

E. Model-based estimates

E. Model-based estimates

17.27. When source data are not timely or otherwise adequate, compilers may also choose to create a data model to estimate statistics, or a portion of them, on the international supply of services.^[1] The data model elements should be based on the compiler's judgment of the specific information available. Some data models can be relatively simple. For example, they may be based on proven relationships between different data sources and variables. An additional straightforward approach for monthly or quarterly estimates of items with strong seasonal fluctuations is to look at relationships of relevant data series from the current period (quarter or month) from each of the preceding predetermined number of years. When quarterly results become available, monthly estimates would then be reconciled so that the three months equal the corresponding quarter. Models can also be more complex, based on mathematical and econometric approaches, including exponential smoothing, linear and logistic regression, polynomial interpolation (or cubic splines) and multilevel models, to account for hierarchical relationships among variables.

17.28. For estimating travel services, a model could be constructed using primarily the number of visitors and other individuals travelling short-term (partially available from tourism statistics or transportation operators) and estimates of per capita expenditure. However, it should be noted that such data are not always accurate, as the transport operator may not always collect information on the residence of the customer. Furthermore, when collected, the information may relate to the nationality of the customer and not the residence. Estimates of per capita expenditure could be obtained from occasional surveys of persons who travel. If necessary, separate data models can be designed to measure education- and/or medical-related travel on the basis of information from relevant institutions or special surveys (e.g., a survey of students).^[2] However, in order to enable the use of tourism statistics, the differences in terms of the coverage between BOP definitions, on the one hand, and tourism statistics as used in IRTS 2008 (as well as in *Tourism Satellite Account: Recommended Methodological Framework 2008*), on the other, need to be understood^[3] (see also chapter 14 of the present *Guide*).

17.29. For estimating the value of mode 4 and the number of persons moving under mode 4, model-based estimates could be developed using existing statistics on the international supply of services and travel information, as well as existing data from tourism, migration and employment statistics, etc. All that information could be used for building an estimation of mode 4. (More information on combining different data sources is provided in chapter 13.) Model-based estimates are a cost-efficient way to use existing data to compile modes of supply data.

17.30. For such an approach to be efficient, compilers need to analyse existing metadata, familiarize themselves with the methodologies from other statistical frameworks, cooperate closely and exchange the relevant microdata with the other statistical domains and institutions involved. Some adjustments to the data or data collection tools will probably be needed, in particular to obtain the relevant breakdowns and to clearly distinguish the population of interest. For example, identifying or estimating the population qualifying as visitors or travelling for short-term employment, including appropriate breakdowns, would be necessary as inputs for data models for estimating travel or the number of mode 2 persons/trips. For mode 4, that would entail distinguishing, within the group of travelling persons, those who are employed, those who have an employer-employee relationship in their country of origin (mode 4) and those who have that relationship in the host economy.^[4]

17.31. For example, tourism data sources could be used for collecting further information on the characteristics of mode 4, while some adjustments on different concepts and terminology in both statistical domains would almost certainly be necessary. Subsequently, an estimation of the number of mode 4 persons and of the value of mode 4 trade could be derived by grossing up the figures on the basis of a research-based model.

17.32. A similar approach could be adopted for other types of sources, such as labour force surveys, household surveys or even business surveys. Adding options in the questions asked to respondents could serve to identify mode 4 persons in such data sources.

Include Page:

Country experience: Austria: a model for estimating monthly balance of payments data

Country experience: India: estimation of non-response in the survey on exports of information technology and information technology-enabled services

Country experience: United Kingdom: experience with modelling outward FATS

Back to [Chapter 17 Estimation and modelling of missing data and forecasting or back-casting](#)

^[1] See *BPM6 Compilation Guide*, paras. 8.17 and 8.25.

^[2] Ibid., para. 12.75.

^[3] Ibid., para. 12.76.

^[4] Actually, defining services contracts versus employment relationships is something that would better serve the compilation of tourism statistics as well as migration statistics.