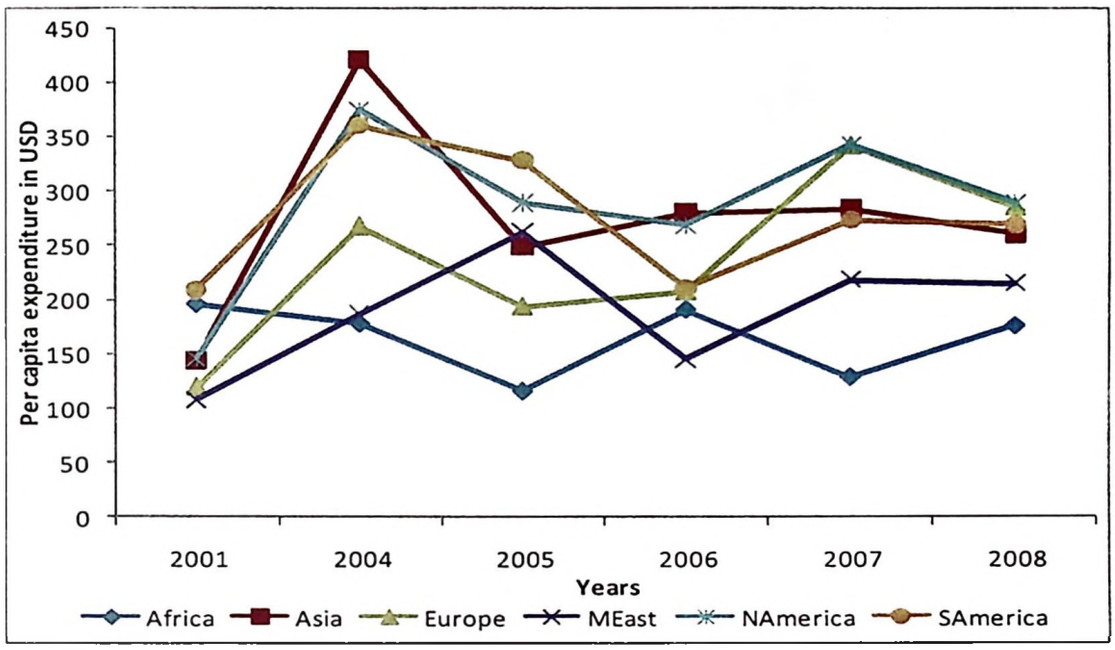


Figure 2.5: Tourist per capita spending in Tanzania by region: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.5 indicates that, in general, tourist per capita spending in each region is stable across the years except for the 2004. With the exceptional of Africa and the Middle East, the rest of the regions had a sharp increase in spending for the year 2004. Apparently no reason can be given for such a scenario because no surveys were conducted for the years 2002 and 2003.It is interesting to note that for nearly all the years, Asia, South America and North America experienced above average spending of USD 243, while the spending in Africa, the Middle East and Europe was below this average.

It would also be important to assess the status of tourist spending against key attributes affecting tourist spending. These include demographic characteristics, such as age, trip-related characteristics such as travel arrangements, purpose of visit and travel party size, as well as destination attributes such as season of travel.

(i) Status of tourists per capita spending against age of tourists

Figure 2.6 indicates average expenditure by tourists across age groups.

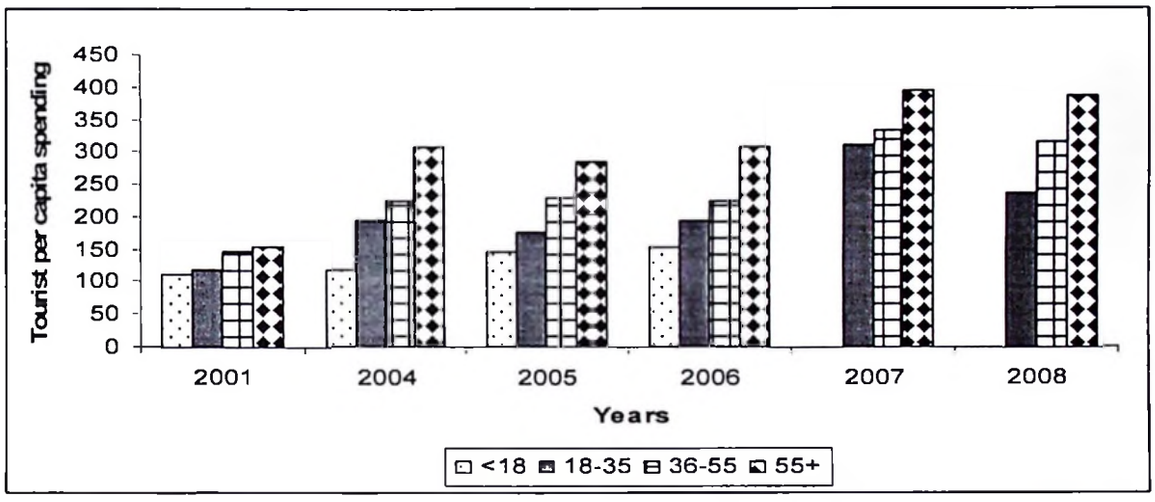


Figure 2.6: Tourist per capita spending in Tanzania by age group: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that tourist spending increases with age.

(ii) Status of tourist spending against travel arrangements

Figure 2.7 indicates average tourist spending between package tourist and non-package tourist from 2001 up to 2008.

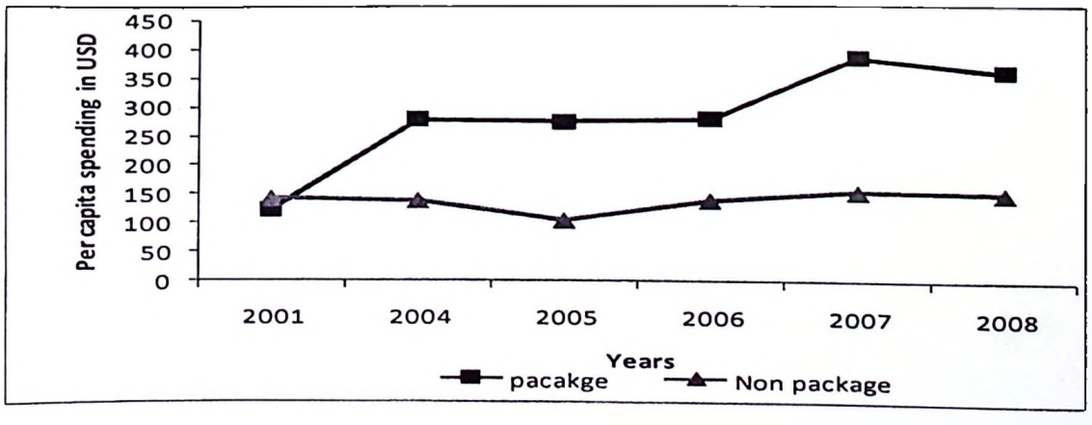


Figure 2.7: Tourist per capita spending across travel arrangements: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure shows that, on average, a tourist on package tour spends more than a tourist on non-package tour.

(iii) Status of a tourist spending against purpose of visit

Figure 2.8 shows average spending of tourists by purpose of visits.

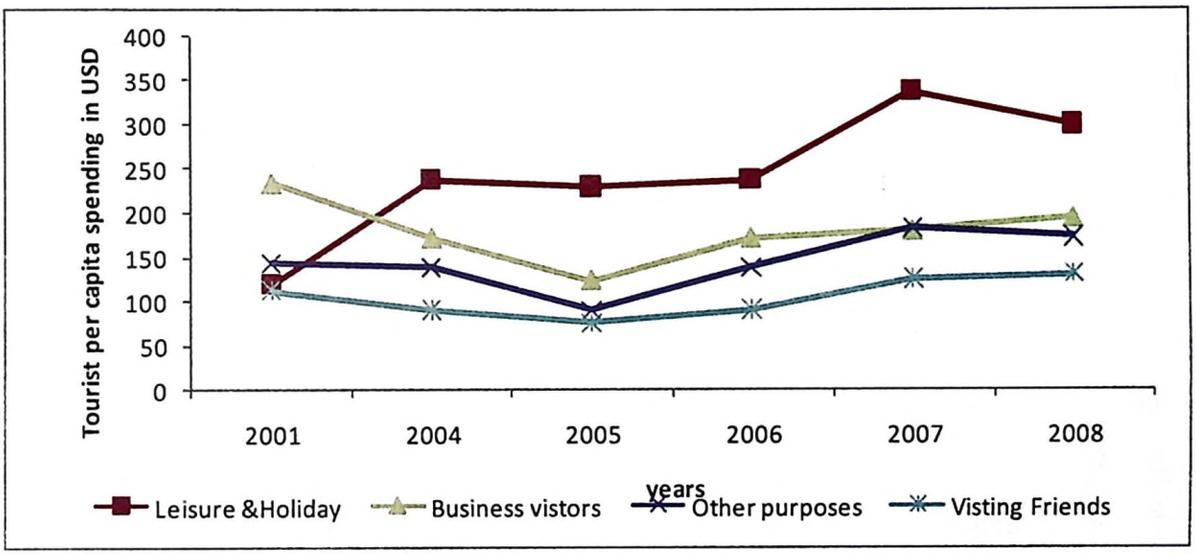


Figure 2.8: Tourist per capita spending against purpose of visit: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that, on average, tourists on leisure and holiday spend more than other categories of tourists. They are followed by tourists on business visits, then tourists on other purposes and lastly tourists visiting friends and relatives.

(iv) Status of tourist spending against travel party size

Figure 2.9 shows average spending of tourists by travel party size.

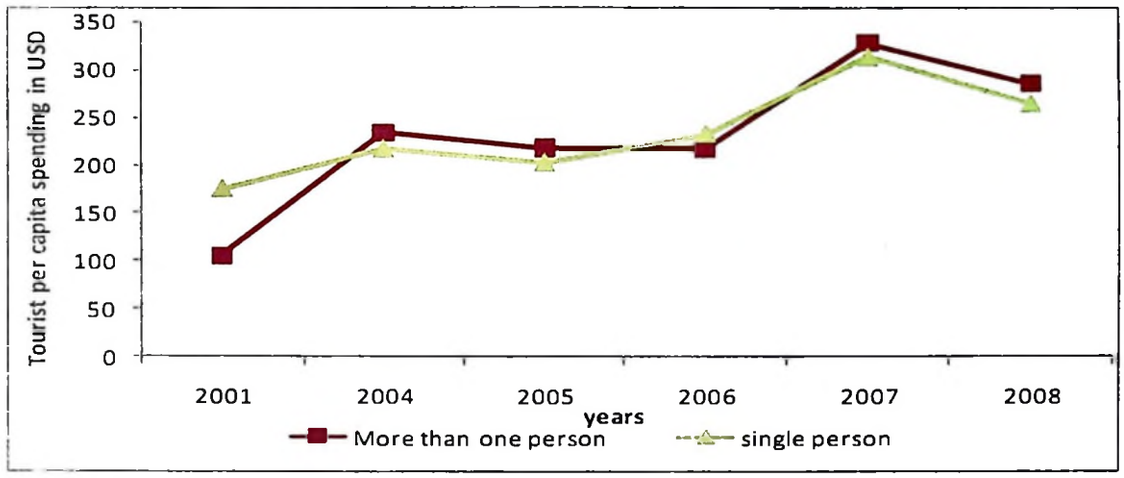


Figure 2.9: Tourist per capita spending against travel party size: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.9 indicates average spending by tourists against travel party size. The figure suggests that there is no significant difference in tourist spending across travel party size.

(v) Tourist spending against presence or absence of a child

Figure 2.10 compares average spending by a tourist accompanied by child/children against the one not accompanied by a child.

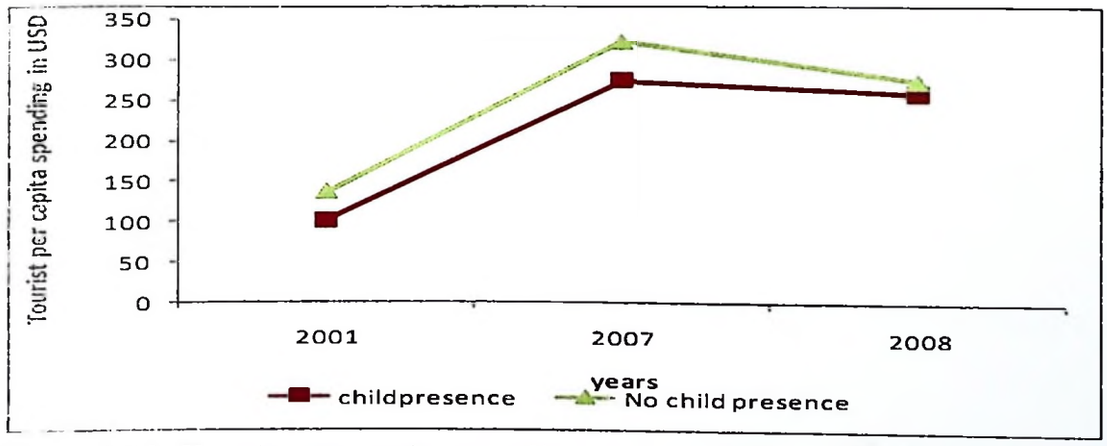


Figure 2.10: Tourist per capita spending against child presence: 2001, 2007, 2008

Source: Own drawing based on statistics from TTSS (2001, 2007, 2008)

Figure 2.10 suggests that a tourist unaccompanied by a child spend more that the one accompanied by a child. The figure further shows that spending is higher in 2007 than in 2001 and in 2008. Given the missing values in the years 2002, 2003, 2004, and 2006 nothing can be said about the scenario in the year 2007.

(vi) Status of tourist spending against seasons of travel

Figure 2.11 indicates the status of tourist spending across seasons of travel (high peak season against mini- peak season).

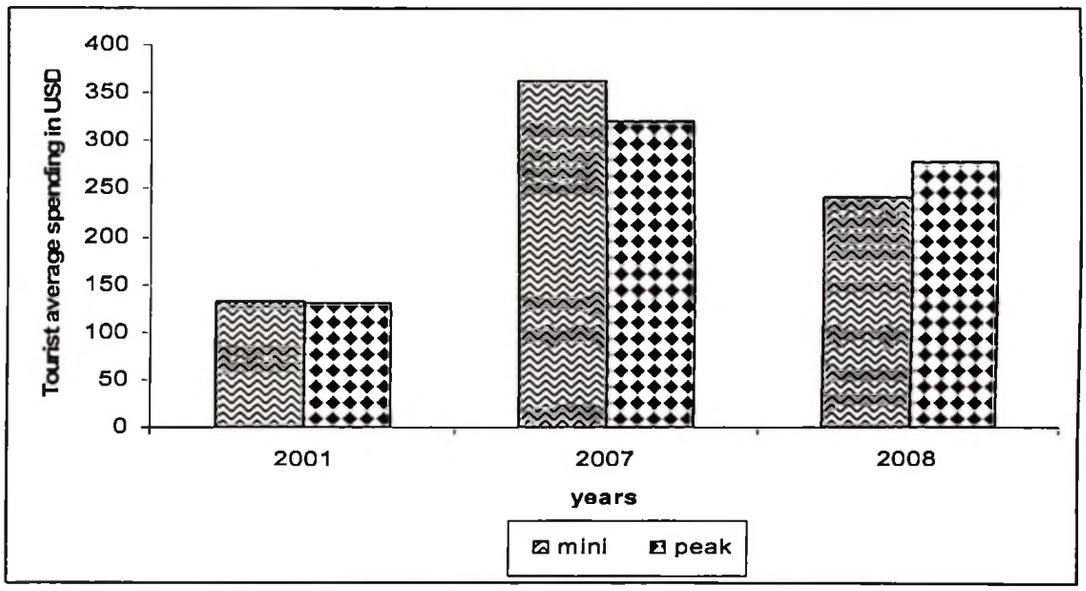


Figure 2.11: Tourist per capita spending against seasons of travel: 2001, 2007, 2008

Source: Own drawing based on statistics from TTSS (2001, 2007, 2008)

Figure 2.11 suggests that tourist per capita spending is less in the peak season than in non-peak season except for the 2008.

2.4.3 The Evolution of Tourist Length of Stay in Tanzania

The average length of tourist stay, according to MNRT (2002), was 8 days from 1995 to 2007. But according to TTSS (2006), the average length of stay is 12 days. Figure 2.12 provides the evolution of tourist length of stay in Tanzania based on the figures from both MNRT (2002) and TTSS (2001-2008).

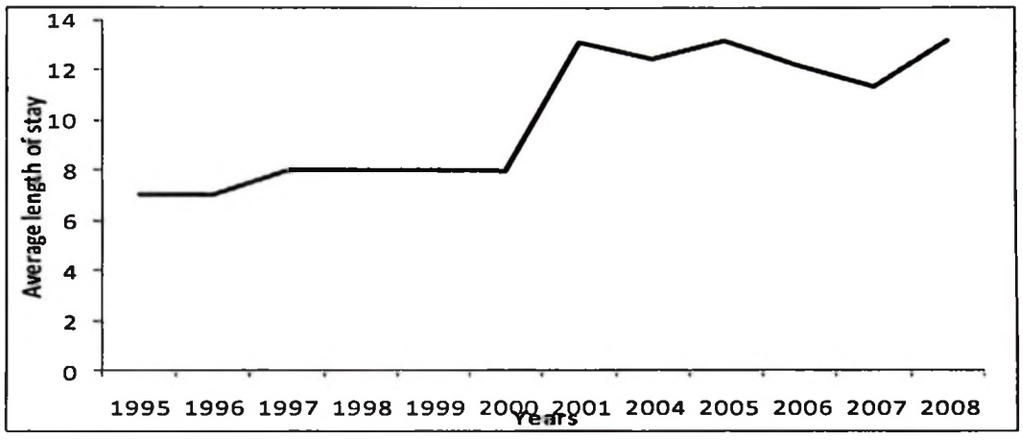


Figure 2.12: Tourist length of stay 1995-2008

Source: Own drawing based on statistics from MNRT (2002) &TTSS (2001-2008)

Figure 2.12 indicates that the average length of stay of tourists has remained stable at 13 days, after initial rise in 2000.The figure shows clearly that according to the data collected by MNRT, the tourist length of stay was fairly constant from 1995-2000. However, from 2001, data from TTSS recorded a much higher length of stay. Given the fact that there is no significant reason to believe that tourists changed their appetite to stay longer in Tanzania so rapidly, it can fairly be urged that this sharp rise from 2001 can be attributed by improvement in data collection. When the length of tourist stay is examined across the regions, it depicts more or less the same trend. Figure 2.13, provides the trend of tourists' length of stay across the regions.

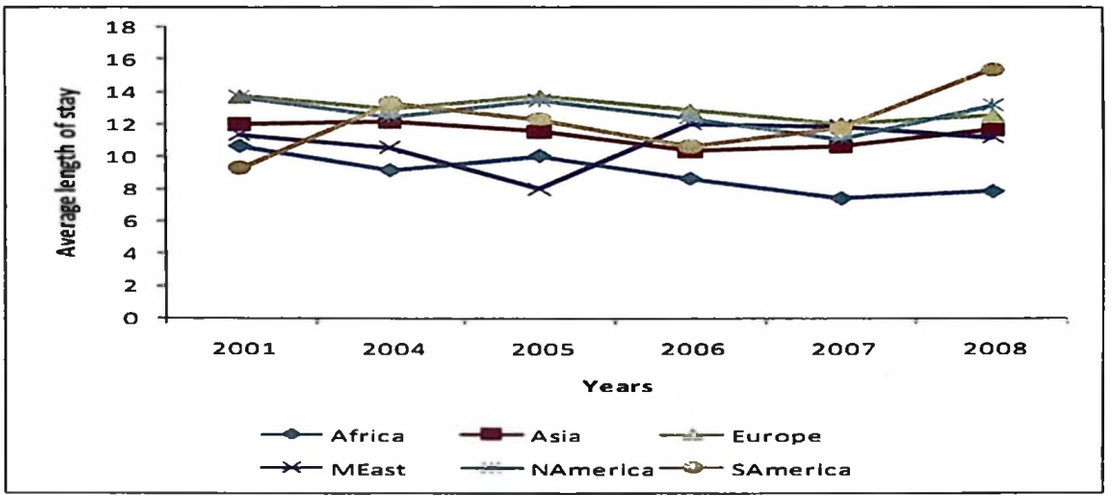


Figure 2.13: Tourist length of stay in Tanzania by region: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.13 indicates the trend in the tourist length of stay across the regions from 2001 to 2008. The figure indicates that, in general, tourist length of stay has remained stable at between 10 and 14 days. The figure also indicates that Europe and North America have an edge over other regions, while Africa has the least average length of stay among all the regions.

As in the case of a tourist spending, it may also be important to examine the status of tourist length of stay across tourists' demographic characteristics, trip-related characteristics and destination attributes.

(i) Status of tourist length of stay against age of tourists

Figure 2.14 depicts trends of a tourist length of stay across age groups over time.

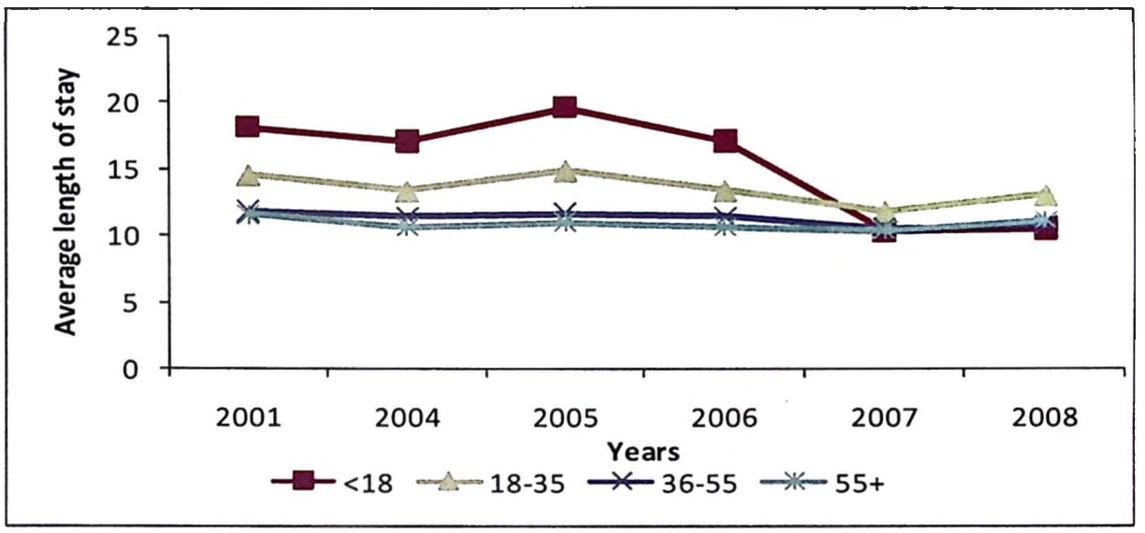


Figure 2.14: Tourist length of stay in Tanzania by Age group: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that tourist length of stay in Tanzania declines with his/her age group.

(ii) Status of tourist length of stay against travel arrangements

Figure 2.15 indicates the status of tourist length of stay against travel arrangements.

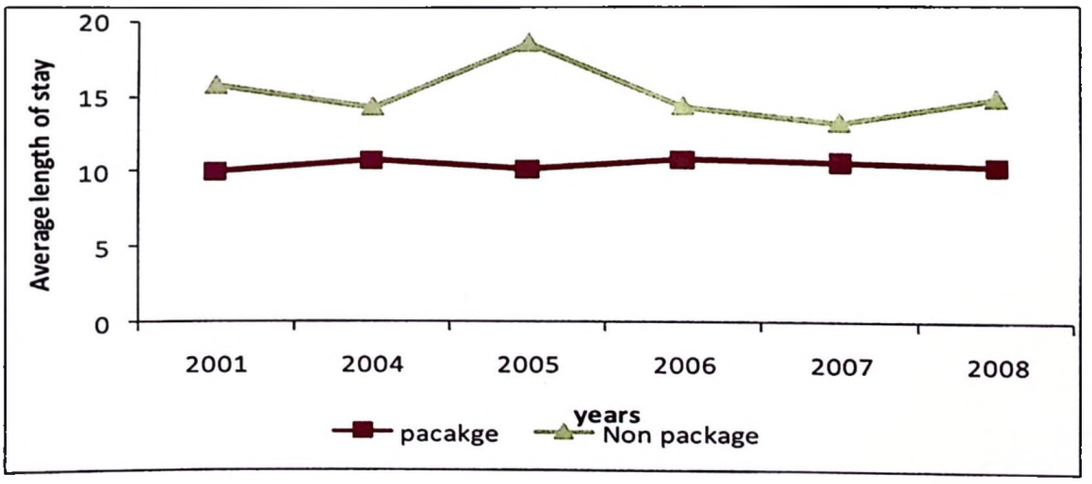


Figure 2.15: Tourist length of stay in Tanzania by Travel arrangements: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that tourists on non-package tour stay longer than those of package tour.

(iii) Status of a tourist length of stay against purpose of visit

Figure 2.16 provides the trend of tourist length of stay across tourist purpose of visit.

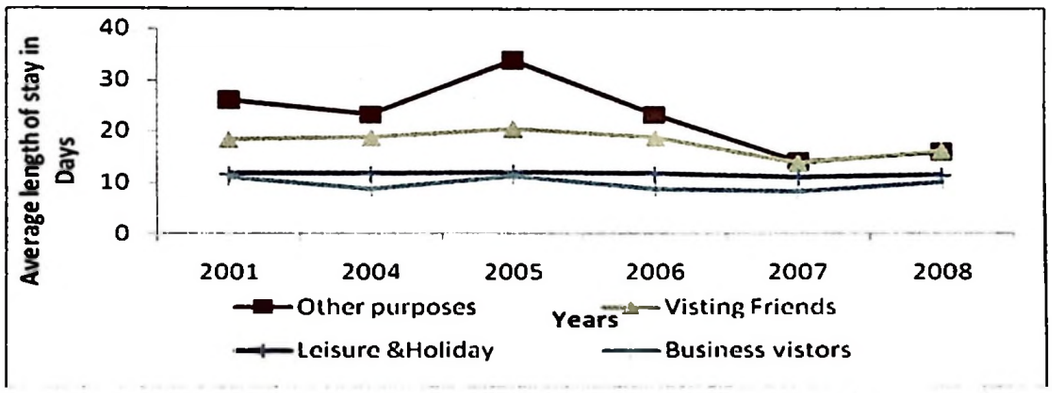


Figure 2.16: Tourist length of stay in Tanzania by purpose of Visit: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure indicates that visitors coming for other purpose stay longer than others, followed by those visiting friends, then those on leisure and holiday and finally business visitors.

(iv) Status of a tourist length of stay against the travel party size

Figure 2.17 provides the status of tourist length of stay against travel party size accompanying a tourist.

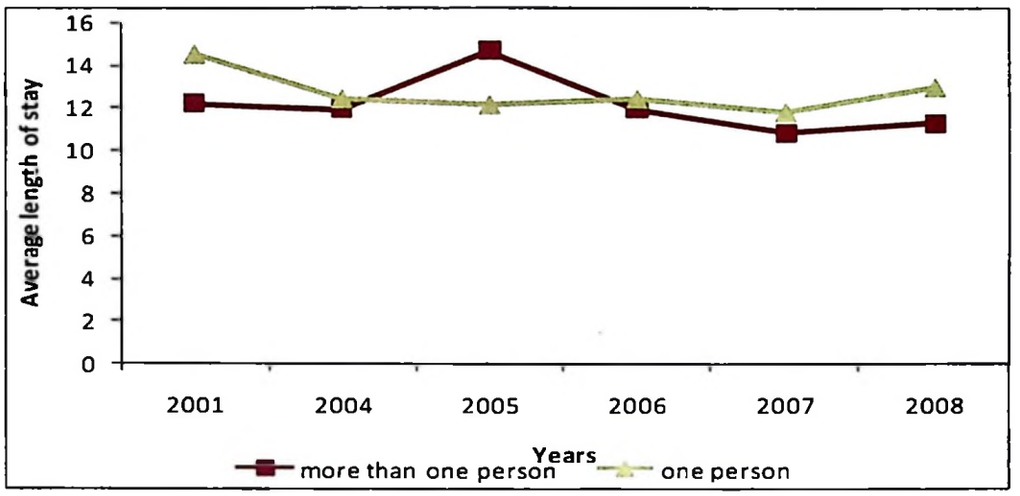


Figure 2.17: Tourist length of stay in Tanzania by travel party size: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

With the exception of the year 2005, Figure 2.17 indicates that, in general, tourists in small travel parties stay longer than those in big travel parties.

(v) Status of tourist length of stay against seasons of travel

Figure 2.18 indicates tourist length of stay against high peak season and mini peak season.

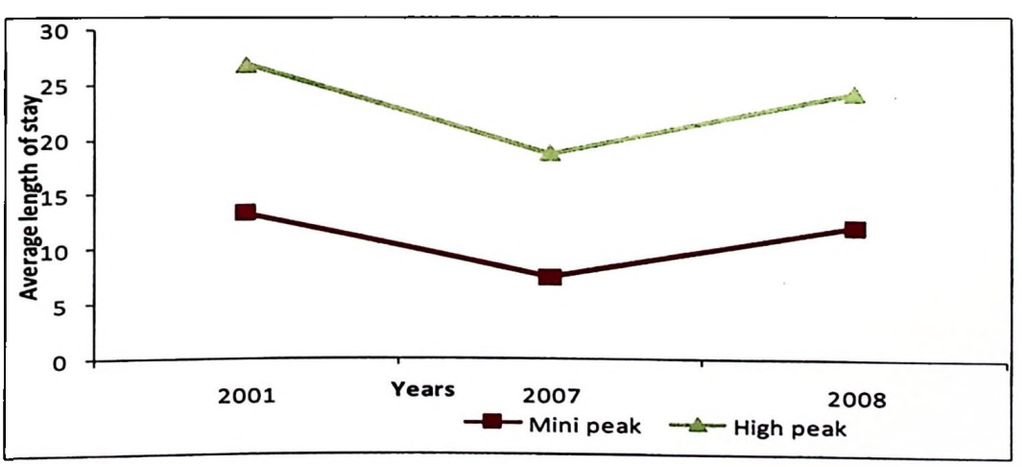


Figure 2.18: Tourist length of stay against season of travel: 2001, 2007, and 2008

Source: Own drawing based on statistics from TTSS (2001, 2007, 2008)

Figure 2.18 suggests that a tourist visiting during the high peak season tends to stay longer than a tourist visiting during the mini-peak season. It is interesting to note that for the year 2007 tourists in both mini and peak seasons stayed for a short time as compared to the year 2001 and 2008. Given the fact that observations in years 2002, 2003, 2004, 2005 and 2006 are missing, there is nothing really that can be concluded as regards the decline in length of stay for the year 2007.

2.4.4 The Status of Package Travel

According to TTSS (2001), Tanzania receives more tourists on package tours than those non-package tours. TTSS (2001), states that this phenomenon is common worldwide. Figure 2.22 shows the percentage of tourists on package tours who visited Tanzania from 2001 to 2008.

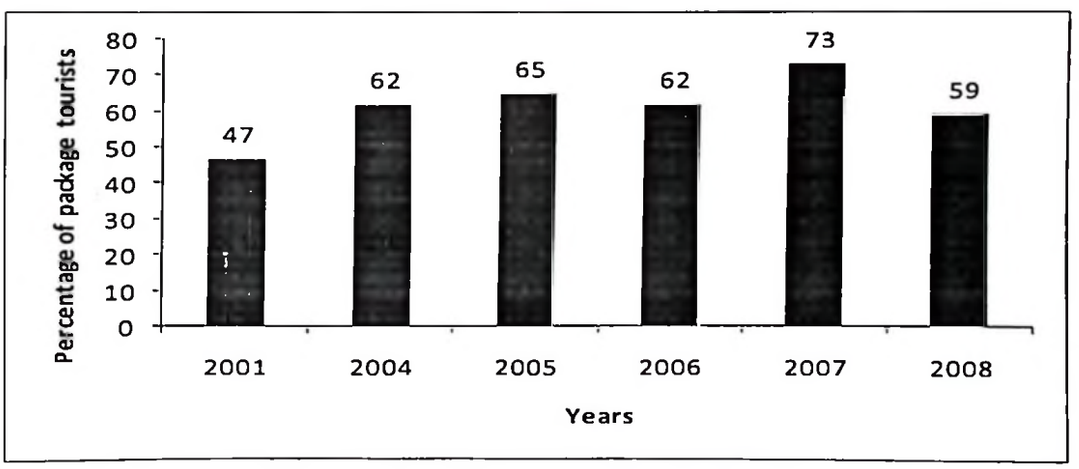


Figure 2.19 Percentage of package tourist: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.19 indicates that, in general, Tanzania receives more tourists on package tours than those on non-package tours. The trend is stable at an average of 60% per annum. This stability can also be observed across all the regions (Figure 2.20)

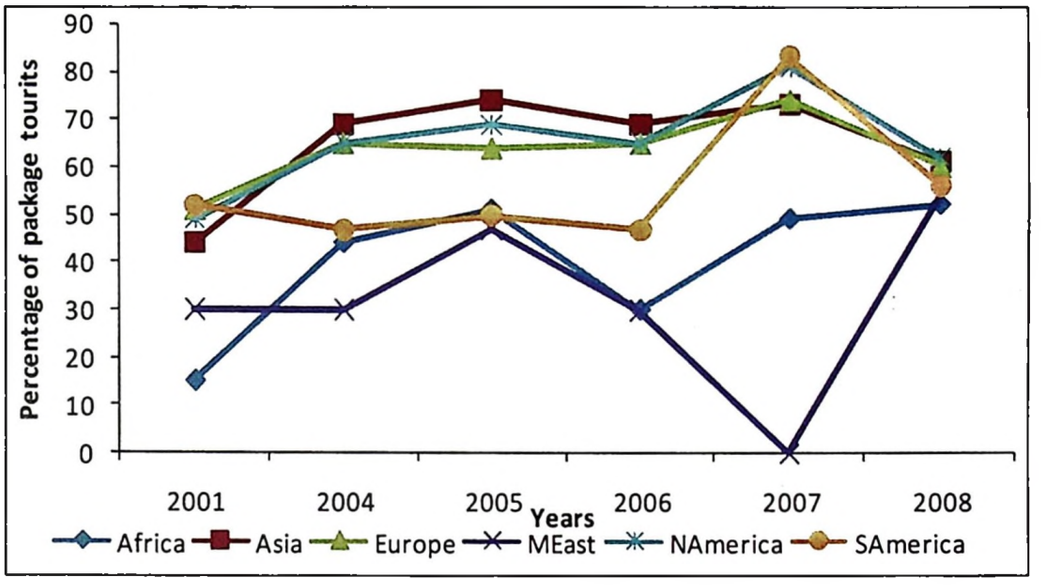


Figure 2.20: Percentage of package tourist across the regions: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.20 shows that the percentage of tourist on package tours was fairly stable over time for all the regions except for the Middle East in the year 2007. This exceptional scenario could have to do more with problems in data collection. Further, the figure indicates that tourists from Asia, North America and Europe prefer package tours whereas the opposite seems to be the case for tourists from South America, Africa and the Middle

Again it may be important to assess the status of package travel against the key demographic and trip-related characteristics of tourists.

(i) Status of package travel against tourist length of stay

Figure 2.21 provides the status of tourist choice of package tour against length of stay in Tanzania for 2001, 2007 and 2008.

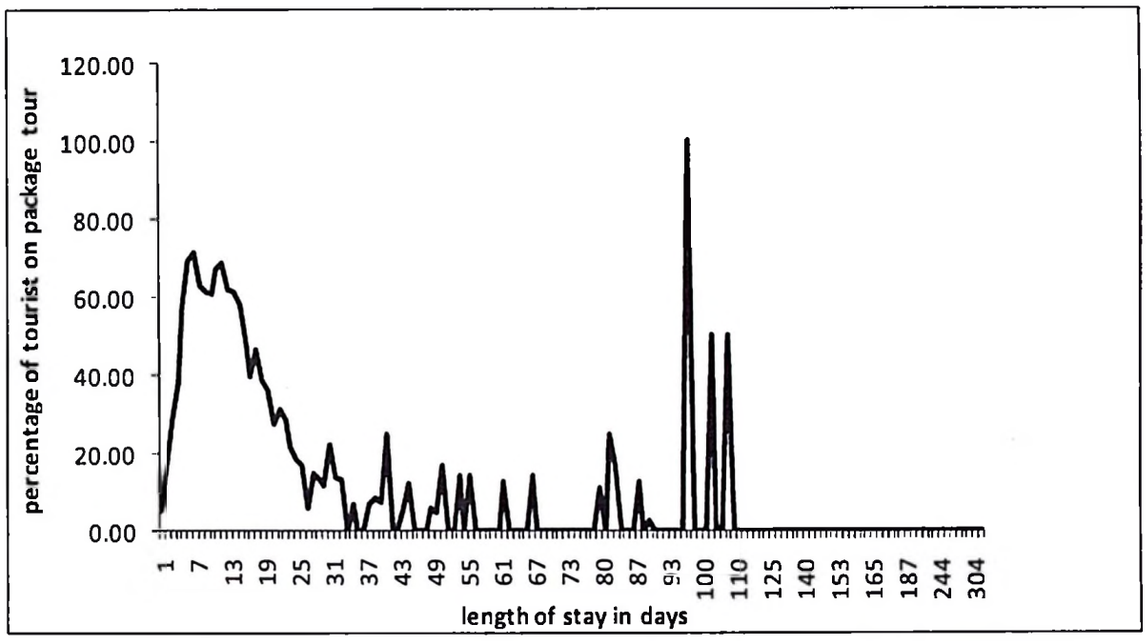


Figure 2.21: Percentage of package tourists against length of stay 2001, 2007, 2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.21 indicates that the percentage of tourist on package tours decline with length of stay. This suggests that, in general, a tourist staying longer is likely to choose non-package tour. It is interesting to note that for tourists staying for less than two weeks this does not hold. Such tourists could be visitors in pursuit of leisure and recreation as well as business visitors much as their stay is around two weeks (Figure 2.16).It is also interesting to note that for tourists staying for more than three months (110 days) it is almost certain that they will opt for non-package travel. Surprisingly there is more or less constant percentage of tourists on package tour staying between 31 days and 93 days as well as between 93 days and 110 days.

(ii) Package tour against age group

Figure 2.22 indicates percentage of package tourist against age group.

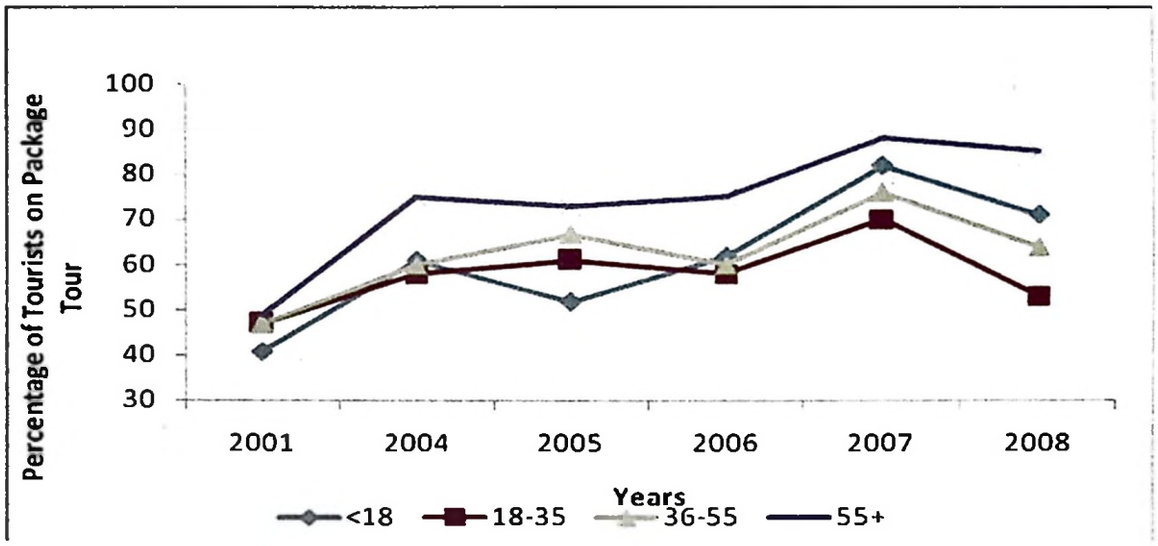


Figure 2.22: Percentage of package tourists against age group: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that old tourists tend to prefer package tour more than young tourists.

(iii) Percentage of tourist on package tours against purpose of visit

Figure 2.23 shows the percentage of tourists on package tours across purpose of visit.

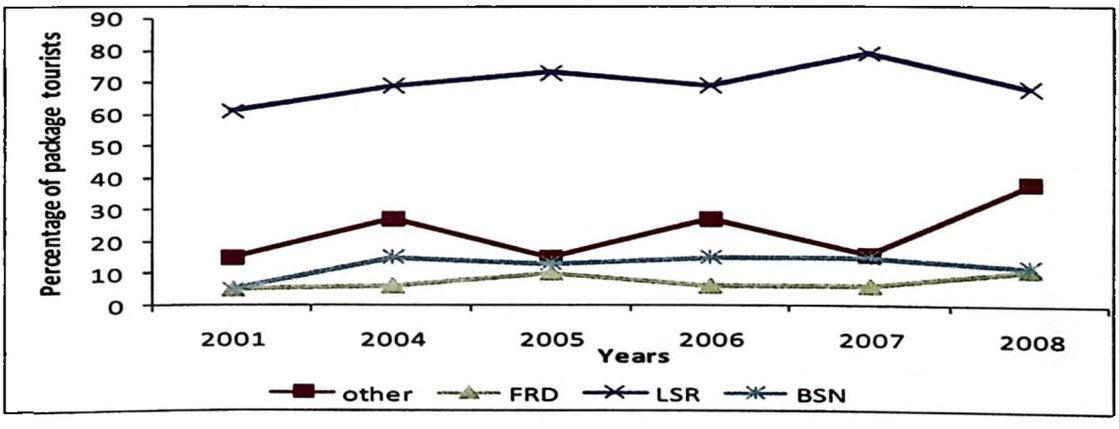


Figure 2.23: Percentage of package tourists against purpose of visits: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

The figure suggests that visitors on leisure and holiday prefer package tours than others.

(iv) Package travel against travel party size

Figure 2.24 shows the percentage of tourists on package tours against travel party size.

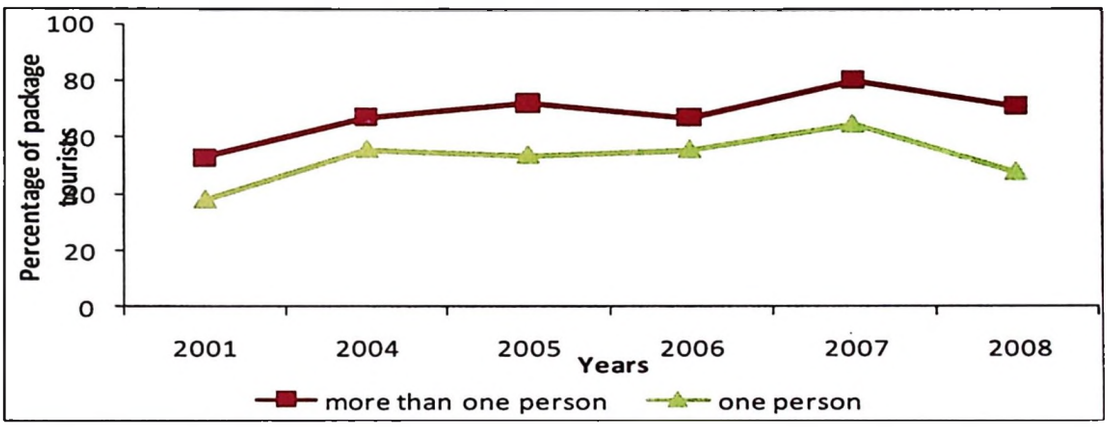


Figure 2.24: Percentage of package tourists against travel party size: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.24 indicates that tourists in large parties prefer package tours, whereas those in small parties do not.

(v) Percentage of tourists on package tours against child presence

Figure 2.25 provides a comparison of the percentage of tourists on package tours who are accompanied by child/children against those not accompanied by a child.

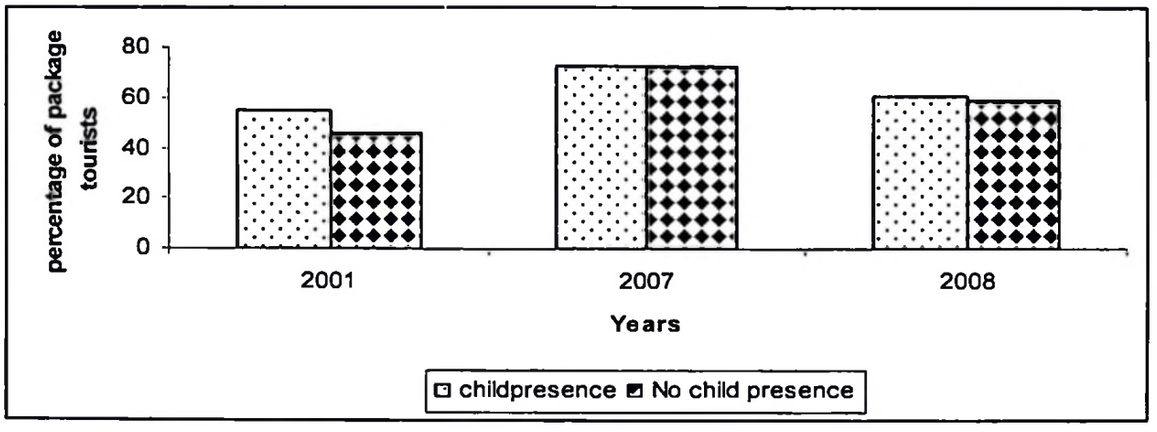


Figure 2.25: Percentage of package tourists against child presence: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.25 appears to suggest that there is no difference in the preference of package travel between tourist accompanied by a child and those not accompanied by a child except for the year 2001.

(vi) Percentage of tourists on package tours against the source of tourist travel information.

Figure 2.26 indicates the percentage of tourists on package tours across the source of travel information (word-of-mouth versus non-word-of-mouth information)

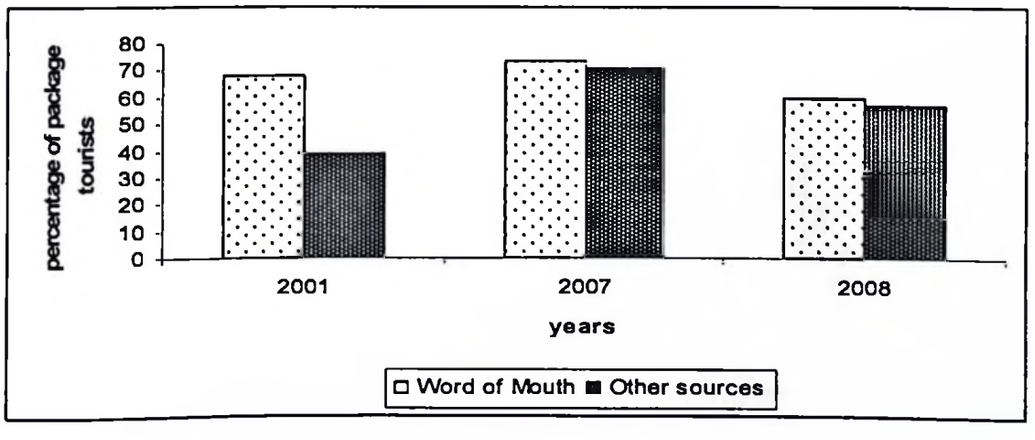


Figure 2.26: Percentage of package tourists versus information source: 2001-2008

Source: Own drawing based on statistics from TTSS (2001-2008)

Figure 2.26 indicates that tourists whose source of information is word-of-mouth prefer package tours more than tourists whose source of travel information is not word-of-mouth.

2.4.5 The Contribution of Tourism to the Tanzanian Economy

Tourism makes big contribution to Tanzania, by generating foreign exchange through tourism exports, by improving the general productivity of the country, both directly and indirectly, and by creating employment both directly and indirectly. Tourism also strengthens social and political ties with other countries as well as cultural interactions. However, for the purpose of this study only the economic benefits have been explored.

(a) Tourism's Contribution to Total Exports

Tourism's contribution to the export sector of Tanzania is remarkably significant. Its contribution increased from 5% in 1980 up to a staggering of 40% in 1995 and 1999 (Figure 2.27).

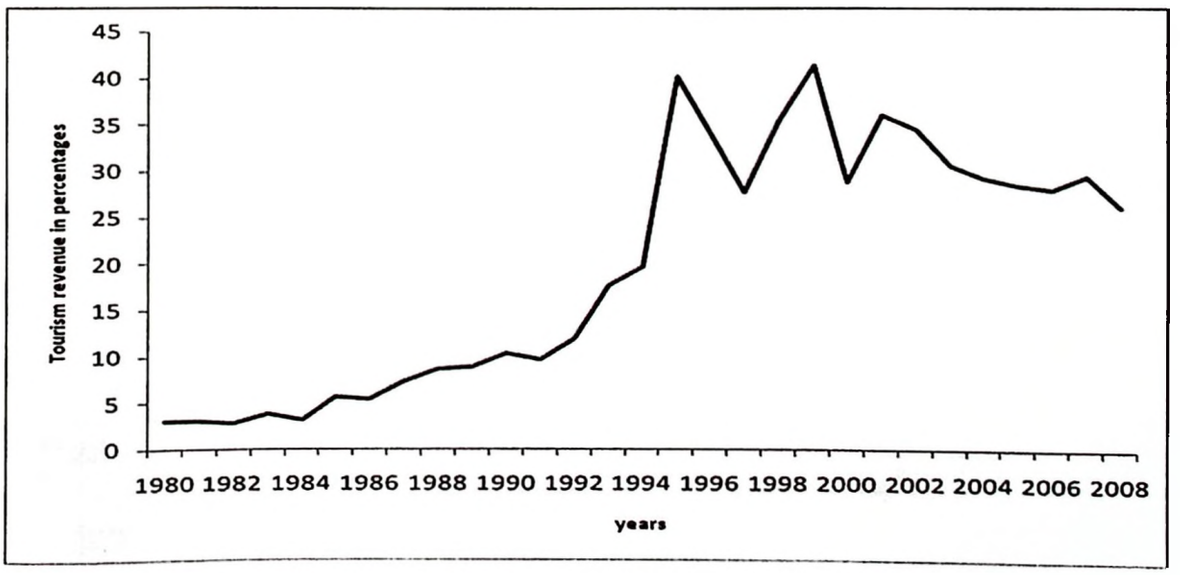


Figure 2.27: Tourism Revenue as a percentage of Total Exports1970-2007

Sources: Own drawing based on statistics from MNRT and WTO

Figure 2.27 indicates the trend in the share of tourism revenue in total exports, has been experiencing a steady increase from the mid 1980's highlighting its importance to the economy. It is worth noting that the contribution of tourism in the country's total exports reached a peak in the mid 1990s and there after experienced a fluctuating behaviour but still remaining higher than the level recorded in the early 1990s.

Table 2.2: Tourism Contribution to Total Exports and Service Exports: 2003-2007

Year	Total Exports	Service Exports	Tourism Exports	% of Tourism in total Exports	% of Tourism in Services Exports
1980	676	165	21	3	13
1981	765	185	23	3	12
1982	565	115	16	3	14
1983	470	106	18	4	17
1984	404	106	13	3	12
1985	353	106	20	6	19
1986	462	101	25	5	25
1987	394	105	29	7	28
1988	392	117	34	9	29
1989	482	117	43	9	37
1990	462	131	48	10	37
1991	484	142	47	10	33
1992	586	170	70	12	41
1993	761	311	134	18	43
1994	930	411	183	20	45
1995	1248	566	502	40	89
1996	1386	602	473	34	79
1997	1223	470	339	28	72
1998	1123	534	399	36	75
1999	1119	576	464	41	81
2000	1309	575	377	29	66
2001	1705	854	615	36	72
2002	1840	860	635	35	74
2003	2116	900	647	31	72
2004	2553	1074	746	29	69
2005	2894	1215	824	28	68
2006	3385	1467	950	28	65
2007	4063	1836	1199	30	65
2008	5173	2136	1354	26	63

Sources: WTO (2009)

Table 2.2 just as Figure 2.30 indicates tourism contribution to both total exports and service exports. The Table indicates that in 1989 tourism share in the service exports sharply increased to 89% and has since remained stable at 70%

(b) Tourism's Contribution to GDP and Employment.

Figure 2.28 provides trends in the percentage contribution of tourism to GDP and employment. These figures are taken from WTTC (2009)[8].

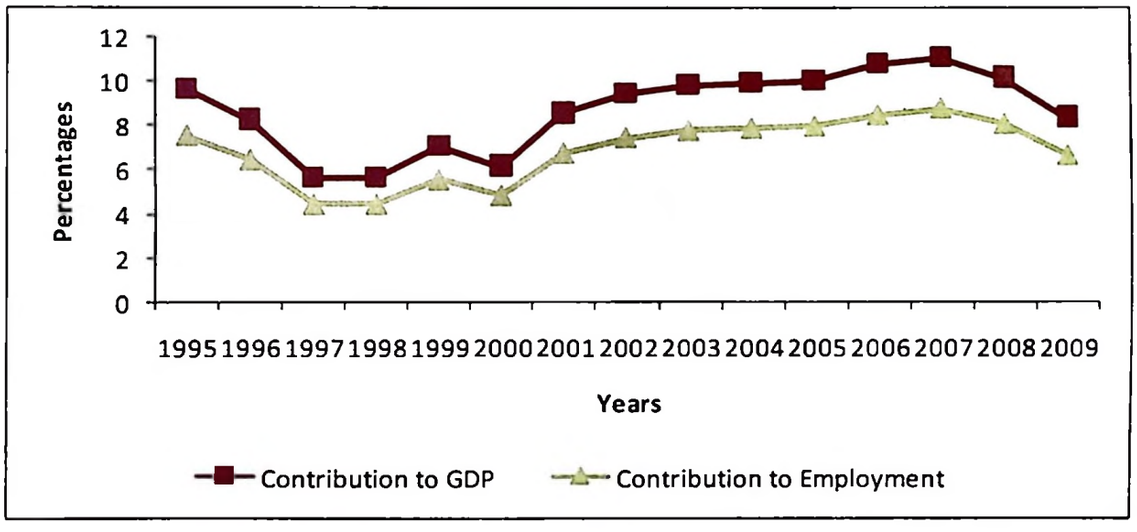


Figure 2.28: Tourism's contribution to GDP and employment: 1995-2009

Sources: Own drawing based on statistics from WTTC (2009)

Figure 2.28 shows that tourism's contribution to both GDP and employment has been increasing since 1999. According to WTTC (2009), tourism contributed US$ 1.7 billion in 2007, both directly and indirectly, which is about 11.1% of GDP. In 2008 its contribution increased to US$ 1.9 billion which is over 10% of GDP. However in 2009

[8] The figures differ from the ones given by the MNRT. But ideally the trend over time should be the same. This study uses the figures from WTTC because there is no systematic and long-term record of tourism contribution to GDP by MNRT. This lack of long-term records is due to the fact that tourism has never been counted as an independent sector in the national accounts. Efforts are now being made to make tourism an independent sector in the national accounts(see section 2.2 and TTSS,2001)

its contribution dropped back to US$ 1.7 billion which is 8.4% of GDP. This decline is likely to have been caused by the 2008 global financial crisis. On the employment side, it is estimated to have contributed 8.8% of total employment in 2007 and 6.7% in 2009.

2.4.6 Performance of Tourism in Tanzania Relative to Neighbouring Countries

Tanzania is not doing well in terms of number of arrivals and tourism revenue compared with some neighbouring countries. Comparatively, South Africa leads in the whole of Sub-Saharan Africa in terms of number of arrivals, followed by Botswana and Kenya interchangeably (WTTC 2009). Before its economic crisis, Zimbabwe had been second, next to South Africa, followed by Botswana and Kenya (WTTC, 2009).Tanzania has always been behind Kenya, both in terms of number of arrivals and revenue (Figure 2.29).

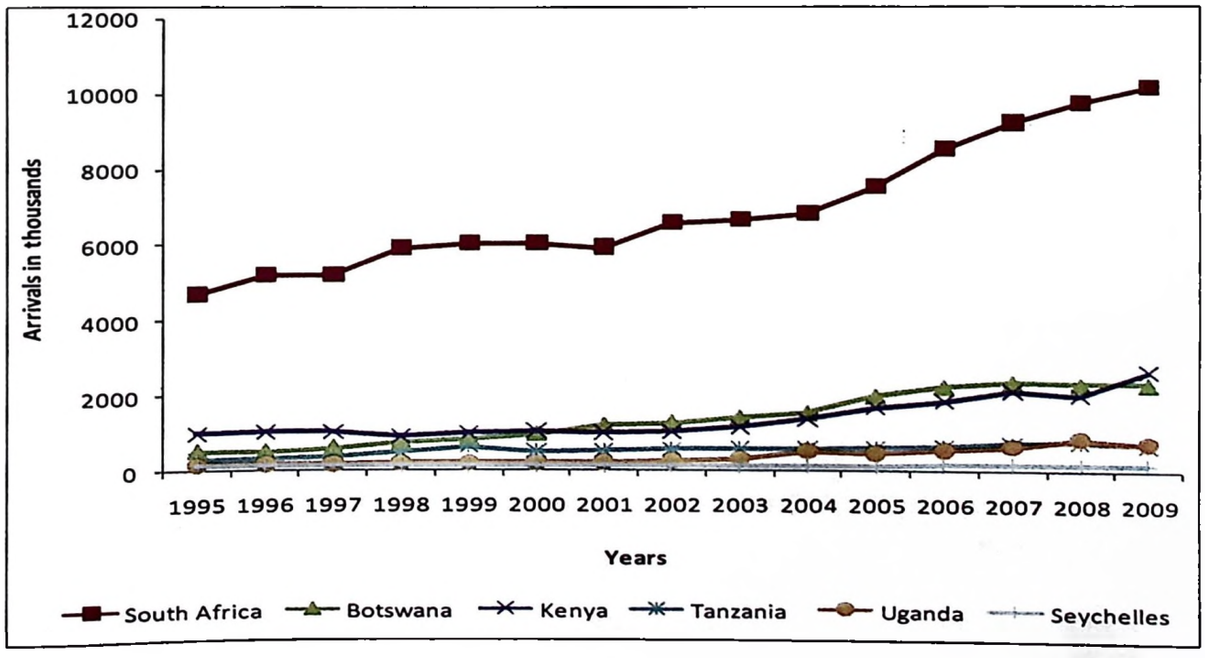


Figure 2.29: Arrivals in Tanzania versus some neighbouring countries: 1995-2009

Sources: Own drawing based on statistics from WTTC (2009)

Apart from showing that Tanzania is trailing behind South Africa, Kenya and Botswana, Figure 2.29 also shows that the country's relative share of the number of arrivals to the region has not been growing. The same picture is observed when comparison is made in terms of tourism revenue. Figure 2.30 provides this comparison, but South Arica and Botswana are omitted for ease of comparison. The key interest is to show how Tanzania is struggling against Kenya.

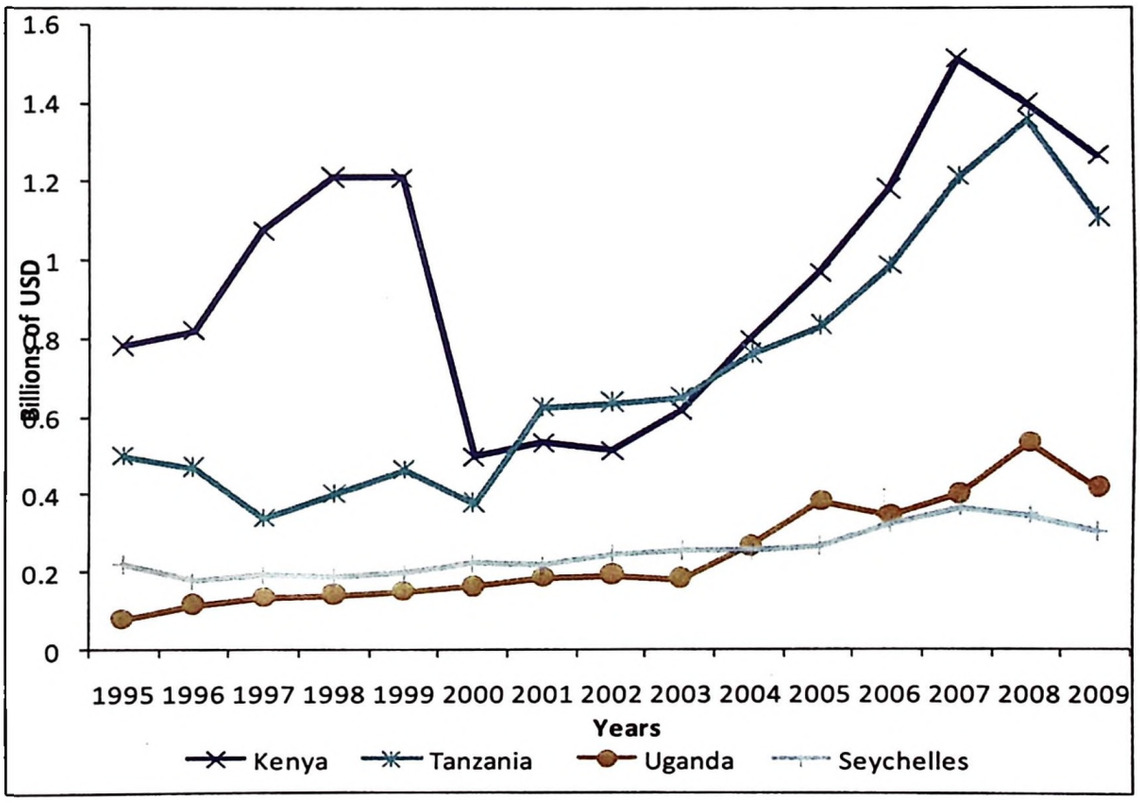


Figure 2.30: Tanzania's Tourism Revenue versus some neighbouring countries: 1995-2009

Sources: Own drawing based on statistics from WTTC (2009)

Figure 2.30 indicates that for most of the years Tanzania has been behind Kenya in terms of revenue, but well above some countries such as Uganda and Seychelles. The reasons as to why Kenya leads against Tanzania could be attributed to the fact that Kenya has more developed infrastructures such as good airports capable of attracting

direct flights from abroad and good standard hotels and restaurants. In general Kenya has invested more in the tourism industry than Tanzania.

It is also worth to note that for Tanzania there was a sharp decline in tourism revenue in the year 2009 while for Kenya a sharp decline began in 2007. As for Tanzania the 2008 global financial crises must have fueled the scenario in 2009 while for Kenya, the Kenyan political crisis was responsible for a decline in revenue in the year 2007 and later the 2008 global financial crisis worsened the situation. It is unclear why there was a sharp decline in tourism revenue for Kenya in the year 2000. However, the reason why Tanzania's revenues were higher than those of Kenya in between 2001 and 2003 could have been attributed by the fact that the year 2002 was a year for election in Kenya.

2.5 Challenges of Tourism Industry in Tanzania

The tourism industry in Tanzanian, as in other countries, faces a lot of challenges. This study has identified twelve main challenges facing the tourism sector, which are accounted in the literature and which should be of a concern to Tanzania.

First, are the environmental challenges to the tourism industry which is a global problem (WTO, 1995; Neto, 2002). The industry faces a number of problems which include pollution of national parks due to tourist activities, such as the wastes discharged from hotels, air pollution from vehicles carrying the tourists and land degradation (Kulindwa et al., 2001; MNRT, 2002, TTB, 2006) as well as the disposal of other waste such as plastic bags. Another serious environmental threat associated with tourism is the depletion of ebony trees, which are used for making carvings for

sale to tourists (Kulindwa et al., 2001). This product has been one of the most preferred by tourists from several countries. These trees take a relatively long time to grow and mature compared to other species, implying that if the situation is left to continue the species might become extinct. The environmental threats to tourism in Tanzania and in the neighbouring countries are of greater concern because the tourism industry is predominantly wildlife based (TTB, 2006).

While due attention has been given environmental concerns in Tanzania mainland, in Zanzibar and in the mainland coastal areas such as Bagamoyo, tourist activities have lead to the destruction of mangrove trees through Hotel construction, the discharges of untreated waste from hotels into the sea and beach erosion (Kulindwa et al., 2001).

The second challenge to the tourism industry, as identified by this study, is the lack of empirical studies linking promotion efforts to the growth of tourisms. For example, TTB (2006) argues that, besides recent efforts and measures to advertise and sell Tanzania as a tourist destination, the policy has not achieved much in attracting more tourists or in giving incentives for activities relating to tourism for it to flourish. In general there is a lack of rigorous demand studies which encompass a number of factors such as the studies in merchandize trade by Rutasitara (1999) and Nyoni (1996).The existing studies in tourism industry in Tanzania lack empirical verification of the pillars upon which tourism demand rests. The lack of such knowledge can lead to theoretically correct but practically unfounded arguments with regards to tourism demand. For example, in the face of the recent global financial crisis how has the sector been affected? This question can only be answered if there is a clear understanding of what determines tourism demand in Tanzania. This study will

attempt to address the issue of promotion strategies, based on the determinants of tourism demand, as well as pointing the implications of the recent global financial crisis for the sector.

A third challenge the tourism industry faces is what tax policy should be in place for maximizing government revenue without hampering the growth of the sector. This concern is crucial because tourism charges in Tanzania, particularly for accommodation, are relatively high compared with those of neighbouring countries (Kulindwa et al., 2001; MNRT, 2002; TTB, 2006).In view of this problem Kulindwa et al.(2001) suggest that the country needs to ascertain the strength of the uniqueness of its attractions by knowing whether or not tourism demand is price elastic.

This study has attempted to address this challenge. The difficulty lies in the fact that tourism demand is broadly defined, with scholars using a variety of response variables, such as number of arrivals, length of stay, and number of occupied rooms per period (Meniz and Moniz, 2006). In order to study tax sensitivity, one has to venture separately into all such areas, including the sensitivity of park fees and other charges. Another complication is the lack of a definitive index for tourism prices. Such indices exist in very a few developed countries, where they are able to distinguish a basket of tourism goods against those of the general population. In most of the studies, such as Eita and Jordaan (2007) or Saayaman and Sayaaman (2008), relative consumer price indices are being crudely used as a measure of tourism prices. This study has discussed price sensitivity to taxation, using these very crude measures, but focusing only on the number of arrivals as one of the elements of tourism demand.

The fourth challenge facing the tourism industry is competition in the use of natural resources between the tourist parks and the indigenous population surrounding the parks (Kulindwa et al., 2001, MNRT, 1999).The Masai people of Serengeti and Ngorongoro areas are good examples of this competition. These people are normally cattle herders with a great need for the grazing areas, which is at the expense of national parks. While tourism has provided some with employment, it has also caused unemployment among cattle herders. There is a need to know how many people have suffered from unemployment and how many are employed due to the expansion of tourism activities.

A fifth challenge facing Tanzanian tourism is the problem of financial leakage that characterizes many developing countries (Sandbrook, 2008). Boo (1990) cited by Sandbrook (2008) account that 55% of tourism revenue in developing countries goes back to the developed countries through the importation of goods and services demanded by tourists. Most of the goods sold in tourist hotels, particularly alcoholic beverages and luxurious foods, are imported from foreign countries. Furthermore, some of these hotels and tour operating companies are either wholly owned by foreigners or jointly by the Tanzanians. The end result may be that some of the accrued profits also go out of the country. For example Anderson (2013) assert that most resorts in Zanzibar are owned managed and operated by non-locals and that only 16 percent of the resort requirements are sourced within Zanzibar. MNRT (2002) gives a rather surprising account that is different from that of Boo (1990), as it shows that only 27% of tourists' expenditure goes out of Tanzania as leakage. However, MNRT (2002) analysis was based on a few hotels and lodges selected in the Northern circuit.

Given the leakage suspicion, there is a need for a comprehensive survey of all tourism investments to ascertain the actual amount of leakage.

The sixth challenge facing the tourism industry in Tanzania is the little benefits of tourism industry to the local population, especially those people surrounding tourist sites. Sandbrook (2008) found that more than 75 percent of the tourism revenue accrued from the local area of Bwindi Impenetrable National Park of Uganda goes out of the area as leakage. Kulindwa et al. (2001) also give an account of the same problem in the areas surrounding Tanzanian National Parks. According to Kulindwa et al. (2001), although TANAPA has so called Community Based Conservation (CBC) programmes which aim to provider social services to the surrounding community, these programmes have not been exhaustive. Interviewed people argue that the money from these programmes is always inadequate and sometimes it is delayed or not given at all. This problem of lack of benefits for the local population is partly contributed to the fact that tourism is dominated mostly by package tours whereby tourists pay their bills for the respective hotels in advance, via the travel agents and so they spend very little money in the communities they pass through. This study has indirectly addressed this problem when studying the determinants of package/non-package tours.

The seventh challenge facing the tourism industry is cultural in nature. The coming of tourists has brought in some cultures viewed as destructive. The growth in tourism has led to the construction of refreshment centres such as casinos (Kulindwa et al., 2001) and beach hotels where prostitution is greatly encouraged by the owners of these investments. The study by Mgani (2007) shows how Arusha has succumbed to prostitution due to its being the core town of tourist activities in the Northern circuit.

Prostitution has led to the spread of venereal diseases, particularly AIDS. Other problems associated with cultural interference include excessive smoking and drug use.

The eighth challenge facing the tourism sector in Tanzania and Africa in general is the lack of domestic tourism (Andrew, 2008). It is true that inbound tourism is the best in the sense that it brings in foreign revenue, but as we have seen it creates several problems. Some problems associated with heavy reliance on inbound tourism include the creation of seasonal unemployment during low peak seasons, inflation due to the money being pumped into the country, appreciating country's exchange rate which negatively affects other export industries (Blake, 2008), risking the country in the case of diplomatic conflicts with key source markets; and intermittent political crises such as the Kenyan election of 2008. These problems can be significantly reduced if domestic tourism is promoted. In other countries, such as Australia, domestic tourism forms the core of the industry and contributes quite substantially to GDP (Athanasopoulos and Hyndiman, 2006).

The ninth challenge facing the tourism industry is the lack of proper statistics for recording and subsequent analysis. For example, in the period before 1995, it was very hard to get statistics on the number of arrivals by country. Nevertheless, this problem is gradually being solved. Currently, it is easy to get number of arrivals by different categories, such as purpose of visits, country of origin and mode of transport. However, the most serious problem concerning tourism statistics at present is not the number of arrivals but rather tourists' expenditure. It is very hard to capture the amount that tourists spend owing to the fact that tourism transactions are now handled

by private commercial banks, and private bureau de change which do not observe Bank of Tanzania (BOT) regulations, requiring them to give an account of which transactions relate to tourism and which do not. This is a global problem, but much more of a problem in developing countries, where individuals/business enterprises do not observe the rules and regulations. In recognition of this problem, TTSS was formed so as to estimate, among other things, the annual revenue from tourism

The tenth challenge facing the tourism industry in Tanzania is the country's poor infrastructure. As it will be seen in chapter three, improvement in the infrastructure and the general performance of the economy matter a lot in attracting tourists. The infrastructure issue is broad. It covers lack of direct international flights to the country, quality accommodation, good tarmac roads, and quality tour operators and guides (MNRT, 2002). The poor quality of these factors has a negative effect both on attracting more arrivals and on their per capita spending. The problem is more dominant in western and southern areas of the country (TTB, 2006), which has led to the concentration of tourist activities in the northern part of the county, leaving other areas unexploited, despite their immense wildlife and cultural resources (TTB, 2006).

The eleventh challenge is the communication problem facing most Tanzanians, this including mastery of the English language. The study views this problem as peculiar to Tanzania, and denies local communities from interacting directly with tourists with a view to tapping their expenditure. More generally, there is a shortage of skilled manpower for the tourism industry, such that investors have been employing skilled

Prostitution has led to the spread of venereal diseases, particularly AIDS. Other problems associated with cultural interference include excessive smoking and drug use.

The eighth challenge facing the tourism sector in Tanzania and Africa in general is the lack of domestic tourism (Andrew, 2008). It is true that inbound tourism is the best in the sense that it brings in foreign revenue, but as we have seen it creates several problems. Some problems associated with heavy reliance on inbound tourism include the creation of seasonal unemployment during low peak seasons, inflation due to the money being pumped into the country, appreciating country's exchange rate which negatively affects other export industries (Blake, 2008), risking the country in the case of diplomatic conflicts with key source markets; and intermittent political crises such as the Kenyan election of 2008. These problems can be significantly reduced if domestic tourism is promoted. In other countries, such as Australia, domestic tourism forms the core of the industry and contributes quite substantially to GDP (Athanasopoulos and Hyndiman, 2006).

The ninth challenge facing the tourism industry is the lack of proper statistics for recording and subsequent analysis. For example, in the period before 1995, it was very hard to get statistics on the number of arrivals by country. Nevertheless, this problem is gradually being solved. Currently, it is easy to get number of arrivals by different categories, such as purpose of visits, country of origin and mode of transport. However, the most serious problem concerning tourism statistics at present is not the number of arrivals but rather tourists' expenditure. It is very hard to capture the amount that tourists spend owing to the fact that tourism transactions are now handled

by private commercial banks, and private bureau de change which do not observe Bank of Tanzania (BOT) regulations, requiring them to give an account of which transactions relate to tourism and which do not. This is a global problem, but much more of a problem in developing countries, where individuals/business enterprises do not observe the rules and regulations. In recognition of this problem, TTSS was formed so as to estimate, among other things, the annual revenue from tourism

The tenth challenge facing the tourism industry in Tanzania is the country's poor infrastructure. As it will be seen in chapter three, improvement in the infrastructure and the general performance of the economy matter a lot in attracting tourists. The infrastructure issue is broad. It covers lack of direct international flights to the country, quality accommodation, good tarmac roads, and quality tour operators and guides (MNRT, 2002). The poor quality of these factors has a negative effect both on attracting more arrivals and on their per capita spending. The problem is more dominant in western and southern areas of the country (TTB, 2006), which has led to the concentration of tourist activities in the northern part of the county, leaving other areas unexploited, despite their immense wildlife and cultural resources (TTB, 2006).

The eleventh challenge is the communication problem facing most Tanzanians, this including mastery of the English language. The study views this problem as peculiar to Tanzania, and denies local communities from interacting directly with tourists with a view to tapping their expenditure. More generally, there is a shortage of skilled manpower for the tourism industry, such that investors have been employing skilled labour from neighbouring countries such as Kenya (TTTB, 2006).

The last challenge facing the tourism industry in Tanzania as well as in most developing countries is uncertainty about health. This problem can be viewed in two ways. One concerns the prevalence of tropical diseases, such as malaria, which is a dangerous disease for visitors from non-tropical areas. Tropical diseases are very hard to eliminate; and malaria in particular has proved to be the most difficult to deal with. The second is the danger that tourists themselves bring; this includes HIV/AIDS.

2.6 National Tourism Policy.

Due to these challenges, as well as some other constraints, the government decided to formulate and adopt a national tourism policy to address some of these issues. The national tourism policy was first adopted in 1991 and later reviewed in 1999 (MNRT, 1999, 2002). According to MNRT (1999), the policy seeks to assist in efforts to promote the economy and the livelihoods of people, essentially through alleviating poverty by encouraging the development of sustainable and quality tourism that is culturally and socially acceptable, ecologically friendly, environmentally sustainable, and economically viable. It has also sought to market Tanzania as a favoured tourist destination for touring and adventure in a country renowned for its cultural diversity and numerous beaches. It is recognized that the private sector will play a major role in the industry's development, with the government playing a catalytic role of providing and improving the infrastructure as well as providing a conducive climate for investment.

2.7 Conclusion and a Way Forward.

The chapter has given the current status of tourism demand in Tanzania. Two major issues were found. First tourism demand is doing well based on any of the indicator

variables. Second when compared with neighbouring countries tourism demand is not doing well, especially bearing in mind the country's vast tourism assets. The challenge ahead therefore lies in the sustainability and promotion of the sector. Apart from this challenge, the chapter has also highlighted some other challenges facing the industry and the national tourism policies devised for such challenges. However, this study has undertaken only one task that of establishing the determinants of tourism demand for effective promotion of the sector. This is indeed the basic problem the study has investigated. But in the pursuit of this problem the study has found itself inevitably touching other key challenges facing the sector, such as issue of taxation, environmental threat, the leakage issue and the tourism statistics problem, which add more value to the overall significance of the study.

CHAPTER THREE

DETERMINANTS OF TOURIST ARRIVALS

3.1 Introduction

Although tourism demand is a broadly defined subject, with some scholars using various indicator variables, such as number of arrivals, tourism revenue, number of nights and occupied rooms(Meniz and Moniz,2006), the coming of tourists is necessary for without it there can be no talk of tourism demand. This chapter examines the factors determining the number of tourist arrivals in Tanzania. Unlike in some other countries, these factors have not been rigorously studied in Tanzania.

According to Munoz (2004), very few studies on tourism demand have included marketing expenditure as one of the explanatory variables in their models. In this study special attention is given to marketing expenditure as well as number of lagged visits. The former assesses the relevance of promotion efforts while the later assesses role played by destination attractiveness and the capacity to meet clients' needs.

Some studies on tourism demand, such as Durbarry (2001), assume that any country neighboring another country is a competitor as regards tourism, but this study does not take this assumption for granted, as it sets out an empirical investigation to establish which countries are real competitors to Tanzania and which are not. The notion of competition in tourism might be somewhat ambiguous, as it would rarely happen that two countries have perfectly substitutable tourist attractions. In other words, contrary to the popular conception of most studies, such as TTSS (2001), which regards countries in East and Southern Africa to be Tanzania competitors it may turn out that

most of them are not, because Tanzania's tourist attractions, such as Mount Kilimanjaro and Ngorongoro Crater, are unique.

Most tourism studies have not addressed the role of trade liberalization in the coming of tourists. According to the summary done by Lim (1997) as well by Naude and Saayman (2005) on tourism demand studies, since the 1960s only 5% of such studies have measured the impact of business trade on arrivals, using proxies such as trade, direct foreign investment and capital flows. One of the possible reasons might be that most such studies have been done in countries with more open economies, unlike countries like Tanzania, which pursued closed economy policies since late sixties up to the mid-1980s.This study has attempted to assess the role of trade liberalization in the coming of tourists to Tanzania.

Besides assessing the impact of marketing expenditure, the role of destination attractiveness as measured by the influence of past number of visitors, the competitiveness of neighbouring countries and the role of trade liberalization, the study also looks at other factors that may impact the number of arrivals in Tanzania. In general, all the factors have been grouped into price and non-price factors, a categorization which is useful for making recommendations. Its usefulness stems from the fact that, whereas price factors, such as a tourist income, the government and other stakeholders cannot directly affect them, for most of the non-price factors, such as infrastructures, the government can be in a position to affect them directly. Therefore the study is set out to test the hypothesis that price factors are more influential than non-price factors in determining the number of tourist of arrivals.

The study uses panel data estimation models, which provide more data points and allow for the estimation of observable country-specific effects, unlike time series models, which have dominated international trade studies in Tanzania.

The remainder of this chapter is organized as follows: Section 3.2.1 provides a theoretical literature review, section 3.2.2 reviews the empirical literature, section 3.3 provides the methodology and data sources, section 3.4 gives the results and discussion and section 3.5 concludes the chapter.

3.2 Literature Review

3.2.1 Theoretical Literature Review

Inbound tourism is an export commodity; therefore analysis of its demand should be formulated in the contest of general international trade models (Eita and Jordaan, 2007). These models have evolved over time along with developments in economic thought; starting with the earliest protectionism view, dominant in the mercantilist era, up to the recent gravity model theory (Mjema, 2004). Salvatore (2010) puts these international trade models chronologically into five main categories, capable of explaining the pattern of trade as follows:

(i) First is the Adam Smith absolute advantage theory, which asserts that when one nation is more efficient than(or has an absolute advantage over) another in the production of one commodity, but is less efficient than(or has an absolute disadvantage with respect to) than the other nation in producing a second commodity ,then both nations can gain by each specializing in the production of the commodity of its absolute advantage and exchanging part of its output with the other nation for a commodity of its absolute disadvantage.

(ii) Second is the Ricadian comparative advantage theory, which asserts that even if one nation is less efficient than (has an absolute disadvantage with respect to) the other nation in the production of both commodities, there is still a basis for mutually beneficial trade. According to the theory, the movement of goods and services between nations is caused by the comparative advantages which nations have in the production of different commodities.

(iii) Third is the Hescscher- Ohlin theory. This theory differs from the former as it considers comparative advantage in terms of factor endowments. According to the theory, countries will export goods that make intensive use of locally abundant factors, and import goods that make intensive use of locally scarce factors.

(iv) Fourth is the new trade theory, which tries to explain several facts about trade, which the previous models have had difficulty with. These include the fact that most trade is between countries with similar factor endowments and productivity levels and a large amount of multinational production (i.e, foreign direct investment) characterizes this approach.

(v) Fifth is the gravity model theory. The model, in its basic form predicts trade based on the distance between countries and the interaction of the countries' economic sizes. The model has been borrowed from the Newtonian law of gravity, which also considers distance and physical size between objects, as the factors for attraction. The model has proved to be strong through econometrics analysis.

With respect to tourism demand, the gravity model is more appealing than the rest of the models in explaining tourist flows to a destination (Durbarry, 2001, Maliugina

2006; Eita and Jordaan, 2007). Explaining why the Hescscher- Ohlin theory cannot explain the flow of tourists from one country to another, Eita and Jordaan (2007) argues that the theory is based on relative factor endowments (labour and capital), which in the case of tourism cannot apply because its most important factors of production are unique to the specific country and are not easy to measure, evaluate or compute.

Analogous to the gravity model, most recent studies have used two types of variables to model tourism demand. First are the conventional price determinants of any export commodity. These are the income of the tourist country of origin, the price of tourism as measured by relative prices and the prices in competing destinations (Durbary2001; Munoz 2004; Munoz and Martin 2006,). Second are the non-price variables deduced from the implication of the gravity model theory. These are variables which hinder or facilitate the smooth flow of tourists from one country to another. They include, distance, infrastructure, advertisement costs, geographical location, preference and other destination characteristics such as peace, security, culture and diplomatic relations (Maliugina 2006; Muchapondwa and Pimhidzai 2008; Eita and Jordaan, 2007).

3.2.2 Empirical Literature Review

Several studies conducted in developed countries have identified price factors to be more significant than non-price factors in attracting tourist visits. .For example Durbarry (2001), using the gravity model and panel data modeling studied tourism demand in the United Kingdom (UK) and found that the appreciation of sterling pound, and a fall in relative prices in other destinations is associated with falling

tourism receipts in the UK. In this study Durbarry (2001) treated a number of neighbouring destinations as competitors to the UK and formed a weighted average of relative prices as a single competing price for the UK.

Similarly Maliugina (2006) established that income from a tourist's country to be the most significant determinant in attracting tourist arrivals in the Ukraine. The study used gravity model theory and unbalanced panel data of 75 sending countries to the Ukraine. The study also identifies longer distance of a tourist country from the Ukraine as one of the barriers to arrival of tourists.

Munoz and Martin (2006), using panel data modeling also found tourism demand in the Balearic Islands to be heavily dependent on the economic activities (GDP) of the tourist country of origin and the relative cost of living. They also find the lagged value of tourist visits to be significant.

Some studies in developed countries have found that non-price factors are significant in attracting tourists. For example, Vietze (2008) studied the influence of culture on tourist flow into the USA, using the gravity model approach. Apart from finding an evidence that gravity model was an adequate instrument for explaining international tourist flow, the study also found that the cultural proximity between countries of origin and country of destination have a positive effect on tourism flows between these countries. The study concluded that travelers from English-speaking and Christian countries greatly prefer to visit the USA than other countries.

Some African case studies have found that non-price factors are more significant. For

countries for the period 1996 to 2000, concluded that typical developed country determinants of tourism demand, such as the level of income in the origin country, the relative prices and the cost of travel, are not significant in explaining the demand for Africa as a tourist destination. The study thus recommends that attention should be given to improving the overall stability of the continent and the availability of tourism infrastructure. These findings are similar to those of Muchapondwa and Pimhidzai (2008), who studied international tourism demand in Zimbabwe. In their study, they found that tourism demand was insensitive to both domestic and foreign prices. The study therefore recommended improvement in international tourism infrastructure in order to attract more international tourists to Zimbabwe.

However, other studies in African countries have found that price factors are most significant. Eita and Jordaan (2007), using the gravity model and panel data modeling, concluded that trading partner's income, depreciation in the exchange rate, improvements in Namibia's infrastructure and sharing a border with Namibia are associated with an increase in the number tourists arrivals. Saayman and Saayman (2008) studied the determinants of inbound tourism in South Africa. The study found income from tourists' countries, relative prices and travel costs to South Africa as the most important determinants of inbound tourism. Similarly Mohamed (2011) in Egypt and Eja et al. (2012) in Nigeria found that tourism prices as measured by relative cost of living matters more for tourist arrivals in the two countries.

Studies in other developing countries also show that both price and non-price factors matter in attracting tourists. For example Zhang et al (2009), using OLS, studied tourism demand in Thailand from 1987-2006 and found that travel demand to

Thailand could be explained more by the exchange rate, promotional budget, and the Asian financial crisis. Likewise Hanafiah and Harun (2010), using the gravity model, analyzed determinants of the number of tourist arrivals in Malaysia, and found the exchange rate, financial crisis, and tourist income to be influential factors. These two studies could imply that price factors are more influential than non-price factors in attracting tourists to Asian countries. The same is confirmed by Kordbacheh et al. (2012), who found tourist country's per capita income as most influential determinant of tourist arrivals to Iran.

Song and Li (2008), reviewing the published studies on tourism demand modeling and forecasting, since 2000, concluded that as far as influential factors are concerned, recent econometric studies on tourism demand have shown that tourist's income, tourism prices in a destination relative to those in the original country, tourism prices in competing destinations (i.e. substitute prices) and the exchange rates are the most important determinants of tourism demand.

From this review, the following conclusions can be drawn: First, whether price factors are more influential than non-price factors in attracting tourists, it depends on the destination. In other words, to ascertain the case in Tanzania, a study on Tanzanian tourism demand is needed. Second, most studies, such as Durbarry (2001), take it for granted that neighbouring destinations compete with each other. This may not necessarily be the case. Third, few studies have examined the role of trade liberalization in attracting tourists. Fourth, few studies have included marketing expenditure as one of the explanatory variables.

Munoz (2004) argues that most tourist organizations consider marketing expenditure a key factor in determining international tourism flow. However, Munoz (2004) argues that very few studies have included marketing expenditure (Barry and O'Hagan, 1972; Papadopoulos and Witt, 1985), and those who included this variable had inconclusive results, first because promotional and marketing campaigns of competing destinations may offset the impact of marketing expenditure by a particular country. Second, the actual impact of the promotional activity can vary over time; and third the impact will vary across the media used. In other words inclusion of marketing expenditure as an explanatory variable should take into consideration of not only the absolute level of expenditure but also expenditure of competing destinations and the type of mass media used.

3.3. Methodology

3.3.1 Model Specification

This study adopts the gravity model approach in modeling the determinants of the number of tourist arrivals in Tanzania. According to Durbarry (2001), the commonly used form of the gravity model in international trade is of the following specifications:

$$PX_{ij} = \beta_o (Y_i)^{\beta 1} (Y_j)^{\beta 2} (D_{ij})^{\beta 3} f(A_{ij}) u_{ij}, \qquad (3.1)$$

where

PX_{ij} = the value of flow from county i to country j,

$Y_i (Y_j)$ = the nominal value of GDP in country i(j),

D_{ij} = the distance between i and j,

$f(A_{ij})$=is a function containing a vector of additional variables either aiding or

resisting the flow between i and j

u_{ij} = a log normally distributed error.

Further Durbarry (2001) argues that following Matyas (1998), one can write the general specification of the gravity model for econometric estimation as

$$Y_{ijt} = \alpha_i + \gamma_j + \lambda_t + \beta_1 x_{ijt} + \beta_2 X_{it}^{*} + \beta_3 X_{jt}^{**} + ...\mu_{ijt}, \qquad (3.2)$$

where

Y_{ijt} = the dependent variable (e.g. volume of trade from country i to j, in time t)

X_{ijt} = independent variables with variations in all three dimensions i, j and t(e.g. exchange rates)

X_{it}^{*} = independent variables with variations in the dimensions i and t (e.g. GDP)

X_{jt}^{**} = independent variables with variations in the dimensions j and t;

α_i = the local country effect;

γ_j = the target country effect;

λ_t = the time effect and

μ_{ijt} = the white noise disturbance term

One of the earliest criticisms of the gravity model, when firstly introduced in economics by Tinbergen in 1962 was the lack of an economic theory to justify its existence (Tayyab et al., 2012) In later years this criticism was subsequently addressed by different scholars. Linneman (1966) showed that the model could be obtained from the partial equilibrium model of export supply and import demand. Anderson (1979), using a traded share expenditure system also derived a model which postulates identical Cobb-Douglas or constant elasticity of substitution preference functions for

all the countries and weakly[9] separable utility function between traded and non-traded goods. Bergstrand (1985, 1989) derived the gravity model from the general equilibrium of demand and supply systems. For each country the model of trade is derived by maximizing a constant elasticity of substitution utility function, subject to the income constraints of importing countries.

While Anderson (1979) derived the general gravity model by considering a weak separable utility function involving tradable and non-tradable goods, this concept could be improved on by showing that the gravity model can be equally well applied to tourism. Tourism as an export commodity differs from other exports due to the fact that a consumer (a tourist) consumes the commodity in the country where it is produced. This means, analogous to Anderson (1979) assumption of weak separability of the traded goods and non-traded goods, a tourist utility function should accommodate the fact that a tourist maximizes his utility to meet the demand for both non-tourism goods consumed in his home country and tourism goods consumed in the country he/she expects to visit.

One of such models which accommodate this fact is the one by Nordstrom (2002), which was also adopted by Naude and Sayaaman (2005) when studying tourism

---

[9] A utility function is weakly separable if the marginal rate of substitution (MRS) between any two goods from one group of goods, say tradable goods, is independent from the quantities of goods outside this group, say non-tradable goods. A more detailed explanation of this concept and why it is needed is given in appendix 3.2, where an example of a utility function comprising tourism and non-tourism goods is given. One can draw an analogy in this case by comparing tradable goods to tourism goods and non-tradable goods to non-tourism goods. The utility separability concept originated with Sono (1945) and Leontief (1947). Appendix 3.2 provides detailed explanations of the concept.

demand for Africa. This model assumes a [10]strong separable utility function where a tourist follows a two-stage utility maximization procedure. In the first stage, the consumer decides how much expenditure to allocate between various consumptions, one of being tourism in Africa. The second stage consists of allocating of tourism expenditure to particular African countries (destinations).

This model by Nordstrom (2002) and Naude and Sayaaman (2005), though not addressing the gravity model, can be modified and shown to contain the usual gravity model specifications by making slight improvements in some of its assumptions.

Naude and Sayaaman (2005) argues that following (Nordstrom, 2002), a two level tourist utility function can be written as,

$$U(q) = F[U^1(q_1), \cdots, U^n(q_n)], \tag{3.3}$$

Where

$U^j(q_j)$= the sub-utility function, which consists of African tourism consumed in the j^{th} country (out of m-number of African destinations to choose from)

Naude and Sayaaman (2005), assuming a Cobb-Douglas utility function for m different goods (African countries) and the fact that past tourism affects the current consumption specifies the sub-utility function as

---

[10]A utility function is strongly separable if the MRS between two goods belonging to two different groups is independent of the quantity of goods which do not belong to any of those two groups. Unlike the weak separability case strong separability demands that each group of goods be separable from the other group of goods. Any of these two conditions makes it possible to consider utility maximization by a tourist of tourism goods alone without having information on prices and quantities of non-tourism goods. All that is needed is the budget of tourism goods. In view of this, the study does not maintain the strong separability assumption, because the weak separability condition is sufficient (see Appendix 3.2)

$$U_i(q_t / q_{t-1}) = \prod_{j=1}^{m}(q_{ijt} - k_{ijt})^{\delta_j}, \sum_{j=1}^{m}\delta_j = 1, \qquad (3.4a)$$

where

$$k_{ijt} = y_{ijt} + \gamma_j q_{ijt-1} \qquad (3.4b)$$

y_{ijt} = positive and represents the minimum consumption requirement in period t,

$\gamma_j q_{ijt-1}$ = consumption based on past consumption.

The subtraction of k_{ijt} from q_{ijt} provides the quantity demanded as a function of price factors (i.e. tourism price as well as a tourist income). In order to invoke a gravity model, this study modifies the specification for k_{ijt} to be dependent, not only on past consumption but also on other factors implied in the gravity model (notably the destination country's GDP and distance between a tourist country and the country of destination). Therefore k_{ijt} in $3.4b$ is now specified as

$$k_{ijt} = y_{ijt} + \gamma_j q_{ijt-1} + O_{ijt} \qquad (3.4c)$$

Where O_{ijt} is a vector of regressors derived from the gravity model theory, which covers among other things distance and the destination's GDP.

The particular African country tourism demand function is obtained by maximizing the sub-utility function given in $3.4a$ subject to budget constraints

$$\sum_j p_{jt} q_{jt} = e_t, \qquad (3.5)$$

where

p_{jt} = the price of tourism in the given j^{th} African country.

e_t = is the tourist income

The resulting individual demand function after maximization can be written as

$$q_{ijt} = y_{ijt} + \gamma_j q_{ijt-1} + O_{ijt} + \sum p_{jt} y_j + \frac{\delta_j}{p_{jt}}\left(e_t - \sum p_{jt}\gamma_j q_{ijt-1}\right) + \varepsilon_{jt} \qquad (3.6)$$

[11]

Equation (3.6) tells us that tourism demand is a function of past consumption q_{ijt-1}, destination prices p_{jt} , tourist income e_t and other factors O_{ijt}. Using (3.6) one can specifically write the tourism demand equation of Tanzania in time *t* by origin *i* as follows:

$$q_{it} = x_{ijt}\beta + c_i + \varepsilon_{it}, \qquad (3.7)$$

where

q_{it} = demand for international tourism by origin i for Tanzanian destination.

x_{it} = a vector of explanatory variables with cross-section and/or time variation including both price and non price factors as given (3.6)

c_i = qualitative factors in origin i, intended to capture unobserved individual specific factors.

In order to interpret the coefficients as elasticities, as preferred in demand studies and also to avoid the impact of highly likely abnormality of observations such as number of arrivals and exchange rates, the study adopts the double logarithmic function model. The use of logs of the variables instead of the variables themselves reduces the

---

[11] Equation (3.6) is obtained by maximizing (3.4a) subject to budget constraints (3.5) using the Lagrange multiplier formulation. Appendix 3.3 gives a detailed procedure on how (3.6) is obtained. Note that the prices of related commodities are inherently implied in (3.6) when one expands "$\sum p_{jt}\gamma_j q_{ijt}$".

extremely huge values into moderate ones, without reducing much the smaller ones, because the logarithmic function increases at a declining rate. Accordingly, one can rewrite equation (3.7) in a more detailed form as

$$\ln TA_{it} = \beta_o + \ln \beta_1 \ln TZ_{it} + \beta_2 \ln ER_{it} + \beta_3 \ln GDP_{it} + \beta_4 \ln DIST_i + \beta_5 \ln AD_i + \beta_6 \ln IFR_t + \\ + \beta_7 TGDP_t + \beta_8 \ln KR_{it} + \beta_9 \ln SR_{it} + \beta_{10} \ln MR_{it} + \beta_{11} UG_{it} + \beta_{12} BW_{it} + \beta_{13} ZW_{it} + \beta_{14} SE_{it} \\ + \beta_{15} \ln TA_{it-1} + B + \mu_{it}, \qquad (3.8)$$

Where [12]

TA_{it}=the number of tourist arrivals in Tanzania from origin i, in year t

TZ_{it}= is the ratio of the consumer price of Tanzania in year t to a tourist country consumer price in the same year, adjusted by the nominal exchange rate (ER)[13] . Its coefficient β_1 is expected to be negative.

ER_{it}= is the nominal exchange rate between Tanzania and the tourist country i, expressed in terms of Tanzanian shillings per a unit of a tourist's country currency. Its coefficient β_2 is expected to be positive[14].

GDP_{it}=the GDP per capita of the tourist country of origin, which is a good proxy for a

---

[12] The variables are included in the model according to the literature review in sections 3.2 and 3.3 which also explain the direction of the impact of each of the regressors on the tourist number of arrivals.

[13] $TZ_{it} = CPI^t_{TZ} / CPI^t_O \; ER_{it}$ Where CPI^t_{TZ}, is the consumer price index in Tanzanian and CPI^t_O is the consumer price index in a tourist country of origin.

[14] This variable is commonly used in studies examining determinants of tourism demand (Lim, 1997).Despite being a common variable one may be tempted to question its relevance because a country's exchange rate may be devalued or overvalued depending on the country's decision. In reality however, countries have little choice on this. Devaluation/overvaluation depends mostly on the country's economic situation and as such there very few poor countries with favorable exchange rates and the vice-versa also holds. In other words, the country's exchange rate may be a reasonable proxy for a country's prosperity, even though in this study it was found to be insignificant (see Table3.6 & 3.8)

trading partner's disposable income, necessary for spending on leisurely activities such as tourism. Its coefficient β_3 is expected to be positive.

$DIST_i$ = the distance between Tanzania and country i as supposed in the gravity model. This variable acts as a proxy for the cost of transport between Tanzania and a tourist country . Its coefficient β_4 is expected to be negative.

AD_t = the value of advertisement expenditure in year t, by the Tanzanian government via its various promotional organs. Its coefficient β_5 is expected to be positive.

IFR_t = the level of infrastructure in Tanzania in year t. There is no unique measure of infrastructure. While some use the length of a paved road, in other studies amount of electricity produced annually has been used as a measure for infrastructure (Eita and Jordaan, 2007). In this study the amount of electricity generated annually was used. The use of electricity as a measure of infrastructure is justified by the fact that electricity is widely used in many sectors, tourism being one of them. It is used for communication by the air lines, travel agents, tour operators and in tourist hotels. Even though most tourist hotels probably have their own sources of electricity, such as generators, the use of such sources is restricted during power shortages. The coefficient β_6 for this variable is expected to be positive.

$TGDP_t$ = [15]Tanzania's per capita GDP. Its coefficient β_8 is expected to be positive.

[15] Tanzania's per capita GDP is included according to the requirements of the gravity model (see equation 3.1 and 3.2). This variable is a proxy for the country's capacity to supply tourism goods and services.

KR_{it} = tourism prices in Kenya defined in a similar way as TZ_{it}

SR_{it} = tourism prices in South Africa defined in a similar way as TZ_{it}

MR_{it} = tourism prices in Mauritius defined in a similar way as TZ_{it}

UG_{it} = tourism prices in Uganda defined in a similar way as TZ_{it}

BW_{it} = tourism prices in Botswana defined in a similar way as TZ_{it}

ZW_{it} = tourism prices in Zimbabwe defined in a similar way as TZ_{it}

SE_{it} = tourism prices in Seychelles defined in similar way as TZ_{it}

The choice of these countries as possible candidates that compete with Tanzania is based on the suggestion by TTSS (2001). Coefficients for each of the variables KR, SR, MR, UG, BW, ZW,SE are indeterminate, depending on whether a country is really a competitor or not.

TA_{it-1} Is the lagged number of tourist arrivals and its coefficient β_{16} is indeterminate, depending on the tourists' experience of their stay in Tanzania. The variable captures tourists' preferences and tastes. If their experiences was positive then β_{16} will be positive otherwise it will be negative. This variable also makes possible the estimation of long-run demand elastic ties. The long-run elasticities can be obtained by setting $TA_{it-1} = TA_{it}$ after estimating equation 3.8 and solving TA_{it}. The idea is that, in the long run when the demand function reaches equilibrium, the number of arrivals is constant.

B = 1 if a country borders Tanzania and 0 if a country does not. Its coefficient is expected to be positive.

Table 3.1 summarizes the discussed variables and their expected signs.

Table 3.1: A priori direction of the relationship between the number of arrivals and the explanatory variables

variable	acronym	Direction
Consumer prices of Tanzania relative to that of the tourist country	TZ_{it}	-
Tanzania nominal exchange rate against that of the tourist country	ER_{it}	+
GDP per capita of the tourist country of origin	GDP_{it}	+
Distance between Tanzania(DSM) and the tourist country's capital	$DIST_i$	-
Annual Advertisement expenditure by public institutions	AD_t	+
Infrastructure measured by the amount of electricity produced	IFR_t	+
Tanzanian per capita GDP	$TGDP_t$	+
Tourism prices in Kenya	KR_{it}	indeterminate
Tourism prices in South Africa	SR_{it}	indeterminate
Tourism prices in Mauritius	MR_{it}	indeterminate
Tourism prices in Botswana	BW_{it}	indeterminate
Tourism prices in Zimbabwe	ZW_{it}	indeterminate
Tourism prices in Seychelles	SE_{it}	indeterminate
Past visits	TA_{it-1}	indeterminate
An indicator of a tourist country bordering Tanzania	B	+

Equation (3.8) can be estimated by OLS. However, the observations (number of tourist arrivals) have both a time dimension and cross-section dimension. It is therefore worth exploiting these advantages and use panel data estimation techniques as suggested in the introduction.

3.3.2 Variables and Data Sources

(i) Tourist number of arrivals.

The number of tourist arrivals has been taken from the data provided by MNRT from 1995-2007, for 121 countries across the world. The choice of this period is based on the fact that in the years preceding 1995 there were no records of tourist arrivals by

country, but rather total number of arrivals in aggregate. This made it impossible to consider the years before 1995, panel data method was to be used as intended in this study. Similarly, the 121 countries were chosen, based on the availability of the number of arrivals from these countries for the entire period as well as the availability of the other variables used in the study, such as exchange rates and inflation rates. The composition by region of the tourist countries used in this study was as follows: Africa (40), South America (18), North America (3), Asia (21), Europe (31), and the Middle East (8).With 13 years (1995-2007) of observation and 121 countries, the data set gives a total of 1573 data points.

(ii) Tourism price

Two variables were used as a proxy for tourism prices: relative cost of living and nominal exchange rates.

The relative cost of living is defined as $TZ_{jt} = CPI^{t}_{TZ} / CPI^{t}_{O}\, ER_{jt}$, where CPI^{t}_{TZ}, is the consumer price index in Tanzania, CPI^{t}_{O} is the consumer price index in the tourist country of origin, and ER_{jt} is the nominal exchange rate defined in terms of Tanzanian shillings per a unit of a tourist country's currency. The consumer price indices of all the countries were sourced from IMF (2009), nominal exchange rates were taken from the Italian National Bank a source which is also used by The Economist.

(iii) Tourist income as measured by per capita GDP

The data on per capita GDP was taken from IMF (2009) and measured in USD fixed at 2000 prices.

(iv) Distances

The figures on distances were obtained from the online distance calculator. These figures indicate the shortest ground distance from the tourist country capitals to Dar-es- Salaam measured in miles. The choice of their capital cities is justified by the fact that most departures start from there. Dar es Salaam was chosen because most tourists coming to Tanzania enter the country by air through Julius Nyerere Airport of Dar -es-Salaam.

(v) Advertising Expenditure

Data on advertising expenditure were sourced from four institutions, which in one way or another are responsible for promoting tourism in Tanzania. These are TTB, TANAPA, Ngorongoro Conservation Area Authority (NCAA), and the Zanzibar Commission for Tourism (ZCT). So for each year, the figure is the aggregate sum of the costs of the four institutions fixed at the 2000 consumer prices. It should be noted that this Advertising expenditure (AD) just assess the impact of marketing expenditure by the Government on tourism demand, because it does not cover expenditure by private stakeholders such as tour agents and air companies, who indirectly advertise tourism.

(vi). Infrastructure

The amount of electricity generated during the years 1995 to 2007 measured in kilowatt hours is used as a proxy for infrastructure. The data were taken from Tanzania National Electrical Supply Company (TANESCO).

(vii). Tanzania's per capita GDP

The figures for Tanzania's per capita GDP were taken from World Economic Outlook 2009 measured in USD at current prices.

(viii) Tourism prices at destinations thought to compete with Tanzania

These prices have been calculated as being analogous to tourism prices in Tanzania. The countries perceived to be competitors of Tanzania are Kenya, South Africa, Mauritius, Uganda, Botswana, Zimbabwe, and the Seychelles (TTSS 2001).

3.4. Results and Discussion

3.4.1 Summary Statistics

The summary statistics of the variables used in this study are given in table3.2a.

Table 3.2: Summary statistics of the variables used in the analysis of number of arrivals

Variable	median	Mean	Std. Dev.	Min	Max	CV
TA	422	4253.68	11,616	1	130,823	273.08
TZ	0.02	1.01	3.54	0.00003	30.70485	350.50
ER(TSHS)	46.13	287.84	514.60	0.02171	4369.53	178.78
GDP(Mil.USD)	2,506.06	8961.37	13281.30	54.62	103590.60	148.21
DIST(Miles)	4039	4231.26	2254.99	419.00	9527.00	53.29
AD	1.84	3.74	4.25	0.29	16.00	113.64
IFR	2748	2748.85	664.09	1877.00	4069.00	24.16
TGDP	303.15	313.99	56.15	213.26	428.37	17.88
KE	0.3	13.50	46.62	0.00076	366.97	345.33
SR	3.07	145.43	507.12	0.00572	4839.79	348.70
MR	0.78	55.67	621.23	0.00137	21863.68	1115.91
UG	0.01	1.00	11.95	0.00002	419.49	1195.00
BW	4.4	317.26	3583.78	0.00798	127017.50	1129.60
ZW	0.77	1838.77	22929.38	0.00340	479804.30	1246.99
SE	3.57	250.22	2714.71	0.00497	95745.23	1084.93
border	0	0.06	0.23	0.0	1.0	383.33

From Table 3.2 it can generally be seen that observations in all the variables have very high variation except infrastructure, which has a coefficient of variation of about 24 percent, and Tanzania's per capita GDP with a coefficient of variation of about 18 percent.

As regards the number of tourist arrivals, the coefficient of variation is 273 percent. This is indeed very high and in line with the fact that tourists visiting Tanzania originate from different continents, which have different geographical and social-economic characteristics affecting their movements abroad. This great variation in the number of arrivals across countries suggests that the use of logarithmic transformation as proposed earlier is unavoidable.

3.4.2 Correlation Analysis

The correlation matrix of the variables used in the analysis is provided in Table 3.11 of appendix 3.1.The matrix shows that most of the variables correlate with the variable number of tourist arrivals in the expected manner, as discussed in the literature review and in section 3.3.1. The matrix also shows that most of the regressosrs are highly correlated suggesting that multcollinearlity may be a problem. More details on the correlation can be found in appendix 3.1

3.4.3 Static Panel Regression Results

The estimation of equation 3.8 requires the use of dynamic panel data estimation techniques, because static panel techniques will produce inconsistent results (Cameron and Trived, 2005). However, for comparison purposes, the study gives the static panel regression results first, which exclude the lagged dependent variable.

Before estimating a static panel model an investigation was made to see whether the individual effects and time effects are really important in the overall set up of the data using the pooled OLS. That is seeing whether the regression intercepts vary with individual and/or with time. After including the 121 country dummies, the increment in the explained variation was very high and significant (new Adj-R^2=0.87, F_{cal}=54, P=0.000), which suggests that the individual effects are significant. As for the time dummies their inclusion led to very small increment in the explained variation (new adjusted R^2=0.2862, F_{cal}=0.10<1, p=0.999). This suggests that time dummies do not matter. Another variable found to make an insignificant contribution was trade openness, whose inclusion led to no increment in the explained variation (new Adj-R^2=0.2862, Fcal=0<1).The insignificance of trade openness is consistent with what was established in the time series analysis (appendix 2.1)

Following what had been established before, the fixed effect model and random effect model were estimated and the results are provided in Table 3.3 along side those of pooled OLS. The F-test for the significance of the individual effects is also provided in Table 3.3. The test indicates that the fixed effects are highly significant (p=0.000), confirming what was found earlier. This finding rules out the consideration of the pooled OLS model. A Hausman test was employed for the choice of either a fixed effect or random effect model. The variance covariance matrix of the difference in coefficients between the fixed effect model and the random effect model was not positive definite, although the null hypothesis was accepted (P=0.98). This lack of positive definiteness could be more of an indication that the variances of the efficient model (random effect) are as such not smaller compared to those of the more robust model (fixed effect). An examination of Table 3.3 does not seem to imply that the

standard errors of the random effect model are smaller than those of the fixed model. To be on the safe side a fixed effect model was opted for, because it is always consistent even under the random effect model, making its choice much safer given the said scenario. The estimated results are given in Table 3.3.

Table 3.3: Static Panel regression results of the log of the number of arrivals

	Pooled OLS		Fixed effect		Random effect	
Var	Coeff	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
TZ	1.47	1.62	1.55**	.72	1.53**	.72
ER	-.23**	.108	-.13*	.08	-.16**	.070
GDP	.32***	.029	.19**	.07	.24***	.059
DIST	-.78***	.095	(dropped)		-.73**	.31
AD	-.24	0.99	-.19	.44	-.21	.44
IFR	.84	1.89	.83	.84	.84	.84
TGDP	-.45	2.67	-.58	1.18	-.53	1.18
KE	.58	1.50	.61	.66	.60	.66
SR	1.85	2.83	2.03	1.25	1.96	1.25
MR	-5.9	9.55	-6.28	4.22	-6.14	4.22
UG	-.96	1.41	-.902	.62	-.93	.62
BW	-1.85	3.02	-1.93	1.32	-1.90	1.33
ZW	.21	.23	.22**	.10	.22**	.10
SE	4.32	6.05	4.56*	2.65	4.47*	2.66
Border	3.21***	.27	(dropped)		2.94***	.899
cons	4.42	10.99	-0.561	4.90	4.74	5.51
R^2	0.28		0.10[16]		0.27	
F/χ^2	F=41.51		F=12.89		χ^2 (16)=216.1	
$P>F/\chi^2$	P>F=0.000		P>F=0.000		$P>\chi^2$ =0.000	
N	1573		1573		1573	
Test for Fixed Effects			F(120,1439)=54.52, P=0.000			

NB: *** significant at 1%, **significant at 5% , * significant at 10%

Table 3.3 indicates the OLS estimates, fixed effect estimates as well as the random effect estimates of equation 3.8 which excludes the lagged number of visits. The overall fit of the models in all cases is very significant, with p=0.000.

[16] The R^2 for the fixed effect within is incorrect because its computation doest not take account of the individual effects .An alternative fixed effect estimator that gives the correct R^2 is the LSDV and it shows R^2 to be 0.87 This estimator yields the same value of slope coefficients just as the fixed effect within. Appendix 3.4 presents the LSDV regressions result (Table3.12).

Though it is too early to comment on the results till after the choice of the right model, it is important to highlight some striking results and their implications for further analysis. For example, the coefficient of exchange rate in all models is significant with a negative sign contrary to the expectations. These unexpected signs can be attributed more to multcollinearlity. Multcollinearity should be expected owing to the nature of the variables included in the model. Some of these variables are highly correlated as shown in the correlation matrix (Appendix 3.1).

An attempt was made to assess the extent of the multcollinearlity problem in the regression estimates by calculating the variance inflation factor (VIF) for each explanatory variable. The result indicates that with the exception of GDP, border and distance, the remaining variables have a very high value of VIF (>10) making them not immune to the multcollinearlity problem. More details for these results are provided in Table 3.13 given in appendix 3.5.

The results in Table 3.3 shows that there few coefficients of explanatory variables which are significant, especially for the pooled OLS and Fixed effect model. However, the estimated models assumed the absence of serial correlation among the errors. But the test for AR (1) indicated strong evidence of the presence of first order autocorrelation ($p = 0.0000$). Failure to take account of such disturbances may greatly bias the standard errors leading to erroneous inferences (Cameron and Trived, 2005). Therefore both the fixed effect and random models were re-estimated accommodating AR (1) disturbance. Table 3.4 gives the regression results of the re-estimated models.

Table 3.4: Static Panel regression results of the log of the number of arrivals accommodating an AR (1) disturbances

	Fixed Effect			Random Effect		
Variables in logs	Coefficient	Std.Err	t	Coefficient	Std.Err	Z
TZ	3.11***	.52	5.96	1.55***	.40	3.91
ER	-.010	.10	-0.10	-.11	.082	-1.37
GDP	.076	.091	0.83	.23***	.061	3.76
DIST	Dropped	Dropped	Dropped	-.72**	.30	-2.40
AD	-.069	.23	-0.30	-.30	.23	-1.32
IFR	.798*	.44	1.82	.94**	.45	2.11
TGDP	3.30***	1.04	3.17	-.28	.61	-0.46
KE	1.25***	.42	3.02	.41	.38	1.08
SR	.52	.83	0.62	1.87**	.77	2.42
MR	-4.32*	2.38	-1.82	-5.38**	2.39	-2.25
UG	-2.66***	.56	-4.78	-.95**	.39	-2.47
BW	-1.61*	.92	-1.76	-1.78*	.93	-1.91
ZW	.11*	.062	1.69	.21***	.058	3.62
SE	3.57**	1.40	2.55	3.95***	1.42	2.78
Border	Dropped			2.92***	.87	3.36
cons	-21.79***	1.62	-13.49	3.67	3.87	0.95
R^2	0.08			0.27		
F/χ^2	F(13,1389)=8.62			χ^2 (16)=194.93		
P>F/ χ^2	P>F=0.0000			P> χ^2 =0.0000		
Test for Fixed effects	F(120,1318)=9.64 and p=0.0000					
N	1452			1573		

NB: *** significant at 1%, **significant at 5% , * significant at 10%

Unlike in the previous case, the results indicate that in both models, most of the coefficients of the explanatory variables are significant and appear with their expected signs except for relative cost of living in Tanzania and the advertising expenditure. The relative costs of living in Tanzania is significant, but with an unexpected positive sign while advertisement expenditure is both insignificant and appears with an unexpected negative sign. These perverse signs, as well as the insignificances as explained earlier, could be attributed more to multcollinearlity problems.

One of the remedies for removing multcollinearlity is to drop some of the highly

correlated variables or differencing them (Gujarat 2003). Differencing seemed to be more appropriate in this study for two reasons. First, dropping some variables would deny the study an opportunity to investigate the impact of those variables. Second, differencing is also a remedy for problems associated with [17]non-stationary variables. [18]An alternative fixed effect model which accommodates differencing is the first difference estimator. This estimator, though less efficient than the within estimator, has the advantage of removing possible multcollinearlity and spuriousness of regression results, as it involves the differencing .Another advantage of this model is that it makes the static regression coefficients comparable to those of the dynamic regression, as the later also involves differencing.

Owing to the said advantages, the first difference estimator was opted for instead of the fixed effect within estimator. Estimation was done for six regions of the world. These are Africa, Asia, Europe, the Middle East, South America, and North America. This comparison is important for regional policy formulation. Before presenting the results a Hausman test is indicated for each region for the choice of either the fixed effect or random effect model (Table3.5)..

[17] Levilin panel root test indicates the variables at levels are stationary except number of arrivals and exchange rates, whereas at their first difference they are all stationary. More details are given in Table 3.14 of appendix 3.5.

[18] The fixed effect estimates which were given in table 3.3 and 3.4 are from the within fixed effect estimator which is default one in stata 9. This estimator is more efficient than the first difference. But for the sake of this study the first difference outshines it.

Table 3.5: Hausman test for the choice between fixed effect model and random effect model in six regions of the World

	AFRICA	ASIA	EUROPE	M.EAST	N.AMERICA	S.AMERICA
chi	2.36	17.60	14.77	15.89	1.1	0.32
P-value	0.99	0.17	0.32	0.26	0.99	0.99
V_b-V_B	Is not positive definite	Is not positive definite	Is not positive definite	Is not positive definite	Is not positive definite	Is not positive definite

From Table3.5, the fixed effect model is opted for all the regions despite the fact that the null hypothesis is accepted in each case. This follows the same reasoning as before that the variance covariance matrix is not positive definite, undermining the main reason for choosing the random effect model. Table 3.6 gives the OLS estimator of the first difference, with panel corrected standard errors, which takes care of possible heteresdasticity and serial correlation in the error term.

Table 3.6: First difference fixed effect regression on the log of the number of arrivals across regions

Var	World	Africa	Asia	Europe	M.East	NAmeri	SAmeri
TZ	-1.7*** (0.34)	-2.8*** (0.32)	-1.3 (1.00)	-1.3*** (0.16)	-.62** (0.80)	-0.086 (0.53)	-1.16 (1.73)
ER	0.02 (0.09)	0.099 (0.14)	-1.63 (1.25)	0.07 (0.37)	-0.19 (2.06)	-0.48 (1.70)	1.25 (2.01)
GDP	0.04 (0.09)	-0.07 (0.15)	1.05 (1.26)	-0.005 (0.13)	1.60 (1.83)	0.89 (1.73)	-0.59 (1.72)
AD	1.01*** (0.16)	2.3*** (0.15)	0.29 (0.46)	.66*** (0.10)	0.11 (0.35)	1.41*** (0.13)	-.001 (0.74)
IFR	3.1*** (0.27)	2.5*** (0.26)	3.6*** (0.80)	1.7*** (0.09)	2.89*** (0.64)	-0.29 (0.32)	7.03*** (1.37)
TGDP	5.6*** (0.58)	9.16*** (0.55)	3.36** (1.6)	4.04*** (0.25)	3.95*** (1.19)	1.87*** (0.60)	4.48 (2.81)
KE	5.1*** (0.43)	8.26*** (0.40)	4.6** (1.38)	2.98*** (0.15)	4.6*** (1.41)	3.68*** (0.71)	3.54 (2.21)
SR	-0.98** (-0.98)	-2.7*** (0.45)	-1.03 (1.26)	-1.42*** (0.18)	1.61 (0.99)	1.18*** (0.46)	0.36 (2.23)
MR	-2.1* (1.19)	-3.03*** (1.15)	-2.57 (3.35)	-1.94*** (0.74)	-1.6 (2.6)	-12.7*** (2.37)	0.63 (5.71)
UG	-3.1*** (0.26)	-4.65*** (0.25)	-3.22** (0.92)	-1.86*** (0.05)	-5*** (1.18)	-1.6*** (0.43)	-1.25 (1.53)
BW	-2.7*** (0.45)	-1.77*** (0.42)	-5.1*** (1.39)	-0.28*** (0.09)	-4.8*** (1.21)	-1.00* (0.54)	-6*** (2.28)
ZW	-.01 (0.03)	-.0.06* (0.03)	0.055 (0.09)	0.02 (0.01)	0.18** (0.08)	0.24*** (0.06)	-0.11 (0.16)
SE	5.4*** (0.68)	6.8*** (0.65)	5.8*** (2.18)	3.9*** (0.43)	5.95** (2.56)	9.73*** (1.30)	4.94 (3.39)
_cons	-0.8*** (0.07)	-1.28*** (0.06)	-0.6*** (0.20)	-0.43*** (0.02)	-0.71** (0.16)	-0.36*** (0.08)	-0.73* (0.32)
R^2	0.09	0.17	0.17	0.04	0.30	0.62	0.16
χ^2	561.97	936.52	105.36	25549.9	379.66	1120.75	83.03
$P>\chi^2$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
n	1452	480	252	372	96	36	216

NB: 1. *** significant at 1%, **significant at 5%, * significant at 10%

2. Values in the brackets are the standard errors.

Table 3.6 provides the first difference fixed effect estimates for the world in general and across regions. As can be seen this time round, all the variables appear with their expected signs. Even though these results are not the final ones it can be clearly seen

that the null hypothesis of price factors being more influential than non-price factors is rejected, as it is observed that Tanzania's per capita GDP(TGDP) has a largest and significant coefficient of 5.6 than all the others(first column of Table3.6). On the regional basis, the hypothesis is rejected in Africa, Europe and South America.

3.4.4 Dynamic Panel Regression

The previous estimations were based on static panel data models, which excluded the influence of a lagged dependent variable. The previous models were estimated only for comparison with the dynamic model. The original equation (3.8) includes the lagged number of visits as one of its independent variables. The lagged number of visits is meant to capture the influence of the destination tastes on the current visits. This equation however cannot be estimated using static panel data techniques, because, TA_{it-1} will be correlated with the error term μ_{it} leading to an inconsistent estimate (Cameron and Trivedi, 2005). In order to consistently estimate Equation 3.8, its first difference needs to be considered, as given in (3.9).

$$\Delta \ln TA_{it} = \beta_1 \Delta \ln TA_{it-1} + \beta_2 \Delta \ln P_{it} + \beta_3 \Delta \ln ER_{it} + \beta_4 \Delta \ln GDP_{it} + \beta_5 \ln AD_{jt} + \beta_6 \Delta \ln IFR_{jt} + \beta_7 \Delta TGDP_{jt} + \beta_8 \Delta \ln KR + \beta_9 \Delta \ln SR + \beta_{10} \Delta \ln MR + \beta_{11} \Delta UG + \beta_{12} \Delta BW + \beta_{13} \Delta ZW + \beta_{14} \Delta SE + \Delta \mu_{it} \quad \cdots(3.9)$$

Note that distance and border which are time invariant variables disappear from equation (3.9). An OLS estimate of this equation can be inconsistent because $\Delta TA_{it-1} = TA_{it-1} - TA_{it-2}$ may be correlated with the differenced error term $\Delta \mu_{it} = \mu_{it} - \mu_{it-1}$ .Anderson and Hsiao (1982) propose using $\Delta TA_{it-2} = TA_{it-2} - TA_{it-3}$ as an instrument for ΔTA_{it-1} because ΔTA_{it-2} is directly correlated with ΔTA_{it-1} and

uncorrelated with the error term $\Delta\mu_{it}$ (Cameron and Trived,2005) .But this method does not necessarily yield efficient estimates because it does not use all possible lags[19]. Arrelano and Bond (1991) proposed a generalized method of moments (GMM), which uses all possible lags of the dependent variable as instruments for ΔTA_{it-1} (Cameron and Trived, 2005) .This method yields both consistent and efficient estimates, incase its key assumptions are fulfilled.

The first key assumption is the absence of second order serial correlation in the equation (3.9). The fulfillment of this assumption guarantees $E(\Delta TA_{it-2}\Delta\mu_{it}) = 0$.The second crucial assumption is the exogeneity of the instruments used in GMM. The instruments cover the lagged dependent variables as well as the lagged predetermined variables. The GMM estimator in stata 9 provides the test for the two assumptions. The test for exogeneity of the instruments is known as the Sargan test for overidentifying restrictions. In both tests acceptance of the null hypothesis guarantees the consistency of the GMM estimates. Table 3.7 gives the results of the two tests for six regions of the World.

Table 3.7: The Sargan test and the test for the Second order serial correlation

Test	World	Africa	Asia	Europe	MEast	NAmeri	SAmeric
Sargan	Chi(472) =382 P>z= 0.99	Chi(285) =272 P>z= 0.70	Chi(175) =155 P>z= 0.86	Chi(175) =181 P>z= 0.36	Chi(65) =61 P>z= 0.61	Chi(65) =19.53 P>z= 1	Chi(165) =109 P>z= 0.99
2nd Order Corr	Z=-0.67 Pr>Z= 0.51	Z=-0.79 Pr>z= 0.43	Z=0.23 Pr>z= 0.82	Z=-0.00 Pr>Z= 0.93	Z=1.32 Pr>Z= 0.19	Z=-.48 Pr>Z= 0.63	Z=-0.72 Pr>Z= 0.47

[19] Despite what has been argued, for the sake of comparison the study produces the IV estimates as suggested by Anderson and Hsiao (1982). These are shown in table 3.15 of appendix3.6.

Table 3.7 indicates that for both tests the null hypothesis is accepted, guaranteeing confidence in the regression estimates. The estimates are provided in Table 3.8

Table 3.8: Dynamic Panel regression on the log of the number of arrivals across regions (First difference GMM estimate)

Differenced	World	Africa	Asia	Europe	MEast	N.A	SAmeric
TA	0.74*** (0.03)	0.52*** (0.1)	0.67*** (0.06)	0.70*** (0.05)	0.48*** (0.10)	0.14 (0.24)	0.62*** (0.07)
TZ	-7.2*** (2.5)	-4.15 (3.4)	-10.2** (5.17)	-3.08 (4.83)	-5.17 (7.2)	3.02 (3.89)	-14.43* (8.75)
ER	0.03 (0.12)	-0.17 (0.15)	-2.31 (0.91)	-0.011 (0.20)	-2.73 (2.12)	2.88 (2.49)	2.26 (2.43)
GDP	0.08 (0.13)	-0.22 (0.16)	2.53*** (0.95)	0.03 (0.13)	0.80 (1.67)	-2.76 (2.46)	-1.59 (2.23)
AD	0.46 (0.54)	3.15*** (0.74)	-1.33 (1.12)	0.50 (1.07)	-0.58 (1.54)	1.73** (0.70)	-3.25 (1.92)
IFR	6.5*** (1.3)	3.44** (1.74)	9.12*** (2.7)	2.95 (2.51)	5.78 (3.66)	-0.90 (1.70)	14.34*** (4.55)
TGDP	12.1*** (3.4)	10.97** (4.7)	13.99** (7.11)	6.23 (6.71)	9.94 (9.95)	-1.15 (4.73)	19.20 (12.1)
KE	7.4*** (1.8)	9.57*** (2.5)	8.86** (0.91)	3.87 (3.57)	9.22* (5.48)	1.14 (2.84)	6.02 (6.52)
SR	-0.27 (1.5)	-4.69** (2)	3.88 (3.02)	-0.98 (2.86)	2.80 (4.18)	1.04 (1.82)	5.66 (5.20)
MR	3.65 (4.6)	-6.12 (6.2)	9.73 (9.63)	-0.88 (9.04)	5.787 (13.06)	-15** (5.82)	21.93 (16.3)
UG	-3.8** (1.5)	3.3 (2.0)	-6.51** (3.12)	-1.78 (2.92)	-7.51* (4.55)	0.91 (2.22)	-5.58 (5.35)
BW	-7.27*** (2.6)	-0.13 (3.4)	-15.*** (5.32)	-2.11 (4.98)	-10.43 (7.51)	1.15 (3.49)	-20.27** (9.05)
ZW	-0.06 (0.01)	-0.09 (0.14)	-0.03 (0.21)	0.015 (0.20)	0.12 (0.28)	.22* (0.13)	-0.099 (0.35)
SE	7.49*** (2.2)	8.90*** (3.0)	9.47** (4.54)	4.92 (4.24)	8.48 (6.64)	6.97* (3.75)	6.25 (7.71)
_cons	-1.5*** (0.4)	-1.6*** (0.5)	-1.57** (0.72)	-0.74 (0.68)	-1.24 (0.99)	-0.17 (0.46)	-1.999 (1.22)
Chi2	648.36	194.01	255.73	220.31	112.04	40.63	123.69
P>chi2	0.0000	0.000	0.000	0.000	0.000	0.0000	0.0000
N	1331	440	231	341	88	33	198

NB:1. *** significant at 1%, **significant at 5%, * significant at 10%
2. Values in the brackets are the standard errors.

Table 3.8 provides the dynamic panel regression results. These are the ultimate results of the study, based on equation (3.9).The table shows that almost all the findings from the static panel regression are valid. With the exception of marketing expenditure and tourism prices in Zimbabwe and Mauritius, the general influence of the rest of the variables remained the same. The inclusion of previous visits appears to affect the size of the elasticities of the covariates, but not the direction of their impact.

The results indicate that the influence of previous visits on current visits is positive and highly significant for the world in general and in all other regions except for North America. However, its elasticity is not the highest among all (0.74). This points to the need by stakeholders in the sector to identify tourists' specific requirements and areas dissatisfactions and deal with them accordingly so as to make them better ambassadors in advertising Tanzania overseas. The stakeholders here range from private tour agents, hotel managers, the Airport authority and the immigration department, to camp sites managers, TANAPA administrators and the government in general

The study reveals that an increase in the relative cost of living in Tanzania reduces the number of tourists coming to Tanzania, with an elasticity of 7.2 percent. This is an indication that tourists visiting Tanzania are highly sensitive to prices as measured by the relative cost of living. This finding implies that lowering our cost of living might be a priority of the Government, particularly taxation on tourism-related goods, such as heavy value added tax on hotels and restaurants preferred by tourists should be done with a great care.

The study revealed two key competitors to Tanzania in the East and Southern African

region. These are Kenya with an elasticity of 7.4 percent and the Seychelles with an elasticity of 7.5 percent. These results are in line with those of the static case. The findings on Kenya as the strongest competitor to Tanzania are consistent with those of (TTSS, 2001) and of the general literature. The price elasticity on arrivals to Tanzania is as high as in these two countries. That is to say Tanzanian policy makers have to be equally worried about price fluctuations in both Tanzania and in these two countries. For example, a decline in the cost of living in Kenya by 1 percent would reduce the number of tourist arrivals in Tanzania by 7.44 percent, where as a 1 percent rise in the cost of living in Tanzania would cut the number of arrivals by 7.28 percent, which is slight lower than the cut caused by a drop in cost of living in Kenya. Therefore this study suggests the need to harmonize Tanzania's tax policy on tourism goods with that of Kenya if Tanzania is to compete effectively with Kenya.

Nevertheless, Kenya's competition is less stiff as regards tourists from the Middle East, while competition from the Seychelles is less stiff as regards tourist from North America. This is a chance for Tanzanian promotion authorities to seize this opportunity and intensify its promotion campaigns in these regions rather than concentrating on the traditional European markets.

The study also reveals interesting results that most of the countries in the region are not competitors to Tanzania, but rather complementary to Tanzania tourism, contrary to the belief of TTSS (2001).With the exception of Zimbabwe, which tends to offer a weak competition in the North American market, the rest which are South Africa, Uganda, Botswana and Mauritius are complementary to Tanzania.

The implication of these this result is that a tourist visiting any region especially a tourist on leisure and recreation would rarely visit a single country. To maximize his visit he would rather visit all the countries within the vicinity of the first destination, unless convinced that neighbouring countries have nothing new to be seen. This will depend on the variety and quality of tourist attractions among the countries in the region. For example, a tourist visiting Kenya or Tanzania would rarely be tempted to go to Uganda, whereas the opposite would not be case, because Tanzania and Kenya can offer the more than Uganda. One of the possible strategies of Tanzanian policy makers should be to strengthen marketing campaigns in neighbouring countries to try to convince the tourists visiting them of the peculiarity of Tanzania's tourist attractions.

The preceding discussion about the pattern of tourist visits can be further supported by the TTSS tourist surveys of 2007 and 2008. [20]Out of 3042 tourists interviewed in 2007, 50% of them visited Tanzanian alone, while the remaining 50% were to visit or had already visited neighbouring countries before coming to Tanzania on the same trip. Out of those who visited other countries, 26% had Tanzania as the country of their first visit. This is about 13% of all visits to Tanzania. Likewise in 2008, out of 8789 interviewed tourists, 48% visited Tanzania alone, while 52% of them were to visit or had already visited other countries before coming to Tanzania on the same trip. Out of those who visited other countries, 27% had Tanzania as the country of their first visit. This is about 14 % of all visits to Tanzania.

[20] The statistics given in this passage are from my own calculations based on the surveys done by TTSS in 2007 and 2008.

These statistics on tourists' travel patterns in the region not only provide evidence of destinations' interdependence but also suggest that Tanzania has a challenge to overcome the situation, because most tourists visiting Tanzania and the neighbouring countries (about 74%) do not start their tours in Tanzania. Tanzania needs to improve its infrastructure and in particular, increase the number of direct flights from various regions of the world.

Tanzania's per capita GDP is the most influential factor on the number of tourist arrivals, with an elasticity of about 12 percent leading again to the rejection of the null hypothesis that price factors are more influential than non-price factors in determining the number of tourist arrivals. The joint test of the influence of price factors against that of non-price factors confirmed that non-price factors are more influential than price factors on the coming of tourists from all regions except North America and South America. The test further confirmed that the individually insignificant tourism prices were also jointly insignificant. However, this does not mean that price factors are unimportant. As can be seen, the cost of living in Kenya and the Seychelles are the second most influential determinants after the Tanzania's per capita GDP. In other words, Tanzanian policy makers, apart from focusing on improving the infrastructure and the country's economy need also to pay attention to what is happening in Kenya and the Seychelles.

The positive influence of Tanzania's per capita GDP is an evidence of the reliability of the gravity models in predicting the flow of tourists from one country to another. On the contrary, the incomes of tourists' countries are generally not significant except for tourists from Asia. This result suggests that tourists' income is not such an important

determinant of their coming to Tanzania.

As to whether the tourism sector in Tanzania could be affected by the recent global financial crisis, it would seem that tourists' incomes would have little effect. However, there two main ways through which the country could suffer. First via the rise in the general cost of living in the entire region, because most of the countries are complementary to Tanzania and second would be if some of the countries in the region have their arrivals very sensitive to income.

Regarding advertising expenditure, the dynamic result gave no conclusive evidence of its influence on the number of tourist arrivals. Although it was generally insignificant, it was significant for tourists from North America and Africa. The significance of the advertising expenditure in North America and Africa is unique in the sense that most studies have found no conclusive evidence of the positive impact of marketing expenditure on tourism demand (Munoz, 2004).

Regarding trade liberalization the study has found no conclusive evidence of its influence on the coming of tourists. The influence of trade liberalization was assessed using dummies for 1986 and trade openness which was measured by the ratio of trade to GDP (appendix 2.1& section3.4.3). An attempt was made to adjust the trade openness measure using the ratio of exports to GDP and imports to GDP separately. None of these yielded conclusive evidence of the influence of trade liberalization on the coming of tourists. These results could imply that the measures used failed to capture the influence of trade liberalization rather than that trade liberation has no influence on the coming of tourists.

Distance and border could not be assessed in both the fixed effect model and the GMM because they are time invariant characteristics. In order to assess the impact of distance and border or any other time invariant characteristics one can use the approach of Eita and Jordaan (2007) as borrowed from Martinez-Zarzoso and Nowak-Lehmann (2003) According to this approach the impacts of time invariant characteristics under fixed effect method/GMM can be assed by regressing the individual specific effects against the time invariant characteristics such as distance and border. Borrowing this idea, Table 3.9, provides the regression results on the log of countries' specific effects against distance and border. The country-specific effects have been predicted from the LSDV regression provided in Table 3.12 of appendix 3.5.

Table 3.9: OLS regression on the log of the countries' specific effects against distance and border (Regional Comparison)

	World	Africa	Asia	Europe	MEast	N.A	SAmeric
Distance	-0.70***	-1.3***	0.26	3.85***	-4.8***	-4.2	-4.2***
Border	2.8***	2.65***	N/A	N/A	N/A	N/A	N/A
R^2	0.22	0.63	0.0014	0.06	0.32	0.06	0.09
F	225.18	431.23	0.39	26.05	47	2.36	26.31
P>chi2	0.0000	0.000	0.53	0.000	0.000	0.13	0.0000
N	1573	520	273	403	104	39	234

The results from Table 3.9 reveal that distance and border are in general very significant. With the exception of Europe, where the longer distance of travel positively influences the coming of tourists, in the rest of the regions the higher the distance of travel the lower the number of arrivals. As regards borders, those countries bordering Tanzania influence positively the coming of tourists. These results suggest the need by the government and other stakeholders to reduce the cost of travelling to

Tanzania. One way of reducing the cost is to involve as many international air companies as possible, and especially those operating direct flights, which would make the overall cost of travel a lot cheaper.

As said before, the inclusion of lagged number of visits would enable the estimation of long-run demand elasticities by setting $TA_{it-1}=TA_{it}$ in the estimated equation 3.9 and solving for TA_{it}. This means that every coefficient of the explanatory variable in equation 3.9 has to be divided by $1-\beta_1$ where β_1 is the coefficient of TA_{it-1}. Table 3.10 gives the long-run demand elasiticies alongside the short-run ones, for comparison purpose. The long-run elasticities are shown in italics. These long-run elasticities are useful for long-term policy planning.

Table 3.10: Long-run demand elasticities for tourism demand

Differenced	World	Africa	Asia	Europe	MEast	N.A	SAmeric
TZ	-7.2*** *-27.7****	-4.15 *-7.98*	-10.2** *-15.2***	-3.08 *-4.4*	-5.17 *-10.77*	3.02 *21.57*	-14.43* *-23.27**
ER	0.03 *0.12*	-0.17 *-0.33*	-2.31 *-3.45*	-0.011 *-0.02*	-2.73 *-5.69*	2.88 *-20.57*	2.26 *3.65*
GDP	0.08 *0.31*	-0.22 *-0.42*	2.53*** *3.78****	0.03 *0.04*	0.80 *1.67*	-2.76 *-19.71*	-1.59 *-2.56*
AD	0.46 *1.77*	3.15*** *6.06****	-1.33 *-1.99*	0.50 *0.71*	-0.58 *-1.21*	1.73** *12.36***	-3.25 *-5.24*
IFR	6.5*** *25****	3.44** *6.62***	9.12*** *13.61****	2.95 *4.21*	5.78 *12.04*	-0.90 *-6.43*	14.34*** *23.13****
TGDP	12.1*** *46.5****	10.97** *21.10***	13.99** *20.88***	6.23 *8.90*	9.94 *20.71*	-1.15 *-8.21*	19.20 *30.97*
KE	7.4*** *29****	9.57*** *18.40****	8.86** *13.22***	3.87 *5.53*	9.22* *19.21**	1.14 *8.14*	6.02 *9.71*
SR	-0.27 *-1.04*	-4.69** *-9.02***	3.88 *5.79*	-0.98 *-1.40*	2.80 *5.83*	1.04 *7.43*	5.66 *9.13*
MR	3.65 *14.04*	-6.12 *-11.77*	9.73 *14.52*	-0.88 *-1.26*	5.787 *12.06*	-15** *-107.14*	21.93 *35.37*
UG	-3.8** *-14.6***	3.3 *6.35*	-6.51** *-9.72***	-1.78 *-2.54*	-7.51* *-15.65*	0.91 *6.50*	-5.58 *-9*
BW	-7.27*** *27.1****	-0.13 *-0.25*	-15.*** *-22.4****	-2.11 *-3.01*	-10.43 *-21.73*	1.15 *8.21*	-20.27** *-32.69***
ZW	-0.06 *-0.23*	-0.09 *-0.17*	-0.03 *-0.04*	0.015 *0.02*	0.12 *0.25*	.22* *1.57*	-0.099 *-0.16*
SE	7.49*** *28.8****	8.90*** *17.12****	9.47** *14.13***	4.92 *7.03*	8.48 *17.67*	6.97* *49.79**	6.25 *10.08*
cons	-1.5*** *-5.8****	-1.6*** *-3.08****	-1.57** *-2.34***	-0.74 *-1.06*	-1.24 *-2.58*	-0.17 *-1.21*	-1.999 *-3.22*
Chi2	648.36	194.01	255.73	220.31	112.04	40.63	123.69
P>chi2	0.0000	0.000	0.000	0.000	0.000	0.0000	0.0000
N	1331	440	231	341	88	33	198

NB: *** significant at 1%, **significant at 5%, * significant at 10%

Table 3.10 indicates that the long-run demand elasticities and the short-run elasticities have the same signs. However, the long-run demand elasticties are much bigger than the short-run demand elasticities. It is worth noting that in the case of Europe, the postulated determinants for number of tourist arrivals does not matter in the long run. The same appears to be the case for the Middle East. These results suggest that the estimated model is more robust for European case; because truly in the long run nothing should really matter because the equilibrium shall have been achieved.

3.5 Conclusion

The study has investigated the determinants of the number of tourist arrivals in Tanzania, dividing them into price and non-price determinants. Using panel data regression analysis, the study has revealed the key determinants of number of tourist arrivals. These include improvement in, the Tanzanian economy, infrastructure and marketing expenditure. Others are the destination taste and travelling costs. All these might be regarded as non-price factors.

Price factors, particularly tourism prices in competing destinations such as Kenya and the Seychelles, also have an effect on tourist arrivals in Tanzania. Quite surprisingly, the study has revealed that South Africa, Mauritius, Uganda and Botswana are generally not competitors to the Tanzanian tourism market, but complementary to Tanzania. A decline in tourism prices in those countries could lead to coming of more tourists to Tanzania and vice-versa.

CHAPTER FOUR

DETERMINANTS OF TOURIST SPENDING

4.1 Introduction

A country's tourism revenue is a product of not only the number of tourist arrivals but also of their per capita expenditure. Even though a country may receive as many tourists as possible it may end up receiving relatively less revenue than other countries unless it undertakes strategies which accommodate the determinants of tourist spending. Moreover owing to the need to protect the environment, it is important to maximize tourist per capita expenditure and reduce/maintain the number of arrivals and/or length of a tourist stay at a destination. To achieve this objective one needs to establish the determinants of tourist spending.

In Tanzania, similar studies were done by TTSS in 2001, 2004, 2005 and 2006. However these were designed to estimate the country's tourism revenue and not to find the determinants of tourist per capita expenditure. TTSS (2001) used the WTO model to estimate the country's total revenue by grouping the tourists into three hierarchal segments thought to be influence tourist per capita expenditure. These were tourist country of origin, purpose of visit, and travel arrangements. For each country tourists were grouped according to the purpose of visit (4) and within each purpose of visit tourists were grouped according to travel arrangements (2). In total there were 8 categories for each country. In each category tourist revenue was estimated as a product of the number of arrivals, average expenditure per person per day and average length of stay. The consideration of those variables by TTSS as key determinants of tourist expenditure was based mainly on the descriptive analysis and previous literature.

This study seeks not only to verify the validity of the variables included by TTSS in the estimation of a country's revenue using conventional econometric methods, but also to assess the impact of some other variables not considered by TTSS in the expenditure model. This assessment could possibly lead to the betterment of the TTSS expenditure model. The TTSS expenditure model is basically the WTO model, which assumes that tourist daily spending is constant across the days of stay. This assumption was taken for granted by TTSS. Unfortunately this assumption is not valid in Tanzania. This study has found that tourist daily spending declines with length of stay, implying that the TTSS expenditure model overestimates the country's tourism revenue. This study has proposed an improvement in the model to accommodate variations in tourist daily expenditure.

This study has considered several variables as possible determinants of tourist per capita expenditure. It derives most of these variables from the surveys by TTSS of visitors to Tanzania in 2001, 2007and 2008 as well as a few macro variables from other sources. Most of these variables can be divided into demographic, trip related characteristics, and destination attributes. The study compared the three groups of variables by testing the hypothesis that destination attributes are more influential than the trip-related and demographic characteristics of the tourists.

Among the variables included in this study is the season of travel, which was made possible because the TTSS surveys were done during different periods in a year. The inclusion of the season of travel in the model made it possible to study the influence of destination attributes on tourist expenditure by using seasons as proxies for different destinations. The logic behind this was that different seasons have different tourist

activities and different weather conditions, which would have an effect on tourists' destinations. This is especially so for Tanzania, which has long tourist seasons and many activities (e.g. game viewing, bird watching) which offer different sightings with the changing seasons (MNRT, 2002).

Details of a tourist observed in a past year could not be retrieved as no tourist was personally identified by TTSS, the reason being its desire to encourage them to reveal confidential information, particularly on expenditure (TTSS 2001). Even if they had been identified, there was no guarantee that they would have visited the country again, and if they had visited calling on them again would have been very difficult, because the TTSS surveys interviewed just some of the visitors and not every visitor who came to Tanzania. Not being able to identify the details of a previous visitor has been a common phenomena in most tourists surveys across the world and as a consequence the data used in a microanalysis of tourism demand across the world has been primarily cross-sectional (see for example Soest and Kooreman, 1987; Sampo and Perez, 2000; Jang et al., 2005; Wang et al., 2006; Litvin, 2007; and Tosun et al., 2007).

Surprisingly most of these studies conducted across the world do not address the problem of endogeneity of the regressors, whereas endogeneity is a more or less expected phenomenon in a cross-section model (Cameron and Trivedi, 2005). The reason is unclear. This study has attempted to examine the endogeneity of the regressors and its implication for the study objectives.

The remainder of this chapter is organized as follows: Section 4.2 reviews theoretical literature on the determinants of tourist per capita expenditure as well as empirical

literature. Section 4.3 describes the methodology, Section 4.4 presents and discusses the estimation results and Section 4.5 concludes the chapter.

4.2. Literature Review

4.2.1Theoretical Literature Review

The conventional determinants of demand for a commodity include the price of the commodity and income of a consumer other things being equal. Equally well these factors should apply as determinants of tourist spending. However in this particular case, the only conventional factor which may apply is tourist income, since tourists visiting the same destination face the same prices. Other factors may include taste and preferences, perception, the environment, society, the situation, etc. In fact these other factors are arbitrary and may not be fully accounted for

The Theory of Reasoned Action (TRA), which attempts to explain why a person pursues an action, can provide a theoretical basis for understanding these other factors. The theory was formally developed by Fishbein and Ajzen(1975). According to Fishbein and Ajzen (1975), TRA has three basic constructs: Behavioural intention (BI), Attitudes (A) and Subjective Norms (SN). TRA suggests that a person's behavioural intention depends on the person's attitude to the behaviour and subjective norms (BI=A+SN). Attitudes consist of beliefs about the consequences of performing the behaviuor multiplied by his or her evaluation of these consequences. Subjective norms are seen as a combination of the perceived expectations of the relevant individual or groups along with intentions to meet these expectations. The two predictive constructs, attitudes and subjective norms, will be influenced by situational variables such as the behaviour observability and personal characteristics and

preferences (Fishbein and Ajzein, 1975).

This theory can be applied to various fields including a study on consumer behaviour When this theory is adopted in relation to consumer behavior, it translates into purchase intention theory in as much as a purchase is also an action (Barros and Correia, 2007; Belleau, 2007; Sheppard et al., 1988). Purchase intention is a function of the attitude towards the purchase, as well as social norms (Barros and Correia, 2007), where attitudes and social norms are as defined in the TRA.

In the context of TRA, the determinants of tourist expenditure at a destination, are to be thought about by conceptualizing the relationship between individual expectations/beliefs of a particular expenditure and his own evaluation of these outcomes, as well as what he expects of society, the environment or situation. For example, for a tourist to buy an item at a destination, apart from having the income he will have to perceive the importance or necessity of such an item. This perception will depend on his /her individual evaluation of the item, as well as the environment in which it is sold, such as its hygienic appearance, and quality (Tosun et al., 2007, Sangpikul 2008).

Unfortunately these attitudes (expectations and beliefs) to the product as well the social norms can not always be measured, but rather are reflected by the individual make-up, characteristics of the purchase and the environment/society. Decrop and Snelders (2005), as quoted by Nzuki (2006) describe this better in the following words: "Fundamentally various studies, e.g. Ball et al., 1992, Dickson, 1992, Hawkins, Best and Coney, 1995) posit that purchase behaviour is influenced by the interaction

of individual and situational characteristics, as well as the product characteristics, with these three characteristics being determined by the environment (social, cultural and geographic)"

The study of tourist purchase behaviour has led to various theories on the tourist decision process. Barros and Correia (2007) applied the theory of the tourist decision process as given by Mathieson and Wall (1984), as the basis for discerning the hypothetical determinants of tourist length of stay. The theory is described be consisting of four determinants of the tourist decision process: tourist profiles, trip features, travel awareness and destination characteristics.

In line with this theory, and the TRA discussed above, we can place the determinants of tourist expenditure at a destination in five categories, based on the interaction among individual characteristics, product characteristics and the environment (social, cultural, geographic). First are the socio-demographic characteristics which stem from individual characteristics. Key variables here would include income, age, sex, education, and nationality (Wang et al. 2006, Jang et al. 2005, Sangpikul, 2008). Second are the trip-related characteristics which stem from the product characteristics. These cover travel party, length of stay, purpose of visit, travel arrangements and travel distance. Third are the variables relating to travel awareness, which stem from an individual's past experience. These include, visiting rates and familiarity with the destination/repeat visits (Clottey and Lennon, 2003).Sometimes the distinction between the trip awareness factor and a trip-related characteristic is arbitrary. For example, Wang et al. (2006) refer to trip awareness variables such as the rate of visits, as merely trip-related characteristics. Fourth are the destination characteristics, which

cover tourist satisfaction with the services and hospitality and destination attractiveness (Tosun et al., 2007, Sangpikul, 2008, Vassiliads, 2008). These characteristics are also known as pull factors (Wang et al., 2006, Sangpikul, 2008). Fifth are the psychographic characteristics, which further explain why a person would prefer to travel to a particular destination rather than purchase an item (Wang et al., 2006 Sangpikul, 2008). These factors are also called push factors (Wang et al., 2006 Sangpikul 2007).

4.2.2. Empirical Literature Review

Sampol and Perez (2000) analyzed the determinants of tourist expenditure in the Balearic Islands and found that people at higher professional levels spent more. The results also indicated that expenditure differed across nationalities, with German visitors spending more than British visitors. Other variables associated with high tourist expenditure were age, i.e. those over 30 years old and people having a positive opinion of their holiday.

TTSS (2001) established descriptively that visitors' expenditure in Tanzania depended on their country of origin, length of stay, purpose of visit and travel behaviour. According to the study, tourists from countries with higher income, tourists staying longer, business visitors, and tourist on package tours were associated with high per capita expenditure. Nevertheless, the study could not establish the relative importance of the variables.

Suh and Gartner (2004) studied the preferences of travellers from Europe, North America and Japan in Seoul. Among the key findings were that travellers from distant

areas (Europeans and Americans) tended to value intangible goods (local culture) more than those from nearby places (Japaneses), who preferred tangible goods (shopping). Probably this difference in preferences between the two groups of tourists could have been attributed to the fact that tourists from nearby have roughly identical cultures to that of the destination and therefore they have no interest in the destination's culture.

Jang et al. (2005), studied the effects of travel activities and travel season on tourist expenditure in France. Using path analysis, the study found that season of travel and travel activity significantly affects tourist per capita expenditure having controlled for income. To maximize revenue, policy makers in the destination areas are advised to identify the type of activities and seasons which contribute most to tourist expenditure at the destinations.

Wang et al. (2006) studied determinants of tourist expenditure in destinations in the western region of the USA. The study used three sets of variables; socio-demographic, trip-related and psychographic variables. Applying OLS estimation technique, the study revealed that among the three sets of variables, income and trip related characteristics were the most influential variables affecting tourism demand. The trip-related variables were number in travel group, number of adults, number of children, number of visits, length of stay and travel distance. Among the trip-related variables, travel distance, length of stay, and number of adults in the travel group were found to affect expenditure positively. The positive influence of distance on expenditure appears to be contrary to what was found by Suh and Gartner (2004), who found that tourists from nearby tend to prefer shopping more than intangible goods. Other

demographic variables studied were gender, age, marital status, and having children under 17. Out of these, only age was significant and affected expenditure negatively, contrary to the findings of most of the previous studies.

Litvin (2007) studied the travel behaviours of first time visitors and repeat visitors, with a focus on the impact which repeat visitors have on paid for attractions in Charleston, South Carolina. The study found that higher tourist expenditure is associated with first-time visitors to paid for attractions. However, in relation to most non-attraction-related activities, such as shopping, walking through the historic district, visiting an art gallery or browsing antique district, the two segments were similar.

Dwyer and Forsyth (2008), while studying the measure for tourism yield and visitor markets to be targeted for Australia, found that tourists' expenditure varies with nationality. The study showed that tourists who injected most expenditure in terms of the total per trip came from the UK, other European countries, the USA and Japan, and those who injected the least came from Canada, Taiwan, Thailand, and other Asian countries. The study also found that the highest spending per visitor night was associated with visitors from Singapore, Hong Kong, Malaysia and Japan and lowest visitor daily expenditure was associated with visitors from other Asian countries, the UK, and Canada.

One may summarize the variables determining tourist per capita expenditure at a destination as being age, gender, income, profession/education level, nationality, length of stay, travel arrangements, purpose of visit/travel activity, travel party size,

number of adults, number of children, frequency of visits/familiarity with the destination, and destination attributes, such culture, season of travel and satisfaction. Among these, age, income, higher profession level, length of stay, travel party size, number of adults and number of children, tour package, business visits, satisfaction and cultural link are expected to have a positive influence on total per capita expenditure. Destination familiarity may have a negative influence on expenditure, while distance, gender and season of travel could be indeterminate. All these variables can accordingly be grouped into socio-demographic and trip-related characteristics, and destination attributes. It is also important to note that none of these studies addressed the possibility of having endogenous regressors.

4.3 Methodology

4.3.1 Model Specification

Following Naude and Sayaaman's specification of the utility function (see equation (3.3) and (3.8) of chapter 3), the study considers the demand function for tourism goods in Africa resulting from the second stage of utility maximization by a tourist visiting a j^{th} destination in Africa. The resulting demand function is given by equation (3.7) of chapter 3, which is also reproduced as equation (4.1a) below;

$$q_{it} = x_{ijt}\beta + c_i + \varepsilon_{it}, \qquad (4.1a)$$

Where

q_{it} = demand for international tourism by origin i for Tanzanian destination.

x_{ijt} = a vector of explanatory variables with both cross-section and time variation including both price and non-price factors.

c_i = qualitative factors in origin i, intended to capture unobserved individual specific factors.

Unlike in Chapter 3, where equation (4.1a) was considered for the number of tourist arrivals as a response variable here it is considered for tourism expenditure in Tanzania. The response variable is expenditure per person belonging to different years of observation and tourists' different countries of origin. As said before, no past details are available for an individual tourist. Therefore this time around panel data estimation cannot be used with an individual tourist as the subject. In other words, if we ignore a tourist's country of origin as well as his/her year of visit and pool the data together we can rewrite equation (4.1a) in a more detailed way while taking into consideration the variables discussed in sections 2 and 3 of the literature review as follows:

$$y_i = \beta_o + \beta_1 Age_i + \beta_2 Females_i + \beta_3 Income_i + \beta_4 Er + \beta_5 El + \beta_6 Childno + \beta_7 Childpresence + \beta_8 Lstay + \beta_9 Tarra + \beta_{10} VistOth + \beta_{11} VistFRD + \beta_{12} VistLSR + \beta_{13} VistBSN + \beta_{14} Adultno + \beta_{15} Tpartyno + \beta_{16} Frvist + \beta_{17} Fadest + \beta_{18} Dist + \beta_{19} Nosites + \beta_{20} Peak + \beta_{21} \text{Price} + Africa + Asia + Europe + MEast + NAmerica + SAmerica + year1 + year2 + year3 + \varepsilon \quad (4.1b)$$

where

y_i = expenditure per day by an i^{th} person

Age= 1 if the respondent's age group is <18, 2 if age group is 18-35, 3 if age group is 36-55 and 4 if age group is over 55 .Its coefficient β_1 is expected to be positive.

Females=number of females in the travel party. Its coefficient β_1 is indeterminate.

According to Wang et al. (2006), there is no reason for gender to affect tourism expenditure since much of the travel is a group activities especially for families. However men and women may differ in preferences and motivation

to travel, which may have varied effects on spending pattern

Income =tourist's income level. In this study per capita GDP of a tourist's country of origin was used as a proxy for income, because tourists were not asked about their incomes during the surveys. They would have been very reluctant to state their actual incomes. Its coefficient β_3 is expected to be positive.

Er =Exchange rate between Tanzania and the tourist country of origin expressed in terms of units of Tanzanian shillings per unit of a foreign currency. This variable, as GDP is meant to assess the impact of income on tourist spending. Its coefficients β_4 is expected to be positive

El=1 if a tourist is from an English speaking country and 0 if otherwise. Its coefficient β_5 is indeterminate .The ability to communicate fluently in English may lead to the ability to bargain, thereby reducing the amount they spend, particularly tourist on non-package tours. On the other hand, the ability to communicate easily may lead to more interaction with the locals at the destination and /or to engage in more tourist activities. This engagement may lead to more spending (Jang et al. 2006).

Childno=number of children in the travel group. According to Wang et al. (2006), past literature has found that the number of children negatively affect food expenditure. On the other hand the number of children may mean a greater requirement for rooms and other facilities which may lead to more spending. Wang et al. (2006) concludes that there is no evidence from past studies of the definite sign of the number of children. Therefore its coefficient β_6 is indeterminate.

Child presence=1 if there is at least one child in the travel party and 0 otherwise. The

inclusion of this variable is justified by the fact that while the intensity as captured by the number of children may not be significant, the mere presence or absence of a child may matter. Therefore its coefficient β_7 expected to have the same sign as children number.

Lstay= tourist length of stay in Tanzania measured in days. Its coefficient β_8 is indeterminate. The longer a person stays at a destination the more familiar he becomes with the environment and probably may spend less per day which might also be contributed to by the dwindling budget. On the other hand the longer he stays might imply his increased satisfaction with the destination and hence spend more per day. Nevertheless, the coefficients β_8 should certainly be positive when an individual's total expenditure is considered rather than expenditure per day.

Tarra= 1 if a tourist is on package tour and 0 if a tourist is on non- package tour. Its coefficient β_9 is expected to be positive. This expectation is built on the intuition that tourists on package tours are wealthier than those on non-package tours (TTSS, 2001).

VistOth= 1 if a tourist is visiting for other purposes, which excludes friends, leisure and business. Its coefficient β_{10} is indeterminate. The literature is not explicit on this variable.

VistFRD = 1 if a tourist is visiting friends and relatives, 0 otherwise. Its coefficient β_{11} is indeterminate.

VistLSR= 1 if a tourist is visiting for leisure and recreation, 0 otherwise. Its coefficient β_{12} is indeterminate.

VistBSN=1 if a tourist is visiting for business purpose. Its coefficient β_{13} is expected to be positive.

Adultno= Number of adults in the travel group. Its coefficient β_{14} is indeterminate. It may be negative if the advantages of a group is utilized by sharing some of the bought goods and facilities, such as rooms and vehicles .

Trpartyno=1 if there are at least 2 members in the travel party, 0 if otherwise. Its coefficient β_{15} is expected to have the same sign as that of the number of adults if the advantages of group spending works. Wang et al. (2006) found that the coefficient for travel party number positively influence per capita expenditure. Its coefficient would certainly be positive if the total expenditure is considered rather than expenditure per person per day.

Frvist = 1 if a tourist has visited at least one African country before Tanzania, 0 if otherwise. Its coefficient β_{16} is expected to be positive. A frequent visitor in this context is likely to be an explorer or a businessman, for these are the kind of people who travel frequently. Therefore he is likely to be a visitor on holiday and /or business, who is presumed by the literature to spend a lot more (Sampol and Perez 2000; TTSS, 2001).

Fadests = 1 if a tourist has visited Tanzania at least once before the current visit and 0 if otherwise. Its coefficient β_{17} is indeterminate. Although Dwyer and Forsyth (2008) found that repeat visitors spend less on paid for attractions than new ones, this cannot lead to the conclusion that repeat visitors spend less daily. Moreover Dwyer and Forsyth (2008) also found that there is no significant difference between repeat and the new visitors as regards non-paid for attractions. On the other hand, those familiar with the destination could be

visitors on a business trip, who are likely to spend more per persona as posited by the literature(TTSS,2001 Sampol and Perez 2000).

Dist = the shortest air distance between Tanzania and the tourist country of origin measured in miles. Its coefficient β_{18} is indeterminate. According to Wang et al.(2006) tourists coming from far away tend to do a lot more shopping than those from nearby, leading to the positivity of β_{18}. But this positivity will depend on the magnitude of other factors. Distant tourists especially those on non package tour may also be worried by the cost of travel and freight charges and hence reduce their expenditure on tangible goods. Sometimes distant tourists are interested in culture than tangible goods (Suh and Mcavoy, 2004), implying that such a tourist may end up spending less than the ones from nearby.

Nosites=1 if a tourist had visited more than one site and 0 if otherwise. Its coefficients β_{19} is indeterminate, depending on whether the number of sites visited implies a relatively longer length of stay. If a tourist takes a few days to visit several sites then the coefficient will be positive, otherwise the sign of the coefficient will be influenced by other daily experiences of the tourist. On the other hand, its coefficient is certainly positive if total expenditure is considered rather than expenditure per day.

Peak=1 if a tourist traveled during the peak season (July- September) and 0 otherwise. Its coefficient β_{20} is expected to be positive.

Price= the relative cot of living between Tanzanian and the tourist's country of origin measured as price= $CPI'_{TZ}/CPI'_{O} * ER_{jt}$ where CPI'_{TZ}, is the consumer price

index in Tanzanian and CPI_O' is the consumer price index in a tourist's country of origin. Its coefficient β_{21} expected to be negative.

Africa, Asia, Europe, MEast, NAmerica, SAmerica, represents regions' dummies for Africa, Asia, Europe, Middle East, North America and South America.

year1, year2 and year3 are dummies for the year 2001, year 2007 and year 2008.

The summary of this discussion is given in table 4.1 next.

Table 4.1: Aprior direction of the relationship between daily tourist spending and the explanatory variables.

Variable	acronym	Direction
Age of the respondent	age	+
Number of females in the travel party	Females	indeterminate
Income of the travel party	income	+
Tanzania nominal exchange rate against that of the tourist's country	ER	+
Tourist familiarity with the English language	EL	indeterminate
Number of children in the travel party	Childno	indeterminate
Presence of a child in the travel party	Childprese	indeterminate
Tourist length of stay(days) in Tanzania	Lstay	indeterminate
Travel arrangements by a tourist(Package)	Tarra	+
A tourist on Other purposes of visits	VistOth	indeterminate
A tourist visiting Friends and Relatives	VistFRD	indeterminate
A tourist coming for Leisure and Recreation	VistLSR	indeterminate
A tourist on Business Visit	VistBSN	+
Number of adults in the travel party	Adultno	indeterminate
Travel party size(number of people in the travel group)	Trpartyno	indeterminate
Tourist frequency of travelling (not to Tanzania only)	Frvist	+
Tourist familiarity with Tanzania(frequency of visiting Tanzania)	Fadests	indeterminate
Distance between Tanzania(DSM) and the tourist country's capital	Dist	indeterminate
Number of sites visited by a tourist during his stay in Tanzania	Nosites	indeterminate
The peak season of travel by a tourist	Season	indeterminate
Consumer prices in Tanzania relative to that of the tourist's country	price	-

4.3.2 Econometric Issues: Consistency of the Regressors

It is not unusual to have inconsistent estimates in a cross-section model like the one given in (4.1b). The literature identifies two main sources of regressors' endogeneity. These are simultaneity and omission of regressors, which correlates with the regressors in the current model (Cameron and Trvivedi, 2005). As far as tourist spending is concerned, simultaneity is ruled out because spending occurs much later after a tourist has dealt with all the other characteristics (demographic, trip-related and destination characteristics) meaning that spending can in no way cause the happening of these characteristics, whereas the vive-versa is possible.

However, in this study an obvious regressor which is endogenous is the tourist length of stay, not because it is automatically supposed to be so rather because of the way it was used to derive other variables (section 4.3.3).Both tourist per capita daily expenditure and tourist party daily expenditure were derived from the reported total expenditure using tourists' length of stay. Therefore length of stay was not included in the models of the two variables (see section 4.3.1 & Table 4.5). In order to assess the influence of length of stay on the two variables, length of stay was instrumented and results are shown in appendix 4.1.The estimated results were not used for inferences on other variables except on length of stay itself because the instruments used which were the only ones available were very weak.

An examination of regressors' endogeneity in a cross-section model is easier said than done, as it requires an appropriate instrument to be used for testing the endogeneity (Cameron and Trivedi, 2005). An instrument has to be exogenous and strongly correlated with the instrumented variable (Cameron and Trivedi, 2005). The first

condition is a necessary one while the second is important if one intends to use the instruments for inference purposes. When an attempt was made to test for the endogeneity of each of the regressors, only a few of them fully met the first condition and somewhat the second condition. These few regressors were successfully tested to be exogenous. The details of these tests are provided in appendix 4.1.

An assessment of the variance inflation factor (VIF) showed that all of the regressors had a very small VIF (<6) (see Table 4.2 in appendix 4.2). This scenario apart from ensuring that the estimates are free from multcollinearlity (Gujarat, 2003), it also guarantees that even if there is an endogenous regressor among those which were not proved to be exogenous, its endogeneity may not affect the exogeneity of the other regressors and hence of the entire model (Cameron and Trivedi, 2005). In addition, among the regressors which were proved to be exogenous was the peak season, which is a proxy for destination attributes and the tourist travel purposes as measured by a dummy of visit to friends (vistFRD). These two by themselves are enough to successfully test the hypothesis that destination attributes are more influential than the demographic and trip-related characteristics. This is because the dummy for a tourist visiting friends was much more influential than the dummy for the peak season, leading to the rejection of the null hypothesis.

4.3.3 Variables and their Sources

The study used survey data from TTSS. Although there were six years of survey by the TTSS (2001, 2004, 2005, 2006, 2007, 2008), the study mainly used data for the years 2001, 2007 and 2008, which had similar and relatively more explanatory variables than the rest of the years. These three years made a total of 30,782

observations. But only 25,880 observations out of 30,782 observations were used, the rest being excluded due to missing and/or unacceptable values.

The model equation (4.1b) contains 19 variables which are described next. The following variables were obtained directly from the TTSS surveys: age, number of females in the travel party (Females), travel party number (Tpartyno), length of stay by a tourist (Lstay), travel arrangements i.e. package or non package tour (Tarra), tourists visiting friends and relatives (VistFRD), tourists visiting for Leisure and Recreation (VistLSR), tourists visiting for Business purposes (VistBSN) and number of sites visited by a tourist in Tanzania (Nosites).

The following variables were derived from the TTSS survey:

(1) (Expenditure per person per day). The survey data reported only total party expenditure. In order to obtain expenditure per person per night, the total party expenditure was divided by the size of the group and by the number of days the party stayed. There were two types of expenditure, depending on whether a tourist was on a package tour or a non package tour. For tourists on a non-package tour no adjustment was made in the data, but the data was adjusted for tourists on package tour, whose bills are paid by travel agents in their home countries. According to TTSS (2001), the actual expenditure accruing to Tanzania can be found by deducting 10% of the package cost as the amount paid to the agents. The same was done in this study. According to TTSS (2001), after deducting the 10% commission, the international travel fare, was also deducted for tourists whose packages include that.

Another consideration was to set the minimum expenditure and the maximum

expenditure as proposed by TTSS (2001). According to TTSS (2001), meaningful expenditure was regarded as not being less than 10 dollars per person per night ,whereas the maximum expenditure per tourist was set at USD1000. The same was adopted by this study. This was important for two reasons. First is the need to compare the findings from this study with those of TTSS. Second is to minimize extreme values in the data.

Consideration was also made of the categories of expenditures. These were expenditure on food, accommodation, transport and shopping. In order to remove unacceptable values, a minimum of 2 US dollars per person per night was set for each category of the expenditure, which was reported for both tourist on package and non-package tours. As regards the former, what was reported on the categories of expenditure represented merely some extra money spent at a destination; but did not include the entire expenditure as a lot of money had already been spent before arrival .Therefore analysis of the categories of expenditure was done exclusively for tourists on non-package tours.

(2) Frequent vistor (Frvist). A frequent visitor was taken to be a person who had visited at least one other African country before visiting Tanzania. There was a question asking tourists to compare the cost of visiting Tanzania against that of neighbouring destinations such as South Africa and Kenya. A tourist responding to this question was regarded as a frequent visitor. This information was used because there was no question asking a tourist directly whether he/she is a frequent visitor.

(3) Familiarity with the destination (Fadest). A tourist who was not on his first visit to

Tanzania was regarded as being familiar with the country. Unfortunately this variable does not capture the intensity of familiarity.

The following variables were taken from other sources:

(1) Income was proxied by a tourist country's GDP. The figures were obtained from the IMF (2009).

(2) Exchange rate (Er) was taken from the Italian Bank, a source also used by the Economist website.

(3) English language proficiency (El). A tourist coming from a country where the official language is English was regarded as an English speaking person. The country's official language was obtained from the internet (www.yahoo.com) by searching the country's profile.

(4) Distance travelled (Dist). This is the shortest ground distance between Tanzania and a tourist's country of origin, measured from Dar es Salaam to the country's capital. The variable was sourced from the internet using the online distance calculator.

(5) Relative cost of living between a tourist country of origin and Tanzania (Price). Figures in the Consumer Price Indexes were obtained from the IMF (2009).

4.4 Results and Discussion

4.4.1 Summary Statistics of the Variables used in the Model

Table 4.3 gives the summary statistics of the variables used in the analysis.

Table 4.2: Summary statistics of the variables used in the analysis of tourist spending

Variable	Mean	Std. Dev.	Min	Max	CV
Pcapita(USD)	191.68	206.92	9.5	1017.978	107.95
Age (Ranks)	3	0.69	1	4	23
Females	0.89	0.90	0	21	101.12
Incomes(Mil.USD)	29123.30	15379.39	54.62	113044	52.81
Er(TSHS)	965.49	603.92	0.03	3229.158	62.55
El	0.54	0.50	0	1	92.59
Childno	0.14	0.58	0	13	414.29
Childprese~e	0.07	0.25	0	1	357.14
Lstay	12.69	13.60	1	360	107.17
Tarra	0.53	0.50	0	1	94.34
VistOth	0.05	0.22	0	1	440
VistFRD	0.08	0.27	0	1	337.5
VistLSR	0.77	0.42	0	1	54.54
VistBSN	0.10	0.30	0	1	300
Adultno	1.75	1.28	1	31	73.14
Tpartyno	0.56	0.50	0	1	89.29
Frvists	0.53	0.50	0	1	94.34
Fadest	0.37	0.48	0	1	129.73
Dist(Miles)	5287.09	2389.07	419	9527	45.19
Nosites	0.65	0.48	0	1	73.85
Peak	0.74	0.44	0	1	59.46
Price	0.05	0.58	3.24E-05	26.00667	1160
Africa	0.09	0.29	0	1	322.22
Asia	0.08	0.27	0	1	337.5
MEast	0.01	0.09	0	1	900
SAmerica	0.01	0.09	0	1	900
year1	0.62	0.49	0	1	79.03
year2	0.10	0.30	0	1	300
year3	0.28	0.45	0	1	160.71

Table 4.3 indicates that most of the variables have a reasonable variation except for age. The poor variation in age is due to the fact that the survey reported only age groups rather than individual ages. It is also worth to noting that tourist per capita expenditure is not normally distributed (skewness=1.7).This suggests that logarithmic transformation is needed before regression.

4.4.2 The Correlation Matrix

It was observed that for most of the regressors there was no strong correlation among them, suggesting little multcollinearlity, except for the number of children and child presence, distance and dummy for North America, as well as dummy for Europe. A decision had to be made to drop some of the variables. Arbitrarily child presence was dropped and number of children left, while dummies for Europe and North America were dropped and distance left. More details can be found in Table 4.4 in appendix 4.3

4.4.3 OLS Regression Results of the Log of Daily Tourist per Capita Expenditure:

Table 4.5 provides OLS estimates for equation 4.1b. Besides estimates of the determinants of tourist per capita expenditure, the table also provides regression results for the determinants of party expenditure per day and total party expenditure which are included for comparison purposes. For ease of presentation, the table shows the standard errors for the main model only. But still the level of significance is provided for each regressor in all the models.

Table 4.5: OLS regression on the log of daily tourist per capita expenditure

Variables	Expenditure per capita per day		Travel Party daily expenditure	Total party expenditure
	Coef	Std. Err.	Coef	Coef
Age	0.15***	0.009	0.15***	0.09***
Females	-0.196*	0.0109	-0.026**	-0.002
GDP	1.4E-06**	8.23E-07	1.32E-06	3.3E-06***
Er	0.00038***	2.01E-05	0.00038***	0.00038***
EL	-0.16***	0.0202	-0.17***	-0.15***
Childno	-0.076***	0.0122	0.18***	0.15***
lstay	NA[21]	NA	NA	0.019***
Tarra	0.35***	0.014	0.35***	0.19***
VistFRD	-0.15***	0.0362	-0.15***	-0.08**
VistLSR	0.08***	0.0300	0.07***	0.03
VistBSN	0.53***	0.035	0.53***	0.19***
Adultno	-0.065***	0.0078	0.13***	0.11***
Ttpartyno	-0.17***	0.0159	0.37***	0.37***
Frvists	0.0707**	0.0136	0.07**	-0.09***
Fadest	0.028***	0.0141	0.03***	0.05***
Dist	0.00005***	4.01E-06	0.00005***	3.4E-05***
Nosites	-0.036***	0.0138	-0.037***	0.23***
peak	-0.17***	0.0150	-0.18***	-0.10***
Africa	0.51***	0.0508	0.51***	0.26***
Asia	0.24***	0.0285	0.24***	0.10***
MEast	0.23***	0.0682	0.22***	0.075
SAmerica	0.43***	0.0782	0.44***	0.399***
year1	(dropped)		(dropped)	(dropped)
year2	0.73***	0.0293	0.72***	0.75***
year3	0.59***	0.0238	0.59***	0.49***
_cons	3.47***	0.0512	3.306***	5.55***
Adjusted R^2	0.24		0.25	0.27
F	353.89		369.65	399.16
P>F	0.0000		0.0000	0.000
n	25,880		25,880	25,880

N:B*significant at 10%, ** significant at 5%, significant at ***1%

Table 4.5 indicates that all three models are highly significant, with P>F=0.000, and with almost all covariates being significant. This more robust estimate may be

[21] Length of stay is an endogenous regressor for both tourists' per capita daily expenditure and tourist party daily expenditures because it was used to derive both the two variables(see sections 4.3.2 & 4.3.3)

attributed by the large number of observations. A histogram plot of the fitted residuals is provided in appendix 4.4 as figure 4.1. The plot indicates that the residuals are fairly normally distributed.

Results from Table 4.5 indicates that destination attributes, as given by the coefficient of peak season, are not the most influential attributes as regards tourist per capita spending, but rather the trip-related characteristics are most influential than the other two categories of explanatory variables (demographic attributes and destination attributes).This leads one to reject the hypothesis that destination attributes matter more than the demographics and trip-related attributes. The implication here is that as long as a tourist has reached a destination he must spend. It is rather the interaction of his personal attributes and the destination attributes that decides his expenditure pattern. Given the fact that all the visitors are at the same destination, what matters more is the variation in their demographic and trip-related characteristics. These findings can be reaffirmed by several other studies. For example Anderson (2011), asserts that holiday experience at a destination, visitor's attributes as well as travelling attributes are most important determinants of all-inclusive tourists' expenditure in Balearic Islands. In the contest of this study all these variables are trip-related characteristics.

However, the rejection of the null hypothesis does not imply the unimportance of the destination attributes in influencing a tourist expenditure, as it was observed that the magnitude of its coefficient (-0.17) is not the least among them all. The negative coefficient for peak season implies that as the season gets progressively near the peak there is likely to be less daily spending by a visitor, which appears to be counter-

intuitive. This is certainly an indication that the destination stakeholders do not appear to be aggressive. It is more likely that, owing to the high influx of tourists during the peak season, the destination service providers are unable to meet their needs both in terms of quality and quantity which may lead to lower per capita spending. The other reason could be that during the peak prices for tourism goods and service becomes vey high making an individual tourist spend less.

Tourist travel purpose as measured by the dummy on the business visitors had a positive and the highest coefficient (0.53) among all the trip-related characteristics. The positive influence of business visitors was also observed on the travel party daily expenditure (0.53) and on total party expenditure (0.19). Similar results had previously been established by Suh and Gartner (2004), who concluded that irrespective of a tourist's country of origin, business travellers spend significantly more than pleasure visitors to Korea. The significance of the purpose of visit reaffirms the use of this variable in the TTSS expenditure model. Analysis of the TTSS survey data indicates that Africa has a higher percentage of business visitors than any other continent (40%), followed by Asia (12%), South America (11%), North America, the Middle East (7%) and finally Europe (6%). This highlights the growing importance of non-traditional markets for Tanzanian tourism. Whereas business visitors had a positive influence on spending, visitors to friends and relatives (VISTFRD) as well as visitors on leisure and recreation (VISTLSR) had a negative influence.

The next most influential trip-related characteristic is the travel arrangements (0.35) as operationalized in terms of package or non-package tours. The variable had already been asserted by TTSS (2001) as one of the most significant determinants of tourist

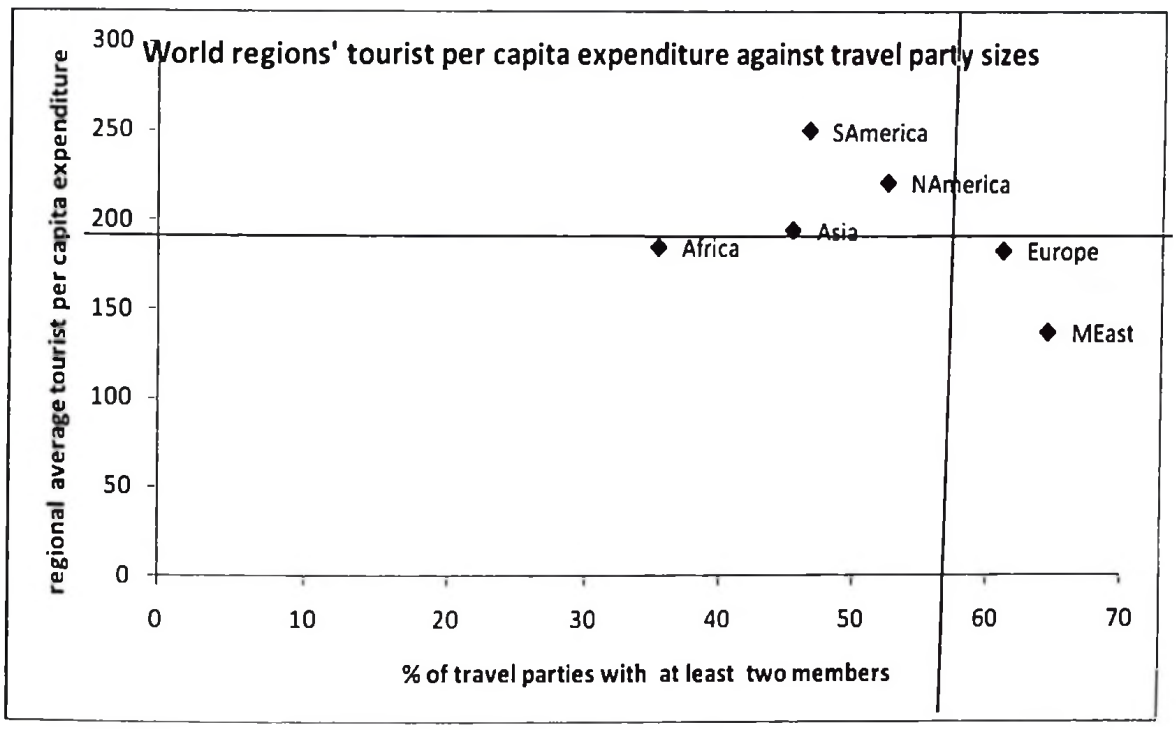
spending, and used in their expenditure model. This finding therefore reconfirms its plausibility. According to TTSS (2001, 2004,2005, 2006), tourists on package tours spend more than those on non-package tours as they are associated with high incomes. Nevertheless this finding is contrary to a number of studies done globally and to those done in Tanzania. For example Anderson (2011) did a study in Zanzibar and found that tourists on package tour spend relatively less than those on non-package tour. This discrepancy could be attributed by the difference in methods of data collection and/or conception of tourist expenditures. Further research into this area is needed. Therefore, according to both TTSS (2001, 2004, 2005, and 2006) and to this study, encouragement of the package tour industry appears to be one possible policy implication. The significance of this variable justifies a study on the determinants of tourist choice of a package tour when coming to Tanzania. This is done in chapter six.

Another influential trip-related characteristic is the travel party number, which had a significant and negative coefficient (-0.17).Its coefficient was positive for travel party daily expenditure as well as total party expenditure. This point to the fact that the negative effect of travel party number on tourist per capita spending has nothing to do with his daily experience but rather with his being on his own. The chances are that when people spend in groups they save a lot more money than when an individual is on his own. For example, a room which could accommodate a single tourist may accommodate more than one person for the same cost. Similar inferences can be drawn when one considers other types of expenditures on things such as food and transport. For example, a tour vehicle hired by one tourist would cost virtually the same if it were to be hired by ten tourists. Economies of scale appear to play a role here. This finding has two policy implications, either, promotion of tourist arrivals

should concentrate on countries which have smaller travel parties or propose ways of limiting the advantages of large parties which lower per capita tourist expenditure. For example, hotel owners could limit the number of occupants in a room, or tour operators could limit the number of tourists in a tour vehicle. The second suggestion has to do more with the private stakeholders such as tour operators, hotel owners, and others while the former has to do directly with the government through its marketing organs, such as TTB. However, the idea of not attracting tourists from countries with large travel parties jeopardizes the country's overall revenue, because travel party number positively influences total party revenue. Moreover, if the latter suggestion can work there would no longer be any need to not attract tourists from regions with huge travel parties. Therefore the second suggestion appears to be more plausible than the former.

Probably the only concern about limiting the number of arrivals from tourists travelling with huge travel parties would be that of environmental pollution. In order to avoid environmental pollution while at the same time allowing huge travel parties we may set an arbitrary threshold for a reasonable size of travel party as well as the expenditure per region and work out priority markets. This is a common approach in marketing studies. This approach was also used by of Dwyer and Forsyth (2008) when studying the measure for tourism yield and visitor markets to be targeted by Australia. Although their concern was not about the environment, the same idea could work. It is clarified in the next graph.

Figure 4.2: World regions' tourist per capita expenditure against travel party size



Source: Own drawing based on TTSS survey data

The graph above indicates a plot of tourist per capita expenditure across the six regions of the world against the travel party sizes. The travel party sizes have been worked out as a percentage of travellers whose travel parties have at least 2 members. Percentages instead of average travel party size were used after observing that average travel party size gives little variability. The graph suggests that regions in the first quadrant are those whose tourist per capita expenditures exceeds the overall tourist per capita expenditure (192 USD) in Tanzania as well as the overall percentage of travel party size (55%). The figures 192 and 55 were worked out from the TTSS survey data (2001, 2007, and 2008). Regions in the second quadrant are those whose per capita expenditure is above the overall average, but whose percentage travel party size is below the average (55). The third quadrant consists of regions where both per capita

expenditure and travel party sizes are lower than the overall average. The fourth quadrant consists of regions where per capita expenditure is lower than average but travel party size is above average. As shown by the graph, priority markets for Tanzania are the regions in the second quadrant, for from these regions the country would achieve both higher revenue and a minimum degree of environmental destruction. These regions are Asia, South America and North America. There is an indication that Africa has the potential to belong in this group. On the other hand, regions in the fourth quadrant should be given least priority by the marketers for from them the country achieves both less revenue and lower degree of environment protection.

Finally it is worth noting that the effect of travel party size is similar to that of adult numbers as suggested earlier. The coefficient for the number of adults is significant (p=0.000) and negative (-0.065). Given that majority of the travellers are adults (the survey for 2001,2007 &2008 indicates that 94% of tourists were adults[22]), then most of what has been recommended as regards travel party sizes applies equally to this variable

The next most influential trip-related characteristic is the English language (-0.16), which shows that there is less spending in the presence of an English-speaking tourist than otherwise. This situation is consistently observed for total party expenditure as well as total party expenditure per day.

[22] My own calculations based on the statistics from TTSS

No obvious deductions can be made from this result. One tends to think that an English-speaking tourist could easily familiarize himself with the environment and hence opt for cheaper spending options, but this is not all that obvious, especially because Tanzanians are not as conversant with English language as their neighbours such as Kenyans and Ugandans.

Another influential trip-related characteristic is number of children (-0.08). When the presence/absence of children was used instead of number of children its influence was much bigger (-0.16). The presence of a child in a travel party significantly reduces a tourist per capita expenditure, which is similar to what was found by Agarwal and Yochum (1999). However its coefficient is positive on travel party daily expenditure as well as on total party expenditure, implying that presence of a child affects the individual tourist directly and not the travel party. The only plausible explanation is that the child's presence limits tourist activities and hence reduces daily expenditure irrespective of his/her length of stay.

Length of stay shows a significant and negative influence on tourist per capita spending (appendix 4.1). This result is different from that of Wang et al. (2006), who found that length of stay has a positive influence on tourist daily expenditure. Appendix 4.1 also show that length of stay has a negative and significant influence on a party daily expenditure but positive and significant influence on total party expenditure (Table 4.5), which suggests that its negative influence has more to do with the daily experiences of tourists. It is not unusual for the spending of tourists/consumers to be unequal across the days. Given the fixed budget tourists are on, they are likely to spend more in the early days and reduce their daily expenditure

as money runs out.

The significance of length of stay on tourist daily expenditure gives credit to the TTSS (2001) expenditure model, which also included length of stay as one of the key determinants. However, the negative influence of length of stay on tourist daily expenditure does not support the WTO/TTSS model. The model assumes that an individual's daily expenditure is constant across the days of stay. On the contrary, the results show that a tourist's daily expenditure declines with his length of stay (appendix 4.1). Therefore it is erroneous to simply use a constant value of daily spending across all the days. In other words, revenue estimates produced by TTSS are likely to be over-optimistic similarly, in an area or country where length of stay positively influences tourist daily expenditure, the WTO/TTSS model may under-estimate the total revenue. I propose an improvement in the WTO/TTSS model as follows:

Given that daily spending varies from day to day, one may consider the following modification:

$$TR_m = A * \sum_{i=1}^{T} E_i \qquad (4.9f)$$

Where

A is the total number of arrivals

TR_m is the modified Tourism Revenue estimates

E_i is the average expenditure by tourists who stays for (i) days.

T is the average length of stay by a tourist in Tanzania.

The new model requires the computation of average tourist expenditure for different days of stay so as to take care of the variation in daily spending. Alternately, if a constant growth rate of daily spending is assumed, the model (4.9f) can be written as

$$TR_m = AE_1\left[1+\sum_{t=2}^{T-1}(1+r)^{T-1}\right] \qquad (4.9g)$$

Where E_1 is the average tourist expenditure in day 1 and r=constant growth rate of daily expenditure, which is the coefficient of length of stay in the OLS estimate of equation 4.1.If revenue estimates are to be computed on an annual basis then for each year a separate OLS estimate of r is needed.

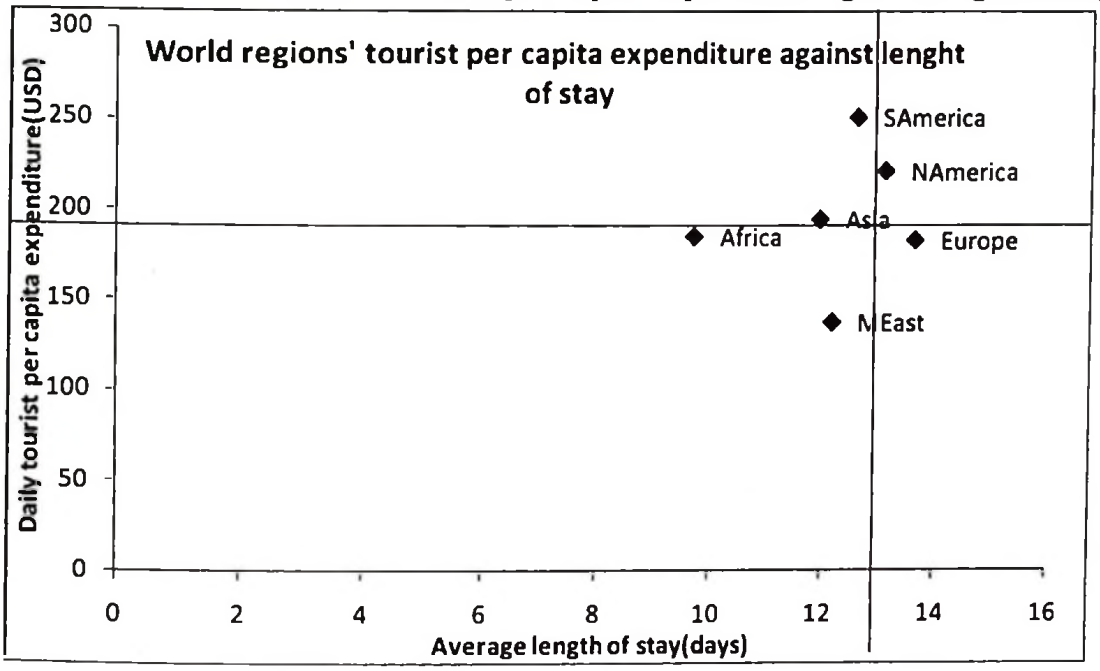
On the other hand, the positive influence of length of stay on total party expenditure should be expected, as previous studies, such as Agarwal and Yochum (1999), found this to be the case. Its positive influence justifies the importance of a tourist staying longer for the country to earn as much revenue as possible. In other words, studying the determinants of a tourist length of stay is needed. This is attempted in the next chapter.

In general, tourists staying longer yields more revenue but might be detrimental to the country's environment. Similarly, although their short stay could enhance environment, it definitely has a negative effect on the country's total revenue. Therefore, what is the way forward? It is unfortunate that there is no minimum number of days which can be thought of as safe from environmental destruction.

The study considers the current averages of the two variables (length of stay and per capita expenditure as the minimum thresholds), as done previously when comparing

tourist travel party size against tourist per capita spending (see figure 4.2). The graph for tourist daily per capita expenditure against average length of stay is given in figure 4.3.

Figure 4.3 World regions' tourist per capita expenditure against length of stay



Source: own drawing based on the statistics from TTSS

As per the figures derived from the TTSS surveys, the overall tourist per capita per day expenditure for 2001,2007 and 2008 was 191.77 USD, while the average length of stay was 12.93~13 days. The first quadrant of the graph represents regions whose tourists' per capita expenditure and average length of stay are both above the overall. The second quadrant consists of the regions whose tourists are above the overall average for per capita expenditure but below the overall average for length of stay. The third quadrant consists of regions whose tourists fall below the average for the two variables. The fourth quadrant comprises regions whose tourists are above the overall average for length of stay but below for per capita expenditure. Therefore as per the graph, regions in the 2nd quadrant are the priority markets for Tanzania if

protection for the environment is a concern. These regions are South America and Asia. From these regions the country can be expected to receive higher revenue and suffer minimum environmental destruction. However, if concern is purely for the environment, then North America and Europe should be avoided.

Other trip-related characteristics included were familiarity with the destination, frequent visiting and distance, all of them exhibited a significant and positive influence. The positive influence of a familiar visitor and a frequent visitor points more to the earlier suggestion that these types of visitors could be explorers or business visitors/consultants, who are generally wealthier and therefore the positive coefficients on these variables would not be accidental. Even if they were not explorers/business visitors, their frequent visits imply their satisfaction with the destination or region, which should positively influence their spending. As regards distance, the result reaffirms what was pointed out by Wang et al. (2006) that visitors from a distance tend to spend more than those from the nearby. For Tanzania these visitors could also represent rich tourists, as most rich countries are further away from Tanzania than the poor ones.

Apart from the trip related characteristic the study also investigated the demographic attributes of age, gender, GDP, exchange rate and language. Of these, only gender coefficient was slightly significant. The reason for the slight significance of the gender coefficient could have to do more with the argument by Wang et al. (2006), that there is no reason for gender to affect tourism expenditure since much of the travel expenditure is a group activity.

What can be deduced from the positive influence of age is that older tourists have a higher income accrued through either long-term saving or through holding lucrative positions. Some previous studies, such as Sampoz and Perez (2000) and Jang et al. (2004) also found that age has a positive influence on tourist spending, although Wang et al. (2006) found that age has a negative influence. However, according to Wang et al. (2006), the finding on the negative influence of age on expenditure could be accidental and warrants further investigation. The significance of age as regards tourist per capita spending could be utilized by the TTSS to improve their expenditure model. The TTSS expenditure model uses three variables: nationality, purpose of visit and travel arrangements. Owing to the fact that tourists can also be stratified across age groups, this variable could therefore be used by TTSS to improve their expenditure model.

The positive influence of GDP and exchange rates is to be expected as they indicate a tourist's country credit worthiness. It is unfortunate that the study could not capture the income of individual tourists. Most studies have found income to be the most significant determinant of tourist spending at a destination (Wang et al. 2006). This finding led Wang et al. (2006) to conclude that demographic attributes are much more significant than other factors. In this particular study the demographic attributes are not the most influential, but rather the region dummies are more influential than other factors. As can be observed, coefficients for the dummies for Africa and South America are the largest of them all. All the regions' dummies indicate a positive influence on spending except, for the European dummy, which has a negative influence (It was excluded alongside the American dummy, for their inclusion led to evidence of multicolinearity).When the dummy for a European country was included,

in the absence of other regions' dummies, and distance with which it highly correlates, it exhibited a negative and significant response of -0.30.

The negative influence of the European dummy is rather surprising and contrary to the common expectation that regions/countries with higher incomes should spend more than others. In the same way being a North American is associated with positive spending (0.11), which is lower compared to being an African or a South American.

The last variable in the model is the number of sites visited, which has a significant and negative effect on tourist daily spending. This result is consistent with the influence of the length of stay on tourist spending. This is so because the more the sites a tourist vist the longer he/she stays in a country leading to a negative influence on tourits per capita spending (appendix 4.1)

It is also worth seeing whether the results found in Table 4.5 would vary across the world regions much as what was presented was on a general tourist, irrespective of his region of origin. Table 4.6 provides regression results for the six regions of the world alongside the general results. Standard errors are omitted for ease of presentation, but the level of significance is still provided for every regressor.

Table 4.6: OLS regression results of the log of a daily tourist per capita expenditure: Comparison across regions

Variable	World	Africa	Asia	Europe	MEast	NA	SA
Age	0.15***	.11***	.20***	.12***	.11	.16***	.27**
Females	-0.0197*	.02	.043	-.03*	.03	-.02	.10
GDP	1.4e-06**	-.1.7e-05	8e-06**	-3.6e-06***	-1.e-05	-1.2e-04	-2.4e-06
Er	3.8e-04***	8.7e04	4.6e04*	2.8e-04***	1.1e04	.005	4.4e04
EL	-0.16***	-.14*	-.32***	-.08***	dropped	dropped	-.34
Childno	-0.076***	-.13**	-.10*	-.07***	-.09	-.06**	-.008
Lstay	NA	NA	NA	NA	NA	NA	NA
Tarra	0.35***	-.24***	.27***	.40***	.31*	.45***	.3*
VistFRD	-0.15***	-.18*	-.15	-.10**	-.49	-.29***	-.28
VistLSR	0.08***	.15*	.15	-.06	-.54*	.13***	.04
VistBSN	0.53***	.42***	.31**	.54***	.25	.47***	.19
Adultno	-0.065***	-.04	-.11***	-.08***	-.14*	-.03	-.50***
Ttpartyno	-0.17***	-.40***	-.11**	-.16***	-.23	-.14***	.50**
Frvists	0.0707**	.12*	-.02	.08***	-.03	.08**	.04
Fadest	0.028***	.05	.18***	.001	-.07	.04	-.03
Dist	0.00005***	3.9e05	1.e04***	1.3e05	-2.7e-04	1.2e-06	7.e05
Nosites	-0.036***	-.11**	-.03	-.02	.13***	-.05	-.32*
peak	-0.17***	.47***	-.14***	-.21***	-.11	-.34***	-.34*
Cons	3.47***	4.1***	4.5***	4.9***	5.1***	4.5***	5.2***
Adj R^2	0.24	0.15	0.22	0.28	0.19	0.26	0.20
F	353.89	21.73	30.65	329.68	4.02	109.44	3.58
P>F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
n	25,880	2379	2028	15560	232	5485	196

NB: *significant at 10%, **significant at 5%, ***significant 1%

From Table 4.6 it can be observed that the trip-related characteristics were still the most influential determinants of tourist per capita expenditure in most regions, leading again to the rejection of the hypothesis that destination attributes are more influential than the demographic and trip-related characteristics. However, contrary to other regions, the peak season coefficient came to be positive and most influential for tourists from Africa (0.47), implying that tourists from Africa spend more during peak seasons than during other seaons. Therefore the null hypothesis is not rejected for

African tourists. Furthermore, the variable tavel arrangement was negative and significant for tourists from Africa unlike tourists from other regions.

The following variables behaved the same way in all the regions: age, language, number of children, adult numbers, and distance. The following variables behaved in the same way except for South America: travel party size and number of sites visited. Whereas travel party size had a negative effect in other regions in South America, it had a positive impact on tourist spending. As regards the number of sites visited, it had a negative effect in South America contrary to other regions.

It would also be important to see how different categories of expenditure respond to the suggested determinants. Four categories of expenditures were identified by TTSS, namely, expenditure on accommodation, food, transport and shopping. Analysis of the first two categories is useful for hotel/restaurant owners, while analysis of the third category is useful tour operators. Analysis of the last category is useful for the general public and retailers in general. The idea is to see whether there are sector-specific considerations apart from the general findings presented in Table 4.5. The results on the determinants of tourist per capita expenditure across the four categories of expenditure are given in Table 4.7, along with the general results of Table 4.5. As in the case of regional comparison standard errors are omitted for ease of presentation. The interest lies in comparing the coefficients of the estimates across expenditure categories, although the level of significance is still provided for every regressor.

Table 4.7: OLS regression results of the log of a daily tourist per capita expenditure: Comparison of the categories of expenditures

Variable	overall	Accommodation	Food	Transport	Shopping
Age	0.12***	.34***	.18***	.22***	.15***
Females	-0.007	-.06***	-.07***	-.038	-.05**
GDP	1.6E-06**	1.6e-06	2.3e-06*	-4.4e-06**	-6.9e-06***
Er	0.00037***	1.5e-04***	6.7e05**	8.0e-05	-1.4e-05
EL	-0.15***	-.014	.01	-.07	.04
Childno	-0.087***	.03	-5.e-05	-.001	.09***
Tarra	0.26***	NA	NA	NA	NA
VistFRD	-0.24***	-.19***	-.04	-.12	.1
VistLSR	-0.08***	.21	.23***	.19**	.15***
VistBSN	0.36***	.69***	.47***	.26**	.27***
Adultno	-0.075***	-.09***	-.03*	-.11***	.09***
Ttpartyno	-0.191***	-.14***	-.22***	-.02	-.25***
Frvists	0.031**	.15***	.10***	-.06*	.11***
Fadest	0.043***	.08***	.06***	.07*	.035
Dist	0.00004***	8.7e-06	-1.3e-05*	3.4e-05***	2.4e-05**
Nosites	0.053***	-.22***	-.14***	.16***	-.12
peak	-0.15***	-.08***	-.09***	-.12***	-.027
Africa	0.41***	.31**	.19**	-.008	.27***
Asia	0.21***	.24***	.20***	-.02	-.036
MEast	0.18***	-.007	-.16	-.28*	-.07
SAmerica	0.42***	.30*	.34**	.12	.18
_cons	3.97***	2.5***	1.97***	2.4***	1.6
Adjusted R^2	0.28	0.21	0.17	0.05	0.11
F	416.59	97.54	73.10	14.10	33.89
P>F	0.0000	0.000	0.000	0.000	0.000
n	25,880	7998	7998	5981	5547

NB: *significant at 10%, **significant at 5%, ***significant 1%

The regression results show that, with the exception of the dummies, the trip-related characteristics are the most influential factors in all the categories of expenditure, leading again to the rejection of the null hypothesis that destination attributes are more influential than demographic and trip-related characteristics. The influence of destination attributes, as measured by peak season, has remained negative in relation to all categories of expenditure.

As regards accommodation expenditure, only number of sites behaved differently, contrary to the general results. It was significant but with a negative influence. Its negative influence could reflect the fact that, the more sites a tourist visit the more time he spends in camps and parks, leading to less expenditure on high quality hotels.

Regarding food expenditure, both distance and number of sites visited came to have a negative influence, contrary to the general result. For distance, no obvious deduction can be made, as there is no reason to believe that travellers from a distance do not like food, because food is a basic necessity. Regarding the negative influence of the number of sites visited, it points to the same idea in relation to accommodation that the more time a tourist spends at sites, the less he spends on expensive hotels.

Concerning transport, frequent visitors had a negative and significant influence, contrary to the general results while visitors coming for leisure and recreation had a positive and significant influence, which is also contrary to the general results. Apparently nothing can be inferred from the negative influence of frequent visitors. One could argue that frequent visitors spend less time at a destination and hence spend less on transport. But this should have equally affected other categories of expenditure. As regards the positive influence of leisure and recreation, the inference is these types of visitors could be visiting relatively more sites than other categories of tourists, hence spending more on transportation[23].

[23] In the TTSS surveys for 2001, 2007 and 2008, more than 70% of visitors coming for leisure and recreation visited more than one site, compared with only 46% of visitors not coming for leisure and recreation. These figures are based on my own calculations using the TTSS survey statistics.

With regard to shopping, number of children had a significant (p=0.000) and positive influence (0.16), contrary to the general results. The idea is that although number of children may constraint overall expenditure owing to the lack of time to spend (Wang et al.2006) on shopping, tourists accompanied by children do a lot of shopping. Another variable which behaved differently from the general results was number of adults which had a positive influence on the expenditure on shopping. Wang et al. (2009) obtained a similar result on the influence of the number of adults on shopping expenditure.

4.5 Conclusion

The study has addressed the determinants of tourist per capita expenditure in comparison with travel party total expenditure and travel party daily expenditure. Three groups of variables were of major interest, destination attributes, demographic characteristics and trip-related characteristics. In each case, the trip-related characteristics tuned out to be the most influential determinant of tourist per capita expenditure, except for tourists from Africa.

The results have shown that as the tourist season moves progressively to the peak there is correspondingly less per capita spending by a tourist. In general, Tanzanian stakeholders need to find out why this happens. One of the possible reasons could be the inability to handle effectively a big influx of tourists during the peak season and/or the fact that prices are high during the peak season.

A number of findings have been obtained with regards promotion and environmental protection. In order to achieve both maximum revenue and a high degree of

environmental protection through key trip-related characteristics such, as travel party size and length of stay, some markets have to be avoided. Priority should shift from the traditional markets of Europe to the less popular markets of Asia, South America and Africa which would cause relatively less environmental pollution and bring in a higher degree of revenue.

The study has also confirmed that variables used in the TTSS expenditure model are appropriate. Among the variables confirmed were tourist length of stay and travel arrangements. The confirmation of these variables as among the key determinants of a tourist per capita spending paves the way for studying further the factors determining these variables (length of stay and travel arrangement choice).This is done in chapters five and six.

CHAPTER FIVE

DETERMINANTS OF TOURIST LENGTH OF STAY

5.1 Introduction

The country's total tourism revenue depends not only on the number of arrivals and tourist per capita expenditures but also on their length of their stay in a country. In other words, studying the determinants of tourist length of stay may be equally as important as studying the determinants of number of tourist arrivals and their per capita expenditure. Therefore this chapter examines the determinants of tourist length of stay in Tanzania. The study is interested in testing the hypothesis that destination attributes as measured by season are more influential than demographic and the trip-related characteristics.

As already argued earlier in chapter four, tourists staying longer is associated with higher total party expenditure on all categories of expenditure such as accommodation, transport and shopping, although tourists staying longer is associated with lower per day expenditure. But the overall total spending is what the country is interested in, which depends very much on tourists staying longer.

However, there is one major setback in advocating longer length of stay by tourists and that is environmental pollution. Although there is no ideal length of stay deemed not to be destructive of the environment, it is generally agreed that the longer the stay the greater the risk of environmental destruction. In other words, the shorter the length of stay the better it is for the environment. This is self-explanatory.

In order to both protect the environment and receive more revenue through tourists'

staying longer, a balance can always be found by considering the current tourist average length of stay as the ideal level for environmental protection and the current tourist average per capita expenditure as the ideal level for tourism revenue. Based on these two averages, tourists from countries who stay longer (above average), but with lower per capita expenditure (below the average), are to be discouraged from visiting the country and/or staying longer. Similarly tourists from countries who stay for a short time, but with high per capita expenditure should be highly encouraged. Therefore knowing the determinants of tourist length of stay greatly provides the marketing strategies required for various destinations.

Apart from emphasizing the need for a tourist to stay longer for providing reasonable revenue, the study has employed survival analysis, which can provide the hazard rate for tourist length of stay. The hazard rate would enable hotel owners and other stakeholders, such as tour operators, to condition their activities on the probability that a tourist will stay beyond a certain time (Barros and Correia, 2007, Menezes et al. 2008).

The use of survival analysis in tourism studies for the Tanzanian setting has not been attempted before. Globally, the first study to use survival analysis in studying tourist length of stay was by Govakali et al.(2007) in Turkey. Other recent pioneers in the field have been Barros and Correia (2007), Garcia and Raya (2008) and Menezes et al.(2008).

Even though such studies have been done globally, all of them overlook the major issue regarding endogeneity of the regressors. For example, the modeling of tourist

length of stay has often encompassed a tourist's travel arrangements (package versus non-package) as one of its explanatory variables (see for example Govakali et al., 2007, Garcia and Raya, 2008).The inclusion of this variable assumes that it is purely exogenous as required by the OLS assumption. However, in some cases, a tourist's choice of a non-package tour might have been motivated by his expecting to stay longer, making the variable endogenous.

A tourist who expects to stay longer is more likely to opt for a non-package tour (see Figure 2.24 of chapter two). The reasons are unclear. But probably it could be that package deals for longer stays are very expensive, or rather because he wants to stay longer he is motivated to choose a non-package tour, which is a more flexible travel arrangement. Similarly, a tourist who expects to stay for a very short time is time-bound and hence less flexible, and so it is convenient for him to choose a package tour (see Figure 2.18 of chapter two). This scenario whereby a tourist's length of stay may affect his choice of travel arrangements and the vice-versa has been overlooked in most such studies.

This study investigates this in detail, by first showing that travel arrangements is an endogenous variable in a log linear model of tourist length of stay. Apart from assessing the endogeneity problem the log linear model gave a hint as to which variables could be significant in determining tourist length of stay. After examining the log liner model on tourist length of stay, the conventional survival models were adapted and a choice was made as to which was the best.

The remaining sections in the chapter are organized as follows: section 5.2.1 reviews

the theoretical literature while section5.2.2 reviews the empirical literature. Section 5.3 describes the methodology and data sources and section 5.4 outlines the findings and discussion. Section 5.5 concludes the chapter.

5.2 Literature Review

5.2.1 Theoretical Literature Review

Like the determinants of tourist per capita expenditure, length of stay is a commodity, which its demand depends among other things on a tourist's income and how much the length of stay costs. The consumer theory puts great emphasis on the price determinants of a commodity. The non-price determinants (other things, being equal) can best be explained from the sociological point of view, using TRA.

TRA attempts to explain why a person pursues a certain action. TRA suggests that a person's behavioural intention depends on the person's attitude to the behavior and the subjective norms. Attitudes consist of beliefs about the consequences of performing the behaviour multiplied by his or her evaluation of these consequences. Subjective norms are seen as a combination of perceived expectations of relevant individuals or groups and intentions to comply with these expectations.

Barros and Correia (2007), in studying the determinants of tourist vacation length of stay, argue that the conceptual problem consists of understanding the relationship between vacation length of stay and expectations and attitudes in the behavioural intentions, as well as the relationship between intentions and subsequent behaviour. In order for a tourist to stay longer, he must have perceived the destination to be an attractive place.

These attitudes and beliefs concerning an action are not tangible, but are reflected in the characteristics of the individual and the characteristics of the action. In terms of consumer behaviour these attributes and subjective norms would be reflected in consumer characteristics, product characteristics and the environment where the purchase was made (Nzuki, 2006).

The application of TRA in tourism studies led to the theory of the tourist decision process, which Barros and Correia (2007) describe as being influenced by four factors: tourist profiles, trip-related characteristics, trip awareness, and destination characteristics. Tourist profile is a reflection of a consumer's characteristics, which essentially refer to social-economic characteristics. Trip-related characteristics are comparable to product characteristics, while destination characteristics are comparable to environment of purchase.

Alegre and Pou (2006) as well as Garcia and Raya (2008) assert that the theoretical determinants of a tourist's length of stay can be looked at two perspectives. One is the determinants of consumer preferences, such as demographics, and the second is the price determinants, covering a consumer's income and the cost of travel time and the holiday time.

5.2.2 Empirical Literature Review

TTSS (2001), used cross-tabulations for establishing the variations of tourists' length of stay in Tanzania. The study found that German and Italian visitors had the highest length of stay of about 13 days, followed by the Netherlands with an average length of stay of 12 days. Kenya was the last with an average length of stay of 6 days. The study

also found variation in average length of stay between tourists on package and non package tours. Those on non-package tours had an average length of stay of 11 days, whereas those on package tours had an average length of stay of 8 days

Alegre and Pou (2007) studied the determinants of a tourist's length of stay in the Balearic Islands using logistic regression. The study included a number of explanatory variables in the model, amongst which were age, type of jobs, nationality, type of accommodation, number of trips, visiting rate, size of party, daily cost of holiday and total party expenditure. Most of these variables were found to be significant with type of job, nationality, single trip and total holiday expenditure having a positive influence on a tourist's length of stay.

Gokovali et al. (2007) analyzed the determinants of a tourism length of stay in Turkey. The study was the first in tourism studies to use survival analysis for examining the determinants of tourist length of stay. The study identified 16 variables, which were significantly associated with a tourist's decision about the length of stay during the summer vacation. More specifically, the study identifies nationality, education, income, traveling experience, familiarity and daily spending as the major determinants of length of stay. An increase or decrease in the values of such variables was accompanied by a significant increase or decease in length of stay.

Barros and Correia (2007) analyzed the determinants of length of stay of Portuguese taking vacations in Latin America using the survival model. The study found that longer stays were associated with tourists on bigger budgets, who were motivated by culture, climate and security. Other variables positively related to longer stays were

age, frequency of visits and information received by word-of-mouth. It ws also established that repeat visits negatively affect tourist length of stay.

Menezes, et al. (2007), analyzed the determinants of the length of stay of tourists in the Azores using count data models (Zero-truncated Poisson model and Zero-truncated negative binomial model). It was found that socio-demographic profiles, such as nationality and Azorean ascendancy, and trip-related attributes, such as repeat visit and type of flights, were important determinants. Other demographic attributes found to increase tourist length of stay were age, being a female tourist and high-level professions. As regard to trip attributes, business visits, repeat visits and number of islands visited were associated with longer stays. Furthermore the study found that destination image and attitudes describing environmental initiatives also influenced length of stay.

Menezes et al. (2008) studied variables contributing to tourist length of stay in the Azores Islands using survival analysis. The findings were remarkably similar to those established using the count data models (Zero-truncated Poisson model and Zero-truncated negative binomial model). The study found that being a repeat visitor and taking charter flights is highly associated with longer length of stay. It was also found that visiting more sites is positively associated with longer length of stay, while higher level of education reduces the expected length of stay. Further, the study found that marketing strategies, promoting the Azores for its nature, landscape, remoteness and weather, could increase tourist length of stay, whereas strategies which promote the island as a cultural heritage site could not.

Barros et al. (2010) studied determinants of length of stay by golf tourists in the Algarve using survival analysis. The study established that tourist length of stay is positively influenced by nationality (especially German and British), education, age, type of hotel where the individual stays, events at the destination, climate and hospitality.

Yang et al. (2011) in China studied the determinants of tourist length of stay using an ordered logit model and established that age, organized tour, transportation, motivation and past visits, were the major determinants of tourist length of stay. The results indicated that traveling distance and the assessment of accommodation are positively associated with the length of stay.

Chaiboonsri and Chaitip (2012) studied determinants of tourist length of stay in India using both Poisson regression and Negative Binomial regression. The use of negative binomial regression was motivated by the fact that Poisson regression could not do well owing to over-dispersion of the response variable. The study established that longer tourist stay is associated with being a frequent visitor, having a higher spending, being less educated and having less income. The findings on the negative impact of income on the length of stay is surprising and contrary to the finding by previous studies such as by Gokovali et al. (2007) as well as by Barros and Correia (2007).Other factors which contributed to longer stay by a tourist as established by the study included the fall in cost of living in India, Indians' understanding of foreigners' culture and less aggressive policy towards environmental developments in urban areas.

In general, one may conclude that tourist length of stay depends among other things on

income, nationality, education/profession, experience/frequency of travel, familiarity with the destination/repeat visit, number of sites visited, similarity of culture between tourist country of origin and the destination, security, distance, information source, climate and satisfaction with the destination attributes. Among these variables, income, age, higher profession level/education, non-package tour, frequency of visits, number of sites visited, cultural similarity, satisfaction, and information received by word-of-mouth are thought to positively affect tourist length of stay. Repeat visits/familiarity with the destination and distance are thought to negatively affect tourist length of stay, while season of travel/climate could be indeterminate. These factors can be grouped into socio-demographic and trip-related characteristics, and destination attributes. It is also evident that among these studies none has attempted to address possible endogeneity of the regessors, which is dealt with in this study.

5.3 Methodology

5.3.1 Model Specification

The Nordstrom (2002) model adopted in chapters three and four does not explain the position of length of stay in the utility function of a tourist. A discrete choice/continuous model by Durbin and Mcfadden 1984 as well as by Hanemann 1984, as described by Alegre and Pou (2006), takes account of the length of stay in the tourist's utility function.

According to the account given by Alegre and Pou (2006), the model assumes that a tourist utility function comprises three goods: q, a vector of consumer goods excluding tourism services, z, the vector of characteristics that define the holiday (the destination, type of accommodation, category of accommodation), t the length of holiday. A

consumer chooses the values of q , z and t which maximize his utility, subject to income constraint Y and time constraint T. The time constraint T consists of time taken to travel to the destination and the time of staying at the destination. Accordingly budget Y constrains the expenditure on travel, expenditure at the destination and the expenditure on non-tourism goods q.

If one assumes weak separability in the utility function[24], the utility function can be maximized separately in the absence of non-tourism goods q. Under this weak separability assumption, the demand for length of stay can therefore be viewed as a function of the holiday characteristics, the price of traveling to a destination, the daily price of the holiday, the total expenditure available for the holiday, the maximum time available for the holiday, the characteristics of the consumer and the unobservable random effects. Following this discussion, Alegre and Pou (2006) specify the following demand function for length of stay:

$$D = f(p_{tour}, z, Y - pq - p_{travel}, T - t_{travel}, \tau, \varepsilon) \qquad (5.1)$$

where

P_{tour} = price of the holiday, z are the trip characteristics, p is the price of the non-tourism goods, $p_{travel,}$ is the price of travelling to a destination, T is the total holiday time, t_{travel} is the travel time to the destination, τ are the consumer characteristics and ε is the random error term.

[24] A utility function is weakly separable if the marginal rate of substitution between any two goods belonging to a group of goods, say tourism goods, is independent of any quantity of goods outside this group. This assumption is important for solving a tourist's maximization problem. When the vice-versa is true then one has a strong separable utility function. More details and clarification as to why weak separability is adopted are given in appendix 3.2

Garcia and Raya (2008) also give an account of this model. According to them, determinants of length of stay emanate from the tourist's preferences as given by his utility function as well as the utility constraint. The former encompasses the determinants of a tourist's preferences, such as demographics (age, education, gender, e.t.c), while the latter encompass the arguments involved in the constraint, which involve price of travel and price of the holiday time, total budget for the holiday and the total holiday time. A tourist therefore chooses time t which maximizes his utility given these constraints. As before, the maximization of the utility function given the length of stay must assume weak separability between the non-tourism good q and length of stay t, in the entire utility function.

In this study the same model was adopted. However, the nature of the data collected contained no information on the total holiday time of a tourist, the budget allocated for the holiday, and the prices of the travel and of the holiday. Nevertheless for total allocated budget, per capita GDP was used as a proxy whereas travel distance was used as a proxy for cost of travel. The missing variables (maximum holiday time and the daily costs of the holiday) were not the key ones in the study as the hypothesis was based on the consumer's demographic characteristics and trip-related characteristics and the destination attributes, for which data were available.

Following this discussion, equation 5.1 can be rewritten in a more detailed manner while considering the variables discussed in sections 5.2 and 5.3 as follows:

$$L_{ict} = \beta_o + \beta_1 Age_i + \beta_2 Females_i + \beta_3 GDP_i + \beta_4 Er + \beta_5 El + \beta_6 Childpresence + \beta_7 childno + \beta_8 Tarra + \beta_9 VistFRD + \beta_{10} VistLSR + \beta_{11} VistBSN + \beta_{12} Adultno + \beta_{13} Tpartyno + \beta_{14} Frvist + \beta_{15} Fadest + \beta_{16} Dist + \beta_{17} Nosites + \beta_{18} \text{Price} + \beta_{19} Peak + \beta_{20} Isource + Africa + Asia + Europe + MEast + NAmerica + SAmerica + year1 + year2 + year3 + \varepsilon \quad \cdots(5.2)$$

where

L_{ict} =the length of stay by a tourist belonging to country c and observed in year t.

Age= 1 if the respondent's age group is <18, 2 if age group is 18-35, 3 if age group is 36-55 and 4 if age group is over 55 .Its coefficients β_1 is indeterminate. Most older people travel on package tours, which are normally associated with shorter stays. Older age could also imply having more commitments back home than younger people. Alegre and Pou(2006) finds tourist aged between 45-60 having a significant and negative influence on tourist length of stay. On the contrary, Menez et al. (2008) found no clear pattern between age and length of stay. Older age could also have a positive influence on length of stay, as older people are likely to have more income than younger ones.

Females=number of females in the travel party. Its coefficient β_1 is indeterminate. According to Menez et al. (2008), female tourists stay longer than male tourists. But there is no obvious reason for this.

Income =tourist's income level. In this study per capita GDP of a tourist's country of origin was used as a proxy for income, because tourists were not asked about their incomes during the surveys. Its coefficient β_3 is expected to be positive.

Er=Exchange rate between Tanzania's currency and the tourist's country of origin. Currencies are expressed in terms of units of Tanzanian shillings per unit of a foreign currency. This variable, like the per capita GDP, is designed to

assess the impact of income on tourist length of stay. Its coefficients β_4 is expected to be positive

EL=1 if a tourist is from an English-speaking country and 0 otherwise. Its coefficient β_5 is expected to be positive .An ability to communicate fluently may lead to the tourist becoming more familiar with the destination leading to longer stay other things being equal.

Child presence=1 if there is at least one child in the travel party and 0 otherwise. Unfortunately, there is no account in the literature of the influence of this variable on length of stay. However, this variable could be equivalent to the marriage status of a tourist of whose influence has been found to be negative and significant by Menez et al.(2008). Therefore its coefficient β_6 is expected to be negative.

Childno= Number of children in the travel party. This variable should follow the same sign as the former one in the sense that more children could imply more family commitments back home, hence leading to shorter length of stay. The only difference from the former one is that, it captures not only the direction but also the intensity of the influence of children on a tourist's length of stay. Therefore its coefficient β_7 is expected to be negative.

Tarra= 1 if a tourist is on package tour and 0 if a tourist is on a non- package tour. Its coefficient β_8 is expected to be negative. A tourist on package tour has limited time and would rarely extend his length of stay as everything is scheduled unlike the tourist on non-package tour. Govakali et al.(2007), Garcia and Raya(2008) find this coefficient to be negative.

VistFRD = 1 if a tourist is visiting friends and relatives, 0 otherwise. Its coefficient

β_9 is indeterminate.

VistLSR= 1 if a tourist has come for leisure and recreation, 0 otherwise. Its coefficient β_{10} is indeterminate.

VistBSN=1 if a tourist is on business purpose. Its coefficient β_{11} is expected to be positive.

Adultno = Number of adults in the travel party. This variable has not been common in the literature examining tourist length of stay. However, most travelers would be adults rather than children leading to the possibility of it behaving more or less the same as the total number of visitors in the party. Therefore its coefficient β_{12} is expected to be negative.

Tpartyno=1 if there are at least 2 members in the travel party, 0 otherwise. Its coefficient β_{13} is expected to be negative. This is because most travellers in a large party are on organized tours, which tend to be shorter than those of single travellers or smaller parties. Alegres and Pou(2007) found that travel party size has a negative influence on length of stay.

Frvist = 1 if a tourist has visited at least one African country before Tanzania, 0 otherwise. Its coefficient β_{14} is expected to be negative. A frequent visitor in this context must be either an explorer or a businessman, for these are the kind of people who travel frequently. These people in much the same way as frequent visitors, would rarely stay for a long time at a particular destination.

Fadests = 1 if a tourist has visited Tanzania at least once before the current visit and 0 otherwise. Its coefficient β_{15} is expected to be positive. The inference is that repeat visitor is very much attracted to the destination for one reason or another

and therefore is likely to stay longer . Barros and Correia (2007) as well as Menes et., al(2008) proved that repeat visitors stay longer.

Dist = the shortest distance by air between Tanzania and the tourist country of origin measured in miles. Based on the model formulation distance would be a proxy for travel cost to the destination and based on the model formulation, distance should reduce the total time to be spent at a destination. Therefore its coefficient β_{16} is expected to be negative, although Menez et al. (208) found no clear pattern between distance and time spent at a destination. So far few studies have included this variable in the model.

Nosites=1 if a tourist had visited more than one site and 0 otherwise. Its coefficients β_{17} is expected to be positive. The logic here is much clearer than anywhere else. For a tourist to visit many sites he needs more time.

Price= the relative cot of living between Tanzania and the tourist's country of origin measured as price= $CPI_{TZ}^{t}/CPI_{O}^{t} * ER_{/t}$ where CPI_{TZ}^{t}, is the consumer price index in Tanzania and CPI_{O}^{t} is the consumer price index in a tourist's country of origin. It coefficient β_{18} is expected to be negative. This variable can also be a proxy for the cost of a tourist's holiday time.

Peak=1 if a tourist travelled during the peak season (July -September) and 0 otherwise. Its coefficient is expected to be positive if a season's abundance of tourists reflects their expectations of the destination's attractiveness especially in terms of tourist goods and activities.

Isource= 1 is source of information is from the word-of-mouth, 0 otherwise. Its coefficient β_{19} is indeterminate, depending on a number of factors. First, if the

returning tourists depict a negative picture of the destination back home, a coming tourist would certainly plan to stay for a short while. But that will also depend on what he finds after reaching the destination, for he/she can always extend his/her length of stay. Therefore its coefficient β_{20} is indeterminate.

Africa, Asia, Europe, MEast, NAmerica, SAmerica, represents the dummies for Africa, Asia, Europe, the Middle East, North America and South America.

year1, year 2 and year3 are dummies for the year 2001, year 2007 and year 2008.

A summary of this discussion is presented in Table5.1.

Table 5.1 A prior direction of the relationship between tourist length of stay and the explanatory variables

Variable	acronym	Direction
Age group of the respondent	Age	indeterminate
Number of females in the travel party	Females	indeterminate
Income of the travel party	income	+
Tanzania nominal exchange rate against that of a tourist country	ER	+
Tourist familiarity with the English language	EL	+
Daily tourist per capita expenditure	Pcapita	-
Number of children in the travel party	Childno	-
Presence of a child in the travel party	Childprese	-
Travel arrangements by a tourist(Package)	Tarra	-
A tourist visiting Friends and Relatives	VistFRD	indeterminate
A tourist coming for Leisure and Recreation	VistLSR	indeterminate
A tourist on Business	VistBSN	indeterminate
Number of adults in the travel party	Adultno	-
Travel party size(number of people in the travel group)	Trpartyno	-
Tourist frequency of travelling (not to Tanzania only)	Frvist	-
Tourist familiarly with Tanzania(frequency of visiting Tanzania)	Fadests	+
Distance between Tanzania(DSM) and the tourist country's capital	Dist	-
Number of sites visited by a tourist during his stay in Tanzania	Nosites	+
Consumer price of Tanzania relative to that of a tourist country	price	-
The peak season of travel by a tourist	Peak	+
Information by word-of-mouth	Isource	indeterminate

Equation 5.2 could be analyzed by OLS, but there a number of defects in using OLS to analyze a time variable. Cameron and Trivedi (2005) addresses the following problems of using OLS in analyzing a time variable. First is the lack of normality as most of the time observation is positively skewed. Second is the fact that in most surveys involving the time to an event, observations are censored. In other words, the individuals are observed before the study was completed or the study comes to an end before the event has occured. The former is a case of left censoring while the latter is a case of right censoring. However, in this study, as it will be observed, tourists are interviewed during their departure and thus censoring is not there. Third there may be the issue that a covariate like age may change during the duration, and the assumption of $E(x'\varepsilon) = 0$ .may be violated, resulting in inconsistent coefficients. If the duration is short change in ages may not be substantial. Fourth is the fact that there is no guarantee that OLS will predict positive values of time. This limitation could be serious in relation to prediction.

Because of the above problems the study opted to use survival analysis instead of OLS. However, OLS was also used for the log of tourist length of stay for exploratory purposes, and indeed it gave some useful suggestions for survival analysis, like assessing regressors' endogeneity as well as hinting the most appropriate survival model. The next section gives a brief review of the theory of survival analysis.

5.3.2 An Overview of Survival Analysis

The review is based on the work by Cameron and Trivedi (2005). One may begin by considering the cumulative distribution of the variable time given as $F(t)$ and its density function given by $f(t)$. The relationship between the two is such that

$$f(t) = dF(t)/dt \tag{5.3a}$$

or

$$F(t) = P(T \leq t) = \int_0^t f(s)ds \tag{5.3b}$$

An equally important concept in duration analysis is the *survival* function which is in fact the greater than or equal cumulative function, defined as

$$S(t) = P(T \geq t) = 1 - F(t) \tag{5.3c}$$

This is the probability that a particular duration equals or exceeds time t. Another key concept is the hazard function. This is an instantaneous probability of leaving a state conditional on survival to time t. It is defined as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T \leq t + \Delta t / T \geq t}{\Delta t} = \frac{f(t)}{S(t)} \tag{5.3d}$$

It follows from (18), that

$$\lambda(t) = -dlin(S(t))/dt \Rightarrow S(t) = \exp(-\int_0^t \lambda(u)du) \tag{5.3e}$$

A final related function is the cumulative hazard function or integrated hazard function define as

$$\Lambda(t) = \int_0^t \lambda(t)dt = -\ln S(t) \tag{5.3f}$$

These functions can be estimated using both non-parametric and parametric approaches. Non-parametric estimation can be carried out as described below:

Let

d_j be number of durations (spells) ending at time j;

m_j be the number of spells censored in (t_j, t_{j+1})

r_j spells at risk at time tj.

Then accordingly the hazard rate is estimated as $\hat{\lambda}_j(t) = \frac{d_j}{r_j}$

and the survival function known as the Kaplan-Meier estimator as

$$\hat{S}(t) = \Pi'_{t_j}(1 - \hat{\lambda}_j(t)) =$$

$$\hat{S}(t) = \Pi'_{t_j}(\frac{r_j - \hat{\lambda}_j(t)}{r_j}) \qquad (5.3g)$$

The parametric estimation involves estimating the hazard function through regression analysis. As these functions are non-linear, the maximum likelihood method is used to estimate them. Among the popular hazard functions used in survival analysis are Exponential, Weibull, and Gompertz distributions, whose hazard functions are respectively, γ, $\gamma\alpha t^{\alpha-1}$ and $\gamma\exp(\alpha t)$.These are examples of proportional hazard models (PH), because their hazard functions can be written in the form $\lambda(t/x) = \lambda_o(t,\alpha)\phi(x,\beta)$, where $\lambda_o(t,\alpha)$ is the baseline hazard expressed as a function of time and $\phi(x,\beta)$ is the relative hazard expressed as function of the individuals' covariates.

It is important to note that the hazard rate for the exponential distribution is constant across time (length of stay.) The hazard rate for the Weibull distribution may increase with time or decrease with time, depending on the value α.When $\alpha > 1$ it increases, whereas when $\alpha < 1$ it decreases. When $\alpha = 1$ it is constant and in fact it reduces to the hazard rate of the exponential distribution. Therefore the exponential hazard is a special case of the Weibull hazard when $\alpha = 1$. As regards the hazard function for the Gompertz distribution, it either monotonically increases when $\alpha > 1$ or monotonically decreases in the case $\alpha < 1$. When $\alpha = 1$ it reduces to that of the exponential distribution.

Other approaches include the log-logistic, log-normal and gamma distributions, which fall under the Accelerated Failure Time model (AFT). They are called the accelerated

time failure rate because, unlike the proportional hazards, the covariates lead to changes in the baseline hazards. The hazards are formed when modelling the natural log of time rather than time itself. In other words, when modelling $\ln(t) = x\beta + \mu$, the hazard will result in either log-logistic, log-normal or gamma, depending on the specification of the distribution of μ.

The hazards for the log-logistic, Gamma and log-normal distributions are respectively:

$$\frac{\alpha\gamma^{\alpha}t^{\alpha-1}}{[1+(\gamma t)^{\alpha}]},\ \frac{\gamma(\gamma t)^{\alpha-1}\exp[-(\gamma t)]}{\Gamma(\alpha)[1-I(\alpha,\gamma t)]},\ \frac{\exp(-(\ln t-\mu)^2/2\sigma^2}{\sigma\sqrt{2\pi}}\Big/[1-\varphi((\ln t-\mu)/\sigma)]$$

One can easily derive these hazard functions by following the steps given in equations (5.3a) to (5.3d). Appendix 5.1 provides the derivations of all the mentioned hazards. Two of the proportional hazards mentioned before also follow under AFT. These are the Exponential and Weibull hazards. The survival analysis of equation (5.2) can now be formulated as follows:

$\lambda(t/x) = \lambda_o(t,\alpha)\phi(x,\beta)$ where

$\lambda_o(t,\alpha)$ = the baseline hazard function

β = a vector of covariates so that

$$x'\beta = \beta_o + \beta_1 Age_i + \beta_2 Females_i + \beta_3 GDP_i + \beta_4 Er + \beta_5 El + \beta_6 Pcapita + \beta_7 Childpresence + \beta_8 childno + \beta_9 Tarra + \beta_{10} VistFRD + \beta_{11} VistLSR + \beta_{12} VistBSN + \beta_{13} Adultno + \beta_{14} Tpartyno + \beta_{15} Frvist + \beta_{16} Fadest + \beta_{17} Dist + \beta_{18} Nosites + \beta_{19} \Pr ice + \beta_{20} Peak + \beta_{21} Isource + Africa + Asia + Europe + MEast + NAmerica + SAmerica + year1 + year2 + year3 + \varepsilon \quad \cdots(5.4)$$

5.3.3 Addressing the Endogeneity of the Regressors.

As said earlier in the introduction, most studies have ignored the possibility of having endogenous regressors when modelling tourist length of stay. This study is particularly interested in assessing the endogeneity caused by tourist travel arrangements. When an

assessment was done using an instrumental variable technique it was proved that travel arrangements is an endogenous variable in the OLS model for length of stay. The variables child presence and dummy for a visit for leisure and recreation were used as instruments in assessing the endogeneity of travel arrangements (see appendix 5.1). Both the Sargan test and Basmann test for overidentifying restrictions/exogeneity proved that the instruments were exogenous (see appendix 5.1). This proof not only provides confidence in the entire test but also in the model specification (Green 2003, Cameron and Trivedi, 2005). The test also provides extra knowledge on the exogeneity of the variables child presence as well visits on leisure and recreation.

In additional to these tests, the variable peak season which is a proxy for destination attributes was also proved to be exogenous (see appendix 5.1). By the virtue of these few variables which were proved to be exogenous (visits for leisure and recreation, peak), it was possible to test the null hypothesis of destination attributes being more influential than demographic and trip-related characteristics, because the variable visits for leisure and recreation was found to be more influential than the variable peak season.

Even though not every variable was proved to be exogenous and some like travel arrangements were definitely endogenous, an examination of variance inflation factors (VIF) showed that all variables had very small VIF (<10), which not only guarantees the fact that regressors are immune to multcollinearlity, but also that any endogenous regressor may not invalidate the exogeneity of other variables and hence of the entire model (Cameron and Trivedi, 2005). The VIF are provided in Table 5.2 of appendix 5.2

Despite what has been said regarding the VIF, to be on the safe side the variable travel arrangements was removed from the model. Alternatively, the said instruments could have been used in its place but the aim was simply to show that travel arrangements is an endogenous variable in the model explaining tourist length of stay and not to estimate it. Besides, the null hypothesis could still be tested even after the removal of the variable travel arrangements using other trip-related characteristics. The instruments were not used so as to avoid other risks associated with instrumental variable techniques, such as model inefficiency (Cameron and Trivedi, 2005).

5.3.4 Variables and their Sources

The study used survey data from TTSS. Although there were six years of survey by the TTSS (2001, 2004, 2005, 2006, 2007, 2008), the study mainly used data for the years 2001, 2007 and 2008, which had similar and relatively more explanatory variables than the rest of the years. These three years made a total of 30,782 observations. But only 25,880 observations out of 30,782 were used, the rest being excluded due to missing and/or unacceptable values.

The following variables were obtained directly from the TTSS surveys: age, number of females in the travel party (Females), travel party number (Tpartyno), length of stay by a tourist (Lstay), travel arrangements i.e. package or non package tour (Tarra), tourists visiting friends and relatives (VistFRD), tourists visiting for Leisure and Recreation (VistLSR), tourists visiting for Business purposes (VistBSN) and number of sites visited by a tourist in Tanzania (Nosites).

The following variables were derived from the TTSS survey:

(1) (Expenditure per person per day). The survey data reported total party expenditure. In order to obtain expenditure per person per night, the total party expenditure was divided by the size of the group and by the number of days the party stayed. There were two types of expenditure, depending on whether a tourist was on a package tour or a non package tour. For tourists on a non-package tour no adjustment was made in the data, but for But the data was adjusted for tourists on package tour, whose bills are paid by travel agents in their home countries. According to TTSS (2001), the actual expenditure accruing to Tanzania can be found by deducting 10% of the package cost as the amount paid to the agents. The same was done in this study. According to TTSS (2001), after deducting the 10% commission, the international travel fare, was also deducted for tourists whose packages include that.

Another consideration was to set the minimum expenditure and the maximum expenditure as proposed by TTSS (2001). According to TTSS (2001), meaningful expenditure was regarded as not being less than 10 dollars per person per night, whereas the maximum expenditure per tourist was set at USD1000. The same was adopted by this study. This was important for two reasons. First is the need to compare the findings from this study with those of TTSS. Second is to minimize extreme values in the data.

(2) Frequent vistor (Frvist). A frequent visitor was taken to be a person who had visited at least one other African country before visiting Tanzania. There was a question asking tourists to compare the cost of visiting Tanzania against that of neighbouring destinations such as South Africa and Kenya. A tourist responding to

this question was regarded as a frequent visitor. This information was used because there was no question asking a tourist directly whether he/she is a frequent visitor.

(3) Familiarity with the destination (Fadest). A tourist who was not on his first visit to Tanzania was regarded as being familiar with the country. Unfortunately this variable does not capture the intensity of familiarity.

The following variables were taken from other sources:

(1). Income was proxied by a tourist country's GDP. The figures were obtained from the IMF (2009).

(2) Exchange rate (Er) was taken from the Italian Bank, a source also used by the Economist website.

(3) English language proficiency (El). A tourist coming from a country where the official language is English was regarded as an English speaking person. The country's official language was obtained from the internet (www.yahoo.com) by searching the country's profile.

(4) Distance travelled (Dist). This is the shortest ground distance between Tanzania and a tourist's country of origin, measured from Dar es Salaam to the country's capital. The variable was sourced from the internet using the online distance calculator.

(5) Relative cost of living between a tourist country of origin and Tanzania (Price). Figures in the Consumer Price Indexes were obtained from the IMF (2009).

5.4 Results and Discussion

5.4.1 Summary Statistics

Table 5.3 provides summary statistics of the variables used in the analysis

Table 5.3: Summary statistics of the variables used in the analysis of tourist length of stay

Variable	Mean	Std. Dev.	Min	Max	CV
Lstay(Days)	12.69	13.60	1	360	107.17
Age(Ranks)	3	0.69	1	4	23
Females	0.89	0.90	0	21	101.12
GDP(Mil.USD)	29123.30	15379.39	54.62	113044	52.81
Er	965.49	603.92	0.030067	3229.16	62.55
El	0.54	0.50	0	1	94.30
Pcapita(USD)	191.69	206.92	9.5	1017.98	107.95
Childno	0.135	0.58	0	13	429.63
Tarra	0.53	0.50	0	1	94.34
VistFRD	0.08	0.27	0	1	337.5
VistLSR	0.77	0.42	0	1	54.54
VistBSN	0.10	0.30	0	1	300
Adultno	1.75	1.28	1	31	73.14
Tpartyno2	0.56	0.50	0	1	89.18
Frvists	0.53	0.50	0	1	94.34
Fadest	0.37	0.48	0	1	129.73
Dist(Miles)	5287.09	2389.07	419	9527	45.19
Nosites	0.65	0.48	0	1	73.85
Price	0.05	0.58	3.24E-05	26.007	1160
Peak	0.74	0.44	0	1	59.46
Isource	0.39	0.49	0	1	125.64
Africa	0.09	0.29	0	1	322.22
Asia	0.08	0.27	0	1	337.5
MEast	0.009	0.09	0	1	1000
SAmerica	0.008	0.09	0	1	1125
year1	0.62	0.49	0	1	79.03
year1	0.62	0.49	0	1	79.03
year2	0.10	0.30	0	1	300
year3	0.28	0.45	0	1	160.71

Table 5.3 indicates that most of the variables have a reasonable variation except for age, which is due to the fact that the survey reported on age groups rather than the individual ages. It is interesting to note that length of stay is positively skewed (skewness=7.9).This observation is consistent with what is described in the literature regarding the distribution of time (Cameron and Trivedi 2005, Greene, 2003).This

skewness justifies the use of logarithmic transformation when modelling length of stay and/or the consideration of survival models which accommodates among other things the skewed nature of time distribution.

5.4.2 The Correlation Analysis

The correlation matrix of the variables used in the analysis is given in Table 4.3 of appendix 4.3.The matrix indicates that most of the variables correlate in the expected manner. It was also observed that for most of the regressors there was no strong correlation among them, suggesting little multcollinearlity. More details can be found in Table 4.4 of appendix 4.3

5.4.3 OLS Estimation Results of the Log Linear Model on Length of Stay

The results of the log-linear model on length of stay are presented in Table 5.4. These results are important in suggesting the influence of the covariates and suggesting the appropriate survival model for length of stay. Therefore the estimates in this model can seldom be taken as final.

Table 5.4: OLS regression results of the log of tourist length of stay.

Variable	Coef.	Std. Err.	t	P>t
Age	-0.10***	0.006	-17	0.000
Female	0.033***	0.007	4.8	0.000
GDP	2.6E-06***	5.19E-07	4.94	0.000
Er	-4E-05***	1.3E-05	-3.24	0.001
El	0.053***	0.013	4.13	0.000
Childno	-0.026***	0.009	-3.6	0.000
VistFRD	0.023	0.023	0.99	0.321
VistLSR	-0.28***	0.019	-14.8	0.000
VistBSN	-0.49***	0.022	-22.3	0.000
Adultno	-0.023***	0.005	-4.72	0.000
Tpartyno2	-0.034***	0.010	-3.38	0.001
Frvists	-0.199***	0.009	-23.2	0.000
Fadest	0.060***	0.009	6.8	0.000
Dist(Miles)	-2E-05***	3.15E-06	-6.89	0.000
Nosites	0.335***	0.009	38.7	0.000
Price	0.091***	0.00949	9.55	0.000
Peak	-0.012*	0.007	-1.68	0.094
Isource	-0.123***	0.0085	-14.5	0.000
Africa	-0.365***	0.032	-11.4	0.000
Asia	-0.175***	0.018	-9.75	0.000
MEast	-0.163***	0.043	-3.79	0.000
SAmerica	-0.037	0.049	-0.76	0.449
year2	(dropped)			
year3	0.02837	0.019	1.52	0.129
cons	-0.1049	0.015	-6.88	0.000
R2	17			
F	227.20			
P>F	0.000			
n	25880			

NB: *significant at 10%, **significant at 5%, ***significant 1%

Table 5.4 suggests that number of females, GDP, exchange rate, familiarity with the destination, number of sites visited and peak season have a positive influence on tourist length of stay. The Table also suggests that older age, number of children, higher per capita income, being a business visitor/a visitor for leisure, more adults /huge travel party , being a frequent visitor, being a tourist from a distance,

expenditure, facing a relatively higher cost of living at a destination and receiving information by word-of-mouth are all associated with negative tourist length of stay.

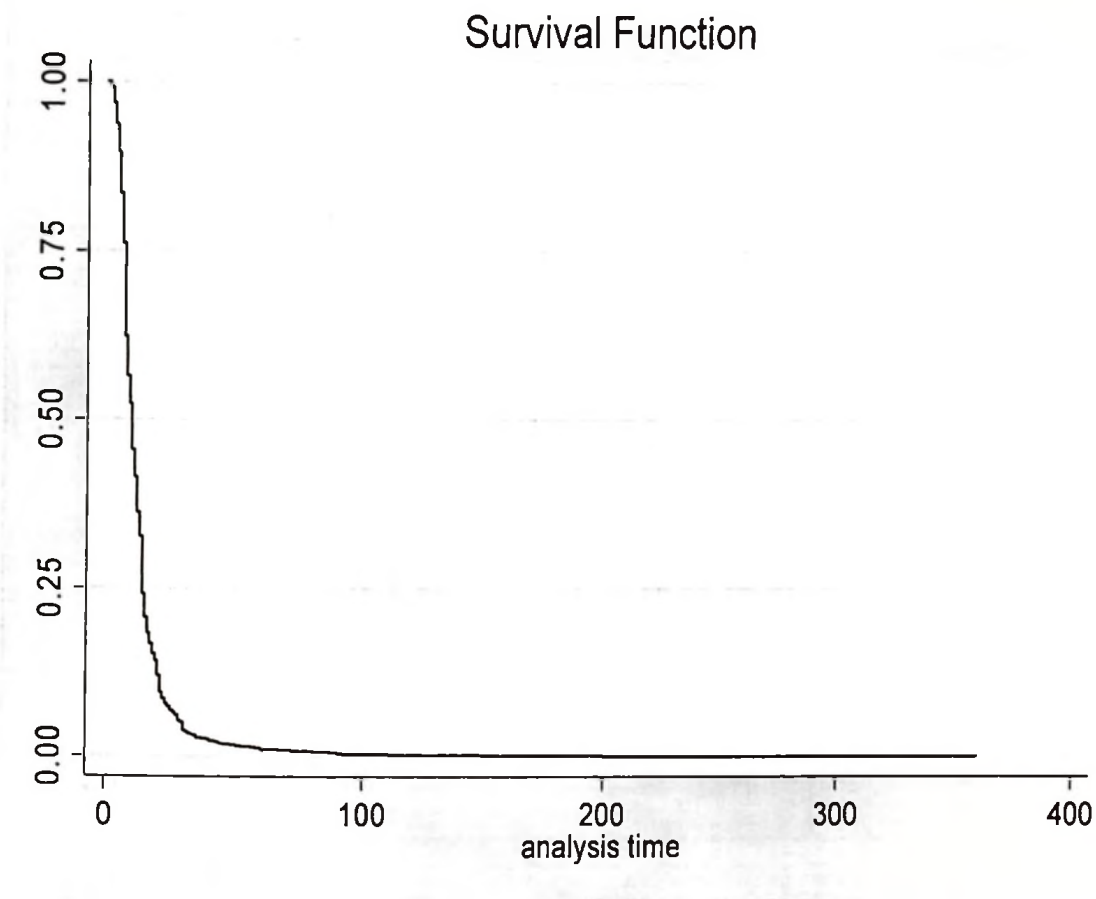
The diagnostic test indicates that the model assumptions are quite satisfactory (see-Appendix 5.4). The histogram of the residuals appears to be symmetrical consistent with the property of the normal distribution, possibly indicating that the log-normal and/or log-logistic hazard functions are the most appropriate survival models for the data as the two of them are also symmetrical.

5.4.6 Results From the Survival Analysis

(i) . Non-parametric estimation

(a) Figure 5.2 gives the Kaplan Meir survival function.

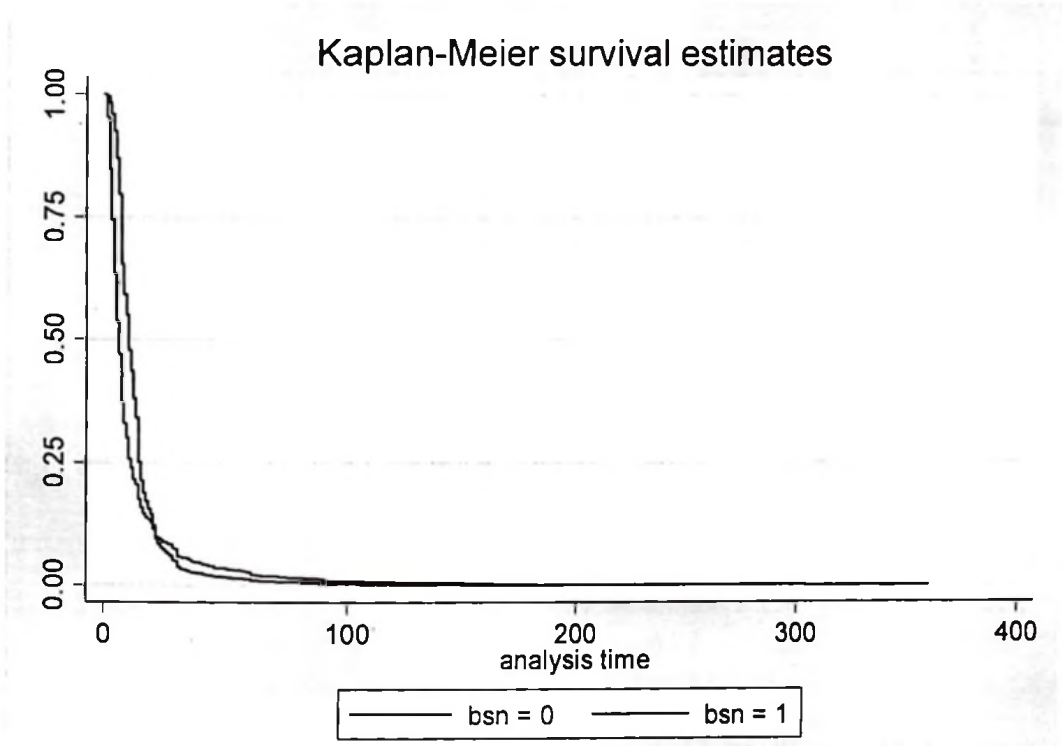
Figure 5.2: Kaplan- Meir survival Function



The figure appears to follow the normal trend of the survival function. From the figure the median survival time is estimated to be 10 days and the mean time is 13.The two figures are quite close to those established by TTSS (2001).

(b) Figure 5.3 gives the Kaplan Meir survival function of tourist on business visit against non-business visit). Note that bsn=1 stands for business visit while bsn=0 stands for non-business visit (refer to section 5.3.1).

Figure 5.3: Kaplan Meier survival function by Business visit vs. non-Business visit



The figure indicates that the two survival curves interact; suggesting that the hazard function of the data under the study is not proportional to the covariates.

(c) Figure 5.4 gives the Kaplan Meir survival functions of tourists by age group. Note that the numbers 1, 2, 3 and 4 stand for age groups <18, 18-35, 36-55, and 55+ respectively (refer to section 5.3.1).

Figure 5.4: Kaplan Meier survival function by age

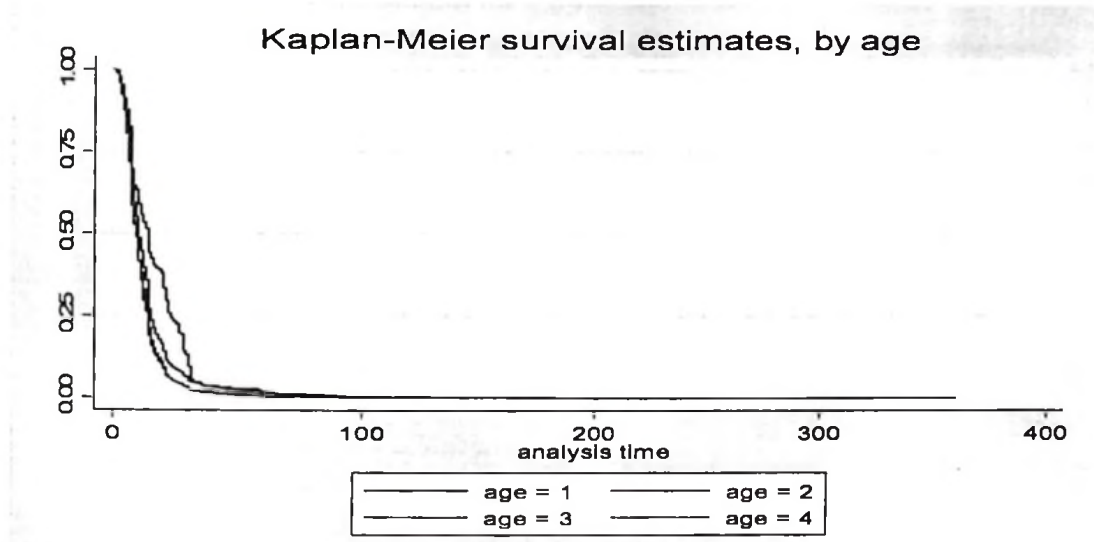


Figure 5.4 like figure 5.3, suggests that the assumption of proportional hazard is not correct as the indicated curves cross each other. This suggestion can be further supported by what was seen when a log-linear model was fitted (Appendix 5.4). The histogram of the residuals from this model depicted an almost perfect normal distribution, which suggested that the log-normal and /or the log-logistic hazard models which all belong to AFT hazards could be the most appropriate. The Schoenfeld test on the residuals of the Cox PH model confirms this suggestion as it rejects the null hypothesis of proportional hazard (chi2 (24) =1057, p=0.000). This statistical test suggests that our consideration should be entirely focused on the AFT hazards models, which covers among others Exponential, Weibull, Log-normal, Log-logistic and Gamma distributions. In order to choose the best models from among these, Akaieke[25] Information Criteria (AIC) is employed, because some of the models

[25] $AIC = \left(-2\ln \hat{L}(M_\beta) + 2P\right)/N$ Where P is the number of parameters in the model, N is the number of observations and $\hat{L}(M_\beta)$ is the likelihood of the fitted model. The smaller the value of AIC the better the fit of the model (Scott long, 1997).

do not nest in each other. According to the AIC, a model estimate with the least value of AIC is the best. Table 5.5 presents the AIC values for the five mentioned models as well as the log-likelihood values.

Table 5.5: AIC and Log-likelihood values from the survival models

Model	Log-like hood	AIC
Exponential distribution Hazard	-31491.08	63032
Weibull distribution Hazard	-28675.64	57403
log-normal distribution hazard	-24841.62	49735
Lo-Logistic distribution hazard	-23995.62	48043
Gamma distribution hazard	-23759.97	49573

Based on the AIC values, the log-logistic model outweighs the rest as it has the least AIC value (48043). Therefore Table 5.6 provides parametric estimation results from the log-logistic along side those of the Cox regression model for comparison purposes.

(ii) Parametric estimation

Table 5.6: Log-logistic and Cox regression of the length of tourist stay

	Cox-Regression		Log-logistic	
Variable	Hazard Ratio	Std.Err	Coefficients.	Std.Err
Age	1.17***	0.011	-0.09***	0.005
Females	0.97***	0.011	0.02***	0.006
GDP	1.00***	0.000	2.66e-06***	0.000
Er	1.00***	0.000	-3.1e-05***	0.000
El	0.94***	0.019	0.05***	0.012
Childno	1.04***	0.013	-0.02***	0.007
VistFRD	0.96	0.035	0.07***	0.023
VistLSR	1.63***	0.049	-0.25***	0.019
VistBSN	1.62***	0.057	-0.52***	0.023
Adultno	1.03***	0.008	-0.02***	0.005
Tpartyno	1.10***	0.018	-0.02**	0.009
Frvists	1.27***	0.017	-0.19***	0.008
Fadest	0.87***	0.012	0.06***	0.008
Dist	1.00***	0.000	-2.3e-05***	2.9e-06
Nosites	0.64***	0.009	0.34***	0.008
Peak	0.91***	0.014	0.08***	0.009
Price	1.01	0.009	-0.01	0.007
Isource	1.21***	0.016	-0.12***	0.008
Africa	1.65***	0.082	-0.33***	0.031
Asia	1.18***	0.034	-0.18***	0.017
MEast	1.29***	0.088	-0.15***	0.040
SAmerica	1.04	0.082	-0.05	0.046
year1	0.99	0.030	-0.03*	0.017
year3	1.14***	0.028	-0.15**	0.014
Constant	NA		2.78	0.040
likelihood	-237029		-22995.62	
Chi	3454.68		5470.21	
P>Chi	0.0000		0.0000	
n	25880		25,880	

NB: *significant at 10%, **significant at 5%, ***significant 1%

Table 5.6 indicates that the two models are highly significant. The results of the Cox-regression are given in terms of the hazard ratio, whereas the log-logistic results are given in terms of the coefficients of the explanatory variables. It is important to note that a hazard ratio smaller than 1 implies a negative impact of the covariate on the hazard ratio, which is equivalently to a positive impact of the covariate on length of

stay and the vice versa.

It is evident that the null hypothesis of destination attributes being more influential than the other attributes is rejected in both models. Rather the trip-related characteristics, as measured by a tourists' frequency of traveling (*frvist*), business visits (*VistBSN*) and leisure visits (*VistLSR*) are more influential than the other attributes.

Nevertheless, destination attributes as measured by season are quite significant (p=0.000, logit coefficient=0.08, hazard ratio=0.91), which points to the necessity of not neglecting the influence of these attributes in increasing tourist length of stay. The positive coefficient of the variable peak season implies that during the peak season (July –September) tourists stay much longer than otherwise. In other words, these tourists must have perceived the peak season as the most ideal period for recreation and other tourist activities. Unfortunately, as observed in chapter four, their per capita expenditure tends to decline with season, implying that their longer stay may not generate more revenue. Therefore stakeholders need to find ways to boost their per capita expenditure so as maximize tourism revenue. One of the possible reasons put forward in chapter four is that, owing to the growing influx of tourists during the peak season, the destination service providers are unable to meet their needs. This is calls for diversification of tourist activities across the country, such as in the western part, the southern higher lands and the lake zone, where tourist activities are less popular. This could be done by improving the appropriate infrastructure, particularly roads and electricity which are greatly lacking in those areas compared to the traditional tourist attraction areas of Northern Tanzania and Zanzibar.

The variable information source (Isource), which can also be regarded as a part of the destination attributes, is significant but with a negative sign. This implies that tourists whose source of information is word-of-mouth do not stay longer as compared to their counterparts. The reason is unclear, but probably the returning tourists do not paint a favourable picture, which discourage tourists from staying longer. This was proved by the survey by TTSS (2006) in which the majority of the tourists complained about the state of the infrastructure and accommodation. The survey further noted that the average length of stay by a tourist has remained at around 12 days since 2001. Nevertheless, the coefficient for this variable (-.12) is smaller than some of the trip-related characteristics leading, again to the rejection of the null hypothesis.

As regards the trip-related characteristics, tourist purpose of visit as measured by business visits and those for leisure purposes are the most influential variables in the three categories of variables. The coefficients of both visits for business visits and leisure visits appear with negative and significant signs, implying that these types of tourists do not stay long unlike tourists visiting for other purpose who stay much longer. Menezes et al. (2008) established a similar results for business visits when studying tourist length of stay in the Azores.

The next most influential trip-related characteristic is frequency of traveling, which has a negative and significant influence (-0.19), indicting that a frequent traveller is associated with shorter stays than an ordinary traveller. In fact this variable increases the hazard ratio for that kind of a tourist (1.3). As previously postulated (refer to section 5.3.1), a frequent visitor would allocate holiday time to various destinations compelling him to spend as little time as possible at a particular destination. This result

however, is contrary to the study by Govakali et al. (2007), who found that an experienced tourist spends much more time than others.

Another most influential trip-related characteristic is the number of children, which has a significant and negative influence (-0.02). There is not much in the literature regarding the influence of this variable on tourist length of stay. But comparably the influence of this variable could be equivalent to the influence of travelling with a family (as the family includes children) which was found to be negative by Menezes et al. (2008).

Other trip-related characteristics which influence tourist length of stay are travel party number (-0.02), number of adults (-0.02) and distance (-0.00002). Large travel parties are organized tours which tend to last for a short time. Similar results were obtained by Alegre and Pou (2006).With regard to distance, the negative coefficient should be expected because the longer the distance of travel the less time is spent at a destination. As said in section 5.3.1, distance is a proxy for the cost of travelling to a destination and indeed it should negatively affect tourist length of stay as the longer the travelling time the less is the time, less time is spent at the destination. The policy implication here is that ways need to be sought to attract direct flights which considerably reduce the travelling time to a destination. This should be undertaken along with the construction of airports of international standard to handle such a big number of direct flights.

Even though excluded from the model, owing to its endogeneity tourist travel arrangements (package vs. non-package) is one of the most influential trip-related

characteristics (refer to section 5.3.3). When included as an associate factor rather than a determinant, it had a significant and negative influence. In fact when instrumented it also yielded a negative influence. This means tourists on a package tour are associated with a shorter length of stay than those on a non-package tour. The same result was obtained by Govakali et al. (2007) as well as by Garcia and Raya (2008).

Apart from the trip-related characteristics, demographic characteristics were also found to be significant, though in general most of them are less influential than the trip-related characteristics.

The first of these is the age of a tourist. This variable negatively affects tourist length of stay. Another demographic variable is income as measured by per capita GDP. This variable positively influences tourist length of stay. This finding emphasizes the need by stakeholders to target markets with high per capita GDP. Exchange rate has also been found to positively influence tourist length of stay. The logic is the same as that of GDP, because in most cases the favourable exchange rate obtained by a tourist implies that he will have more income at a destination, giving him the opportunity to stay longer, given other things being equal. Lastly is the number of females which has a positive influence on tourist length of stay. This finding is consistent with that of Menezes et al.(2008), who found that female tourists to have a positive and significant influence on length of stay .

Some variables included in the model do not belong to the stated categories, namely, number of sites visited, relative cost of living (price) and regional dummies. Among these number of sites visited has a significant and positive influence on the length of

stay. The same result was obtained by Menezes et al. (2008) on the influence of the number of sites on tourist length of stay. One of the implications for policy from this result is that more advertisements of tourist attractions are needed, especially those found out of the Northern tourist zone as they are rarely visited. By so doing tourist length of stay based on the number of sites, will definitely increase. The variable price was insignificant. As regards the regional dummies, all of them had a negative influence except the dummies for Europe and North America (which do not appear in Table 5.6 because their inclusion led to evidence of multcollinearlity. Therefore to establish their influence the two dummies were assessed in the absence of the other four regional dummies and depicted positive influence.

The discussed results give a general account of the factors determining tourist length of stay in Tanzania, irrespective of tourist country of origin. Table 5.7 provides regional specific regression results. For ease of presentation standard errors/z-values are omitted. The primary aim is to compare the influence of the coefficients across the regions, but the significance level for each regressor is still indicted in all the regions.

Table 5.7: Log-logistic regression results of the length of tourist stay: A Comparison by region.

	World	Africa	Asia	Europe	MEast	NA	SA
Age	-0.09***	-.06**	-.10***	-.09***	-.15***	-.09***	-.06
Females	0.02***	.09***	-.02	.02***	.07	.004	-.09
GDP	2.66e-06***	8.8e-06	-2.5e-06	3.8e-06***	-8.6e-07	.0007	1.e-06
Er	-3.1e-05***	.00004	-5.8e-05	1.4e-05	.0001**	-.03	-.0004
El	0.05***	.06	.20**	-.02	dropped	dropped	.15
Childno	-0.02***	.024	-.02	-.02*	-.041	-.03**	.02
VistFRD	0.07***	.36***	.06	.06*	.6**	-.11***	.28
VistLSR	-0.25***	.17***	-.16*	-.30***	.39*	-.37***	-.18
VistBSN	-0.52***	-.21***	-.24**	-.62***	-.07	-.41***	-.47*
Adultno	-0.02***	-.05***	-.01	-.02***	-.01	-.003	.17***
Tpartyno	-0.02**	-.12***	-.03	-.001	-.11	-.08***	-.35***
Frvists	-0.19***	-.045	-.24***	-.19***	-.27***	-.24***	-.11
Fadest	0.06***	.02	.04**	.07***	.04	.08***	-.011
Dist	-2e-05**	.0002***	6.7e-05	1.2e-05	-.0003	.0008***	-.00005
Nosites	0.34***	.26***	.36***	.35***	.45***	.31***	.33***
Peak	0.08***	.20***	.03	.08***	.15	.05***	.26**
Price	-0.01	.02**	-.011	-0.001	.19	-.13***	-.07***
Isource	-0.12***	-.09***	-.04	-.11***	-.17**	-.14***	-.13
Constant	2.78	1.4***	1.1***	2.6****	2.8***	4.1***	2.3**
Likelihood	-22995.6	-2679.8	-2150.6	-13407.5	-193.64	-4907.03	-166.40
Chi	5470.21	400.50	298.57	3175.55	94.27	884.22	73.48
P>Chi	0.0000	0.0000	0.000	0.0000	0.0000	0.0000	0.0000
N	25,880	2379	2028	15560	232	5485	196

NB: *significant at 10%, **significant at 5%, ***significant 1%

Table 5.7 provides the estimated coefficients of the log-logistic regression across the worlds' regions. As before, a negative covariate affects tourist length of stay

negatively which is equivalent to affecting the hazard rate positively (hazard rate>1).Vice-versa is also true.

The results from Table 5.7 still indicate that the trip- related characteristics (frequency of visit, travel party number, purpose of visit) are the most influential determinants of tourist length of stay in all the regions. This leads again to the rejection of the null hypothesis that destination attributes are more influential than the demographic and trip-related characteristics.

The results also indicate that most of the variables behaved similarly in all the regions, with the exception of a few variables. Distant tourists from Africa behaved differently from the general observation, as they tend to stay longer unlike in the general results. Furthermore, tourists from North America who visit friends and relatives tend not to stay long, contrary to the general results.

Apart from suggesting that there is not much difference in policy across the regions, these results provide confidence in the overall results, particularly when population heterogeneity is taken into account. One of the key issues in survival analysis is how to generalize the hazard rate to a heterogeneous population (Cameron & Trived, 2005).Therefore one needs not worry too much about applying the general results of the hazard rate to different categories of tourists.

5.5 Conclusion

The chapter has highlighted the determinants of tourist length of stay in Tanzania. Several variables have been examined as possible determinants of tourist length of

stay, among which trip-related characteristics appear to be the most influential thus requiring the immediate attention of stakeholders in the industry.

Unlike some other works, this study has attempted to address the possibility of having endogenous regressors when modelling tourist length of stay that is not emphasized in previous studies. It has been found that one potential endogenous variable when modelling tourist length of stay is tourist travel arrangements (packages vs. non-package).

CHAPTER SIX

DETERMINANTS OF A PACKAGE TOUR

6.1 Introduction

It was shown in chapter four that among the key determinants of tourist per capita expenditure is tourist travel arrangements (package or non-package tour).Tourists on package tours are associated with higher per capita expenditure than those on non-package tours[26].

The two types of tourists have different policy implications. For example, though tourists on package tour in general spend more per person, their expenditure does not directly benefit the local population. Their expenditure is channeled to Tanzania as transfer payments by the travel agent to the hotel owners and tour agents. In other words, the local community which tourists visit would benefit more from tourists on non-package tours. Those on package tours, especially those whose package includes everything they spend their money on, including their hotels and services (enclave tourism). This implies that for policy analysis the government and other stakeholders need to understand what determines a tourist's choice of package or non-package tour in Tanzania.

This study therefore seeks to establish what determines a tourist's choice of a package tour. The study considers a number of socio-economic variables as well as trip-related characteristics as possible candidates for influencing a tourist's choice of a package tour. The aim is to test the hypothesis that destination characteristics are much more

[26] This argument should not be taken to be automatic. Some studies find the opposite to be the case. Nevertheless, either of the arguments justifies an investigation of factors determining the two modes of travelling.

influential than demographic and the trip-related characteristic as regards tourist choice of package or non-package travel.

In Tanzania, a similar study was done by TTSS (2001) and by Nzuki (2006). The study by TTSS (2001) did not aim at finding out what determines a tourist's choice of a package tour, but at producing a tourist expenditure model in Tanzania. In the course of this pursuit TTSS (2001) described the factors associated with tourists on package tour in Tanzania. The study by Nzuki (2006) examined the influence of young children on the purchase of package travel to Tanzania. His study was much more specific in the sense that the presence /absence of children were a key target. This study however goes beyond that, as it seeks to find out the determinants of a package tour in general.

Globally there exists a number of studies on the determinants of a package tour, such as those by Hsie et al.(1993), Corcoran et al. (1996), Te Lang et al.(1997) and by uIO and Hallo(2009). Most of these studies address the determinants of tourist choice of a package tour using cross-sectional data. There is always the possibility of having inconsistent estimates in cross-sectional data if there are regressors which correlate with omitted variables (Cameron and Trivedi, 2005).

Surprisingly, in most of these studies this problem has not been addressed (see for example- Hsie et al. (1993), Corcoran et al. (1996) and Te Lang et al.(1997)). For example, the modelling of choice of package or non-package travel may encompass tourist length of stay as an explanatory variable. See for example (Hsie et al. (1993), Te Lang et al.(1997), Io and Hallo (2009)). The inclusion of this variable in the model assumes that the variable is exogenous. However, in some cases, a tourist staying

longer can be influenced by travel arrangements. When a tourist visits a particular destination, is given a visa which exceeds the planned length of stay and therefore there is always a room for the extension of at least a few days if the need arises. This is especially possible for tourist on non-package tours, who are visiting friends and relatives or any other purpose, excluding leisure and business visits (see Figures 2.18 and 2.19 in chapter two).

One of the ways to avoid this endogeneity is to exclude extended length of stay by interviewing tourists when they enter the country. Unfortunately, in most studies, such as by TTSS (2001), tourists are interviewed at the time of their departure and therefore the reported length of stay may always include the extended length of stay. Alternatively, tourists could still be interviewed when at departing, but with a question asking them both about the actual length of stay and the intended length of stay. This study has made an attempt to show that length of stay is indeed an endogenous variable in a model for choice of travel arrangements.

The rest of the sections in this chapter are organized as follows: Section 6.2.1 reviews the theoretical literature, section 6.2.2 reviews the empirical literature review and section 6.3 describes the methodology. Section 6.4 presents the results and discussion and section 6.5 concludes the chapter.

6.2 Literature Review

6.2.1. Theoretical Review.

Package or non-package tour is a commodity like any other commodity in consumer studies. Conventionally the price of a package tour as well as the income of the buyer (a tourist) would be among the determinants of its demand. Price and income

determine the demand for a commodity other things being equal. These other things can be explained by the theory of reasoned action (TRA) which when borrowed in consumers studies translates into purchase intention theory. TRA as proposed by Fishbein and Ajzen(1975) suggests that an individual behaviour intention is a function of his own attitudes about the action as well as the subjective norms . The subjective norms refer to the society opinion and expectation of such an action.

In the context of TRA, the choice of a package or non package tour could be explained from a number of perspectives. One of the advantages of a package tour is the avoidance of risks(Nzuki,2006).A buyer of a package tour would have to perceive among other things the expected risks of a non-package tour. This perception will be affected by his personal feelings about the potential and actual risk (Lee and Yi, 2008). Similarly, a buyer might be motivated by a non-package tour by virtue of its flexibility as opposed to the package tour. This flexibility is certainly both real and the attitude of the mind. In addition, the purchase of either a package or non package tour may be influenced by society members such as friends, spouses and children.

Hsie et al. (1993) while studying what determines a traveller's choice between package and non package travel argues that many factors affect the consumer decision process. These factors can come from marketing (e.g. product quality, price, distinctiveness), social sources (e.g. family, reference group), individual differences (e.g. socio- demographics, lifestyle, personality), or psychological processes (e.g. motivation, perception). He further argues that, since travel behaviour is a special form of consumption, it may be affected by additional factors, such as travel characteristics, destination attributes and past experience.

Barros and Correia (2007) argue that a tourist's action can be explained using the theory of the tourist decision process given by Mathieson (1984). The theory asserts that the tourist decision process is motivated by the four characteristics of tourist profile, trip-related characteristics, trip awareness and destination attributes. This theory is an extension of TRA in the study of tourist purchase intention. This is because individual attitudes to an action can be reflected his physical and psychological makeup such as age, gender and income which in essence are elements of individual profiles. The subjective norms which encompass society and the environment are the trip-related characteristics and destination attributes.

In summary, the theory of the tourist decision process and/or the TRA suggest that a tourist's choice of a package or non-package tour is a function of his socio economic attributes, trip-related characteristics and the destination attributes.

As regards destination attributes, tourist source of information (word-of-mouth versus other sources) is used as a proxy. The variation in choice of package or non-package tour due to tourist sources of information should reflect the picture perceived by a tourist of the destination attributes, such as security and the availability of services as well as of the mode of travel arrangement (package versus non-package). Assuming that the information provided is genuine, then this variable should reasonably proxy for destination attributes. In fact there is no reason to doubt the information provided by returning tourists because they cannot conspire to give wrong information about the destination. Alternatively the season can also be used. But the choice of a package or non-package tour is greatly affected by the desire to avoid risks encountered at a destination (Nzuki, 2006), which are unlikely to vary with season.

6.2.2. Empirical Literature Review

Hsie et al. (1993) studied the determinants of travel choice between package and non package tour among the Australian outbound travellers. Using logistic regression, the study established that older travelers and those in large parties for touring cities and resorts, going on cruises and seeking the "being and seeing" benefit are more likely to take a package tour.

Corcoran et al. (1996) undertook a study in Ireland with the purpose of exploring the types of individual who go for a package holiday in the sun and the reason why they prefer to do so. With regard to what variables influence the choice of a package tour the study found that a hypothetical "average" tourist on package tour going to a destination in the sun was a young, middle class, urban female who was travelling in a group of three and who was visiting the sight for the first time.

Te Lang et al. (1997) using discriminant analysis studied determinants of a choice between inclusive package and non-inclusive package among Taiwanese outbound travellers. The study found that in general, travellers on inclusive package tours tended to be female, older in terms of age, and with lower incomes and educational levels. Travellers on non-inclusive package tours were somewhat younger and most of them were single. In terms of trip-related characteristics, travelers on inclusive package tour had a shorter trip length and larger party sizes than their counterparts.

TTSS (2001) studied a link between travel arrangements and purpose of visit. The study found that most visitors on business, visiting friends and for other purposes tended to travel on non-package arrangements while the majority of holiday makers

travel on the package tour arrangement. The study also postulated a relationship between tourist's country of origin and travel arrangements. Visitors from Kenya, Norway, the Netherlands, Sweden and Canada preferred non-package tours while those from Italy, France, Spain, Germany, Australia and Belgium preferred package tours. Length of stay was also implied as one of the factors. Through cross tabulations, a tourist on non-package tour had an average length of stay of 13 days whereas a tourist on package tour had an average length of stay of 10 days. Nevertheless, these findings can seldom be reliable, as they were more descriptive than analytical.

Nzuki (2006) studied the situational influence of young children on the choice of package or non-package travel of travellers to Tanzania. Alongside the presence or absence of young children, Nzuki (2006) included a number of socio economic and demographic variables. Using binary logistic regression, the study findings indicated that tourists accompanied by young children are more likely to choose package travel than those unaccompanied by young children.

Io and Hallo (2009) studied the determinants of the choice of a package tour against non-package tour of Macao outbound tourists. The study found that satisfaction with the previous mode of tour had the greatest influence in choosing the current mode of tour. Demographics had very little influence on choosing a mode of tour. One of the demographic variables, which appeared to predict well the choice of the two modes of tour, was the language fluency of the visited destination. Tourists on a package tour were found to be less fluent in the language spoken at the destination.

To sum up the review on the variables determining travel behaviour, it can be

concluded that the choice between the two modes of travel depends, among other things, on age, income, country of origin, gender, purpose of visit, length of stay , travel party size, presence/absence of children, familiarity with the destination and frequency of travelling. Among these variables, age, being a female, travel party size, presence of children and number of children are thought to be positively associated with the choice of a package tour. Familiarity with the destination and frequency of visits are thought to negatively affect tourist choice of a package tour, while season of travel, source of information and distance could be indeterminate. All these variables can accordingly be grouped into socio-demographic and trip related characteristics and destination attributes. It is also worth noting that none of the mentioned studies discusses the possibility of regressor' endogeneity. This issue is dealt with in this study.

6.3. Methodology

6.3.1 Model Specification

The package or non-package tour is a non-continuous commodity, makes it undesirable to be modeled under the conventional utility functions that require this condition as a prerequisite. The general approach in such circumstances has been to consider an index function formulation (Cameron and Trivedi2005, Scott long 1997, and Baltagi, 2005).

According to Cameron and Trived (2005), in index function formulation the interest lies in explaining a continuous random variable y^*, which is unobservable as in this case. All that is observed is a binary response y taking the value 1 or 0 depending on whether the continuous variable y^* cross a specified threshold. For example, in this

case one observes $y=1$ if a tourist chooses a package tour. This is to say there is an unobserved continuous utility level which must exceed a certain minimum for a tourist to choose a package tour. This can be expressed mathematically as follows:

First one considers a latent variable $y^* = x\beta + \mu$ whose realization is such that:

$$y = \begin{cases} 1 \ \text{if} \ y^* > 0, \\ 0 \ \text{if} \ y^* \leq 0, \end{cases} \qquad 6.1$$

In the context of probability 6.1 can be reconsidered in the following formulation:

$$\Pr[y = 1/x] = pr[y^* > 0] = \Pr[x'\beta + \mu > 0] = \Pr[\mu > -x'\beta] = F(x'\beta) \qquad 6.2$$

[27]The last expression assumes that the distribution of μ is symmetrical around the mean and F is its cumulative density function (Baltagi, 2005; Cameron and Trived, 2005).An example of such a distribution is the standard normal distribution denoted by $\phi(\mu)$ whose cdf is denoted by Φ (μ). Another is the logistic distribution given as

$\lambda(\mu) = \dfrac{\exp(\mu)}{[1+\exp(\mu)]^2}$. And its cumulative density function given as $\Lambda(\mu) = \dfrac{\exp(\mu)}{1+\exp(\mu)}$

If one assumes the logistic distribution of the error term and carries out further manipulations, equation (6.2) can be written as,

$$P = F(x'\beta) = \Lambda(x'\beta) = \frac{\exp(x'\beta)}{1+\exp(x'\beta)} \Rightarrow \log\left[\frac{p}{1-p}\right] = x'\beta \qquad 6.3a$$

The LHS of the last equation is called the logit, which is the logarithm of the odd ratio of y being equal to1 as opposed to being equal to 0. The modelling of this equation guarantees the probability that p to lie between 0 and 1.

Similarly, if one assumes the probit distribution of the error term and carries out further manipulations equation (6.2) can be written as

$$P = F(x'\beta) = \Phi(x'\beta) = \int_{-\infty}^{x'\beta} \phi(z)dz \Rightarrow \Phi^{-1}(p) = x'\beta \qquad 6.3b$$

where $\phi(z)$ is the pdf of the standard normal distribution with mean 0 and variance 1. The last equation in 6.3b gives the inverse relationship between the covariates and the probability p. As in the logistic case the modelling of this equation guarantees the probability that p will lie between 0 and 1.

The estimation of covariates in a binary response variable is done by the method of maximum likelihood by maximizing the log likelihood of the binary density function (Bernoulli distribution)

$$f(y_i / x_i) = p_i^{y_i} \quad (1-p)^{1-y_i} \qquad 6.3c$$

Where

$$p = F(x'\beta)$$

It follows that (6.3a) or (6.3b) can be written in a more detailed manner, while taking into consideration the variables discussed in sections 2 and 3. For no apparent reason one may simply consider the logistic and write it in detail as follows:

$$\log\{p/(1-p)\} = \beta_o + \beta_1 Age + \beta_2 Females + \beta_3 Income_i + \beta_4 Er + \beta_5 El + \beta_6 childno + \beta_7 Childpresence + \beta_8 Lstay + \beta_9 VistOth + \beta_{10} VistFRD + \beta_{11} VistLSR + \beta_{12} VistBSN + \beta_{13} adultno + \beta_{14} Tpartyno + \beta_{15} frvist + \beta_{16} fadest + \beta_{17} Dist + \beta_{18} Season + \beta_{19} Isource + Africa + Asia + Europe + MEast + NAmerica + SAmerica + year1 + year2 + year3 + \varepsilon \qquad (6.3d)$$

where

y_{ict} = expenditure per day by an i^{th} person belonging to country c observed in the year t

Age= 1 if the respondent's age group is <18, 2 if age group is 18-35, 3 if age group is 6-55 and 4 if age group is over 55, .its coefficients β_1 is expected to be positive.

Females=number of females in the travel party. Its coefficient β_1 is expected to be positive.

Income =tourist's income level. In this study per capita GDP from a tourist's country of origin was used as a proxy for income. Its coefficient β_3 is expected to be positive.

Er=Exchange rate between Tanzania and the tourist country of origin expressed in terms of units of Tanzanian shillings per unit of a foreign currency. This variable like GDP, is designed to assess the impact of income on tourist purchase of a package tour. Its coefficients β_4 is expected to be positive

El=1 if a tourist is from the English-speaking country and 0 otherwise. Its coefficient β_5 is expected to be negative, implying that tourists who are fluent in English would prefer non-package tours as observed in the literature review.

Pcapita = per capita tourist expenditure in Tanzania. This variable like GDP and exchange rate, is a proxy for a tourist income .Therefore its coefficient β_6 is expected to be positive, implying that wealthy tourists prefer to use package tours which are relatively more expensive.

Childno=number of children in the travel group. Its β_7 coefficient is expected to be positive, implying that tourists with a large number of children will find it more comfortable using a package tour.

Child presence=1 if there is at least one child in the travel party and 0 otherwise. The

inclusion of this variable is justified by the fact that, while the intensity as captured by the number of children may not be significant, the mere presence or absence of a child may matter. Therefore its coefficient $\beta_{8\,is}$ expected to have the same sign as the number children.

Lstay= tourist length of stay in Tanzania measured in days. Its coefficient β_9 is expected to be negative, implying that tourists staying longer prefer non-package tours.

VistOth= 1 if a tourist is visiting for other purposes, which excludes friends, leisure and business. Its coefficient β_{10} is indeterminate. The literature is not explicit on this variable.

VistFRD = 1 if a tourist is visiting friends and relatives. Its coefficient β_{11} is indeterminate.

VistLSR= 1 if a tourist is coming for leisure and recreation. Its coefficient β_{12} is indeterminate.

VistBSN=1 if a tourist is coming for business purposes. Its coefficient β_{13} is indeterminate

Adultno= Number of adults in the travel group. Its coefficient β_{14} is expected to be negative. Although in the literature review no account of this variable was given, it can be viewed as the opposite of child-presence/number of children. Hence its coefficient is expected to be negative.

Trpartyno=1 if there are at least 2 members in the travel party, 0 otherwise. Its coefficient β_{15} is expected to be positive as recounted in the literature review, the main reason being that large travel parties are more likely to be considered

for a price discount owing to the economies of scale, than a single traveller . Additionally they need a convenient way of maintaining their togetherness.

Frvist = 1 if a tourist has visited at least one African country before Tanzania, 0 otherwise. There is no account of this variable in the literature. But a frequent visitor in this context is person conversant with international travel and may be visiting other countries as well asTanzania. This means the chances are that he will choose a non-package tour due its flexibility. Therefore its coefficient β_{16} is expected to be negative.

Fadests = 1 if a tourist has visited Tanzania at least once before the current visit and 0 otherwise. Its coefficient β_{17} is expected to be negative, implying that such a tourist would prefer a non-package tour.

Dist= the shortest distance by air between Tanzania and the tourist country of origin measured in miles. Its coefficient β_{18} is indeterminate.

Peak=1 if a tourist traveled during the peak season (July- September) and 0 otherwise. Its coefficient β_{19} is expected to be positive if most of the tourists travel on package tour. Therefore we shall regard it as uncertain.

Isource= 1 is source of information is from a word-of-mouth, 0 otherwise. Its coefficient is indeterminate, depending on how the returning tourists depict the destination attributes in terms of security, infrastructure, hospitality and the quality of services. If such attributes are genuinely portrayed, then tourists would opt for any mode of travel depending on whether the destination attributes are positive or negative.

Africa, Asia, Europe, MEast, NAmerica, SAmerica, represent regions' dummies for Africa, Asia, Europe, the Middle East, North America and South America.

year1, year2 and year3 are dummies for year 2001, year 2007 and year 2008.

The discussion is summarized in Table 6.1.

Table 6.1: A priori direction of the relationship between the choice of package tour and the explanatory variables

Variable	acronym	Direction
Age group of the respondent	Age	+
Number of females in the travel party	Females	+
Income of the travel party	income	+
Tanzania nominal exchange rate against that of a tourist country	ER	+
Tourist familiarity with the English language	EL	-
Number of children in the travel party	Childno	+
Presence of a child in the travel party	Childpresence	+
Tourist length of stay(days) in Tanzania	Lstay	-
Other purposes of visit	VistOth	indeterminate
Visiting friends	VistFRD	indeterminate
Visiting for leisure and recreation	VistLSR	indeterminate
Visiting for business purposes	VistBSN	indeterminate
Number of adults in the travel party	Adultno	+
Travel party size(number of people in the travel group)	Trpartyno	+
Tourist frequency of travelling (not to Tanzania only)	Frvist	-
Tourist familiarly with Tanzania(frequency of visiting Tanzania)	Fadests	-
Distance between Tanzania(DSM) and the tourist country's capital	Dist	indeterminate
The peak season of travel by a tourist	Peak	indeterminate
Information source from the word of mouth	Isource	indeterminate

6.3.2 Addressing the Consistency of the Model Estimates

The consistency of the model estimates requires the correct specification of the model. i.e. $E(y/x)=F(x'\beta)$ in this case(Cameron and Trived,2005). Although this is a necessary condition it is not a sufficient one. The sufficient condition for consistency entails the satisfaction of $E(x'\mu) = 0$ when the former condition is met.

When the variable length of stay was instrumented and assessed using Hausman tests it was proved to be endogenous, as suggested in the discussion. Two instruments were

used, child presence and source of tourist information (see appendix 6.1).In assessing the endogeneity of length of stay, an OLS model was used even though the variable travel arrangement is a binary response variable. One of the key problems with the linear probability model is the lack of constant variance (Gujarat,2003), an assumption that is crucial in the Hausman tests for endogeneity(Cameron and Trivedi,2005).In order to overcome this problem all the OLS models used for assessing endogeneity including that of length of stay considered the robust variance of the error term. More details can be found in appendix 6.1.

Some variables were proved to be exogenous. These were tourist visiting friends and relatives (VistFRD) and tourists visiting for business purposes (VistBSN). Equally well these variables are a binary response and therefore a robust consideration of the variance was important when estimating using OLS (see appendix 6.1)

By the virtue of the variables proved to be exogenous, it was possible to test the null hypothesis of destination attributes being more influential than the demographic and trip-related characteristics, the reason being that the variables VistFRD and VistBSN, which are both trip-related characteristics, were far more influential than the variable Isource (a destination attribute).

An examination of the variance inflation factors (VIF) using an OLS model of a tourist choice of package tour indicated that each variable had a very small VIF (<5)(see Table 6.2 in appendix 6.2). Apart from ensuring that the coefficients are immune from multcollinearlity, this insipres confidence that even if there is an endogenous regressor among those unproved to be exogenous, its endogeneity may not affect the exogeneity

of other regresors and hence of the entire model (Cameron and Trivedi,2005).

Despite what has been argued regarding the VIF, to be on the safe side the study does not include the variable length of stay in the model. Alternatively instruments could have been used instead, but the primary objective was simply to show that length of stay is an endogenous variable in a model pertaining to travel arrangements and not to estimate it. Besides, the null hypothesis can still be tested even with the removal of the variable length of stay by using other trip-related characteristics. The instruments were not used, essentially to avoid other risks associated with instrumental variable techniques such as model inefficiency (Cameron and Trivedi, 2005).

6.3.3 Variables and their Sources

The study used survey data from TTSS. Although there were six years of survey by the TTSS (2001, 2004, 2005, 2006, 2007, 2008), the study mainly used data for the years 2001, 2007 and 2008, which had similar and relatively more explanatory variables than the rest of the years. These three years made a total of 30,782 observations. But only 25,880 observations out of 30,782 were used, the rest being excluded due to missing and/or unacceptable values.

The following variables were obtained directly from the TTSS surveys: age, number of females in the travel party (Females), travel party number (Tpartyno), length of stay by a tourist (Lstay), travel arrangements i.e. package or non package tour (Tarra), tourists visiting friends and relatives (VistFRD), tourists visiting for Leisure and Recreation (VistLSR), tourists visiting for Business purposes (VistBSN) and number of sites visited by a tourist in Tanzania (Nosites).

The following variables were derived from the TTSS survey:

(1) (Expenditure per person per day). The survey data reported on total party expenditure. In order to obtain expenditure per person per night, the total party expenditure was divided by the size of the group and by the number of days the party stayed. There were two types of expenditure, depending on whether a tourist was on a package tour or a non package tour. For tourists on a non-package tour no adjustment was made in the data, but for But the data was adjusted for tourists on package tour, whose bills are paid by travel agents in their home countries. According to TTSS (2001), the actual expenditure accruing to Tanzania can be found by deducting 10% of the package cost as the amount paid to the agents. The same was done in this study. According to TTSS (2001), after deducting the 10% commission, the international travel fare, was also deducted for tourists whose packages include that.

Another consideration was to set the minimum expenditure and the maximum expenditure as proposed by TTSS (2001). According to TTSS (2001), meaningful expenditure was regarded as not being less than 10 dollars per person per night, whereas the maximum expenditure per tourist was set at USD1000. The same was adopted by this study. This was important for two reasons. First is the need to compare the findings from this study with those of TTSS. Second is to minimize extreme values in the data.

(2) Frequent vistor (Frvist). A frequent visitor was taken to be a person who had visited at least one other African country before visiting Tanzania. There was a question asking tourists to compare the cost of visiting Tanzania against that of neighbouring destinations such as South Africa and Kenya. A tourist responding to

this question was regarded as a frequent visitor. This information was used because there was no question asking a tourist directly whether he/she is a frequent visitor.

(3) Familiarity with the destination (Fadest). A tourist who was not on his first visit to Tanzania was regarded as being familiar with the country. Unfortunately this variable does not capture the intensity of familiarity.

The following variables were taken from other sources:

(1) Income was proxied by a tourist country's GDP. The figures were obtained from the IMF (2009).

(2) Exchange rate (Er) was taken from the Italian Bank, a source also used by the Economist website.

(3) English language proficiency (El). A tourist coming from a country where the official language is English was regarded as an English speaking person. The country's official language was obtained from the internet (www.yahoo.com) by searching the country's profile.

(4) Distance travelled (Dist). This is the shortest ground distance between Tanzania and a tourist's country of origin, measured from Dar es Salaam to the country's capital. The variable was sourced from the internet using the online distance calculator.

(5) Relative cost of living between a tourist country of origin and Tanzania (Price). Figures in the Consumer Price Indexes were obtained from the IMF (2009).

6.4 Results and Discussion

6.4.1 Summary Statistics of the Variables used in the Analysis

Table 6.3 gives the summary statistics of the variables used in the study.

Table 6.3: Summary statistics of the variables used in the analysis of package tour

Variable	Mean	Std. Dev.	Min	Max	CV
Tarra	0.5289	0.4992	0	1	94.38
Age(Ranks)	3	0.69	1	4	23
Females	0.8940	0.8992	0	21	100.58
GDP(Mil.USD)	29123.3	15379.39	54.62	113044	52.81
Er (TSHS)	965.49	603.92	0.0301	3229.158	62.55
El	0.5389	0.4985	0	1	92.50
Childno	0.1353	0.5794	0	13	428.23
Childprese~e	0.0694	0.2542	0	1	366.09
Lstay(Days)	12.69	13.60	1	360	107.17
VistFRD	0.0765	0.2659	0	1	347.58
VistLSR	0.7722	0.4194	0	1	54.31
VistBSN	0.1016	0.3021	0	1	297.34
Adultno	1.75	1.28	1	31	73.14
Tpartyno2	0.5570	0.4967	0	1	89.18
Frvists	0.5285	0.4992	0	1	94.46
Fadest	0.3726	0.4835	0	1	129.76
Dist(Miles)	5287.09	2389.07	419	9527	45.19
Peak	0.7387	0.4393	0	1	59.47
Price	0.0525	0.5792	3.E-05	26.00667	1103.24
Isource	0.3846	0.4865	0	1	126.50
year1	0.6184	0.4858	0	1	78.56
year2	0.1001	0.3002	0	1	299.90
year3	0.2815	0.45	0	1	159.86

Table 6.2 indicates that most of the variables have a reasonable variation except age.

6.4.2 The Correlation Matrix

The correlation matrix of the variables used in the analysis is given in Table 4.4 of appendix 4.3. The matrix indicates that most of the variables correlate with the

variable package tour in the expected manner as discussed in the literature review as well as in section 6. 3.1. In addition, the regressors have a very low correlation among themselves, suggesting the absence of multcollinearlity (in fact VIF confirms this –see appendix 6.2). More details on the correlation can be found in appendix 4.4.

6.4.3 The Logit and Probit Regression Results

Table 6.4 provides both the logistic and probit regression results. Practically the choice of either model is matter of convenience (Scott long, 1997; Cameron and Trived, 2005).However, Cameron and Trived (2005) suggest that the log-likelihoods can be used to choose the best model even though in most cases the log-likelihoods for the two models are very close. Based on this argument the logit model appears to be more appropriate as it has a slightly larger value of the log-likelihoods (-13660.34) than the probit model (-13674.17). In essence this criteria is the same as the [28]McFadden pseudo R^2,which suggests that the logistic model is the best (Table 6.4).This is also suggested by the AIC (Table 6.4). Another comparative test is the count R^2, which indicates the percentage of correct predictions. In both models the count R^2 indicates that 73% of the observations have been correctly predicted. The count R^2, apart from indicating that any of the models can be adequate, inspires confidence in the predictive power of the two models.

---

[28] The McFadden $R^2 = 1 - \ln \hat{L}(M_\beta) / \ln \hat{L}(M_\alpha)$ where $\hat{L}(M_\beta)$ is the likelihood of the estimated model and $\hat{L}(M_\alpha)$ is the likelihood of the null model(Scott long, 1997)

Table 6.4: Logit and Probit regression results of tourist choice of package tour

	Logit regression			Probit regression		
variable	dy/dx	z	P>|z|	dy/dx	z	P>|z|
Age	.10***	16.72	0.000	.09***	16.72	0.000
Females	.03***	4.54	0.000	.03***	4.38	0.000
GDP	-4.2e-06***	-9.36	0.000	-3.8e-06***	-9.34	0.000
Er	.00009***	10.85	0.000	.00008***	10.26	0.000
El[1]	-.10***	-11.73	0.000	-.09****	-11.48	0.000
Childno[1]	-.04***	-5.89	0.000	-.04***	-5.76	0.000
VistFRD[1]	-.32***	-16.71	0.000	-.29***	-14.80	0.000
VistLSR[1]	.37***	27.21	0.000	.37***	27.23	0.000
VistBSN[1]	-.25***	-11.46	0.000	-.21***	-10.08	0.000
Adultno	-.02***	-5.11	0.000	-.02***	-4.89	0.000
Tpartyno[1]	.07***	7.49	0.000	.07***	7.62	0.000
Frvists[1]	-.02**	-2.06	0.032	-.01*	-1.89	0.058
Fadest[1]	-.12***	-14.09	0.000	-.11***	-13.87	0.000
Dist	.00003***	11.33	0.000	2e-05***	11.08	0.000
Nosites[1]	.12***	15.86	0.000	.12***	16.01	0.000
Peak[1]	.02***	2.55	0.000	.02***	2.62	0.001
$Isource_1$	.18***	22.79	0.000	.16***	22.42	0.000
year1[1]	-.23***	-15.41	0.000	-.22***	-15.51	0.000
year3[1]	-.10***	-7.05	0.000	-.10***	-6.91	0.000
likelihood	-13660.17			-13674.17		
χ	8469.99			8442.44		
p>Chi	0.000			0.000		
Pseudo R^2	0.2367			0.2359		
AIC	27360.78			27388.33		
Count R^2	0.73			0.74		
n	25880			25880		

NB: *significant at 10%, **significant at 5%, ***significant 1%
(1) dy/dx is for discrete changes of dummy variable from 0 to 1

Table 6.4 shows the marginal effects[29] of the regressors from the two models. In both models, the most influential variables are the trip-related characteristics as measured by tourist purpose of visit. The destination attributes as measured by the source of

[29] It can be shown that the marginal effects are given as $dp/dx = \hat{p}(1-\hat{p})\beta$ where β measures the change in the logit per unit change in the covariate and $\hat{p}$ is the predicted probability at the average values of the independent variables. In this particular case p=0.50 for the logistic and p=0.50 for the probit. In other words, the relative sizes of the magnitude of the marginal coefficients entirely depend on the relative sizes of the β's. That is to say for comparing the relative influence of the covariates, the β's can equally well be used.

information rank second in terms of influence which leads to the rejection of the null hypothesis that destination attributes have the greatest influence on a tourist's choice of package tour.

Following the fact that both models lead to the rejection of the null hypothesis, the question of which model to choose appears to be of secondary importance. However, for convenience purposes and due to the fact that in a probit model one needs to ascertain the results from the problem of neglected heterogeneity, this study adopts the logit model and the forthcoming discussions is based entirely on the logit model.

The results indicate that the coefficient of the source of information is positive, which implies that a returning tourist has a positive opinion of package travel to Tanzania. One of the reasons why tourists opt for package travel is to avoid unnecessary risks. Given the peace and security of Tanzania, they should not worry about security but rather about the poor infrastructure and disorganized services, which when not booked in advance could lead to delays and other disruptions. This view seems to be supported by Io and Hallo (2009), who investigated the determinants of package tours for outbound tourists in Macao. Their findings indicated that past experience of travel arrangements counts a lot in the current choice of a package or non-package tour.

The coefficient of the variable information source (0.18) is clearly not the largest compared to the coefficients of some trip-related characteristics such as tourists coming for leisure (0.37), tourist visiting friends (-0.32) and tourist on business purpose (-0.25). This scenario leads to the rejection of the null hypothesis of destination attributes being more influential in the choice or non-choice of package

tour compared to other attributes. The results further shows that tourist coming for leisure and recreation prefer to use package tours whereas those visiting friends and on business purpose prefer to use non package tours. The observation is in line with what was established by TTSS (2001).

The next most influential trip-related characteristic after the purpose of a tourist visit is the familiarity with the destination (-0.12).Tourists who are familiar with Tanzania tend to opt for non-package tours. This observation is in line with most studies, such as by Corcoran et al.(1996).

Number of adults came to have a negative influence (-0.02) on the choice of package tour while travel party number had a positive influence (0.07). These results should be expected, as discussed in section 6.3.1. More adults in a travel party means that, they have the ability to withstand the problems associated with a non-package tour compared more children, other things equal. As regards travel party number, tourists in big numbers find it cheaper to travel on a package tour because of economies of the scale. In most cases their prices are discounted, unlike for a single traveler. The positive influence of travel party number has been accounted by several studies, such as by Hisie et al. (1993), Corcoran et al. (1996) and Te Lang et al. (1997).

Other trip-related characteristics which had a significant influence were number of children (-0.04) and distance (0.00002). The observation on the number of children is quite surprising and contrary to expectations. It is assumed that a tourist accompanied by more children would be more likely to opt for a package tour than non-package travel to avoid unnecessary problems. Indeed some studies, such as Nzuki (2006),

have found this to be the case. When the presence of children was considered the influence was still negative (-0.06). One of the possible reasons for this unexpected result could be due to the poor distribution of this variable, as the majority (93%) of travellers were not accompanied by children. The other reason could be that the age of a child may also matter than the mere number. In fact Nzuki (2006) considered age also. But either way, this result warrants further investigation.

As regards distance, there is no obvious reason for its positive influence on package tour. Probably the further the distance from the destination the greater the worry of a traveler regarding the risks. Another way of assessing the influence of distance is to think in terms of how it affects tourist length of stay. Chapter five showed that distance of travel reduces a tourist's length of stay and given that length of stay negatively influences package tour then the overall impact of distance should positively influence a tourist's choice of a package tour.

Even though excluded from the model owing to its endogeneity, length of stay appears to have a significant and negative influence on choice of package tour when included as an associated factor. In fact even when instrumented it yielded a significant and negative influence. Its negative influence is consistent with the discussion in section 6.3.1.The longer the length of stay the greater the motivation to opt for a non-package tour owing to its flexibility.

The study included six demographic characteristics: age, gender, GDP, exchange rate, language and tourist per capita expenditure. Among these all had significant positive influence except language and GDP, which had significant but negative influences.

The observation on language is similar to that by Io and Man (2009) who found that outbound tourists of Macao on a package tour were unfamiliar with the language spoken at the destination of their trip. In other words tourists who are familiar with the language spoken at a destination are more likely to opt for a non-package tour, as found by this study. The result on GDP is unexpected in the sense that wealthy tourists should travel using a package tour. But probably the variable GDP is very crude in the sense that it does not account for an individual's wealth but rather the country's prosperity. The tourist per capita expenditure can be a better proxy for his income. Unfortunately an inclusion of this variable in the model would lead to model inconsistency much as tourist spending partly depends on his mode of travel tour (chapter four). Nevertheless when it was included as an associate factor it had a significant and positive influence on his choice of a package tour (0.003). Exchange rate also had a significant and positive influence reaffirms what was found on tourit per capita Another variable which also measured the same thing is the tourist's country exchange rate, which had a significant and positive influence on the choice of a package tour (0.0009).

The remaining demographic variables (age and gender) are very popular in the literature and their influence has been significant and positive on the choice of a package tour as recounted in most studies such as by Hisie et al. (1993) and Corcoran et al. (1996).

Given what has been discussed regarding the three types of variables (demographic, trip-related and destination attributes), it remains in the hands of a particular stakeholder to utilize their influence for his/her own good. Whether to encourage

package tour or non-package tours is not an easy decision, because each type of tour has its own advantages and disadvantages. As a matter of fact finding out which group is more appropriate for the country could be another area of research interest.

While the argument for the package tour is it brings in more revenue (TTSS 2001, Table 4.5), it is yet unknown how much of the revenue goes out of the country through leakage when compared to revenue from tourists on non-package tours. Likewise, whereas the local population is more likely to benefit from tourist on non-package tours, owing to their direct spending on local goods, their relatively longer stay could be detrimental to the environmental.

Whereas what has been raised is of more relevance to the government as one of the key stake holders in the sector, for private stakeholders, and in particular tour operators, the findings of this study are of ultimate importance. Neither leakages nor environmental pollution is of concern to them. What they need to know are markets are associated with the factors that positively influence tourists to opt for package tours, such as age, gender, travel party number, leisure visits and distance. Some markets such as Europe and North America are obvious candidates for the tour operators in Tanzania. as their dummies had a positive influence on package travel (not indicated in Table 6.4).

Finally, one may like to know whether or not the results apply generally to all the regions of the world. This could be important for specific regional policy formulation. Table 6.5 provides logist regression results across the regions, along with the general results of Table 6.4. For ease of presentation the standard errors and/or z-values are

omitted from the Table. The primary aim is to compare the marginal effects across regions. However, the significance level (p-values) for each regresssor is still provided in all the regions.

Table 6.5: Logit regression of a tourist choice of package tour: Comparison across regions.

variable	World	Africa	Asia	Europe	MEast	NAmer	SAmeri
Age	.10***	.03*	.04*	.08***	.16**	.15***	.15*
Females	.03***	.02	-.01	.02***	.03	.03*	-.02
GDP	-4.2e-06***	-9.e-06	-3.4e-06*	-4.2e-06***	7.7e-06	-.0003	.00002
Er	.00009***	.0006**	.0003***	.0001***	-.00002	.013	.0004
El[1]	-.10***	.05*	.04	-.12***	Dropped	Dropped	.11
Childno[1]	-.04***	-.013	.002	-.03***	-.03	-.04**	.03
VistFRD[1]	-.32***	-.14***	-.19**	-.24***	-.14	-.42***	
VistLSR[1]	.37***	.19***	.38***	.45***	.09	.27***	0.07
VistBSN[1]	-.25***	-.07**	-.25***	-.24***	-.28***	-.33***	-0.19
Adultno	-.02***	-.001	.01	-.03***	.02	-0.01	0.04
Tpartyno[1]	.07***	.07***	-.02	.05***	.12	0.05***	0.03
Frvists[1]	-.02**	.009	.12***	-.012	-.03	-0.03	0.10
Fadest[1]	-.12***	-.05***	-.10***	-.13***	-.15**	-.07***	0.01
Dist	.00003***	.0003**	.0002	-.0001***	.0006	-.0003	0.00004
Nosites[1]	.12***	0.06***	.24***	.07***	.14*	.32***	0.35***
Peak[1]	.02***	.03	-.04	.03**	0.19***	.01	0.08
Isource[1]	.18***	.04**	.10***	.20***	0.33***	.10***	0.14
likelihood	-13660.17	-978.36	-986.82	-8333.47	-99.98	-2731.81	-88.21
χ	8469.99	756.11	835.98	4690.44	96.12	2024.21	67.95
p>Chi	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pseudo R^2	0.2367	0.28	0.30	0.26	0.32	0.27	0.28
AIC	27360.78	1996.72	2013.65	15890.8	235.96	5501.62	214.44
Count R^2	0.73	0.82	0.76	0.72	0.78	0.75	0.77
n	25880	2379	2028	15560	226	5485	185

NB: *significant at 10%, **significant at 5%, ***significant 1%
(1) dy/dx is for discrete changes of dummy variable from 0 to 1

Table 6.5 indicates that almost all the facts established before apply across the regions. The trip-related characteristics as measured by tourist purpose of visit are still the most influential across the regions. This leads again to the rejection of the null hypothesis that destination attributes are more influential than the demographic and trip-related

characteristics in all the regions .Other results indicate that frequent visitors from Asia prefer package tours contrary to the general results.

6.5 Conclusion

The chapter has addressed the determinants of tourist choice of a package tour. Among the variables considered, the trip-related characteristics as measured by the purpose of visit have been the most influential, leading to the rejection of the null hypothesis that destination attributes are the most influential.

Unlike in most of the previous work this study has shown that tourist length of stay is an endogenous variable in a model for the choice of a package tour. Though not estimated in this study, the study has highlighted the need to consider its endogeneity for studies intending to capture its real influence.With regard to policy implications, tourist on package tours can be encouraged by considering markets associated with older age, big travel parties, and tourists coming for leisure and recreation. These are none other than the traditional markets of Europe and America. However this encouragement would be based on the assumption that tourist on package tours are more beneficial to the country, which is yet to be established.

CHAPTER SEVEN

OVERALL CONCLUSION, POLICY IMPLICATIONS AND FURTHER WORKS

6.1 Overall Conclusion and Policy Implications

The study has analyzed the determinants of inbound tourism demand in Tanzanian, both at the macro and micro level. At the macro level the study examined the determinants of tourist arrivals in Tanzania, whereas at the micro-level the study examined the determinants of tourist per capita spending, length of stay and choice of package travel.

With regard to number of arrivals, it has been found that non-price factors, such as the country's economic growth and infrastructure development matter more than other factors. Other factors as well such as tourism prices in Kenya, the Seychelles and Mauritius also counts a lot in determining the number of arrivals.

As regards tourist per capita expenditure, tourist length of stay, and choice of package tour, trip-related characteristics, particularly travel party number, tourist purpose of visit (business visit, leisure and recreation) and frequency of visits are the more influential determinants than demographic and destination attributes.

The study therefore implies that the government and the private sector should invest more in non-price factors such as infrastructure, marketing and the general improvement of the economy so as to attract more tourists to Tanzania. As regards tourist spending, length of stay and choice of a package tour the government as well as

private stakeholders should condition their marketing strategies on the trip-related characteristics and demographic attributes of the tourists more than on other factors.

6.2 Further Works

As said before few studies on tourism have been done in Tanzania. This implies that quite a lot of further works can be done. The following are a few of such possible studies.

One would be to establish the amount of water and air pollution in the national parks and other tourist sites. Such research should make it possible to establish the linkage between the actual amount of pollution and the coming of tourists to such areas. It is suspected that the coming of tourists leads to environmental destruction without any quantification. The environmental scientists need to establish the amount of such damage per tourist.

Second area of research would be to establish the amount of leakage in tourism revenue. This would be done both at the local and national level, while paying attention to leakage by category of tourists. The starting point in such a work would be to visit the tourism related businesses, especially, hotels, national parks, tour operators and guides, airports and seaports, and the immigration department. In each of these places, there would be a need to establish the amount of foreign currency spent on importing goods such as vehicles, aero planes, foods and drinks against the revenue generated by tourism.

The third area of research interest would be to establish the economic value of tourism

assets at national level, with a view to seeing whether or not tourists are being undercharged. Such researches have been done at particular areas of the country, but not in the entire country.

Fourth would be to establish the determinants of domestic tourism, both at macro and micro level. As previously said, heavy reliance on inbound tourism puts the country at risk of possible diplomatic conflicts with key source markets. Therefore domestic tourism could serve to maintain the industry if such conflicts were to happen.

Fifth is would be to establish the determinants of outbound tourism to see whether it should be promoted or discouraged. Apparently a lot of Tanzanians go abroad for various purposes. The key interest here would be to know how much revenue the country loses owing to their going and how much does it gains owing to their staying there (remittances).

Finally one could look into the theoretical problems associated with inbound tourism, such as inflation, health status of the areas frequently visited by tourists, and the effect of tourism on the supply of merchandize exports.

REFERENCES

Agarwal, V.B and Yochum,G.R (1999), "Tourist Spending and Race of Vistors", *Journal of Travel Research*, Vol 38 , page173

Algieri, B. and Kanellopoulou, S (2008), "Determinants of demand for exports of tourism, an unobserved component mode", *Tourism and Hospitality Research advance online publication. Posted on 24 November 2008*

Alegre, J and Pou, L (2006) "The length of stay in the demand for tourism", *Tourism Management* , Vol 27

Andrew, F (2008), "Envoy urges African countries to promote domestic tourism" *The Guardian, 03/07/2008,*

Anderson, J.E (1979), "A Theoretical Foundation for the Gravity equation" *The American Economic Review Vol69.No.1*

Anderson, W. (2010). "Determinants of All-Inclusive Travel Expenditure" *Tourism Review, 65(3): 4 – 15*

Anderson, W. (2011). "Enclave Tourism and its Socio-Economic Impact in emerging Destinations" *Anatolia-An International Journal of Tourism and Hospitality Research, 22(3), 361-377*

Anderson, W. (2011). "Leakages in the Tourism Systems: Case of Zanzibar" *Tourism Development Journal, 9(1): 3-13*

Athanasopoulos, G and Hyndiman, R (2006), "Modeling and Forecasting Australian domestic tourism", *Working paper No.19/06, Department of Economics, Monash University.*

Bashagi and Muchapondwa (2009), "What actions could boost international tourism demand for Tanzania" *School of Economics, University of Cape Town.Working*

paper number 1952

Baltagi, B.H (2005), **Econometric Analysis of Panel data, 3rd edition**, West Sussex, John Wiley& Sons Ltd,

Barros, C and Correia, A (2007), "Survival Analysis in Tourism Demand: The length of stay in Latin American destinations", *working paper series No.2007/03 from the Department of Economics of the School of Economics and Management (ISEG),technical University of Lisbon.*

Barros, C., Butler, R and Correia, A (2010), "The length of stay of golf tourism: A survival analysis', *Tourism management*

Belleau, B (2007), "Theory of Reasoned Action: purchase Intention of Young Consumers", *Clothing and textiles Research Journal*, vol.25, No.3, 244-257

Bergstrand, J.H (1985), "The Gravity equation in International Trade. Some Microeconomic Foundations and Empirical evidence" *The Review of Economics and Statistics*, Vol. 67: 474-81..

Bergstrand, J.H (1989), "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade", *The review of Economics and Statistics,Vol.71,No.1 pages 143-153*

Blake, A (2008), "Tourism and Income Distribution in East Africa", International Journal of Tourism Research, Vol10, 511-524, published online on 10th November 20008 in Wiley InterScience (www.interscience.wiley.com)

Cameron, A.C and Trivedi, P.K. (2005), **Micro econometrics methods and Applications,** *New* York, Cambridge University press, 40 West 20 th Street, NY 10011-4211, USA

Chaiboonsri, C and Chaitip, P (2012), "The Modeling International Tourism Demand

for Length of Stay in India: Social Development, Economics Development", *Journal of Knowledge Management, Economics and Information Technology*

Chenjah, C (1998), " Promotion of Tourism in Tanzania: A critical approach" *Unpublished MBA dissertation from the University of Dar es Salaam.*

Clottey, B and Lennon, R (2003), "Transitional Economy Tourism: German Travel Consumers' Perceptions of Lithuania", *International Journal of Tourism and Research*, Vol 5, pages 295-303

Corcoran, L M; Gillmor, D.A and Killen.J.E (1996), "An analysis of summer sun tourists-outbound package holidays from Dublin airport", *Irish geography,* volume 29(2).

Coupland, R, (1968) , **Exploitation of East Africa 1856-1890 2nd edition**, London, Latimer Trend and Co Ltd Whitstble.

Durbarry, R (2001), "Tourism Expenditure in UK: Analysis of competitiveness using a gravity model". *Christel Dehaan Tourism and Research Institute Nottingham University Business School. Working paper*

Dwyer, L and Forsyth, P (2008), " Economic measure of tourism yield: What markets to Target*?* ", *International Journal of Tourism Research*, Vol.10, page 155-168

Eita, J and Jordaan, A (2007), "Estimating the tourist Potential in Namibia", *MPRA Paper No. 5788, posted 16. November 2007*

Eja, E.I; Ukwayi, J.K; Ojong, F.E (2012), "Success Factors Determining Nigeria as a Tourist Destination", *Journal of Emerging Trends in Educational Research and Policy Studies (JETERAPS) 3(4): 426-432*

Ekanayake, E.M; Halkides, H; Ledgerwood, J.R (2012), "Inbound International

Tourism to the United States: A Panel Data Analysis" *International Journal of Management and Marketing Research, Vol. 5, No. 3, pp. 15-27*

Fishebein, M and Ajzen, I (1975), **Belief, Attitude, Intention and Behavior: An introduction to theory and Research** . Reading, MA: Addison-Wesley

Fletcher, J and Morakabati, Y (2008), "Tourism Activity, Terrorism and Political Instability within the commonwealth: The cases of Fiji and Kenya", *International Journal of Tourism and Research,* Vol.10, pages 537-556

Garcia, E and Raya, J (2008), ":Length of stay for low-cost tourism", *Tourism Management volume,* Vol.29

Goldman,S.M and Uzawa, H (1964), "A Note on Separability in Demand Analysis". *Econometrica*, Vol.32, No.3

Gorman, W. N. [1959], "Separable Utility and Aggregation," Eccnomeirica, July 1959, 469-481

Govakali,U., Bahari, O and Kozak, M (2007), ":Determinants of length of stay: A practical use of Survival analysis", *Tourism Management, Vol28*

Greene, W.H (203) **"Econometric Analysis: fifth edition"** Pearson Education, Inc., Upper Saddle River, New Jersey, 07458

Gujarat, D.N (2003), "**Basic Econometrics: fourth edition**" McGraw-Hill Companies, inc.

Hanafiah, M.H and Harun, M.F (2010), "Tourism Demand in Malaysia: A cross-sectional-pool time-series analysis", *International Journal of Trade, Economics and Finance, vol.1,No1*

Hsie, S., O'Leary, J.T., Morrison, A.M and Chang, P.S. (1993), "Modeling the Travelers mode of Choice for Australian Outbound Travellers" *The Journal of Tourism studies* vol.4,No.1,M

Hore, E.C (1892), **Tanganyika: Eleven years in Central Africa**, London, Edward Stanford 26&27 Cockspur Street, Charing Cross, S.W

International monetary Fund (IMF) (2009), "Word Economic Outlook", IMF

Io, M and Hallo, L (2009), "A study of Macao outbound tourists' behavior and their choice of tour mode for pleasure travel", *South Asian Journal of Tourism and Heritage vol. 2, no. 1, pp. 13-26 0974-5432*

Jaffe, H (1988), **History of Africa, 2nd edition,** London, Zed Books Ltd, 57 Caledonian Road, N19BU,UK.

Jang,S., Bai, B., Hong, G and O'Leary, J (2004), " Understanding travel expenditure patterns: A study of Japanese pleasure travelers to the United States by income level", *Tourism management*, Vol.25, pages 331-341

Jang, S., Cai, L., Morrison, A. and O' Learry, J (2005), "The effects of travel Activities and Seasons on Expenditure" *International Journal of tourism Research* 7, 335-346.

Kulindwa, K., Sosovele , H and Mashindano, O (2001), Tourism growth for sustainable development in Tanzania, Economic research Bureau, *University of Dar es Salaam, Dar es Salaam University press.*

Kweka, J., Morrissey, O and Blake, A (2003), "The Economic Potential of Tourism in Tanzania", *Journal of International development.* Vol15, 335-351.

Kordbacheh, H; Mehregan, N; Shourekandi, A (2012), "A Study of Foreign Tourism Demand for Iran Destinations: Short-and-long run Elasticity Estimates", *DOI: 10.7763/IPEDR. 2012. V50. 9*

Lee, G and Yi, Y (2008), "The effect of Shopping Emotions and perceived Risk on Impulsive Buying: The Moderating Role of Buying Impulsiveness Trait",

Seoul Journal of Business, Vol. 14, No.2.

Leontief, W. [1947), "Introduction to the Internal Structure of Functional Relationships," Econometrica, 1947, 361-373.

Lim,C (1997), " An Econometric Classification and Review of International Tourism Demand Models", *Tourism Economics*, 3:69-81

Linneman, H (1966) "An **Econometric study of International Trade flows**" North Holland Publishing Co., Amsterdam

Litvin, S (2007), "Marketing visitor attractions: A segmentation Study" *International Journal of Tourism Research* vol.9 pages 9-19.

Maliugina, K (2006), "Inbound tourism in Ukraine" *Unpublished MA thesis, National University "Kyiv-Mohyla Academy" Economics Education and Research Consortium*

Martinez-zarzoso, i. & Nowak-lehmann, F.(2003) " augmented gravity model: an empirical application to MERCOSUR-European Union trade flows". *Journal of Applied Economics,* 6(2):291-316

Mathieson, A. & Wall, G. (1984): **Tourism economic, physical and social impacts.** Third edition, Longman Group limited, N.Y, USA

Matyas, L (1998), "The Gravity model: Some Econometric consideration", *The World Economy: 387-401*

Menezes, A.G., Moniz, A and Viera, J.C (2007), "The Determinants of length of stay of Tourists in the Azores: A count model Approach", *presented at Ponta Delgada, Portugal Observatorio Regional do Tourismo.*

Menezes ,A. G., Moniz, A and Viera, J.C (2008), " The determinants of length of stay of tourists in the Azores", Tourism Economics, 2008, Vol.14(1), pages 205-

222.

Mgani, J (1997), "Tourism: A male dominant Industry: A case study of tourism in the Municipality of Arusha Tanzania" *Unpublished M.A thesis, University of Dar es Salaam.*

Ministry of Finance and Economic Affairs (MFEA) (2008), **Review of National Economy**, Ministry of Finance and Economy Affairs

Ministry of Finance and Economic Affairs (MFEA) (2009), "**National Economic Survey**", *Ministry of Finance and Economic Affairs*

Ministry of Information and Tourism (MIT) (1972), "Tourism Ministry hits back at critics: Second year ease leader. Vol3" *Institute of Development studies, University of Dar es Salaam.*

Mjema, G (2004), "Trends in Tanzania's External Trade and Implications on the East African community", *Tanzanian Economic Trends*, A *Biannual Review of the Economy* Vol.17 No.2.

Mohamed, A.M (2011), "The Determinants of International Tourism Demand for Egypt: Panel Data Evidence" *European Journal of Economics, Finance and Administrative Sciences Issue 30 (2011)*

Morishima, M (1961), "A Historical note on Professor's Sono's Theory of Separability", *International Economic Review*. Vol2, No3, PP 272-275

MNRT (1999), "**The National Tourism Policy**", The Ministry of Natural Resources and Tourism

MNRT (2002), "**Tourism in Tanzania: Investment for Growth and Diversification**", Washington Multilateral Investment Guarantee Agency/World Bank Group.

Munoz, T (2004), "Madrid as a Tourist Destination : Analysis and Modelization of Inbound Tourism", *International Journal of Tourism Research,* Vol6, 289-302

Munoz, T and Martin. L (2006), "Tourism in the Balearic Islands: A dynamic model for international demand using panel data", *Tourism management* ,Vol28, pages 1224-1235

Muchapondwa,E and Pimhidzai, O (2008), "Modeling international tourism demand for Zimbabwe" *ERSA working paper* No.10

Muhammed, H.M (1971), "**The life of Hamad Bin Muhammedi El Murjebi, 4th Edition"**, Nairobi, Litho Limited Cardiff Road.

Naude, W and Saayman, A (2005), "The determinants of tourist arrivals in Africa: A panel data regression Analysis", *Tourism Economics , Vol. Touris, No. 3 : pp. 365-391.*

Neto, F (2002), "Sustainable Tourism, Environmental Protection and Natural Resources Management" *Paper submitted to the International Colloquium on Regional Governance and Sustainable Development in Tourism driven Economies, Cancun, Mexico, 20-22 February 2002*

Nordström, J. (2002), "Dynamic and Stochastic Structures in Tourism Demand Modelling", *Umeå Economic Studies, No 596, Umeå University, Umeå*

Nyangwine, N and Maluka, S (2008), **History paper 1Notes, Major events in African history**, Dar es Salaam, Nyambari Nyangwine Publishers.

Nyoni, T(1996), "The Determination of Export Supply and Import Demand in Tanzania" *Unpublished PhD dissertation, University of Dar es Salaam.*

Nzuki, A (2006), "A consumer situational Analysis: "The influence of young children on the purchase of Travel to overseas destinations", *Unpublished PhD*

Dissertation, Graduate School of Clemson University.

Oliver,R and Gervase,M(1963), **History of East Africa**, London, Oxford University Press, Amen House E.C.4

Ouma, J.BP.M (1969), Evolution of Tourism in East Africa (1900-1969), East Africa University Social Sciences Council Conferences, Nairobi.

Rutasitara, L (1999), "Real Exchange rate Behavior and Merchandize Exports in Tanzania" Unpublished PhD dissertation, University of Dar es Salaam.

Saayaman, A and Saayaman, M (2008). "Determinants of inbound tourism for South Africa", *Tourism Economics*, volume 14, No.1, pp 81-96.

Salvatore, D (2010), **International Economics Trade and Finance, tenth edition**, John Wiley &Sons,

Sampol, C.J and Perez, E.A (2000), "Tourist Expenditure determinants in cross-section data Model", *Annals of Tourism Research*, vol.27, Number3.

Sandbrook, C.G (2008), 'Putting leakage in its place: The significance of retained tourism revenue in the local contest in the rural Uganda *'Journal of international development*, Jon Wiley and Sons, Ltd published online at (www.interscience.wiley.com)

Sangpikul, A (2008), "Travel Motivations of Japanese Senior Travellers to Thailand" *International Journal of Tourism Research*, Vol10, pages 81-94, published online 2 November 2007 at (www.interscience.wiley.com)

Scott long (1997), **Regression models for Categorical and limited dependent Variables: Advanced Quantitative Techniques in the Social Sciences**, California, SAGE Publications, Inc. 2455 Teller Road, Thousand Oaks, California 91320

Sheppard B.H; Hartwick, J and Warshaw, P.R (1988), "The Theory of Reasoned

Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research", The Journal of Consumer Research, Vol.15, No3, pp325-343

Soest, A and Kooreman (1987), "A micro-econometric Analysis of Vacation Behavior", *Journal of Applied Econometric*, Vol.2.pages 215-226

Song, H and Li, G (2008), "Tourism demand Modelling and Forecasting". *Published at Surrey Scholarship online (http://epubs.surrey.ac.uk/tourism/31) University of Surrey.*

Strotz, R. H. [19571, "The Empirical Implications of a Utility Tree," Econometrica, April 1957, 269-280

Suh, Y.K and Gartner, W.C (2004), "preferences and trip expenditures-a conjoint analysis of visitors to Seoul, Korea", *Tourism management*, Vol25, pages 127-137

Swayne (1868), **A narrative of explorations in search of the source of the Nile**, Edinburg and London, William Blackwood and Sons

TANAPA (2002), "Annual report of the Tanzania national Parks Authority for the financial year 2001/2002", *TANAPA*

Tayyab,M; Tarar,A; Riaz, M (2012), "Review of Gravity Model Derivations", *Mathematical Theory and Modeling, Vol2, No.9, 2012*

Telang, C., O'leary, T., and Morrison (1997), "International travelers from Taiwan: Travel Arrangements, Philosophy, and Benefits" *Journal of International Hospitality, Leisure and Tourism Management* Vol. 1 (1).

Tosun, C., Temizkan, P., Timothy, J and Fyall, A (2007), "Tourist Shopping experiences and Satisfaction", *International Journal of Tourism Research.*

Vol9, 87-102.

TTB (2006), "**The Medium Term Strategic Plan 2006/7-2010/11**", TTB.

TTSS (2001), "The 2001 Tanzania Tourism Sector Survey", *TTSS*

TTSS (2004), "The 2001 Tanzania Tourism Sector Survey", *TTSS*

TTSS (2005), "The 2001 Tanzania Tourism Sector Survey", *TTSS*

TTSS (2006), "The 2001 Tanzania Tourism Sector Survey", *TTSS*

Vassiliads, Chris A (2008), "Destination product Characteristics as useful Predictors for repeat Visiting and Recommendation segmentation variables in tourism: a CHAID Exhaustive Analysis" *International Journal of Tourism Research.*10, 439-452(2008).

Vietze , C (2008), " Cultural Effects on Inbound Tourism into the USA: A gravity Approach" *Jena Economic Research Papers 2008-037*

Wang, Y., Rompf, P., Severt, D and Peerapadit, N (2006) ."Examining and Identifying the Determinants of Travel Expenditure Patterns", *International Journal of Tourism Research.* Vol8, pages 333-346.

Ward, E.F and White, L.W (1971), **East Africa: A century of Change 1870-1970**, London, George Allen & Unwin Ltd, Ruskin house Museum Street

World Bank Group (2005) "Cooperate Responsibility and the tourism sector in Cambodia". *World Bank.*

WTO (1995), "**Collection of Domestic Tourism Statistics**" *WTO, Madrid*

World Tourism Organization (WTO) (1995), "**Technical manual no.2: The collection of Tourism Expenditure Statistics**" *WTO, Madrid*

World Tourism Organization (WTO) (2009), "WTO Tourism Barometer", interim update April 2009, *WTO*

World Tourism Organization (WTO) (2009), "WTO Tourism Barometer", Vol7, No.1, *WTO*

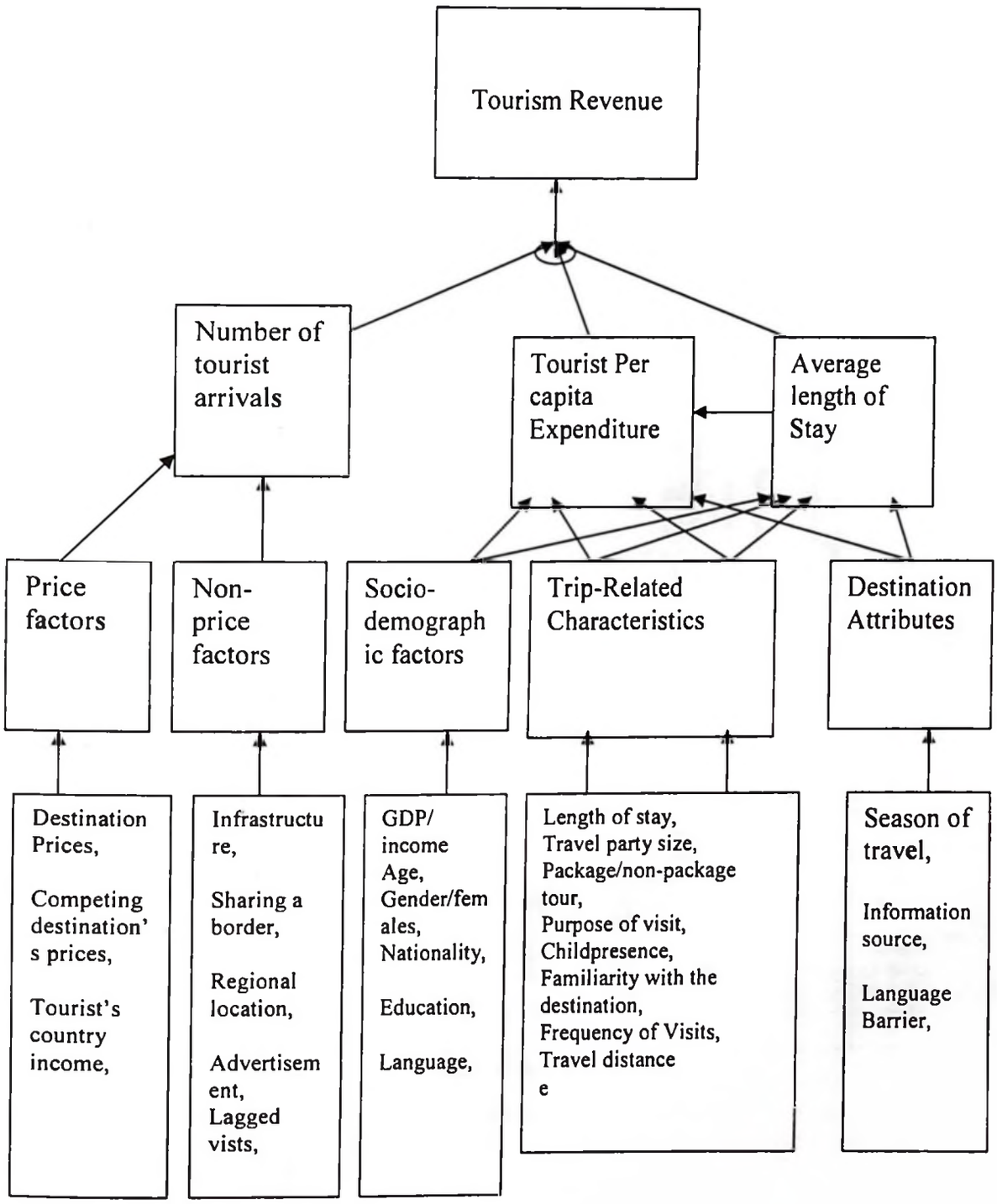
World Travel and Tourism Council (WTTC) (2009), "The 2009 Tourism Research" *Accessed from http://www.wttc.org/ on 6th December 2010*

Yang, Y; Wong, K.F; Zhang, J (2011), "Determinants of Length of *Stay* for Domestic Tourists: Case Study of Yixing", *Asia Pacific Journal of Tourism Research Volume 16, Issue 6, 2011*

Zhang,Y., Qu, Hailin and Tavitiyaman, P (2009), "The Determinants of the Travel Demand on International Tourist Arrivals to Thailand", *Asia Pacific Journal of Tourism Research, Vol14, No.1*

Appendix 1.1:

Figure 1.4: The diagrammatic link of the study's objectives



Appendix 2.1

Time series analysis on the determinants of tourist number of arrivals

Table 2.3 provides time series regression results on the determinants of number of tourist arrivals from 1970 up to 2007 with a view to assessing the role of trade liberalization in the coming of tourists. The dummy for the year 1986 and trade openness (ratio of trade to GDP) are used as proxies for trade liberalization. In both cases none of these variables is significant. The results are shown in Table 2.3

Table 2.3. Time series regression of the number of tourist arrivals: 1970-2007

Variable	Coefficient.	Std. Err.	t-values
Tourism price in Kenya	1.4909	.4545	3.28
Tourism price in Mauritius	.7154	.9986	0.72
Tourism process in Uganda	-.3435	.1512	-2.27
Tourism price in Botswana	-1.3585	.7167	-1.90
Tourism prices in Seychelles	1.4073	.8895	1.58
Tourism price in Tanzania(TZ)	-.4848	.38665	-1.25
Tanzanian exchange rate against the dollar	.27975	.08373	3.34
Trade openness	.06747	.4207	0.16
Year 1986	-.34181	.3085	-1.11
_cons	15.295	3.1468	4.86
R^2	0.83		
F(9,28)	21		
P>F	0.0000		
d statistic(10, 38)	1.03		
Dickey-Fuller unit root test on the predicted residual/Co integration test	P=0.0000 and Z(t)=-4.03		
n	38		

Tourism prices for the given countries have been calculated as a ratio of the cost of living (CPI's) in Tanzania to the cost of living in United States (CPI's). The cost of living in the United States was taken as a proxy for the cost of living by a tourist in any country of the world. This was done because the study considered an aggregate of

tourists from several countries per annum as a response variable which makes it hard to consider individual tourist cost of living. However, in chapter three where panel data regression is done individual tourist countries' cost of living were used. Table 2.4 provides the variance inflation factors (VIF) for every explanatory variable. The Table shows that the dummies for the year 1986 as well as trade openness have reasonable VIF (<10). Therefore their insignificance cannot be attributed to multcollinearlity.

Table 2.4 VIF of the variables used in the time series regression

Variable	VIF.
Tourism price in Kenya	2.24
Tourism price in Mauritius	6.44
Tourism process in Uganda	7.23
Tourism price in Botswana	5.13
Tourism prices in Seychelles	5.68
Tourism price in Tanzania(TZ)	13.44
Tanzanian exchange rate against the dollar	11.42
Trade openness	7.25
Year 1986	8.09

Appendix 3.1:

Table3.11: Correlation matrix of the variables used in the analysis of number of arrivals

	arrivals	P	ER	GDP	DIST	AD	IFR	TGDP
arrivals	1							
P	-0.06	1						
ER	0.16	-0.16	1					
GDP	0.22	-0.61	0.62	1				
DIST	-0.14	0.00	0.04	0.24	1			
AD	0.06	-0.05	0.06	0.003	0	1.00		
IFR	0.07	-0.03	0.14	-0.01	0	0.90	1.00	
TGDP	0.07	-0.02	0.13	-0.01	0	0.90	0.96	1.00
KR	-0.06	0.96	-0.16	-0.61	0.00	0.03	0.03	0.03
SR	-0.06	0.97	-0.16	-0.61	0.00	-0.02	-0.02	-0.02
MR	-0.02	0.18	-0.03	-0.61	0.00	-0.03	-0.04	-0.06
UG	-0.02	0.16	-0.02	-0.61	0.00	-0.05	-0.05	-0.06
BW	-0.02	0.18	-0.03	-0.61	0.00	-0.01	-0.04	-0.05
ZW	-0.01	0.19	-0.04	-0.50	0.00	0.37	0.16	0.16
SE	-0.02	0.19	-0.03	-0.61	0.00	-0.04	-0.04	-0.06
border	0.45	0.16	-0.14	-0.20	-0.38	0.00	0.00	0.00

	KR	SR	MR	UG	BW	ZW	SE	border
KR	1.00							
SR	0.97	1.00						
MR	0.18	0.18	1.000					
UG	0.16	0.17	0.999	1.00				
BW	0.18	0.18	1.000	1.00	1			
ZW	0.36	0.24	0.060	0.05	0.06	1.00		
SE	0.19	0.19	1.000	1.00	1.00	0.05	1.00	
border	0.16	0.16	0.024	0.02	0.02	0.05	0.03	1.00

Appendix 3.2:

The Concept of Utility Separability function

The following discussion is based on the works by Morishima (1961), Godman and Uzawa (1964).

The idea of separable utility function owes its origin to the Japanese mathematician Masazo Sono (Morishima, 1961). In the middle of the World War II Sono published his work on utility separability in Japanese, but his work remained unknown for a long time without attracting the attention of the western economists (Morishima, 1961). According to Morishima (1961), four years later after his work was published, Leontief (1947) discussed the same work, completely independently and other pioneers since then have been Strotz (1957) and Gorman (1959)

The concept of utility separability assumes that goods can be separated into several groups of homogenous units. For example a consumer may have a basket of food items, a basket of home assets, a basket of ornaments, or a basket of tourism goods. A utility function is separable if a consumer in the first stage only thinks about how much money to allocate to each basket of goods and later independently thinks about how to maximize each of the baskets separate from the others. This is a case of a strong separable utility function. When he only thinks about how to maximize one basket independently of the other goods and not vice versa then this is a case of a weak separable utility function. In the former case the marginal rate of substitution of good x belonging to group say A(x) by good y belonging to group say B(y) is independent of any quantity of goods belonging out of those two groups. In the latter case the Marginal rate of substitution between good xi and good xj belonging to the same group say A(x) is independent of any quantity of goods belonging out of the group A(x). This is the assumption used in chapter s 3, 4 and 5 that is A (x) is separable from the other but not necessarily the vice-versa.

Assumption of separability is important in this study because no details are available on the quantities and prices of non-tourism goods. To make this point clearer let us consider the following example.

so that good 1 and 2 can be considered as forming one group and good 3 and 4 forming another group resulting in basket $I^1 = \{x_1^1, x_2^1\}$ and basket $I^2 = \{x_3^2, x_4^2\}$.

Suppose a utility function is given as $U(x_1^1, x_2^1, x_3^2, x_4^2) = x_1^{\frac{1}{3}} x_3^{\frac{1}{3}} x_4^{\frac{1}{3}} + x_2^{\frac{2}{5}} x_3^{\frac{1}{5}} x_4^{\frac{1}{5}}$..... (a)

And let

$$u_2 = u^2(x_3, x_4) = x_3^{\frac{1}{3}} x_4^{\frac{1}{3}}. \qquad (b)$$

Then upon substituting (b) into *a*, the entire utility function can now be written as

$$U(x_1, x_2, u_2) = U(u^1(x_1, x_2), u^2(x_3, x_4)) = x_1^{\frac{1}{3}} u^2 + x_2^{\frac{2}{5}} \left(u^2\right)^{\frac{3}{5}} \qquad (c)$$

It can easily be shown that

$$\frac{\partial U(x_1 x_2 u_2)}{\partial x_3} = \left[x_1^{\frac{1}{3}} + \frac{3}{5} x_2^{\frac{2}{5}} \left(u^2\right)^{-2/3} \right] \frac{\partial u^2}{\partial x_3} \qquad (d)$$

Likewise it can also be shown that

$$\frac{\partial U(x_1 x_2 u_2)}{\partial x_4} = \left[x_1^{\frac{1}{3}} + \frac{3}{5} x_2^{\frac{2}{5}} \left(u^2\right)^{-2/3} \right] \frac{\partial u^2}{\partial x_4} \qquad (e)$$

By dividing (d) by (e) one gets the marginal rate of substation between good *x3* and *x4* belonging to group 2. This MRS is certainly independent of any quantity of good x_1 and good x_2 belonging to group 1. Therefore goods from group 2 are separable from group 1. On the contrary goods from group1 are not separable from those of group 2(it can be shown that $\frac{\partial U}{\partial x_1} \Big/ \frac{\partial U}{x_2}$ is not independent of $u^2\left(x_3^2, x_4^2\right)$). This is a case of weak separability in the utility function..

Now consider group 2 as consisting of tourism goods x_3 and x_4 and group1 as consisting of

non-tourism goods x_1 and x_2.

Suppose a tourist is maximizing $U(x_1^1, x_2^1, x_3^2, x_4^2) = x_1^{\frac{1}{3}} x_3^{\frac{1}{3}} x_4^{\frac{1}{3}} + x_2^{\frac{2}{5}} x_3^{\frac{1}{5}} x_4^{\frac{1}{5}}$

Subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 = I$

where

p_1 and p_2 are the prices of non-tourism goods x_1 and x_2

p_3 and p_4 are the prices of tourism goods x_3 and x_4

One can form the following langrage equation

$$l(x_1, x_2, x_3, x_4) = U(x_1^1, x_2^1, x_3^2, x_4^2) + \lambda(I - p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4)$$

Upon taking the first order condition one would have $MRS_{(x_4 x_3)} = \frac{\partial U}{\partial x_3} \Big/ \frac{\partial U}{x_4} = \frac{p_3}{p_4}$

Using (d) and (e) the MRS can further be written as $MRS_{(x_4 x_3)} = \frac{\partial u^2}{\partial x_3} \Big/ \frac{\partial u^2}{x_4} = \frac{p_3}{p_4}$

From the income constraint $p_1 x_1 + p_2 x_x + p_3 x_3 + p_4 x_4 = I$, one need only to have information on the budget allocated to tourism goods and simply use the constraints $p_3 x_3 + p_4 x_4 = m$ to solve completely for x_3 and x_4, where m is the total budget allocated to tourism expenditure and is equal to $I - p_1 x_1 - p_2 x_2$. This demonstration shows why the assumption of utility separability was important in chapters three, four and five. The demonstration further shows that we need not have non-tourism goods equally separable from tourism goods as our focus is on tourism goods. This explains why there is no apparent necessity to use the strong separable utility function in chapters three, four and five.

Now let us demonstrate the case of a strong separable utility function. Suppose there is a third group of goods consisting of good x_5 an x_6, so that the utility function can now be written as a sum of three separate utility functions as

$$U(x_1^1, x_2^1, x_3^2, x_4^2, x_5^3, x_6^3) = x_1^{\frac{1}{3}} x_2^{\frac{2}{3}} + x_3^{\frac{1}{5}} x_4^{\frac{4}{5}} . + x_5^{\frac{1}{7}} x_6^{\frac{6}{7}} \qquad .(f)$$

It can easily be shown from (f) that both non-tourism goods as well tourism goods are separable from each other. This is so because MRS between two goods from two different groups shall entirely depends on the quantities of goods from those two groups only. This is the case of a strong separable utility function. Formal statements and proofs of the concepts discussed here can be found in Goldman and Uzawa (1964).

Appendix 3.3:

The Derivation of the Demand function for Tanzanian tourism

One starts by considering the utility function comprising African tourism goods, with each African country considered as a commodity. In total there are *m* African countries/tourism goods. This group is separable from non-tourism goods in that a tourist having allocated his budget can in the second stage maximize his utility and hence derive the demand function for any county i. Therefore considering the log transformation of the utility function given in equation (3.4) of chapter 3 and its budget constraint (3.5), one can formulate the following langrage multiplier.

$$l(q_{jt}, p_t, \lambda) = \sum_j^m \delta_j \log(q_{jt} - k_{jt}) + \lambda(e_t - \sum_j^m q_{jt} p_t) \qquad (a)$$

From the first order conditions one establishes that

$$\frac{\partial l}{\partial q_{jt}} = \frac{\delta_j}{q_{jt} - k_{jt}} = \lambda p_{jt} \qquad (b)$$

Manipulating (b) and considering the fact that $e_t = \sum_j p_{jt} q_{jt}$, as given by equation (3.5), equation *b* becomes;

$$\delta_j = \lambda q_{jt} p_{jt} - \lambda k_{jt} p_{jt} \rightarrow \sum_j^m \delta_j = \lambda \sum_j^m q_{jt} p_{jt} - \lambda \sum_j^m k_{jt} p_{jt} \rightarrow \sum_j^m \delta_j = \lambda (e_t - \sum_j^m k_{jt} p_{jt})$$

$$\rightarrow \lambda = \frac{\sum_j^m \delta_j}{(e_t - \sum_j^m k_{jt} p_{jt})} \qquad (c)$$

From (b) one can write

$$q_{jt} - k_{jt} = \frac{\delta_j}{\lambda p_j} \qquad (d)$$

Upon substituting for λ using equation(c), equation(d) becomes

$$q_{jt} = k_{jt} + \frac{\delta_j (e_t - \sum_j^m k_{jt} p_j)}{p_j \sum_j^m \delta_j} = k_{jt} + \frac{\delta_j (e_t - \sum_j^m k_{jt} p_j)}{p_j} \qquad (e)$$

[Note that $\sum_j^m \delta_j = 1$ under the Cobb-Douglass utility function]

Substituting $k_{jt} = y_{jt} + \gamma_j q_{jt-1}$ in (e), one gets the following:

$$q_{jt} = y_{jt} + \gamma_j q_{jt-1} - \frac{\delta_j \sum_j^m y_j p_j}{p_j} + \frac{\delta_j (e_t - \sum_j^m p_j \gamma_j q_{jt-1})}{p_j} \qquad (f)$$

For estimation purpose the negative sign is ignored and one obtains the following equation:

$$q_{jt} = y_{jt} + \gamma_j q_{jt-1} + \frac{\delta_j \sum_j^m p_j \gamma_j}{p_j} + \frac{\delta_j}{p_j}\left(e_t - \sum p_j \gamma_j q_{jt-1}\right) + \varepsilon_{jt} \qquad (g)$$

Equation (*g*) tells us that demand by a tourist is a function of his budget, tourism prices in Tanzania, prices in the competing destinations and tastes. This is the equation (3.6) in chapter three.

Appendix 3.4:

Table 3.12. LSDV regression results of the log of number of arrivals

Var	Coef.	Std. Err.	t
TZ	1.55**	.72	2.16
ER	-.13*	.08	-1.70
GDP	.19**	.07	2.53
DIST	(dropped)		
AD	-.19	.44	-0.43
IFR	.83	.84	1.00
TGDP	-.58	1.18	-0.49
KE	.61	.66	0.93
SR	2.03	1.25	1.62
MR	-6.28	4.22	-1.49
UG	-.902	.62	-1.45
BW	-1.93	1.32	-1.46
ZW	.22**	.10	2.22
SE	4.56*	2.65	1.72
Border	(dropped)		
cons	-0.561	4.90	-0.11
R^2	**0.87**		
F/χ^2	F(13,1439)=12.89		
$P>F/\chi^2$	P>F=0.0000		
Test for Fixed effects	F(120,1439)=54.52 and P=0.0000		
N	1573		

Table 3.13 indicates the LSDV estimator of the fixed effect. The estimation results are equal to those of the Fixed effect within, except the R^2 value (0.87) which is remarkably higher compared with that of the Fixed effect within. This indicates that the unobservable individual effects play a significant role in determining the number of tourist arrivals. For prediction purposes one must not ignore them.

Appendix 3.5

The variance inflation factors of the explanatory variables used in the OLS estimate of the determinants of number of tourist arrivals

Table 3.13. VIF of the variables used in the analysis of number of arrivals

Variable	VIF	1/VIF
TZ	8150.57	0.000123
ER	36.06	0.027732
GDP	1.74	0.575984
DIST	1.55	0.647170
AD	53.29	0.018766
IFR	78.99	0.012660
TGDP	90.35	0.011068
KE	6645.57	0.000114
SR	24835.35	0.000040
MR	281921.53	0.000004
UG	6122.38	0.000163
BW	27807.76	0.000036
ZW	230.66	0.004335
SE	11705.76	0.000036
Border	1.57	0.635381
Mean VIF	31199.53	0.647170

Table 3.13 indicates that, with the exception of income and distance, the remaining variables have a very high value of VIF (>10), meaning that they are not immune to the multcollinearlity problem (Gujarat 2003).

Table 3.14: Panel root test for the variables used in the regression of the log of the number of arrivals

	Levels			First difference		
	t-value	t-star	p>t	t-value	t-star	p>t
TA	-0.11.041	3.77	0.999	-33.44	-3.96	0.000
TZ	-1240.9	-1349.7	0.000	-118.22	-122.1	0.000
ER	8.000	15.71	1.000	-107.16	-109.60	0.000
GDP	-14.2	-8.05	0.000	-27.79	-7.51	0.000
DIST*	NA	NA	NA	NA	NA	NA
AD*	NA	NA	NA	NA	NA	NA
IFR*	NA	NA	NA	NA	NA	NA
TGDP*	NA	NA	NA	NA	NA	NA
KE	-76	-78.25	0.000	-118.22	-122.01	0.000
SR	-76	-78.25	0.000	-118.224	-122.098	0.000
MR	-0.84	-70.6	0.000	-118.224	-122.098	0.000
UG	-75.903	-78.244	0.000	-118.226	-122.10	0.000
BW	-75.909	-78.25	0.000	-118.224	-122.098	0.000
ZW	-76	-78.25	0.000	-118.22	-122	0.000
SE	-75.909	-78.25	0.000	-118.22	-122	0.000

* No test conducted

Table 3.14 shows the panel root test for variables used in the regression analysis of the number of tourist arrivals. The test indicates that the null hypothesis of a non-stationary variable is rejected in all the variables in levels except the log of the number of arrivals (TA) and the log of the exchange rate(ER). On the other hand the test indicates that all the variables in their first difference are stationary. Similar results were obtained when the Im-Pesaran-Shin test was used. These results further justify the use of the first difference estimator instead of the fixed effect within estimator in the static panel estimation (Chapter three).

It is worth noting that Levilin test and other panel root tests require a variety of observations both across time and individuals. The test cannot be carried out for distance because it does not have time a dimension. Likewise, advertising expenditure, infrastructure and Tanzania's per capita GDP have no cross-sectional dimension and thus the test cannot be carried out as either(see foot notes in Table 3.4d).

The test is based on the following equation: $\Delta y_{it} = \alpha_i + \delta_{it} + \theta_t + \rho_i y_{it-1} + \varepsilon_{it}$

Where

$i = 1,2,3...N,\ t = 1,2,3...T$.

The test, tests the hypothesis $H_0 : \rho_i = 0$ for all i against the alternative hypothesis $H_A : \rho_i = \rho < 0$ for all i.

The null hypothesis assumes that a variable is non-stationary across all the panels against the alternative that it is stationary at least in one panel.

Appendix3.6:

Table 3.15: IV regression estimate of the log of the number of arrivals using ΔTA_{it-2} as an instrument for ΔTA_{it-1}

	Coef.	Std. Err.	t
TA	.44***	.10	4.33
TZ	-.29	.89	-0.32
ER	-.001	.14	-0.07
GDP	.05	.11	0.43
AD	1.36***	.39	3.50
IFR	3.43***	.69	5.01
TGDP	3.74**	1.51	2.48
KE	4.15***	1.09	3.81
SR	-.98	1.10	-0.89
MR	-5.52*	2.98	-1.85
UG	-1.17	.78	-1.49
BW	-1.46	1.1	-1.27
ZW	.01	.08	0.11
SE	5.24***	1.7	3.16
_cons	-0.67***	0.17	-4.00
F	7.77		
P>F	0.000		
N	1452		

Table 3.15 indicates the IV estimates using the twice differenced log of number of arrivals as an instrument for the first lag of the log of the number of arrivals as suggested by Anderson and Hsiao (1982).There is great discrepancy between these results and those of the GMM estimate produced in Table 3.14 of chapter three. The coefficients are much smaller in magnitude and the standard errors are much higher than those of the GMM estimate.

Appendix 4.1

Consistency of the model estimates

1. Testing for the consistency of the peak season variable

The variable peak season is instrumented by the variables child presence and price whereas the rest of the variables are assumed to be exogenous.

(i) First the tests for instruments exogeneity/over identifying restrictions are provided. These are the Sargan score test and the Basmann test.

Tests of over identifying restrictions:
Sargan (score) chi2 (1) = .194439 (p = 0.6592)
Basmann chi2 (1) = .194245 (p = 0.6594)

Both Sargan and Basmann tests for instruments exogeneity /over identifying restrictions accept that the instruments are exogenous. The two tests also provide confidence on the model specification.
(ii) The peak season variable is predicted after regressing it against its own instruments as well as on other regressors assumed to be exogenous in equation 4.1b

(iii) Finally the predicted value of the peak season variable is included in the OLS regression of tourist per capita expenditure alongside the original value of the peak season. The insignificance of the predicted peak season proves its exogeneity .The results are given in the Table below:

2. Testing for the consistency of the variable visiting friends (VistFRD)

The variable visiting friends is instrumented by two variables: child presence and price. As before the test for the instruments exogeneity is undertaken
(i)**Tests of over identifying restrictions:**
Sargan (score) chi2 (1) = .109034 (p = 0.7412)
Basmann chi2 (1) =108925 (p = 0.7414)
(ii) The VistFRD variable is predicted after regressing it against the instruments.
(iii) OLS regression of the log of tourist daily expenditure including both the original VistFRD and the predicted one are undertaken. As before the insignificance of this variable proves its exogeneity. The results are given in the Table below:

3. Testing for the consistency of the variable visitors coming for leisure and recreation (VistLSR)
This variable is instrumented by the variables child presence and price as in the previous case. However this time around, its predicted value is highly correlated with the original value, making it impossible to be estimated in the same equation (co linearity). To avoid this problem a comparison is made directly between the 2sls estimate equation against the non-instrumented equation using the Hausman test. The Hausman test does not reject the null hypothesis of no systematic difference in the coefficient (Chi2 (20) =0.67, P=1.000). The value of p=1 tends to reflect what has just been highlighted that the predicted value is almost the same as the original one.

Results for exogeneity of Peak season				Results for exogeneity of Vist FRD			
variable	Coef.	t	P>t	variable	Coef.	t	P>t
Age	0.104012	6.07	0.000	Age	0.123585	12.74	0.000
Females	-0.01559	-1.13	0.258	Females	0.007258	0.44	0.658
GDP	1.74E-06	2.02	0.044	GDP	1.84E-06	2.1	0.036
Er	0.000423	8.81	0.000	Er	0.000378	17.76	0.000
EL	-0.16282	-7.13	0.000	EL	-0.14073	-6.71	0.000
Childno	-0.09953	-6.14	0.000	Childno	-0.08579	-7.33	0.000
Lstay	-0.01706	-15.37	0.000	Lstay	-0.02022	-9.7	0.000
Tarra	0.275847	13.37	0.000	Tarra	0.189555	3.24	0.001
VistFRD	-0.33597	-3.79	0.000	**VistFRDhat**	-2.45798	-1.22	**0.223**
VistLSR	-0.10037	-2.84	0.004	**VistFRD**	-0.23939	-6.42	**0.000**
VistBSN	0.251739	2.59	0.009	VistLSR	-1.55248	-1.28	0.199
Adultno	-0.07186	-7.41	0.000	VistBSN	-1.14918	-0.93	0.353
Ttpartyno	-0.17164	-6.99	0.000	Adultno	-0.09585	-4.72	0.000
Frvists	0.020062	1.22	0.223	Ttpartyno	-0.14545	-3.47	0.001
Fadest	0.097417	2.02	0.043	Frvists	0.01581	0.86	0.392
Dist	4.04E-05	7.89	0.000	Fadest	0.088599	2.2	0.028
Nosites	0.02312	0.8	0.423	Dist	3.05E-05	3.28	0.001
Peak hat	-0.76265	-1.19	**0.235**	Nosites	0.084117	2.88	0.004
peak	-0.1524	-10.62	**0.000**	peak	-0.20608	-4.5	0.000
Africa	0.413678	7.86	0.000	Africa	0.352396	5.15	0.000
Asia	0.225403	7.6	0.000	Asia	0.26071	5.44	0.000
MEast	0.275579	2.66	0.008	MEast	0.172543	2.75	0.006
SAmerica	0.477403	5.24	0.000	SAmerica	0.423617	5.48	0.000
year1	(dropped)			year2	0.71157	21.15	0.000
year2	0.945142	5.23	0.000	year3	0.537356	14.13	0.000
year3	0.68477	7.05	0.000	cons	5.558738	4.25	0.000
_cons	4.481697	10.26	0.000				
R2	0.28			R2	0.28		
F	425.10			F	425.13		
P>F	0.000			P>F	0.000		
n	25,880			n	25,880		

The Table provides the results for endogeneity test of the variable peak season (Peak) and the variable visitors on friends (VistFRD). As said earlier both variables are exogenous much as their corresponding predicts are insignificant in a regression of a log of per capita tourist expenditure.

4. Assessing the influence of the variable length of stay on a tourist per capita daily spending as well as on a tourist party daily spending

The variable length of stay was instrumented by the variables child presence and nosites whereas the rest of the variables are assumed to be exogenous. Both Sargan and Basmann tests for instruments exogeneity /over identifying restrictions accept that the instruments are exogenous.

Tests of over identifying restrictions for the case of tourist per capita dailiy spending
Sargan (score) chi2 (1) = .1.92786 (p = 0.1650)
Basmann chi2 (1) = .192621 (p = 0.1652)

Tests of over identifying restrictions for the case of tourist party daily spending
Sargan (score) chi2 (1) = .051613 (p = 0.8203)
Basmann chi2 (1) = .051563 (p = 0.8204)

variable	Tourist per capita daily expenditure		Tourist party daily expenditure	
	coef	Std.Err	coef	Std.Err
lstay	-0.01***	0.003	-0.01***	0.003
age	0.14***	0.01	0.14***	0.01
gender	-0.01**	0.011	-0.02*	0.011
gdp	0.00*	0	0.00*	0
er	0.00***	0	0.00***	0
language	-0.16***	0.02	-0.16***	0.02
childno	-0.08***	0.012	0.17***	0.012
tarra	0.31***	0.019	0.31***	0.019
frd	-0.18***	0.038	-0.19***	0.038
lsr	0.01	0.037	0	0.037
bsn	0.46***	0.044	0.45***	0.044
adultno	-0.07***	0.008	0.12***	0.008
tpartyno2	-0.18**	0.016	0.36***	0.016
frvists	0.05**	0.015	0.05***	0.015
fadest	0.03**	0.014	0.03**	0.014
dist	0.00***	0	0.00***	0
peak	-0.16***	0.015	-0.17***	0.015
af	0.47***	0.052	0.46***	0.052
as	0.23***	0.028	0.23***	0.028
me	0.21***	0.068	0.20***	0.067
sa	0.42***	0.077	0.43***	0.077
year1	-0.73***	0.029	-0.73***	0.029
year3	-0.15***	0.024	-0.15***	0.024
_cons	4.4	0.101	4.24***	0.101
R2	0.26		0.27	
Wald chi(23)	8430		8814	
P	0.0000		0.0000	
n	25880		25880	

Appendix 4.2

Table 4.2: Variance inflation factors of the regressors used in OLS estimate of equation 4.1b

Variable	VIF	1/VIF
Africa	5.54	0.180549
VistLSR	4.12	0.242794
GDP	4.08	0.244937
Er	3.79	0.26373
Dist	3.69	0.271344
year3	2.95	0.338727
VistBSN	2.89	0.345627
El	2.62	0.381967
Adultno	2.57	0.389216
Female	2.48	0.403977
VistFRD	2.38	0.419532
year2	1.98	0.50395
Tpartyno2	1.61	0.621185
Asia	1.5	0.664559
Tarra	1.36	0.733152
Childno	1.3	0.771862
Fadest	1.2	0.834426
Frvists	1.19	0.841048
SAmerica	1.17	0.851964
Peak	1.13	0.887351
Nosites	1.11	0.89822
Age	1.1	0.907737
MEast	1.06	0.941421
Mean VIF	2.3	

Appendix 4.3

Table 4.4: The correlation matrix of the variables used in the study

	lstay	age	feme l	GDP	er	El	Childpres e	Childn o	tarra	pvis t	Adultn o	Tpart y
Lstay	1.00											
Age	-0.08	1.00										
Fem	-0.01	0.01	1.00									
GDP	0.03	-0.02	0.01	1.00								
Er	0.02	-0.08	0.01	0.63	1.00							
El	-0.01	0.08	-0.05	0.08	0.14	1.00						
Childpres e	-0.02	0.14	0.26	-0.05	-0.04	-0.02	1.00					
Childno	-0.02	0.12	0.31	-0.05	-0.04	-0.02	0.85	1.00				
Tarra	-0.19	0.04	0.09	0.16	0.18	-0.06	0.02	0.01	1.00			
Pvist	-0.18	0.04	-0.08	-0.08	-0.05	0.01	-0.03	-0.02	0.06	1.00		
Adultno	-0.03	0.00	0.70	-0.03	-0.03	-0.06	0.02	0.02	0.05	-0.03	1.00	
Tparty	-0.07	0.05	0.45	-0.02	0.03	-0.11	0.24	0.21	0.16	-0.01	0.52	1.00
Frvists	-0.09	0.01	-0.04	-0.03	-0.01	0.10	-0.01	0.00	-0.02	0.09	-0.03	-0.03
Fadest	0.03	0.07	-0.07	0.05	0.03	0.07	-0.02	-0.02	-0.17	0.01	-0.05	-0.13
Dist	0.03	0.12	0.04	0.52	0.17	0.37	0.01	0.00	0.11	-0.10	0.01	0.00
Nosite	0.13	0.05	0.10	0.13	0.09	0.03	0.04	0.04	0.19	-0.02	0.05	0.08
Satfy	-	-	-	0.0	0.0	0.0	0.00	0.00	-	-	-0.03	-0.04

	0.01	0.01	0.01	7	6	7			0.01	0.01		
Season	0.04	-0.10	0.01	0.22	0.23	-0.02	-0.01	-0.03	0.13	-0.08	0.00	0.05
Price	-0.02	-0.01	-0.03	-0.15	-0.14	-0.01	-0.01	-0.01	-0.06	0.02	-0.01	-0.04
Pcapit	-0.21	0.06	-0.10	0.23	0.25	0.06	-0.08	-0.08	0.21	0.10	-0.12	-0.12
Isourc	-0.10	-0.04	0.03	0.11	0.09	-0.01	0.00	0.00	0.20	0.04	0.00	0.02
Af	-0.07	0.00	-0.09	-0.54	-0.47	0.24	-0.03	-0.03	-0.17	0.11	-0.05	-0.12
As	-0.02	-0.02	-0.03	-0.11	-0.25	0.05	-0.02	-0.02	-0.01	0.02	-0.02	-0.06
Eu	0.05	-0.10	0.05	0.12	0.42	-0.53	0.02	0.02	0.07	-0.02	0.04	0.13
Me	-0.01	-0.03	0.00	-0.06	-0.07	-0.10	0.01	0.01	-0.03	0.00	0.01	0.02
Sa	0.00	-0.01	0.00	-0.13	-0.10	-0.08	0.00	0.01	0.01	0.00	0.00	-0.02
Na	0.01	0.14	0.02	0.36	0.03	0.48	0.01	0.00	0.05	-0.07	-0.01	-0.04
year1	0.04	0.14	0.05	-0.58	-0.49	-0.09	0.11	0.11	-0.15	0.03	0.09	0.11
year2	-0.03	-0.09	-0.03	0.23	0.26	0.01	-0.04	-0.05	0.13	0.01	-0.04	-0.03
year3	-0.02	-0.09	-0.04	0.47	0.35	0.09	-0.09	-0.09	0.07	-0.03	-0.07	-0.10

	frvists	fades	dist	nosit	satf	seaso	price	pcapita	isource	af	as	Eu
frvists	1.00											
fadest	0.16	1.00										
dist	-0.05	-0.05	1.00									
nosites	-0.03	-0.11	0.19	1.00								
satfy	0.03	0.04	0.03	0.01	1.00							
season	-0.01	0.05	0.06	-0.04	0.00	1.00						
price	0.03	0.05	-0.10	-0.06	0.01	-0.03	1.00					
pcapita	0.07	0.08	0.07	0.00	0.04	0.10	0.00	1.00				

lsource	0.03	-0.01	0.00	-0.03	0.03	0.07	0.01	0.10	1.00			
af	0.11	0.14	-0.53	-0.18	0.02	-0.11	0.17	-0.01	0.03	1.00		
as	0.07	0.02	0.13	0.02	0.00	-0.02	0.02	0.00	0.01	-0.09	1.00	
eu	-0.07	-0.09	-0.45	0.02	-0.04	0.06	-0.09	-0.06	-0.04	-0.39	-0.36	1.00
me	-0.02	0.00	-0.11	-0.04	0.00	0.02	0.00	-0.02	0.01	-0.03	-0.03	-0.12
sa	0.00	0.01	0.08	0.00	0.01	0.01	0.06	0.02	0.02	-0.03	-0.03	-0.11

Appendix 4.4

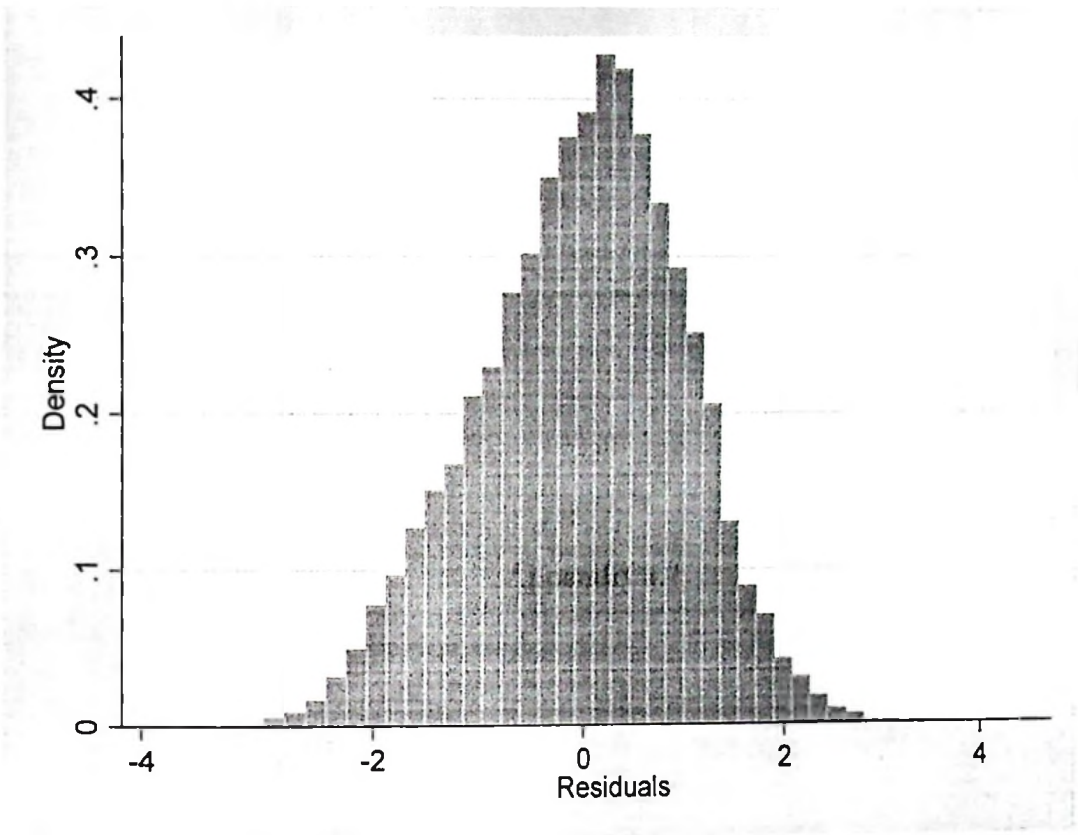


Figure 4.1 Histogram of the residuals from the OLS estimate of daily tourist per capita expenditure

Figure 4.1 above indicates that the errors are somewhat normally distributed consistent with normality assumption of OLS. This result gives confidence on the reliability of the t-tests on the significance of the coefficients as well as the F test on the significance of the overall model.

Appendix5.1

The consistency of the model estimate

1. Testing for consistency of the variable peak season.

The variable peak season is instrumented by the variables child presence and price, whereas the rest of the variables are assumed to be exogenous.

(i) The tests for instruments' exogeneity/over identifying restrictions are provided. These are the Sargan score test as well as the Basman test.

Tests of over identifying restrictions:

Sargan (score) chi2(1)= 2.492 (p = 0.1144)
Basmann chi2(1) = 2.491 (p = 0.1145)

Both Sargan and Basman tests for instruments exogeneity /over identifying restrictions accept that the instruments are exogenous. The two tests corroborate the model specification.

(ii) The peak season variable is predicted after regressing it against its own instruments as well as on other regressors assumed to be exogenous in equation 5.2

(iii) The OLS regression of tourist length of stay including both the original peak season variable and the predicted one is provided. The insignificance of the predicted peak season proves its exogeneity .The results are given in the Table below;

2. Testing for the consistency of the variable Travel arrangement (Tarra)

The variable travel arrangement was instrumented by two variables: child presence and visits on leisure and recreation (VistLSR).

(i) As before the test for the instruments, exogeneity is undertaken.

Tests of over identifying restrictions:

Sargan (score) chi2(1)= .4169 (p = 0.5185)
Basmann chi2(1) = .4165 (p = 0.5187)

(ii) The variable VistLSR is predicted after regressing it against the instruments.

(iii) Finally OLS regression of the log of tourist length of stay including both the original travel arrangement variable and the predicted one is provided. The significance of the predicted variable proves its endogeneity. The results are given in the table below;

Results for testing exogeneity of Peak season				Results for testing the exogeneity of TARRA			
variable	Coef.	t	P>t	variable	Coef.	t	P>t
Age	-0.07721	-7.55	0.000	Age	-0.02767	-4.05	0.000
Females	0.028589	3.6	0.000	Females	0.04759	6.62	0.000
GDP	2.77E-06	5.21	0.000	GDP	6.60E-07	1.16	0.244
Er	6.94E-05	2.44	0.015	Er	6.94E-05	5.32	0.000
EL	0.004092	0.29	0.775	EL	-0.03514	-2.65	0.008
Pcapita	-0.00099	-50.04	0.000	Pcapita	-0.0007	-19.27	0.000
Childno	-0.04874	-5.09	0.000	Childno	-0.05787	-8.33	0.000
VistFRD	-0.05557	-0.97	0.333	**Tarrahat**	-0.53899	-7.25	**0.000**
VistLSR	-0.24859	-9.29	0.000	**Tarra**	-0.16958	-19.95	**0.000**
VistBSN	-0.47681	-7.48	0.000	VistFRD	-0.09858	-2.53	0.011
Adultno	-0.0306	-6.2	0.000	VistBSN	-0.50078	-13.06	0.000
Ttpartyno	-0.0459	-3.5	0.000	Adultno	-0.04159	-8.25	0.000
Frvists	-0.19611	-18.26	0.000	Ttpartyno	-0.01333	-1.26	0.208
Fadest	0.08967	3.21	0.001	Frvists	-0.20033	-23.86	0.000
Dist	-1.1E-05	-3.52	0.000	Fadest	0.002231	0.21	0.836
Nosites	0.322151	21.29	0.000	Dist	-7.8E-07	-0.24	0.808
Peak hat	-0.45379	-1.21	**0.226**	Nosites	0.402255	35.26	0.000
peak	0.091678	9.58	**0.000**	price	-0.01255	-1.38	0.168
Isource	-0.12377	-14.21	0.000	peak	0.105209	11.05	0.000
Africa	-0.24572	-6.85	0.000	Isource	-0.027	-2.16	0.031
Asia	-0.12175	-6.27	0.000	Africa	-0.25103	-7	0.000
MEast	-0.07453	-1.2	0.229	Asia	-0.12532	-6.74	0.000
SAmerica	0.05721	1.03	0.304	MEast	-0.23331	-5.67	0.000
_cons	3.043458	11.42	0.000	SAmerica	-0.01943	-0.4	0.686
				year2	0.228913	11.59	0.000
				year3	0.015608	0.93	0.354
				_cons	2.683536	77.44	0.000
R2	0.24			R2	0.26		
F	320.41			F	325.25		
P>F	0.000			P>F	0.000		
n	25,880			n	258880		

As it can be seen from the Table, the predicted value of **peak season** variable is insignificant, suggesting that the variable peak is exogenous. On the contrary the predicted value of travel arrangement is significant, suggesting that **Tarra** is endogenous.

Appendix 5.2

Table 5.2: Variance inflation factors of the regressors used in the OLS estimate of the log of tourist length of stay.

Variable	VIF	1/VIF
Africa	VIF	1/VIF
GDP	5.5	0.18046
VistLSR	4.1	0.24231
Er	4	0.25152
Dist	3.8	0.2639
year3	3.7	0.27219
VistBSN	3	0.32867
EL	2.9	0.34591
Adultno	2.6	0.38334
Female	2.6	0.38962
VistFRD	2.5	0.40425
year2	2.4	0.4206
Tpartyno2	2	0.49179
Asia	1.6	0.62261
Tarra	1.5	0.66455
Childno	1.3	0.77255
Fadest	1.2	0.83957
Frvists	1.2	0.84106
SAmerica	1.2	0.85086
Iisource	1.1	0.9025
Peak	1.1	0.8875
Nosites	1.1	0.90466
Age	1.1	0.91847
MEast	1.1	0.94195
Price	1.1	0.95262
Mean VIF	2.2	

Appendix5.3

1. The hazard function for the exponential distribution

Consider the exponential distribution given as $f(t) = \gamma \exp(-\gamma t)$, for $0 \le t \le \infty$

$$S(t) = 1 - F(t) = 1 - \int_0^t \gamma \exp p(-\gamma s) ds = 1 - \left[-\exp(-\gamma s)\right]_0^t = 1 + \exp(\gamma t) - 1 = \exp(-\gamma t)$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\gamma \exp(-\gamma t)}{\exp(-\gamma t)} = \gamma$$

2.. The hazard function for the weibull distribution

Consider the Weibull distribution given as $f(t) = \gamma \alpha t^{\alpha-1} \exp(-\gamma t^{\alpha})$, for $0 \le t \le \infty$

[note that when $\alpha = 1$, the Weibull distribution reduces to an exponential distribution]

$$S(t) = 1 - F(t) = 1 - \int_0^t \gamma \alpha t^{\alpha-1} \exp(-\gamma s^{\alpha}) ds = 1 - \left[-\exp(-\gamma s^{\alpha})\right]_0^t = 1 + \exp(-\gamma t^{\alpha}) - 1 = \exp(-\gamma t^{\alpha})$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\gamma \alpha t^{\alpha-1} \exp(-\gamma t^{\alpha})}{\exp(-\gamma t^{\alpha})} = \gamma \alpha t^{\alpha-1}$$

3. The hazard function for the Gompertz distribution

4. The hazard rate for the log-normal distribution

Consider the log-normal distribution given as $f(t) = \exp\left(\frac{\ln t - \mu}{-2\sigma^2}\right) \Big/ t\sigma\sqrt{2\Pi}$,

for $-\infty \le \ln t \le \infty$

$$S(t) = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \text{ where } \Phi\left(\frac{\ln t - \mu}{\sigma}\right) = \int_{-\infty}^{t} \exp\left(\frac{\ln s - \mu}{-2\sigma^2}\right) ds \Big/ t\sigma\sqrt{2\Pi}$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\exp\left[(\ln t - \mu)/-2\sigma^2\right]}{t\sigma\sqrt{2\Pi}\left[1 - \Phi(\frac{\ln t - \mu}{\sigma}\right]}$$

5. The hazard rate for the log-logistic distribution

Consider the logistic distribution given as $f(t)=\frac{\alpha\gamma^{\alpha}t^{\alpha-1}}{\left[1+(\gamma t)^{\alpha}\right]^{2}}$ for $0\le t\le\infty$

[Note that the term $(\gamma t)^{\alpha}$ is analogous to e^{x} in the pdf of the logistic distribution given as $f(x)=e^{x}/\left(1+e^{x}\right)$]

Therefore

$$s(t)=1-F(t)=1-\int_0^t f(s)ds=1-\int_0^t\frac{\alpha\gamma^{\alpha}s^{\alpha-1}}{\left[1+(\gamma s)^{\alpha}\right]^{2}}ds=1-\left[-1\left(1+(\gamma s)^{\alpha}\right)^{-1}\right]_0$$

$$=1-\left[\frac{-1}{1+(\gamma t)^{\alpha}}+1\right]=\frac{1}{1+(\gamma t)^{\alpha}}$$

$$\lambda(t)=\frac{f(t)}{S(t)}=\frac{\alpha\gamma^{\alpha}t^{\alpha-1}}{1+(\gamma t)^{\alpha}}$$

6. The hazard rate for the Gama Distribution

Consider the Gama distribution given as $f(t)=\frac{\gamma^{\alpha}t^{\alpha-1}\exp(-\gamma t^{\alpha})}{\pi(\alpha)}$ for $0\le t\le\infty$

Appendix 5.4

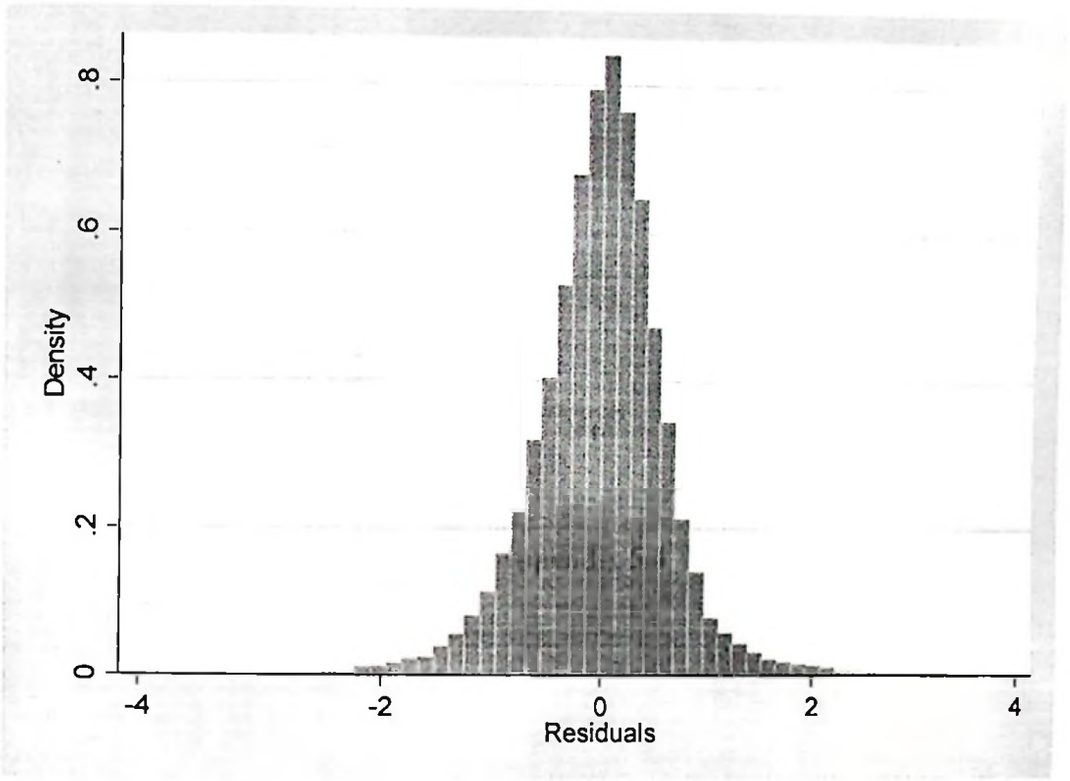


Figure 5.4: Histogram of the residuals from the log linear model of length of stay

The figure indictaes that the residuals from the log-linear model on the length of stay are almost normally distributed with a very symmetricl distribution, suggesting that the log-logistic or log-normal distrbution could be the apprpriate survival model for length of stay.

Appendix6.1

The consistency of the model estimate

1. Testing for consistency of the variable length of stay.

The variable length of stay (Lstay) is instrumented by the variables child presence and information source (Isource) whereas the rest of the variables are assumed to be exogenous.

(ii) Tests of over identifying restrictions:

Sargan (score) chi2(1)= .514346 (p = 0.4733)
Basmann chi2(1) = .513859 (p = 0.4735)

Both Sargan and Basmann tests for instruments exogeneity /over identifying restrictions accept that the instruments are exogenous. The two tests provide confidence on the model specification.

(ii) The variable Lstay is predicted after regressing it against its own instruments as well as on other regressors assumed to be exogenous in equation 6.3d

(iii) Finally a linear probability model of the variable travel arrangement including both the original Lstay variable and the predicted one is provided. The significance of the predicted variable proves its endogeneity. The results are given next.

2. Testing for the consistency of the variable visiting friends (VistFRD)

The variable visiting friends is instrumented by two variables: child presence and price As before
(i) As before I provide the test for the instruments exogeneity

Tests of over identifying restrictions:
Sargan (score) chi2(1) = .310714 (p = 0.5772)
Basmann chi2(1) = .3104 (p = 0.5774)

(ii) The variable VistFRD is predicted after regressing it against the instruments.

(iii) Finally a linear probability model of the variable travel arrangement including both the original VistFRD and the predicted one is provided. The insignificance of this variable proves its exogeneity. The results are given in Table below.

Results of testing the endogeneity of the variable length of stay (Lstay)				Results of testing the endogeneity of the variable visitors on friends (VistFRD)			
variable	Coef.	t	P>t	Variable	Coef.	t	P>t
Age	-0.065	-9.47	-0.065	Age	0.06	15.25	0.000
Female	0.066	12.58	0.066	Female	0.02	2.88	0.004
GDP	0.000	-3.19	0.000	GDP	0.00	-7.08	0.000
Er	0.000	-1.10	0.000	Er	0.00	7.65	0.000
El	0.014	1.42	0.014	El	-0.08	-8.48	0.000
Childno	-0.062	-11.00	-0.062	Childno	-0.03	-5.88	0.000
Lstay	-0.005	-16.41	-0.005	Lstay	-0.01	-16.4	0.000
Lstayhat	-0.069	-21.67	-0.069	**VistFRDhat**	1.07	1.13	0.257
VistFRD	-0.477	-24.68	-0.477	**VistFRD**	-0.18	-13.6	0.000
VistLSRr	-0.435	-11.76	-0.435	VistLSRr	0.94	1.64	0.100
VistBSN	-0.782	-24.17	-0.782	VistBSN	0.51	0.88	0.379
Adultno	-0.051	-12.71	-0.051	Adultno	-0.01	-1.38	0.167
Tpartyno2	-0.038	-4.72	-0.038	Tpartyno2	0.03	1.42	0.157
Frvists	-0.172	-19.31	-0.172	Frvists	-0.02	-2.90	0.004
Fadest	0.011	1.53	0.011	Fadest	-0.09	-4.67	0.000
Dist	0.000	-4.33	0.000	Dist	0.00	4.81	0.000
Nosites	0.423	27.63	0.423	Nosites	0.11	12.84	0.000
Peak	0.089	12.81	0.089	Peak	0.05	2.17	0.030
Africa	-0.405	-14.31	-0.405	Isource	0.11	9.84	0.000
Asia	-0.111	-8.42	-0.111	Africa	0.03	1.08	0.281
MEast	-0.291	-9.72	-0.291	Asia	-0.01	-0.37	0.709
SAmerica	-0.056	-1.73	-0.056	MEast	-0.15	-4.97	0.000
year1	(dropped)		(dropped)	SAmerica	-0.04	-1.36	0.174
year2	0.140	11.54	0.140	Year1	(dropped)		
year3	-0.001	-0.07	-0.001	Year1	0.17	8.41	0.000
cons	2.031	22.19	2.031	Year2	0.10	4.52	0.000
				_cons	-0.61	-1.03	0.305
Adj-R^2	.30			R2	0.30		
F	786.77			F	755.01		
P	0.000			P	0.000		
n	25,880			n	25,880		

As it can be seen from the table, the predicted value of the variable length of stay variable is significant, suggesting that the variable length of stay is endogenous. On

the contrary the predicted value of the variable visitors (VistFRD) on friends insignificant, suggesting that **VistFRD** is exogenous.

3. Testing for the consistency of the variable visitor on Business purposes (VistBSN)
The variable visiting for business purposes was instrumented by two variables: child presence and price just as before.

(i) **Tests of over identifying restrictions:**
Sargan (score) chi2 (1) = 1.55655 (p = 0.2122)
Basmann chi2 (1) = 1.55514 (p = 0.2124)

(iii) The variable VistBSN is predicted after regressing it against the instruments.

(iv) Finally a linear probability model of the variable travel arrangement including both the original VistBSN value and its predicted value is provided. The insignificance of the predicted variable proves its exogeneity. The results are given next:

	Coef.	t	P>t
Age	0.0816005	4.5	0.000
Females	-0.009313	-0.34	0.731
GDP	-1.80E-06	-2.34	0.019
Er	0.0000775	7.75	0.000
EL	-0.067239	-5.87	0.000
Childno	-0.029617	-5.1	0.000
Lstay	-0.005465	-16.41	0.000
VistFRD	-1.230704	-1.46	0.144
VistLSR	-0.721948	-0.89	0.374
VistBSNhat	-1.636028	-1.25	0.211
VistBSN	-0.145365	-11.13	0.000
Adultno	-0.011229	-1.41	0.157
Tpartyno	0.027604	1.72	0.085
Frvists	0.0082414	0.31	0.756
Fadest	-0.033393	-1.05	0.295
Dist	7.02E-06	0.86	0.392
Nosites	0.0952036	5.1	0.000
Peak	-0.020562	-0.56	0.576
Isource	0.1345915	11.55	0.000
Africa	0.1348856	1.33	0.183
Asia	0.0698027	1.51	0.132
Meast	-0.144856	-4.94	0.000
SAmerfica	0.0252175	0.41	0.681
year1	(dropped)		
year2	0.0897495	1.67	0.095
year3	0.016806	0.35	0.725
_cons	1.095617	1.33	0.185
Adjusted R2	0.30		
F	754.71		
P	0.000		
N	25880		

Appendix 6.2

Table 5.2: Variance inflation factors of the regressors used in the OLS estimate of Package tour.

Variable VIF	1/VIF	
lsr	4.08	0.245237
GDP	3.35	0.298774
bsn	2.94	0.340188
adultno	2.57	0.388758
gender	2.47	0.40408
frd	2.38	0.419483
dist	2.08	0.481024
year3	2.01	0.497918
er	2.01	0.498504
year2	1.64	0.610928
tpartyno2	1.61	0.62069
language	1.37	0.729026
childno	1.3	0.771299
pcapita	1.27	0.78502
fadest	1.19	0.838982
frvists	1.19	0.843403
lstay	1.15	0.870238
nosites	1.13	0.883545
peak	1.13	0.887552
isource	1.11	0.899341
age	1.11	0.904493
Mean VIF	1.86	