Determinants of tourist arrivals in Africa: a panel data regression analysis

Naudé, Wim and Saayman, Andrea

2005

Tourism Economics, 2005, 11 (3), 365–391

Determinants of tourist arrivals in Africa: a panel data regression analysis

WILLEM A. NAUDÉ AND ANDREA SAAYMAN

Workwell: Research Unit for People, Policy and Performance, North-West University, Potchefstroom Campus, Private Bag X 6001, Potchefstroom 2520, South Africa. Tel: +27 18 2991440. Fax: + 27 18 2991360. E-mail: ebnwan@puk.ac.za.

Africa's tourism potential is acknowledged to be significant but underdeveloped. This paper uses both cross-section data and panel data for the period 1996–2000 to identify the determinants of tourism arrivals in 43 African countries, taking into account tourists' country of origin. The results strongly suggest that political stability, tourism infrastructure, marketing and information, and the level of development at the destination are key determinants of travel to Africa. 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 so significant in explaining the demand for Africa as a tourism destination. It is therefore recommended that attention should be given to improving the overall stability of the continent and the availability and quantity of tourism infrastructure.

Keywords: Africa; tourism; tourism arrivals; panel data regression

JEL Classification: L83, D12, O55, C33

The tourism[1] industry is the largest in the world, with receipts from international tourism expenditure totalling US$474 billion in 2002 (WTO, 2003a). Besides its ability, as a labour-intensive sector, to create jobs for relatively unskilled workers, it is also an important earner of foreign exchange.[2] For these reasons it is often promoted by less developed countries (Williams and Shaw, 1992). Tourism has the potential to contribute significantly to economic growth and development in Africa (Kester, 2003, p 203). Eilat and Einav (2003, p 1) state that tourism is 'profoundly' important for economic development through its effects on employment, exports, stimulation of infrastructure provision, generation of tax income and the promotion of world peace.

An earlier version of this paper was presented at the International Conference on Growth, Poverty Reduction and Human Development in Africa, hosted by the Centre for the Study of African Economies at St Catherine's College, University of Oxford, 21 March 2004. The authors are grateful to the conference participants, as well as to two anonymous referees, for their helpful suggestions. The usual disclaimer applies.

Africa's cultural and natural resource endowment is such that it ought to be benefiting hugely from tourism. Christie and Crompton (2001, p 1) describe Africa's potential for tourism as 'exceptional', recognizing that:

> Africa has a lot to offer that can no longer be found elsewhere. Africa holds sympathy and a certain romanticism as the continent of the explorers and as a place for adventurers. There are unique places, some of the greatest views in the world and natural attractions that few other regions can match. This is true not only for its natural resources, but also for its culture, traditions and customs.

The tourism sector is already a growing contributor to GDP and exports in more than half of all African countries[3] (Christie and Crompton, 2001). In particular, since the early 1990s there has been significant growth in tourist arrivals in Africa (Chen and Devereux, 1999, p 209). According to the World Travel and Tourism Council (WTTC), the travel and tourism industry generated US$39.8 billion of economic activity in Sub-Saharan Africa in 2003, contributed 2.4% to the region's GDP and provided 5.4% of all its employment (WTTC, 2003, p 10).

Despite its positive endowments and the good growth in tourism over the past decade, it remains true that Africa's tourism potential is underexploited and undeveloped. In 2002 Africa attracted fewer than 4% of total international tourists and received less than 2% of international tourist expenditure (WTO, 2002). More specifically, in 2001 it received about 27.7 million international tourists and total tourism receipts of US$11.7 billion. This boils down to 3.7 arrivals per 100 of population, which compares poorly with the world average of 11 and the 44 per 100 in Europe (WTO, 2002). In Sub-Saharan Africa, only South Africa is among the top 40 global tourist destinations and only 13 of the 315 'Leading Hotels of the World' are situated in Africa.

The economic dimensions of tourism to Africa, and specifically the determinants of the demand for Africa as a tourist destination, are neglected in the economic research literature. Lim (1997a) reviewed more than 70 studies on international tourism demand, none of which focused in detail or exclusively on African countries. Also, as Eilat and Einav (2003, p 5) point out, a weakness of the current international empirical literature on tourism demand is the absence of 'rigorous panel data analysis'. The lack of appropriate empirical research on tourism to Africa undoubtedly contributes to the 'limited policy guidance' to the sector noted by Christie and Crompton (2001).

So far, most research on tourism demand and the international flow of tourism has focused on explaining tourism demand and flows in developed countries, with little attention to developing countries and even less to explaining tourism in Africa. It has also focused more on the influence of the exchange rate and income on tourism receipts rather than on certain country-specific determinants of tourism arrivals.

This research attempts to fill these voids and uses panel data econometrics to explain the determinants of tourism to Africa, taking into account typical factors associated with the continent, such as political and social instability and structural and institutional weaknesses, which might affect the demand for Africa as a tourism destination. Both single-period cross-section data as well as panel data (largely obtained from the World Tourism Organization) for the period 1996–2000 are used to determine the relevance of these developing-

Table 1. Tourist arrivals in Africa (millions).

	1990	1995	2000	2001	2002	Share
World	455.9	550.4	687.3	684.1	702.6	100
Africa (total)	15.0	20.0	27.4	28.3	29.1	4.1
North Africa	8.4	7.3	10.1	10.6	10.3	1.5
West Africa	1.4	1.9	2.6	2.7	2.9	0.4
Central Africa	0.4	0.4	0.7	0.7	0.7	0.1
East Africa	2.8	4.5	5.9	6.2	6.3	0.9
Southern Africa	2.0	6.0	8.2	8.2	8.9	1.3

Source: WTO (2003b).

Table 2. Tourist arrivals and receipts for top African destinations (2002).

Country	Arrivals (thousands)	Receipts ($million)
Kenya	838	297
Mauritius	682	612
Morocco	4193	2152
South Africa	6550	2719
Tunisia	5064	1422
Egypt	4906	3764

Source: WTO (2003b).

country issues to tourism arrivals in 43 African countries, taking into account the origin of tourists.

The remainder of the paper is structured as follows. The next section provides an overview of tourism trends in Africa and discusses the determinants of and obstacles to tourism growth. The subsequent section sets out the modelling approach and describes the variables and data used. Then, the regression results are presented and the determinants of tourist arrivals in Africa are discussed. The final section draws conclusions.

Tourism in Africa

Overview

WTO figures (2003a; 2003b) indicate that tourism to Africa has grown significantly since 1990 (see Table 1), especially tourism to Southern Africa, which grew by 94% between 1990 and 2002. Still, North Africa remains the most popular regional destination, capturing 1.5% of the total international tourism market share. It is also interesting to note that tourism to Africa increased during 2001, a year in which world tourism decreased. Early 2003 results (see WTO, 2003c) indicated a growing trend in tourism to Africa, despite the SARS virus (which affected tourism to Asia and the Pacific and North America during 2003) and the Iraq conflict.

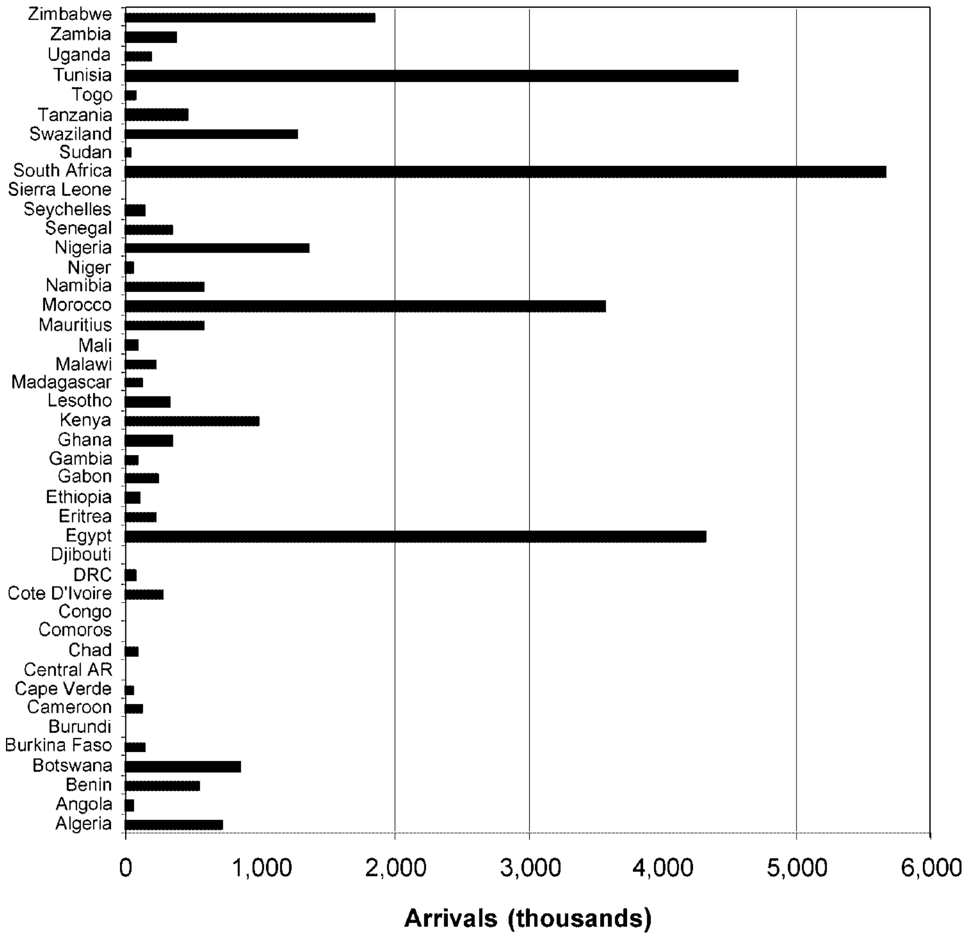

Figure 1. Total tourist arrivals, Africa (annual average, 1996–2000).
Source: WTO (2002).

An important feature of tourism to Africa is that a mere handful of countries is attracting the majority of tourists to the continent. Table 2 presents the number of tourist arrivals in 2002 and Figure 1 depicts average annual arrivals by country during 1996–2000. It can be seen that a relatively small number of countries receives the bulk of tourists to the continent.

As can be seen from Figure 1 and Table 2, South Africa, Tunisia, Egypt, Morocco and Zimbabwe were by far the most important tourism destinations in Africa during 1996–2000. Most tourists come from European countries, followed by tourists from within Africa itself. Figure 2 shows the origin of tourists to Africa between 1996 and 2000.

From Figure 2 it can be seen that between 1996 and 2000 about 44% of all international tourist arrivals in Africa were from European countries, and 41% were from Africa. Only 4% of tourists originated from the Americas, including the USA.[4]

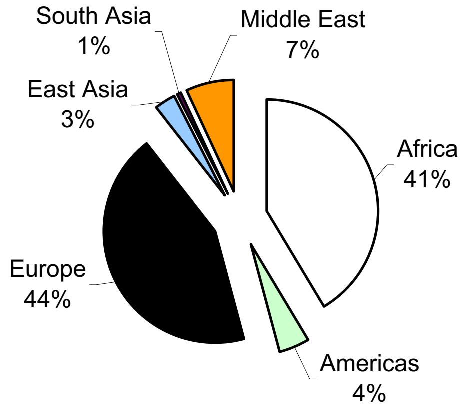

Figure 2. Tourist arrivals in Africa by country of origin (%).

Determinants and obstacles

Why do some destinations attract more visitors than others? This question has been asked by various researchers and has attracted numerous studies since the 1970s (only four attempted to provide answers to the question during the 1960s). It has been found that the responsiveness of demand for international travel varies, depending on the nationality of the tourist and the specific destination involved (see, for example, Divisekera, 2003). Thus, demand elasticity for international tourism varies by country of origin and country of destination. The demand for tourism is therefore a function of the tourist's country of origin, since cultural differences affect travel behaviour (Witt and Witt, 1995).

Lim (1997a; 1997b) summarizes some of the variables used in the analysis of tourism demand since the 1960s. As the dependent variable, tourist arrivals and/or departures is the most popular (used in 51% of studies), followed by tourist expenditure and/or receipts (49% of studies). The number of various independent variables used ranges from one to nine: the most popular, listed from most-used to least-used, in previous research have been:

- income, which affects the ability to pay for overseas travel, and the proxies used include nominal or real per capital personal, disposable or national income, or GDP and GNP (84%);
- relative prices of goods and services purchased by tourists at the destination, compared with the origin and competing destinations as measured by the CPI ratio (73%);
- transportation cost, which refers to the cost of round-trip travel between the destination and the origin country (55%);
- dynamics, often included to account for lagged effects (26%);
- the exchange rate between the currencies of the destination and origin countries (25%);
- trends, which capture secular changes in tourist tastes (25%);

- competing destinations/goods, which lead to substitution when costs associated with travel and tourism increase (15%);
- seasonal factors, often captured in dummy variables (14%);
- marketing expenditure to promote the country as a destination (7%);
- migration and ethnic factors, which capture tourists visiting friends or relatives (5%);
- business trade/travel, as measured by proxies such as trade, direct foreign investment and capital flows (5%);
- economic activity indicators, such as unemployment and income distribution (3%);
- various qualitative factors, such as tourists' attributes, household size, population in the origin country, trip motive or frequency, destination attractiveness, events at the destination (60%); and
- other factors, such as supply/capacity constraints on tourism accommodation, exchange rate reforms or foreign currency restrictions, cross price elasticity of vacation goods and the average propensity to consume tourism goods (27%).

Coshall (2000) indicates that 'there are many financial, perceptual, cultural, social and environmental factors that could be used to try and explain international tourism flows'. The research on which these statistics were compiled was based mainly on tourism demand in developed countries, with little reference to developing countries and none to African countries. Certain factors not included in previous studies but which certainly affect tourism to Africa (see Kester, 2003; Ahmed *et al*, 1998; Gauci *et al*, 2002) need to be identified.

The World Tourism Organization's 'Tourism 2020 Vision' recognizes that the tourism potential of African countries is significant, but that there are serious obstacles that need to be addressed if this potential is to be realized.

Christie and Crompton (2001, pp i–ii) argue that the greatest obstacle to Africa's tourism sector's growth is its lack of price and quality competitiveness. They point out that the worldwide tourism industry and its structures and operation consist of tour operators, travel agents and transport services that sell integrated tour 'packages' to tourists. Whether competitive tour packages can be put together for a particular destination will depend on relative prices, the safety of the destination and the quality and type of product offered. Within such a package, the airfare can have a significant impact on price – the more so for shorter trips, for which the impact of hotel costs on the overall package price is lower. According to Christie and Crompton (2001, p 9) airfares on scheduled flights in Africa are among the highest in the world.

Infrastructure and facilities in a country can negatively affect both relative prices and the quality of products. According to Kester (2003, pp 204–205), the major obstacles to tourist arrivals in Africa in this regard are insufficient air transport, a deficiency in facilities and accommodation, a lack of image and poor perceptions, poverty, disease and conflict. Gauci *et al* (2002, p 4) include among the obstacles to tourism undeveloped public health services and fears of personal safety. The latter are often caused by political instability. Eilat and Einav (2003) find that political risk has a significant impact on tourism demand in both developed and developing countries.

Cleverdon (2002, pp 10–11) points further to the difficulty of access to Africa's tourism endowments,[5] the lack of quality tourism products, weak marketing, fragmentation among tour operators and the lack of banking and communication facilities as factors constraining Africa's tourism development. As far as communication facilities are concerned, the digital divide must be seen as an important obstacle to tourism development in Africa. Only around 1.5 million of the world's more than 300 million Internet users are in Africa (with more than 60% of these in South Africa alone) (Cleverdon, 2002, p 24). The Internet plays an increasingly important role in tourism through marketing, information and online booking and electronic commerce, and has significantly increased competition between tourist destinations (Christie and Crompton, 2001, p 7).

It is also often noted that there are 'neighbourhood effects' of instability, when one country affects perceptions of the region as a whole, with 'potential tourists often unable to distinguish between individual countries' (Kester, 2003, p 204). On the other hand, tourism seems to be sensitive to good economic growth and macro-economic stability (both of which influence the price competitiveness of a country's tourism products). Tourism is thus a significant industry in all five of Africa's strongly performing economies: Morocco, Mauritius, Tunisia, South Africa and Botswana (Gauci *et al*, 2002, p 5).

Finally, there is clearly a geographical pattern to tourism flows to Africa. Northern Africa (Tunisia, Morocco and Egypt) forms a northern node and Southern Africa (South Africa and Mauritius) a southern node. West and Central Africa have the lowest numbers of arrivals on the continent (Gauci, *et al*, 2002, p 19). These are also the areas closest to the tropics, with the highest incidences of malaria and other tropical diseases. Only a few African destinations can offer 'sun and beach' holidays to international tourists, since many countries are landlocked and the cold Benguella sea current makes for unpleasant conditions on most of the south-west coast.

Modelling the determinants of tourist arrivals

Empirical economic research in tourism has focused mainly on five areas:

(1) the economic impact of domestic and/or international tourism on a local economy (see for example Archer, 1977; Kottke, 1988; Zhou *et al*, 1997; Wang, 1997; Vaughan *et al*, 2000; Saayman *et al*, 2000);
(2) the importance of tourism for development (see for example Diamond, 1976; Piga, 2003; Saayman *et al*, 2001);
(3) the economic impact of identified events (Randall and Warf, 1996; Gelan, 2003) and facilities (Chen and Hsu, 2001; Walpole and Goodwin, 2000);
(4) efforts to incorporate the explanation of tourism demand and international tourism flows (Crouch, 1995; du Preez and Witt, 2003; Coshall, 2000; Smeral and Weber, 2000; Jud and Joseph, 1974; Divisekera, 2003; Eilat and Einav, 2003); and
(5) forecasting tourism demand (Witt and Witt, 1992; Song *et al*, 2003; Song and Witt, 2000).

Econometric approaches

Various methods have been used to estimate the demand for tourism (that is, tourist arrivals) and to forecast international tourism arrivals. These methods can be divided into two broad groups: (a) those that focus on non-causal, mainly time series modelling and (b) those that focus on causal, econometric techniques (Song *et al*, 2003). Chu (2004) indicates that the main distinction between these groups is that causal models identify and measure both economic and non-economic variables affecting other variables such as price and quantity, while time series models identify stochastic components (such as autoregressive and moving average components) in each time series. With the non-causal methods, exponential smoothing and the Box–Jenkins procedure is very popular and has been used by, among others, Lim and McAleer (2001), Chu (1998), Turner *et al* (1997) and Sheldon (1993). Single-equation models are normally used and the equation is written in two functional forms – linear and log-linear regression models (Lim, 1997a).

The non-causal time series models are useful tools for tourism demand forecasting, but they have the limitation that they cannot be used for policy purposes, since they are not based on the theory that underlines the tourist's decision-making process. Therefore econometric models are superior to time series models, since they are carefully constructed based on economic theory and thus allow the researcher to assess the manner in which tourists would respond to changes in the determining factors by examining the estimated demand elasticities (Song *et al*, 2003).

However, Martin and Witt (1989) compared the Box–Jenkins approach to other modelling techniques and found that the approach including the naïve no-change model outperformed the causal models based on traditional regression techniques. Research since then has widely focused on implementing these models (see for example Chu, 1998; Lim and McAleer, 2002; Lim and McAleer, 2001; Lim, 2004) and on improving the non-causal, time series techniques by extending univariate autoregressive integrated moving average (ARIMA) models to multivariate ARIMA models (see du Preez and Witt, 2003) or using a cubic polynomial approach (see Chu, 1994).

Less work has been done on improving econometric modelling techniques, and this literature survey found only a couple of papers that could provide some insight. Divisekera (2003) uses a price-independent generalized log-linear utility function and a multivariate regression technique (the maximum likelihood method) to estimate demand. Kulendran and Witt (2001) indicate that using more up-to-date econometric techniques might rectify the problems associated with econometric models in forecasting tourism demand. They further indicate that using cointegrated techniques (such as error correction modelling) may overcome the conceptual problems associated with the least square regression approach, but that the results are still relatively inaccurate compared to the time series results.

Kim and Song (1998) and Song *et al* (2000) found that econometric models outperformed simple time series models. Song *et al* (2003) apply the autoregressive distributed lag model (ADLM) to capture the dynamics of economic activities. Another paper by the same authors (2003a) tests six econometric models that are all special cases of the ADLM model, namely a long-run cointegration regression, two error correction models, a reduced ADLM, an unrestricted vector autoregressive (VAR) model and a TVP model. The results indicate that the TVP model generate the most accurate forecasts, followed by the static regression model. The only paper that uses panel data analysis is that by Eilat and Einav (2003) in which the pooled logit regression technique is used to identify the determinants of tourism to different regions.

Estimating equation

As discussed above, the bulk of empirical research in tourism demand modelling has made use of time series approaches. In this paper we use cross-section and panel data given that cross-section data tends to give better estimates of long-run relationships, whereas time series tend to estimate short-run relationships (see Kennedy, 2003, p 308). Given the challenges facing Africa and the need for sound policy advice for tourism promotion, it seems more appropriate to identify the long-run determinants of tourist arrivals. The use of fixed effects estimators, however, will allow us to pick up short-term effects since they focus on time series components of the data. The methodology that we follow in using cross-section and panel data is driven by a desire to address traditional econometric problems in cross-country regressions such as unobserved country effects, outliers, endogeneity, dynamics and model uncertainty. Thus various estimators and specifications are used: namely, OLS, LAD, GLS (random-effects) and GMM. This section describes the methodology in greater detail.

First, the estimating equation is specified. The demand by an international tourist to travel to a particular country or destination can be derived from microeconomic principles and can be shown to depend positively on income (expenditure) in the country of origin and negatively on relative tourism product prices, which are affected by travel costs (Lim, 1999, p 274). We broadly adopt Eilat and Einav's (2003) approach by modelling the flow of tourists to African countries as based on the demand system for differentiated products. Thus different African countries can be treated as supplying different tourist products.

A strongly separable utility function is assumed, in which the individual consumer (tourist) follows a two-stage utility maximization procedure. In the first stage, the consumer decides how much expenditure to allocate between various consumption goods, one of which is tourism to Africa. Once this decision has been made, the second stage of utility maximization consists of allocation of total African tourism expenditure among African countries (destinations).

This two-level utility function can be written as follows (see Nordström, 2002, p 3):

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

Where $U^j(q_j)$ is the subutility function which will consist of African tourism consumed in the j^{th} country (out of m-number of African destinations to choose from).

A feature of tourism as a consumption good is that it is characterized by taste formation. Past consumption of tourism in a particular country/destination will therefore affect present consumption. Following Nordström (2002), this taste formation can be incorporated by including past consumption of tourism in the utility function and denoting q_{jt} to be the consumption of tourism 'produced' in country j in period t. In the second stage of the utility maximization procedure, subutility is maximized conditional on expenditures allocated to African tourism goods in the first stage, and denoted e_t. More formally, the demand function for tourism from a particular African country is obtained by maximizing

$$U_i(q_t \mid q_{t-1}) = \prod_{j-1}^{m} (q_{jt} - k_{jt})^{\delta_j}, \sum_{j-1}^{m} \delta_j = 1 \quad (2)$$

Subject to the budget constraint that

$$\sum_j p_{jt} q_{jt} = e_t \quad (3)$$

Where p_j = the price of tourism produced in African country j and $k_{jt} = y_{jt} + \gamma_j q_{jt-1}$. Here y_{jt} is positive and represents the minimum consumption requirement in period t, and $\gamma_j q_{jt-1}$ is the consumption based on past consumption.

The resulting individual demand function can be written as

$$q_{jt} = y_{jt} + \gamma_j q_{jt-1} + \Sigma p_j y_j + \frac{\delta_j}{p_j} (e_t - \Sigma p_j \gamma_j q_{jt-1}) + \varepsilon_{jt} \quad (4)$$

The market demand function will have the same broad functional form with q and e representing total demand. A distinction is made for differences in demand for African tourism from different countries of origin. The demand for tourism in Equation (4) above can be simplified and written as follows (see also Lim, 1997b):

$$q_{ijt} = x_j \beta + c_j + u_j \quad (5)$$

where:

- q_{ijt} = demand for international tourism by origin i for African destination j;
- x_j = vector of explanatory variables that will include past tourism consumption, expenditures, e_t, (proxied by incomes) and relative prices (p_j) that influence the travel decisions of tourists when choosing between destinations. The prices will include transport cost between destination j and origin i, the cost of living in destination j (often the ratio of prices in destination j to prices in origin i), and the relative price of tourism products between African countries j and k. Finally, qualitative and other factors in destination j that will influence the demand for that country's tourism, such as health and personal safety as well as image and marketing efforts, are included in light of the earlier discussion on the obstacles to tourism growth in Africa.

Equation (5) can be estimated using Ordinary Least Squares (OLS). However, using OLS and a single cross-section of data has significant limitations (this is discussed in more detail later). It is therefore preferable to use panel data and employ more appropriate estimation methods. Equation (5) can be written in the following manner to illustrate the different estimation options when a panel of data (as opposed to a single cross-section) is available (showing that panel data models have complex error structures):

$$q_{jt} = x_{jt}\beta + c_j + u_{jt} \tag{6}$$

For $j = 1,\ldots.m$ and $t = 2,\ldots.T$ and where we use for q_{jt} = tourist arrivals in country j in period t, x_{it} = a $1\times K$ vector of the explanatory variables – for example, as enumerated above. Note that in the case of panel data these can vary over t and j; c_j = unobserved heterogeneity (country individual effects) with variance σ_c^2. This can be viewed as unobserved country characteristics, for example due to natural environments, unique attractions (such as the pyramids) and/or climate, which are constant over the time period and influence q_{jt}. u_{jt} = an idiosyncratic error term with variance σ_u^2. with the usual properties.

From Equation (6) the so-called 'between' estimator[6] is OLS applied to the following equation:

$$\bar{q}_j = \alpha + \bar{x}_j\beta + c_j + u_j \tag{7}$$

where

$$\bar{q}_j = T^{-1} \sum_{t=1}^{T} q_{jt}.$$

It should be noted that the 'between' estimator is not consistent because $E(x_j c_j) \neq 0$.

The fixed effects (or 'within') estimator[7] is obtained by using OLS to estimate:

$$(q_{jt} - \bar{q}_j) = (x_{jt} - \bar{x}_j)\beta + (u_{jt} - u_j) \tag{8}$$

The random effects estimator is a weighted average of the estimates produced by the between estimator (7) and the within estimator (8):

$$(q_{jt} - \theta\, \bar{q}_j) = (1 - \theta)\alpha + (x_{jt} - \theta x_{jt})\beta + \{(1 - \theta)c_j + (u_{jt} - \theta\, \bar{u}_j) \tag{9}$$

where

$$\hat{\theta}_j = 1 - \sqrt{\frac{\hat{\sigma}_u^2}{T_i\hat{\sigma}_c^2 + \hat{\sigma}_u^2}}$$

In the next section the variables and data used in the econometric estimation are discussed.

Data and variables

At the outset it must be recognized that the estimation of tourist demand functions in the case of Africa (but also elsewhere) is beset with problems of data availability. If a panel data approach is to be followed, one necessarily has to make use of proxies to model the effect of time-varying factors such as tourism and travel prices on tourism demand.

As a dependent variable (q), the total number of tourist arrivals per year to a particular destination, is used to measure the demand for tourism to Africa. A distinction is made between tourist arrivals from the European Union (EU), the Americas (mainly the USA and Canada) and from within Africa itself. This is done to capture the differences in consumption patterns between various origin countries and thus to determine the extent to which the African tourist market is segmented by types of clientele (see Nordström, 2002, p 2). In the dynamic panel estimation the lagged values of (q) are included in order to capture the quality of the experience of the tourist to a particular destination, which will also serve as an indicator of how suitable the tourism products in that country are for the particular market segment.

The key independent variables in Equation (5) are total tourism expenditures (e_t) and relative tourism prices. Real GDP per capita in countries of origin (EU, the Americas and Africa) is used as proxy for total expenditure on tourism, following Nordström (2002, p 2).

As far as relative prices are concerned, it is common in tourism demand studies to use the CPI of a destination country adjusted by the $-exchange rate as a proxy for relative tourism prices. The inverse of this shows how many 'baskets' of goods a tourist has to give up in his or her home country to buy a basket of goods in the destination country. This measure of relative prices captures changes in the real exchange rate over time as well as cross-sectional variation in the cost of travel (Eilat and Einav, 2003, p 12).

Because of the importance of transport/travel costs in the overall 'tour packages', we proxy travel costs by the distance of a country to the origin of its tourists. This distance variable takes the location of a country to be at its capital. It captures the cross-sectional variation in transport costs. However, a weakness of using distance as a proxy for travel costs is that it does not measure changes in travel costs over time. For this, time dummies are included in all the specifications, following Eilat and Einav (2003, p 13).

From the above discussion of factors that may determine tourism to Africa, political stability, personal safety, health risks and available infrastructure and tourism marketing efforts were identified as potentially important. It also has to be recognized that geography can be important, not only in determining the disease burden (for example, by determining the areas of malaria prevalence) but also in determining whether or not a country can offer beach holidays as an attraction (Africa has many landlocked countries). Variables measuring each of these factors were consequently included in the regression analyses. In most instances, finding a suitable proxy (or actual variable) is straightforward. In the case of tourism marketing, however, the number of Internet users in a country is used as a proxy to capture the effects of networks and information on tourist flows. In the case of tourism infrastructure, the relative number of hotel rooms available was used as an indication. Hotel rooms in a country are an indicator of the capacity of the tourism sector: the more rooms, the higher the capacity and therefore the more competitive the country's tourism sector (and it will be able to offer cheaper prices, *cæteris paribus*). Furthermore, hotel accommodation size is needed for a destination to reach so-called 'critical mass' (Christie and Crompton, 2001, p 26). For instance, a certain volume of hotel rooms may be necessary to convince airlines to establish routes or to justify investment in complementary infrastructure such as roads. Thus through its critical mass function we expect the relative number of hotel rooms in a country to be a good indicator of the tourism competitiveness of the destination.

To proxy for health risk, the prevalence of malaria was included as an explanatory variable. Malaria has been identified as a health risk that lowers tourism[8] (Gallup and Sachs, 2000, p 10). Malaria also limits internal movements, the development of transport systems and the formation of skills – all vital for a growing tourism industry. It lowers entrepreneurial productivity and thus limits the extent to which Africa can benefit from its substantial tourism potential. The incidence or extent of malaria can be measured using the malaria index complied by Gallup *et al* (1999) from World Health Organization (WHO) data. This index is the product of land area subject to malaria times the fraction of falciparum malaria cases in 1994. The disease burden is also proxied using the number of frost days experienced on average over a year in a country. Higher numbers of frost days have been argued to be correlated with less malaria and fewer diseases in agriculture – and also with better climatic conditions for open-air tourist attractions.

To estimate Equations (5) and (6), annual data were obtained from the WTO (2002), the *World Development Indicators Online* of the World Bank, and the Penn World Tables (PWT), version 6.1 (October 2002). Most time-variant data covers the period 1996 to 2000. A panel was compiled on 43 African countries.[9] Countries such as São Tomé, Libya, Mauritania and Guinea-Bissau were omitted due to lack of sufficient data on tourist arrivals.

The variables used in Equations (5) and (6) are summarized in Table 3, which shows that in the present sample of African countries the average annual number of tourist arrivals during 1996–2000 was 723,000, most of which came from the EU or from within Africa itself. On average, only about 30,000 tourists per year visited the typical African country over the period. It can also be seen that there are on average about 12,000 hotel rooms available in the typical African destination and that the average price per person per night is around US$131. It is noticeable from the summary statistics in the table that there is substantial variation in these variables among African countries. Note, for instance, the large standard deviations in the numbers of Internet users, hotel rooms, hotel room prices, and the number of tourist arrivals.

Caveats

Before reporting on the results from the various regression models, it is necessary to point to some shortcomings in the data, variables and definitions used in this study.

First, the aggregation of tourist arrivals, without consideration of the purpose of travel, can obscure important aspects of the decision to travel to Africa. It may also affect the sizes of the elasticities obtained in the following sections. WTO identifies five reasons for travel: leisure tourism, visiting friends and relatives, business and professional travel, travel for religious purposes, and travel for other purposes. Aggregating these into total tourist arrivals may obscure the possibility that business travellers may be less sensitive to price changes than, for instance, leisure tourists. Unfortunately, sufficiently detailed data on the purposes of travel to African countries are not available.

Second, in the present study annual data are used. This may obscure potentially important and interesting seasonal effects (Nordström, 2002, p 2). High frequency data on tourism in Africa is, however, lacking.

Third, the problem of weak tourism data is particularly acute in Africa. The Tourism Satellite Account (TSA), developed by the UN in 1993, has not yet been implemented successfully in Africa. There were attempts to establish a

Table 3. Summary of core variables used in regression analyses.

Variable	N	Mean	Standard deviation	Min	Max	Data source
Number of Internet users	43	48532	209970	320	1380800	World Development Indicators
Index of political stability/ lack of violence	34	–0.52	0.87	–2.5	1.14	Kaufmann *et al* (1999)
Number of frost days on average per year	39	1.91	3.37	0	17.47	Gallup *et al* (1999)
Air distance (km)	36	5930	2049	1675	9590	Gallup *et al* (1999)
Number of telephone lines per employee	35	46	31	8	132	World Development Indicators
Prevalence of malaria in 1994	43	0.71	0.41	0	1	Gallup *et al* (1999)
Total number of tourist arrivals (thousands), 5-year average	43	723	1316	13	5656	WTO (2002)
Tourist arrivals from African countries (thousands), 5-year average	40	268	676	1	4083	WTO (2002)
Tourist arrivals from the Americas (thousands), 5-year average	38	30	58	1	270	WTO (2002)
Tourist arrivals from the EU (thousands), 5-year average	41	275	660	3	3091	WTO (2002)
Number of hotel rooms available	37	11883	23043	227	90586	WTO (2002)
Death rate	43	14.4	4.9	5.6	25.3	World Development Indicators
GDP per capita	43	1071	1471	106	7029	World Development Indicators
Life expectancy	43	51	9.5	37.2	71.9	World Development Indicators
Urbanization rate	42	38	17	8.4	84	World Development Indicators
Average hotel room price in 2002 (US$)	26	131	33	78	209	WTTC (2003)
Adjusted CPI	43	31	11	16	85	Penn World Tables 6.1 (2003) (Heston *et al*, 2002), see http://pwt.econ.upenn.edu)

TSA in South Africa in 1998 and in Namibia more recently, but there is still no annual TSA in these countries. This limits any analyses on the effects or impacts of tourism on economic growth and development in Africa.

Another problem was encountered in gathering data on the cost of travel. Since travelling cost is often used as a proxy for the cost of tourism, it is a key determinant. Historical data concerning travelling costs in Africa are not easily obtainable and the proxy used in this research is therefore the distance between destinations. As already mentioned, this proxy has the limitation that it remains constant over time.

Results

In this section the regression results, using OLS, random effects/fixed effects and first-step GMM, are set out and compared to identify the determinants of tourist arrivals in Africa. Four dependent variables are used: total tourist arrivals, arrivals from the Americas, arrivals from Europe, and arrivals from Africa. This allows us to identify whether or not tourists from different countries of origin differ in their determinants.

Cross-section regression results

The limitations of using a single-equation OLS cross-sectional regression model are well-known (see, for example, Kennedy, 2003). The most serious limitations for the present purpose are that simple cross-section may produce biased and inconsistent estimates since it may not take into consideration the endogeneity of some of the regressors, it ignores dynamics, it throws away information (Attanasio *et al*, 2000) and it may suffer from omitted variable bias, causing different intercepts for each country (see, for example, Naudé and Krugell, 2003). Eilat and Einav (2003, p 3) also state that use of cross-sectional data is not theoretically appealing since the most important factors of 'production' in tourism tend to be unique (such as the pyramids), so that it may be more important to investigate the effect of variables that vary over time. To overcome these shortcomings, panel data techniques are advised.

However, this paper still reports, at least for comparative purposes and to get a broad industry overview, standard OLS estimates on a single period (averages taken over the period 1996 to 2000) cross-section. The dangers can be limited by choosing the independent variables in such a way as to minimize simultaneity and interdependence (the two major sources of endogeneity) and to correct the standard errors of the OLS regression by the White procedure.[10] This procedure adjusts for the presence of heteroskedasticity in the data.

The cross-section OLS regression results, using STATA 8.0, are reported in Table 4. The results indicate that political stability, Internet usage, urbanization rate and whether the country is landlocked are all significant determinants of total tourism to Africa (the first two at the 1% level and the last two at the 5% level). Political stability is especially relevant for tourists from America and less significant for those from Europe. Greater stability coincides with more tourist arrivals and this result confirms the finding of Eilat and Einav (2003). The results show that tourists from Africa are perhaps less sensitive towards political risk, which may be attributed to a better knowledge of political risk in the destination country.

Internet usage is significant for tourists from all countries, but more so for international tourists (not from Africa). While Internet usage is an indication of communication structures, and development, the positive relationship may also indicate the important role that the Internet plays in marketing the

Table 4. OLS cross-section regression results (dependent variables: number of tourist arrivals, 1996–2002, five-year averages).

	Total tourist arrivals	**Arrivals from the Americas**	**Arrivals from Europe**	**Arrivals from Africa**
Constant	–2.92	0.03	3.06	–0.05
	(–0.42)	(0.00)	(0.35)	(–0.00)
Adjusted CPI (relative tourism prices)	–0.78	–1.50	–0.75	–0.30
	(–1.41)	(–1.92)*	(–1.40)	(–0.41)
Hotel capacity	0.17	0.51	0.48	0.63
	(0.78)	(1.51)	(1.59)	(1.50)
Malaria	–0.09	0.13	0.26	0.31
	(–0.55)	(0.51)	(1.37)	(1.06)
Air distance (air travel costs)	0.99	0.17	0.35	0.11
	(1.10)	(0.18)	(0.56)	(0.06)
Political stability (lack of violence)	0.75	0.57	0.40	0.68
	(4.26)***	(2.46)**	(1.95)*	(1.79)
Internet usage	0.42	0.45	0.42	0.40
	(3.48)***	(2.46)**	(2.81)**	(1.83)*
Urbanization rate	0.85	0.94	0.82	0.63
	(2.23)**	(1.98)*	(2.09)**	(1.40)
Death rate	–1.25	0.73	–1.11	0.81
	(–1.59)	(0.71)	(–1.18)	(0.55)
Landlocked (dummy)	0.93	0.39	0.44	0.81
	(3.00)**	(1.10)	(1.52)	(1.80)*
Border with South Africa (dummy)	–	–	–	1.9
				(2.10)*
Adjusted R^2	*0.86*	*0.67*	*0.77*	*0.78*
N	*21*	*21*	*21*	*21*

***Significant at the 1%, **at the 5%,* and at the 10% level. The quantities in parentheses are the heteroskedasticity-robust t-values.

destination. Internet bookings and information on the Internet make a destination more easily accessible. The sign and size of the coefficient corresponds well with the marketing variables analysed by Crouch (1995).

The urbanization rate is usually an indicator of development in a country and the results in Table 4 suggest that a higher urbanization rate is consistent with more tourist arrivals, especially from Europe and America. This may again indicate that European and American tourists prefer more developed African destinations. It may also reflect the fact that the tourism industry itself is service-intensive and benefits from agglomeration economies in urban environments.

The sign and significance of the landlocked dummy is interesting, since it is especially significant for tourists from Africa. Interpreting this result increases our understanding of why people travel to Africa. Sun, sea and beaches seem less important to European and American tourists, on average (in the case of certain countries, such as Mauritius, Tunisia and Morocco, it may be important). For African tourists, sun, sea and beach holidays are not important, as the statistically significant positive relation between being landlocked and attracting tourists from Africa suggests.

The 'health' variables, malaria and the death rate, did not show any significant relationships. In more recent years the incidence of HIV/AIDS in Africa has also become a risk factor for tourists, but arguably less so than, say, malaria. The lack of reliable data is, however, a shortcoming in modelling the possible impact of HIV/AIDS on tourism arrivals in Africa.

The distance variable, the proxy for travel cost, is also insignificant for tourists from all source countries. The adjusted CPI is significant only for tourists from America, which indicates that American tourists are more price-sensitive than tourists from other countries. Although insignificant, the sign and size of the coefficient indicates an almost unitary elasticity. This again corresponds with the results of Crouch (1995), who indicates that the price elasticity often falls within the range of unitary, and Eilat and Einav (2003), who found that travelling to less developed destination was less price-sensitive. The second dummy, a border with South Africa, shows a significant positive result. This indicates that there are higher tourism flows between South Africa and its neighbouring countries, reflecting the spillover effects from South Africa's tourism industry into the subcontinent (South Africa has the largest tourism industry in Africa). For many neighbouring countries, South Africa is an important shopping destination. Many South Africans also work in neighbouring states (everyone has family in Namibia/Botswana/Zimbabwe/Mozambique) and many people from neighbouring states work in South Africa, all rendering trips to visit family and friends a major type of tourism.

Given the potential problem of outliers (particularly South Africa, Tunisia, and Egypt in the present case), the above regressions were repeated using the semi-parametric technique of median regression. The least absolute deviation (LAD) estimator is not sensitive to outliers on the dependent variable, as it results from a regression that minimizes the absolute deviation around the median of the distribution of the dependent variable. The results from the LAD estimation are presented in Table 5.

When the effect of outliers is removed, the most significant changes are in the determinants of African tourists. The Americas is still the most price-sensitive region, but political stability and even Internet usage become less significant there. Only Internet usage becomes a significant indicator for tourists from Europe. What is particularly interesting, however, is the huge difference in the demand of African tourists. Political stability becomes highly significant, as do Internet usage and the landlocked and the border with South Africa dummies. The high significance of the malaria dummy is also difficult to interpret. Do Africans travel more to places where there is a high incidence of malaria (as suggested by the results)? Certainly, the limitations of cross-sectional analysis described earlier are present in these results, which lead to the use of panel data to improve the reliability of the findings.

Static panel data regression results

The use of panel data allows us not only to investigate dynamic relations but also to control for unobserved cross-section heterogeneity. With panel data, the issue is whether to use a random effects or a fixed effects estimation approach. The random effects approach to estimating β exploits the correlation in the composite error in Equation (6), $v_{jt} = c_{jt} + u_{jt}$. The approach puts c_i in the error

Table 5. LAD cross-section regression results (dependent variables: number of tourist arrivals, 1996-2002, five-year averages).

	Total tourist arrivals	Arrivals from the Americas	Arrivals from Europe	Arrivals from Africa
Constant	–4.34	–8.75	5.9	–5.6
	(–0.20)	(–0.63)	(0.47)	(–0.76)
Adjusted CPI (relative tourism prices)	–1.07	–2.35	–0.98	–0.13
	(0.86)	(–2.24)**	(–1.41)	(–0.33)
Hotel capacity	–0.04	0.42	0.43	0.59
	(–0.05)	(0.67)	(0.92)	(2.89)**
Malaria	0.06	0.32	0.36	0.38
	(0.12)	(0.83)	(1.05)	(3.49)***
Air distance (air travel costs)	1.45	1.31	0.52	1.11
	(0.62)	(0.81)	(1.23)	(1.30)
Political stability (lack of violence)	0.77	0.61	0.44	0.32
	(1.18)	(1.42)	(1.27)	(4.32)***
Internet usage	0.54	0.55	0.47	0.39
	(1.28)	(1.88)*	(1.87)*	(4.32)***
Urbanization rate	0.82	1.40	0.68	0.33
	(1.40)	(1.77)*	(0.91)	(1.14)
Death rate	–2.60	0.29	–2.40	–0.25
	(–1.05)	(0.16)	(–1.68)	(–0.36)
Landlocked (dummy)	0.96	0.32	0.51	0.88
	(1.01)	(0.50)	(1.18)	(3.40)***
Border with South Africa (dummy)	–	–	–	1.80
				(3.85)***
Pseudo R^2	*0.75*	*0.59*	*0.69*	*0.75*
N	*21*	*21*	*21*	*21*

***Significant at the 1% level; **Significant at the 5% level;*Significant at the 10% level. The quantities in parentheses are the t-values.

term assuming that $\boldsymbol{c}_i$ is orthogonal to $\boldsymbol{x}_{jt}$ and uses a Generalized Least Squares (GLS) estimator to take into account serial correlation in the composite error v_{jt}.

There can, however, be many instances in which this assumption may be violated. Specifically, c_j can be correlated with x_{jt} in the present model if the c_j influences the price, infrastructure and income variables. In such a case, the fixed effects estimator may be more appropriate. Wooldridge (2001, p 266) shows that a fixed effects estimator is more robust than a random effects estimator. A shortcoming of the approach, however, is that time-constant elements, such as geographical factors and summary measures of political stability, cannot be included in x_{jt} – otherwise there would be no way to distinguish the effects of these variables from the effects of the unobservable c_j. Another shortcoming of the fixed effects estimator is that it is less efficient than the random effects estimator – it has less degree of freedom and takes into calculation only the variation 'within' units, not between units.

Accordingly, to determine which estimator is more appropriate in the present case, both fixed effects (FE) and random effects (RE) estimators were initially

Table 6. FE versus RE estimator: diagnostic results.

Dependent variable (model)	Breusch–Pagan LM test	Hausman specification test
Total tourist arrivals	$\chi^2(1) = 4.50$*	$\chi^2(6) = -21.46$
Arrivals from the Americas	$\chi^2(1) = 0.02$	$\chi^2(6) = 1.42$
Arrivals from Europe	$\chi^2(1) = 0.66$	$\chi^2(6) = 3.98$
Arrivals from Africa	$\chi^2(1) = 5.58$*	$\chi^2(6) = 1.60$

*Null hypothesis rejected.

Table 7. RE static panel data regression results.

	Total tourist arrivals	Arrivals from the Americas	Arrivals from Europe	Arrivals from Africa
Constant	–17.7	65.8	63.84	–373.7
	(–0.33)	(0.92)	(1.18)	(–0.51)
Adjusted CPI (relative tourism prices)	0.09	–0.89	0.20	0.93
	(0.19)	(–1.21)	(0.32)	(1.26)
Hotel capacity	0.43	0.98	0.59	1.06
	(1.63)*	(3.00)***	(1.85)*	(2.85)*
Malaria	–0.57	–0.1	0.07	0.16
	(–2.15)**	(–0.36)	(0.26)	(0.40)
Air distance (air travel costs)	0.80	0.02	0.38	–0.08
	(0.84)	(0.02)	(0.43)	(–0.07)
Political stability (lack of violence)	0.86	0.89	0.22	0.56
	(2.34)**	(2.53)**	(0.62)	(1.25)
Internet usage	–0.05	0.11	0.21	–0.00
	(–0.51)	(0.60)	(1.41)	(–0.00)
Urbanization rate	0.83	0.83	0.56	0.11
	(1.59)	(1.58)*	(1.06)	(0.17)
Death rate	0.31	0.15	0.07	0.35
	(0.60)	(0.17)	(0.10)	(0.42)
Landlocked (dummy)	0.42	1.05	0.14	0.81
	(0.87)	(2.20)**	(0.29)	(1.40)
Income in origin (GDP per capita)	1.76	–5.7	–6.32	60.1
	(0.27)	(–0.78)	(–1.11)	(0.51)
Border with South Africa (dummy)	–	–	–	2.33
				(2.46)**
Overall R^2	*0.76*	*0.79*	*0.78*	*0.87*
Within R^2	*0.32*	*0.27*	*0.40*	*0.52*
Number of observations	*30*	*30*	*30*	*30*
Number of groups	*18*	*18*	*18*	*18*

***Significant at the 1% level; **Significant at the 5% level;*Significant at the 10% level. The z-values are shown in parentheses in the case of the RE-estimation.

used to estimate Equation (6), and the Hausman specification test was performed to evaluate the assumption in the RE model that c_j is orthogonal to x_{jt}. In addition the Breusch–Pagan Lagrange Multiplier (LM) test was used to test whether the variance of the intercept components of the composite error term were zero. Rejection of the null in both these cases would lead to rejection of the RE estimator. The results of the Hausman Specification and Breusch–Pagan LM tests are summarized in Table 6. As can been seen from the table, the Hausman Specification test cannot reject the null hypothesis that the difference in coefficients between the FE and RE estimators is not systematic. These findings suggest that the RE estimator can be used without fear of producing biased estimates. The results are shown in Table 7.

Table 7 shows that total tourist arrivals in Africa are sensitive to malaria, political stability and hotel capacity. This time the sign of the coefficient is more than would be expected (that is, negative), indicating the risk of malaria as a deterrent to tourism. The variable 'hotel capacity', the proxy used for tourism infrastructure, becomes more significant and indicates that increased hotel capacity is associated with more tourism arrivals. The determinants of tourism to Africa for American tourists seem to be the most complex, with tourism infrastructure (hotel capacity), political stability, the landlocked dummy and the urbanization rate all significant at various levels. Clearly, the American tourist wants an 'African safari, with the royalties that come with higher levels of development'. Note that the 'border with South Africa' dummy is again significant at a 5% level for African tourists.

Dynamic panel data regression results

Despite the strengths of fixed and random effects estimators based on panel data, two further shortcomings remain to be addressed. These are the potential endogeneity of the x_j and the loss of dynamic information. If there are persistence/reputation effects that apply over time in the tourist's decision on holiday destinations – for example, in the decision to return to a particular destination following a good experience there – this might constitute a serious omission.

The incorporation of dynamics into the model allows Equations (5) and (6) to be rewritten as an AR(1) model as follows:

$$q_{jt} - q_{jt-1} = \alpha_t + \phi q_{jt-1} + x'_{jt}\beta + c_j + u_{jt} \tag{7}$$

Where $q_{jt} - q_{jt-1}$ is the log difference in tourist arrivals over a period; q_{jt} is the log of tourist arrivals at the start of that period; x_{jt} is the vector of explanatory variables as described above; α_t represents period-specific intercept terms to capture changes common to all countries; c_j is the unobserved country-specific and time-invariant effects (unobserved fixed effects); and u_{jt} is the time-variant idiosyncratic error term.

Equivalently, Equation (7) can be written as:

$$q_{jt} = \alpha_t + (\phi + 1)q_{jt-1} + x'_{jt}\beta + c_j + u_{jt} \tag{8}$$

Writing (8) in first differences eliminates the time-invariant components, c_j. This solves the problem of omitted variable bias:

$$\Delta q_{jt} = \alpha_t + (\phi + 1)\Delta y_{jt-1} + \Delta x'_{jt}\beta + \Delta u_{jt} \tag{9}$$

However, it creates another problem, that of endogeneity, since it is clear that q_{jt-1} is endogenous to the error terms through u_{jt-1}. It will therefore be inappropriate to estimate (9) by OLS. To overcome the endogeneity problem, an instrumental variable needs to be used for Δq_{jt-1}. Two approaches, Anderson and Hsiao's (1982) instrumental variable (IV) and Arellano and Bond's (1991) two GMM estimators (first-step and second-step, respectively), have been used in this regard. Anderson and Hsaio (1982) proposed using Δq_{jt-2} or y_{it-2} as instruments. Arellano and Bond (1991) show that using the lagged level q_{jt-2} as instrument is superior and that in fact the list of instruments can be extended[11] to include further $q_{jt-3}, q_{jt-4}, \ldots q_{jt-k}$. Moreover, the Anderson–Hsaio IV approach can be seen as a special case of two GMM estimators proposed by Arellano and Bond (1991) to combine the list of instruments efficiently. These two GMM estimators are preferable as they gain efficiency by using additional moment restrictions.[12]

The results from estimating Equation (9) using the Arellano–Bond (1991) first-step[13] GMM estimator are presented in Table 8. The results indicate that total tourist arrivals in African destinations are sensitive to tourist infrastructure and political stability at a 5% level. The lagged arrival variable is also significant, but the sign of the coefficient is negative, suggesting that African destinations to not generate repeat visits. This may reflect negative experiences of tourists, perhaps due to substandard service, the incidence of violence or high transaction costs ('hassle'). The American tourist is more concerned about the available infrastructure and political stability, while tourists from Europe consider the distance of travel (and associated travel cost), infrastructure and malaria to be important factors when travelling to Africa. The sign of the coefficient of political stability is difficult to explain, since it translates into more instability leading to an increase in tourism from the Americas. This is a result that cannot yet be explained, and needs further exploration.

Other development indicators and whether or not the country is landlocked are insignificant indicators. The insignificance of income in the country of origin is contrary to expectations, since studies such as those completed by Crouch (1995) and Einat and Einav (2003) find a strong positive relation between tourism and income, even for less developed countries. As expected, tourism to Africa is not very sensitive to price changes, but the relationship between lagged income and arrivals from Europe is interesting. It thus seems as if tourism to Africa is an inferior good for tourists from Europe, since the income elasticity measure is negative and significant. The picture differs significantly for tourists from Africa, since there is a positive relationship between income and arrivals, as well as between previous arrivals and current arrivals. For Africans, tourism in Africa is therefore a luxury good and the experience is in general positive, leading to repeat behaviour.

The constant (or intercept) is also interesting, since it is positive for tourists from Europe but negative for tourists from the Americas. This could be an indicator of the view of Africa and tourism to Africa in general, with Europeans having a more positive image of Africa and Americans a more negative image.

Conclusions

The aim of this paper has been to explain the determinants of tourism to Africa, taking into account typical factors present within the continent, such as political and social instability and structural and institutional weaknesses. The

Table 8. Dynamic panel data regression results (first-step GMM estimator).

	Total tourist arrivals	Arrivals from the Americas	Arrivals from Europe	Arrivals from Africa
Constant	–0.98	–11.1	4.23	–2.58
	(–1.08)	(–2.14)**	(2.35)*	(–1.19)
Arrivals (lagged)	–0.68	0.26	–0.49	0.92
	(–1.65)*	(0.70)	(–1.03)	(2.26)**
Urbanization rate	2.66	10.5	–1.43	7.5
	(0.66)	(0.95)	(–0.22)	(0.84)
Internet users	–0.00	0.07	0.15	0.19
	(–0.10)	(0.34)	(1.43)	(1.24)
Adjusted CPI	–1.00	–0.09	0.43	–0.22
	(–0.32)	(–0.10)	(0.83)	(–0.32)
Adjusted CPI (lag)	–0.38	0.49	0.05	–0.22
	(–1.20)	(0.58)	(0.10)	(–0.29)
Hotel capacity	0.39	1.11	0.74	–0.31
	(2.03)**	(2.12)**	(1.77)*	(–0.64)
Hotel capacity (lag)	–0.09	0.59	0.15	–0.76
	(–0.46)	(0.93)	(0.35)	(–1.27)
Income in origin	–10.03	172.02	–37.2	22.88
	(–1.61)	(0.77)	(–1.34)	(0.67)
Income in origin (lag)	2.82	106.5	–45.14	55.5
	(0.29)	(0.86)	(–1.99)**	(1.60)*
Malaria	0.005	–0.06	–0.07	–0.05
	(0.19)	(–0.86)	(–1.76)*	(–0.74)
Air distance	0.14	0.30	–0.25	0.24
	(1.33)	(1.01)	(–1.61)*	(0.97)
Political stability	0.18	–0.28	0.05	–0.08
	(2.21)**	(–1.70)*	(0.55)	(–0.56)
Landlocked	–0.07	0.03	0.02	0.17
	(–1.35)	(0.19)	(0.26)	(1.08)
SA border	–	–	–	–0.1
				(–0.36)
Diagnostics				
Sargan test of over-identifying restrictions	*0.13*	*0.04***	*0.65*	*0.75*
Arellano–Bond test of first-order autocorrelation	*0.48*	*0.48*	*0.31*	*0.14*
Arellano-Bond test of second-order autocorrelation	*0.12*	*0.43*	*0.34*	*0.92*
Number of observations	*40*	*40*	*40*	*40*
Number of groups	*16*	*16*	*16*	*16*

***Significant at the 1% level; **Significant at the 5% level;*Significant at the 10% level. The heteroskedastic-robust z-values are shown in parentheses. The p-values are reported.

key variables, as identified by authors such as Lim (1997a; 1997b), were also taken into account to determine the significance of these variables for tourism to Africa.

The variables used were grouped into seven categories – income, relative prices, air travel cost, infrastructure and marketing, political stability and personal safety, geography and health. In all of the variables, there was substantial variation for different African countries. Some of the main problems encountered centred on the issue of obtaining data in Africa and proxies had to be used to obtain a complete data set. As suggested by Crouch (1995), the countries of origin were also split to determine whether there were any significant variation in the determinants of arrivals from different continents.

The methodology employed included single-equation regressions using OLS and first-step GMM. Because of the limitations of cross-sectional data, panel data techniques were also employed to estimate the determinants of tourist arrivals in Africa. Both static panel data regressions, using a generalized least squares method (GLS) and dynamic panel data regressions, using the Arellano–Bond first-step GMM estimator, were applied.

The results show that different determinants are important to different markets. The cross-sectional data results indicate that political stability seems to be more relevant for international tourists (coming from the Americas and Europe), and especially so for American tourists. Communication infrastructure and marketing (measured by Internet usage) are important considerations for tourists from all continents. The level of development in a country (as measured by the urbanization rate) also shows a positive relation with arrivals from international markets. Clearly tourists to Africa are not interested solely in sun, sea and beach holidays but are travelling to experience Africa's culture, wildlife and diversity. The literature indicates that 'sunlust' destinations are more price-sensitive than 'wanderlust' destinations (see Crouch, 1995). Therefore, tourism to Africa is also not very price-sensitive – a result that confirms the results of Eilat and Einav (2003) and Crouch (1995), who state that tourism to less developed countries is less price-sensitive. The significance of sharing a border with South Africa is also interesting, indicating a dynamic African tourism market in the southern part of the continent.

The panel data results differ somewhat from the cross-sectional results. Again, political stability comes to the fore as a key determinant of tourism to Africa, especially for tourists from the Americas. In the panel data analysis, tourism infrastructure, as measured by hotel capacity, is a strong determinant of tourism to Africa and the health risk (malaria) also becomes more significant. Again there is an indication that tourism to Africa is not solely for sun, beach and sea holidays, but rather to experience African diversity. Again the determinants for tourists from different continents vary (as suggested by Crouch, 1995). Tourism infrastructure, political stability and even higher levels of development are important to tourists from the Americas; tourism infrastructure, travel cost and health issues are more relevant for tourists from Europe; tourism infrastructure and a border with South Africa are key determinants of tourists from Africa.

In the dynamic panel data analysis, lagged income also becomes significant and, while the sign of the coefficient is positive for tourists from Africa, it is negative for tourists from Europe, which may be an indication that Africa does not yet benefit from a positive reputation or from habit-forming effects in its tourism products. The dynamics concerning arrivals indicate a similar pattern with tourists from Europe not planning a return visit to Africa.

To conclude, there are certain structural and institutional weaknesses that influence tourism to Africa, with factors such as tourism infrastructure, the level of development and Internet usage (marketing and information) being especially significant for tourists. The political and social instability in Africa is also a serious deterrent to growth in tourism arrivals. Typical 'developed country indicators', such as income in the origin country, cost of travel and prices, are less significant for the decision to travel to Africa and any analysis of the demand for African tourism should focus rather on the developing-country aspects of tourism demand.

Endnotes

1. The WTO defines tourism as 'the activities of persons travelling to and staying in places outside their usual environment for not more than one consecutive year'. International tourists are 'tourists who stay at least one night in a country where they are not residents'.
2. International tourism is the world's largest export earner, with foreign currency receipts already exceeding US$423 in 1996 (Eilat and Einav, 2003, p 1).
3. Tourism is estimated to contribute more than 2% of GDP and 5% of export receipts in the following African countries (using 1996 data): Benin, Botswana, Burkina Faso, Cape Verde, Comoros, Djibouti, Eritrea, Ethiopia, Gambia, Kenya, Madagascar, Mauritius, Namibia, Niger, Senegal, São Tomé and Principe, the Seychelles, Sierra Leone, South Africa, Swaziland, Tanzania, Togo, Uganda and Zimbabwe (Christie and Crompton, 2001).
4. The country of origin of tourists can be important, since it has been found that tourists from outside Africa tend to spend more per arrival than tourists from within Africa (Cleverdon, 2002, p 12).
5. The lack of roads is particularly problematic. For instance, it has recently been estimated that in Tanzania alone about 500 km of new roads are needed to enable tourism to expand significantly (Cleverdon, 2002, p 14).
6. The between estimator uses only the variation between the cross-section observations.
7. The within estimator uses the time variation within each cross section of observations.
8. Although HIV/AIDS also poses a significant health risk in Sub-Saharan Africa, it is more preventable than malaria. The lack of reliable time series data precludes the inclusion of HIV/AIDS in the present model.
9. The countries are: Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Republic of Congo, Democratic Republic of Congo, Côte d'Ivoire, Djibouti, Egypt, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mauritius, Morocco, Namibia, Niger, Nigeria, Senegal, the Seychelles, Sierra Leone, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia and Zimbabwe.
10. White (1980) proposed the heteroskedasticity-robust variance matrix estimator to adjust the standard errors of a regression in the presence of heteroskedasticity.
11. It is assumed that there is no second-order autocorrelation in the differenced idiosyncratic error term.
12. The IV approach leads to consistent but not necessarily efficient estimates of the parameters because it does not make use of all the available moment conditions (see Baltagi, 1995, p 126).
13. The first-step GMM estimator will be used since it has been shown to result in more reliable inferences. The asymptotic standards errors from the two-step GMM estimator have been found to have a downward bias (Blundell and Bond, 1998).

References

Ahmed, Z.U, Heller, V.L., and Hughes, K.A. (1998), 'Tourism in South Africa', *Cornell Hotel and Restaurant Administration Quarterly*, December, pp 80–90.

Anderson, T.W., and Hsiao, C. (1982), 'Formulation and estimation of dynamic models using panel data', *Journal of Econometrics*, Vol 18, pp 47–82.

Archer, B. (1977), 'Tourism multipliers: the state of the art', University of Wales Press, Cardiff.

Arellano, M., and Bond, S. (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, Vol 58, pp 277–297.

Attanasio, O.P., Picci, L., and Scorcu, A.E. (2000), 'Saving, growth, and investment: a macroeconomic analysis using a panel of counties', *Review of Economics and Statistics*, Vol 82, pp 182–211.

Austin, N.K. (2003), 'Entrepreneurship development and tourism in rural African communities', *Journal of African Business*, Vol 4, No 1, pp 87–98.

Baltagi, B.H. (1995), *Econometric Analysis of Panel Data*, John Wiley, New York.

Blundell, R., and Bond, S. (1998), 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, Vol 87, No 1, pp 115–143.

Chen, J.S., and Hsu, C.H.C. (2001), 'Developing and validating a riverboat gaming impact scale', *Annals of Tourism Research*, Vol 28, No 2, pp 459–476.

Chen, L.L., and Devereux, I. (1999), 'Tourism and welfare in Sub-Saharan Africa: a theoretical analysis', *Journal of African Economies*, Vol 8, No 2, pp 209–227

Christie, I.T., and Crompton, D.E. (2001), *Tourism in Africa,* Africa Region Working Paper Series, No 12, World Bank, Washington, DC.

Chu, F-L. (1998), 'Forecasting tourism demand in Asian–Pacific countries', *Annals of Tourism Research*, Vol 25, No 3, pp 597–615.

Chu, F-L. (2004), 'Forecasting tourism demand: a cubic polynomial approach', *Tourism Management*, Vol 25, pp 209–218.

Cleverdon, R.G. (2002), 'Tourism development in the SADC region: the opportunities and challenges', *Development Southern Africa*, Vol 19, No 1, pp 7–28.

Coshall, J. (2000), 'Spectral analysis of international tourism flows', *Annals of Tourism Research*, Vol 27, No 3, pp 577–589.

Crouch, G.I. (1995), 'A meta-analysis of tourism demand', *Annals of Tourism Research*, Vol 22, No 1, pp 103–118.

DEAT (1996), *The Development and Promotion of Tourism in South Africa*, White Paper, Department of Environmental Affairs and Tourism, Government Printer, Pretoria.

DFID (1999), *Tourism and Poverty Elimination: Untapped Potentials*, Department for International Development, London.

Diamond, J. (1976), 'Tourism and development policy: a quantitative appraisal', *Bulletin of Economic Research*, Vol 28, No 1, pp 36–50.

Divisekera, S. (2003), 'A model of demand for international tourism', *Annals of Tourism Research*, Vol 30, No 1, pp 31–49.

Du Preez, J., and Witt, S.F. (2003), 'Univariate versus multivariate time series forecasting: an application to international tourism demand', *International Journal of Forecasting*, Vol 19, No 3, pp 435–451

Eilat, Y., and Einav, L. (2003), 'The determinants of international tourism: a three-dimensional panel data analysis', unpublished working paper.

Gallup, J.L., and Sachs, J.D. (2000), *The Economic Burden of Malaria*, CID Working Paper No 52, Harvard University, Cambridge, MA.

Gallup, J.L., Sachs, J.D., and Mellinger, A. (1999), *Geography and Economic Development*, CID Working Paper No 1, Harvard University, Cambridge, MA.

Gauci, A., Gerosa, V., and Mwalwanda, C. (2002), *Tourism in Africa and the Multilateral Trading System: Challenges and Opportunities*, Background Paper for the Economic Commission for Africa, Addis Ababa.

Gelan, A. (2003), 'Local economic impacts: the British Open', *Annals of Tourism Research*, Vol 30, No 2, pp 406–425.

Heston, A., Summers, R., and Aten, B. (2002), *Penn World Tables 6.1*, CICUP, http://pwt.econ.upenn.edu.

Jud, G.D., and Joseph, H. (1974), 'International demand for Latin American tourism', *Growth and Change*, January, pp 25–31.

Kaufmann, D., Kraay, A., and Zoido-Lobaton, P. (1999), *Aggregating Governance Indicators*, Policy Research Department Working Paper No 2195, World Bank, Washington, DC.

Kennedy, P. (2003), *A Guide to Econometrics*, 5th edition, Blackwell, Oxford.

Kester, J.G.C. (2003), 'International tourism in Africa', *Tourism Economics*, Vol 9, No 2, pp 203–221.

Kim, S., and Song, H. (1998), 'Analysis of inbound tourism demand in South Korea: a cointegration and error correction approach', *Tourism Analysis*, Vol 3, pp 25–41.

Kottke, M. (1988), 'Estimating economic impacts of tourism', *Annals of Tourism Research*, Vol 15, pp 122–133.

Kulendran, N., and Witt, SF. (2001), 'Cointegration versus least squares regression', *Annals of Tourism Research*, Vol 28, No 2, pp 291–311.

Ledesma-Rodriguez, F.J., Navarro-Ibanez, M., and Perez-Rodriguez, J.V. (1999), *Panel Data and*

Tourism Demand: the Case of Tenerife, FEDEA Documento de Trabajo 99-17, Fundación de Estudios de Economía Aplicada, Madrid.

Lim, C. (1997a), 'Review of international tourism demand models', *Annals of Tourism Research*, Vol 24, No 4, pp 835–849.

Lim, C. (1997b), 'An econometric classification and review of international tourism demand models', *Tourism Economics*, Vol 3, No 1, pp 69–81.

Lim, C. (1999), 'A meta-analytic review of international tourism demand', *Journal of Travel Research*, Vol 37, pp 273–284.

Lim, C. (2004), 'The major determinants of Korean outbound travel to Australia', *Mathematics and Computers in Simulation*, Vol 64, pp 477–485.

Lim, C., and McAleer, M. (2001), 'Forecasting tourist arrivals', *Annals of Tourism Research*, Vol 28, No 4, pp 965–977.

Lim, C., and McAleer, M. (2002), 'Time series forecasts of international travel demand for Australia', *Tourism Management*, Vol 23, pp 389–396.

Lim, C., and McAleer, M. (2003), 'Modelling International Travel Demand from Singapore to Australia', CIRJE-F-214, Faculty of Economics, University of Tokyo.

Martin, C.A., and Witt, S.F. (1989), 'Accuracy of econometric forecasts of tourism', *Annals of Tourism Research*, Vol 16, pp 407–428.

Naudé, W.A., and Krugell, W.F. (2003), 'An inquiry into cities and their role in subnational economic growth in South Africa', *Journal of African Economies*, Vol 12, No 4, pp 476–499.

Nordström, J. (2002), *Dynamic and Stochastic Structures in Tourism Demand Modelling*, Umeå Economic Studies, No 596, Umeå University, Umeå.

Piga, C.A.G. (2003), 'Territorial planning and tourism development tax', *Annals of Tourism Research*, Vol 30, No 4, pp 886–905.

Poirier, R.A. (1997), 'Political risk analysis and tourism', *Annals of Tourism Research*, Vol 24, No 3, pp 675–686.

Pollak, R.A. (1970), 'Habit formation and dynamic demand functions', *Journal of Political Economy*, Vol 78, pp 745–763.

Randall, J.E. and Warf, B. (1996), 'Economic impacts of AAG conferences', *Professional Geographer,* Vol 48, No 3, pp 272–284.

Saayman, A., Saayman, M., and Naudé, W.A. (2000), 'The impact of tourist spending in Southern Africa: spatial implications', *South African Journal of Economic and Management Sciences*, Vol 3, No 3, pp 369–386.

Saayman, M., Saayman, A., and Rhodes, J.A. (2001), 'Domestic tourist spending and economic development: the case of the North West Province', *Development Southern Africa*, Vol 18, No 4, pp 443–455.

Sachs, J.D., and Warner, A.M. (1997), 'Sources of slow growth in African economies', *Journal of African Economics*, Vol 6, No 3, pp 375–377.

Sheldon, P. (1993), 'Forecasting tourism: expenditure versus arrivals', *Journal of Travel Research*, Vol 22, No 1, pp 13–20.

Smeral, E., and Weber, A. (2000), 'Forecasting international tourism trends to 2010', *Annals of Tourism Research*, Vol 27, No 4, pp 982–1006.

Song, H., and Witt, S.F. (2000), 'Tourism demand modelling and forecasting: modern econometric approaches', Pergamon, Oxford.

Song, H., Romilly, P., and Liu, X. (2000), 'An empirical study of outbound tourism demand in the UK', *Applied Economics*, Vol 32, pp 611–624.

Song, H., Witt, S.F., and Jensen, T.C. (2003), 'Tourism forecasting: accuracy of alternative econometric models', *International Journal of Forecasting*, Vol 19, pp 123–141.

Turner, L.W., Kulendran, N., and Fernando, H. (1997), 'Univariate modelling versus period and non-periodic analysis: inbound tourism to Japan, Australia and New Zealand compared', *Tourism Economics*, Vol 3, No 1, pp 39–56.

Vaughan, D.R., Farr, H., and Slee, R.W. (2000), 'Estimating and interpreting local economic benefits of visitor spending: an explanation', *Leisure Studies*, Vol 19, pp 95–118.

Walpole, M.J., and Goodwin, H.J. (2000), 'Local economic impacts of Dragon tourism in Indonesia', *Annals of Tourism Research*, Vol 27, No 3, pp 559–576.

Wang, P.C. (1997), 'Economic impact assessment of recreation services and the use of multipliers: a comparative examination', *Journal of Park and Recreation Administration*, Vol 15, No 2, pp 32–43.

Warner, A.M. (2002), *Institutions, Geography, Regions, Countries and the Mobility Bias*, CID Working Paper No 91, Harvard University, Cambridge, MA.

White, H. (1980), 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica*, Vol 48, pp 817–838.

Williams, A.M., and Shaw, G. (1992), 'Tourism and economic development: Western European experiences', 2nd edition, Belhaven, London.

Witt, S.F., and Witt, C.A. (1992), *Modelling and Forecasting Demand in Tourism*, Academic Press, London.

Witt, S.F., and Witt, C.A. (1995), 'Forecasting tourism demand: a review of empirical research', *International Journal of Forecasting,* Vol 11, pp 447–475.

Wooldridge, J.M. (2001), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge MA.

WTO (2000), *Tourism 2020 Vision, Vol 1: Africa*, World Tourism Organization, Madrid.

WTO (2002), *Compendium of Tourism Statistics (1996-2000)*, World Tourism Organization, Madrid.

WTO (2003a), *Tourism Highlights*, World Tourism Organization, Madrid.

WTO (2003b), *World Tourism Barometer*, Vol 1, No 2, October, World Tourism Organization, Madrid.

WTO (2003c), *World Tourism Barometer*, Vol 1, No 1, October, World Tourism Organization, Madrid.

WTTC (2003), *Sub-Saharan Africa: Travel and Tourism*, World Travel and Tourism Council, London.

Zhou, D., Yanagida, J.F., Chakravorty, U., and Leung, P. (1997), 'Estimating economic impact from tourism', *Annals of Tourism Research*, Vol 24, No 1, pp 76–89.