A JOURNEY THROUGH TOURISM STATISTICS: ACCURACY AND COMPARABILITY ISSUES ACROSS LOCAL, REGIONAL AND NATIONAL LEVELS

Serena Volo
Dipartimento di Metodi Quantitativi per le Scienze Umane
Università degli Studi di Palermo,
Palermo, Italy
serenavolo@unipa.it

ABSTRACT
Tourism statistics collection is a main concern for many national tourism organizations, however a lack accuracy and comparability is revealed at both international and national level. In this paper, Italian official tourism statistics are investigated during their journey through the national offices responsible for collecting them. The effect of redundancy created by the presence of a variety of domestic tourism statistics and the lack of reliability of the official statistics are addressed and the contribution of statistics unreliability to hidden tourism discussed.

INTRODUCTION
Tourism is a leading sector in the economy of many countries around the world, contributing significantly to GDP, employment, economic and social development. The growth of this industry, while creating benefits for many destinations, is starting to present a issue of sustainability in some regions, both in terms of the physical environment as well as the social and cultural environment (Middleton, 1998). Tourism, when well managed, creates numerous benefits for the hosting community, and compared to most other industries does not create as much damage to the environment. However, peaks in tourist arrivals above the carrying capacity are often associated with pollution, overcrowding, difficulties in managing

---

1 This paper is finacially supported by the MIUR research fund: PRIN Anno 2003 “Metodi e tecniche per la valutazione statistica del Sommerso. quando la vacanza diviene turismo”.
public services and a consequent loss of fashion of the destination, decline in market share and abandonment of tourism facilities. These typical signs of stagnation and decline of a destination support a general hypothesis of the unsustainability of tourism in certain places at certain levels. Addressing the needs of tourists is one of the responsibilities of local governments and tourism offices, and these local officials should be able to maintain an equilibrium between residents’ and tourists’ social and economic necessities, and undertake, when necessary, actions to effect a change in seasonality of arrivals. Much tourism activity goes unreported because of the way these statistics are gathered and aggregated. This unreliability of data collection and reporting systems contribute to the problem of hidden tourism and the attendant problems it creates for local, regional and national planners.

In these scenarios, tourism statistics are one of the key sources of information for (a) governments to evaluate tourism contribution to the economy; (b) economic planners to forecast and manage the impact of tourism in a destination; and (c) industry’s decision-makers to strategically plan their marketing activities (Burkart and Medlik, 1981; Wöber, 2000). In particular, tourism statistics, collected through national tourism organizations, represent the primary source of information upon which economists, public officials and many tourism operators in these countries forecast demand and plan supply of tourism-related services. Although the collection of tourism statistics is a main concern for many national tourism organizations, a lack accuracy and comparability is revealed at both international and national level.

Great emphasis can be found in the literature concerning the difficulties encountered in comparing tourism statistics at international level, and this is mainly because of the complexity of tourism activities and the lack of clear and standardized industry definitions. Corresponding difficulties exist when focusing on comparisons among tourism statistics defined and collected by different public and private agencies within a nation. Interestingly enough, even the so called official tourism statistics, measured and classified by different levels within the same organization, often appear to exhibit similar problems of accuracy and comparability. In the present study, Italian official tourism statistics are investigated with a focus on: (a) the effect of redundancy created by the presence of a variety of domestic tourism statistics; and (b) the lack of reliability of the official statistics. In addition, arrivals and overnight stays, as recorded by hoteliers and accommodation owners, are investigated for their journey through the recording system.

WHEN AND WHERE DO THE TOURISTS DISAPPEAR: THE ITALIAN VERSUS THE WTO DEFINITION

Considerable effort has been undertaken by the World Tourism Organization (WTO), the United Nations, the Organization for Economic Co-operation and Development (OECD) and the Eurostat to harmonize the definitions and standards for collecting tourism data. The WTO recommendations on tourism statistics, approved by the United Nations, summarised the efforts towards international tourism statistics comparability and aspire to define the System of Tourism Statistics. The effort started in 1993 with the "Recommendations on Tourism Statistics" and continued in 2000 with the "Tourism Satellite Account: recommended methodological framework" which attempts to move the industry from the existing collage of statistical sources to a unifying system of collecting data in which a guidance process and a new statistical tool are combined.
The World Tourism Organization taxonomy can serve as a useful scheme by which economic planners and tourism development officials can classify tourism activity so as to reveal its economic significance and the required supporting infrastructure. This taxonomy can be applied in any tourism market and thereby bring a standardization across markets to the way different tourism activity gets defined. However, the data capture and reporting systems of different tourism markets vary greatly in their methods, thoroughness and quality. Tourism statistics are typically collected through (a) counts and/or interview at the frontiers; (b) household and destination surveys; or (c) arrivals and nights spent as recorded at accommodation establishments (Edwards, 1991; Page, Brunt, Busby and Connell, 2001). Accuracy, reliability and adequacy problems are associated with each of these methods, creating underestimates or overestimates of tourism flows and tourism earnings and suggesting the possibility of statistically-created misinformation upon which economic planning may be based.

Space, time and motivation are the key constructs used by the World Tourism Organisation (1994) to define the movement of people as tourism. While their concept of "usual environment" is meant to exclude daily or weekly commuters or any other routine movement, the time restriction is intended to disqualify long-term migration as tourism. Moreover, tourism comprises leisure, recreation and holidays, visiting friends and relatives, business and professional travel, health treatment, religion, pilgrimages and other purposes. However, migration for temporary work is clearly excluded when motivation is taken into consideration. Widely accepted and apparently easy to use, this definition, when applied in the context of each country’s system of tourism statistics, can creates ambiguities. In particular, the WTO definition is frequently constrained by national statistical offices that introduce limits to the recording system. Latham (1989) indicated trip purpose, length of stay and type of accommodation as some of these constraints and, moreover, pointed out that statistical measurement of tourism might be affected by the mobile nature of the tourist population, the moment in their travel at which tourists are surveyed as well as a host of other variables.

The variety of statistical systems in Europe is revealed by an Eurostat analysis of official tourism statistics, that showed: (a) all countries record accommodation statistics; (b) 15 out of 17 countries record other supply side data and expenditure figures; (c) 13 out of 17 count arrivals at border and execute household surveys. Additionally, the variety of methods, complexity of tourism activities, the lack of clear and standardized industry definitions and the different coverage system complicates international comparison among tourism statistics countries (Lickorish, 1997).

Implementation of the WTO/UN recommendations in Italy is the responsibility the Italian statistical office (ISTAT), in cooperation with the Italian Office for Exchange. These agencies are responsible for collecting tourism statistics throughout the country by: (a) sample surveys; (b) counts and survey at frontiers; and (c) accommodation records. The Italian system of tourism statistics can be divided in the so-called ISTAT-demand and ISTAT-supply. ISTAT-supply refers to: (a) type, number and location of official accommodation establishments; (b) arrivals and nights spent as recorded at officially registered accommodation establishments; and (c) “rapid survey” on the hotel accommodation flows (with reference to peak seasons). ISTAT-demand, reorganized in accordance with the EU Directive on tourism statistics, refers to the Italian domestic or international demand of tourism measured by: (a) Italian household sample surveys; (b) counts of foreigners at frontiers, and (c)
survey at frontiers for both Italians and foreigners. Together with these official statistics there are numerous other sources of tourism statistics collected either at regional or national level by associations of tourism establishments or by various national institutes of research (Gismondi and Mirto, 2003).

In this plethora of tourism statistics, there is clear evidence of unreported or misreported consumption of tourism products. In particular, in Italy the main purpose for registering arrivals and overnights stay by hoteliers and accommodation owners is to fulfill administrative and taxation requirements. Consequently, there is little or no motivation to register precise data on nights spent, and, by underreporting them, intentionally create an underground component of tourism earnings that might drive differences large enough to underestimate the contribution to the national economy. Such data collection systems create on the one hand a conflict of interest between public agencies responsible for tourism development and private tourism operators, particularly with regard to the reporting of official accommodations statistics, and on the other hand estimation and forecast problems for economic planners and tourism developers.

Moreover, tourists staying with friends and relative, renting an apartment, or staying in any of the non-officially registered accommodation establishments are not represented by the official tourism statistics. Strangely enough, the ISTAT-demand tourism statistics based on surveys (and thus, affected by sampling errors) provides more detailed information when compared to the ISTAT-supply statistics, which are gathered by a census. As noticed by Lickorish (1997), a limited coverage is one of the consequences of using supply side data to measure tourism demand, as certain categories of tourists typically do not stay in officially registered accommodations.

What is described shows the imprecision of tourism statistics accruing from a lack of clarity in defining tourism and technical issues related to the specific-country’s statistical systems. Unfortunately, in addition to this imprecision, statistics on arrivals and nights spent appear to lose reliability after being recorded by accommodation operators as they journey through the local, regional and national offices responsible for collecting tourism statistics. Data recorded as official statistics by the local, regional and national offices, even when examined at identical levels of aggregation, exhibit variances across levels of aggregation and responsibility, suggesting they may contain statistically created misinformation upon which economic planning may be based.

The case of arrivals and overnight stay in Sicily was investigated, and a difference in tourism arrivals and overnight stays were found when comparing data at the different levels of aggregation. As mentioned before, tourism statistics start their journey at the level of individual accommodation establishments. Then, data are communicated to the statistical office of the local tourism organization. These data are collected monthly and sent to the provincial tourism office, which then communicates them to the tourism regional office that is then responsible for communicating them to ISTAT.

From a comparison of statistics across these levels of aggregation and level of responsibility, it was concluded that this journey of tourism statistics creates an even worst case of unreported or misreported consumption of tourism products and arrivals and overnight stays. Facts are rarely the same when comparing the four sources: local, provincial, regional or national. These differences, although representing a small amount of unreported or misreported tourism activity, emphasize the lack of accuracy and the incapacity of holding the official statistics to certain standards.
Recent Italian literature is full of comparison and integration of tourism data from different sources and shows substantial differences between official data and tourism estimates otherwise gathered. Through a comparison of ISTAT’s different tourism statistics (Mercury, 2003), the estimated Italian overnights in 2001 were more than twice the value recorded by official accommodations. By integrating these data with a rough estimate of real tourism made by the local tourism offices, it was estimated that Italian tourism revenues were understated by 18.3% (Gismondi, 2003). A review of past literature revealed visitor specific data and indicators based on aggregate variables have been the most common data used in Italy when reporting estimates of “true” tourist flows in small tourism destinations. Traditional regression approaches are the most common analytic method of choice when these estimates are made from aggregate variables. Estimates based on auxiliary variables include studies on: the Venetian coast (Manente and Scaramuzzi, 1999) in which the authors focused on private houses used as holiday resorts; the Leghorn province (Irpet, 1998) in which newspapers sales, garbage production and electricity consumption were used as independent variables to produce estimates of “true” tourist flows; the Massa-Carrara coastal area in which the so-called Becheri-Gambassi model was used (Pasetti, 2002) where the true overnight stays were estimated by comparing three different methods: (a) an exploratory survey on tourists present in the area, (b) a complex model that used garbage production and electricity consumption and number of vacation houses as indicators, and (c) a simple model in which the garbage production was used as a lone variable to estimate unregistered overnights stay. The model showed the real overnight stays in the Massa-Carrara coastal area were six times the official data, and pointed out the similarity of results when estimating using either the simple or the complex model.

CONCLUSIONS

This unreported tourism activity contributes significantly to what has been referred to as “hidden tourism”. Hidden tourism refers generally to unreported or misreported consumption of tourism products. In the first case economically significant tourism consumption activity does not get systematically recorded and reported in the standard statistical reporting system. In the second case, economically insignificant tourism consumption gets reported as significant, because it is attributed to the wrong class of consumers. Both cases of hidden tourism present a problem for economic planners and tourism developers. The first case leads to understated planning forecasts and a consequent supply that fails to meet demand. The second case leads to overstated demand forecasts for certain products, and often times, by extrapolation, to all product, and a consequent supply that exceeds demand.

When interest is in estimating the real economic contribution of tourism to the destination, one of two “data availability” scenarios typically exists for economic planners and tourism policy makers: visitor specific or aggregate data. Visitor specific data are typically obtained by surveying individual visitors to a destination. Aggregate data are typically obtained from official statistics routinely gathered by a governmental agency such as ISTAT. Each of these scenarios presents different challenges to analytic methods for detecting hidden tourism. Visitor specific data can be very rich in information that can distinguish “hidden” from “official” tourism, but workable only for small areas, very time consuming and expensive. Aggregate data on the other hand, typically contain only indirect indicators of hidden tourism (e.g., electricity consumption, garbage production). In addition, hidden tourism’s
contribution to the total tourism can easily go undetected in all but the most sensitive of statistical procedures.

These different data scenarios call for different analytic approaches. Common regression models or other statistical measures of association can typically unveil predictor variables that can, with adequate reliability, correctly classify visitors as “hidden” and do so with relatively few predictor variables that can be routinely and unobtrusively gathered. Such regression models however, typically do not perform well with aggregate data, yet aggregate data is usually all that local planners and policy makers have available to them. Further, local planners and policy makers often must make demand forecasts for their own local area without the benefit of time or resources to conduct elaborate statistical studies. Typically they have neither the time nor the resources to conduct surveys to gather visitor-specific data, and forecasts made from aggregate data will eventually be abandoned if they do not prove to produce reliable forecasts. A simple but reliable forecast model that can be applied by local planners and policy makers to aggregate data would have substantial practical value. Economic planners and public agencies with responsibility for tourism development are in need of better methods to detect hidden tourism and to more reliably attribute its economic significance. Relatively recent data mining algorithms provide one promising approach to solving this data problem.

REFERENCES