CHAPTER VII. SHORT-TERM DISTRIBUTIVE TRADE STATISTICS

A. Background

7.1. Short-term distributive trade statistics are an important source of information for developing and monitoring effectiveness of economic policy and carrying out business cycle analysis. Whereas annual statistics, such as structural type statistics described in previous chapters, have details, comprehensiveness and accuracy as priorities, short-term statistics' priority is to produce monthly and quarterly indicators of dynamics of distributive trade sector in the most timely manner, but with likely lower accuracy, less details and reduced scope. These statistics are usually produced to a strict timetable, and they are required as soon as possible by policy makers. Sometimes this means that initial figures are subsequently revised or adjusted as more data is collected and analyzed.

7.2. Most commonly, short-term statistics in general and short-term distributive trade statistics in particular, are presented in the form of indices (relative to a base period) or as growth rates, although it is a widespread practice for absolute figures to be compiled and disseminated as well.

7.3. Although there are many users of short-term distributive trade statistics with many different motivations for using the data, the analyses performed with them generally fall into one of two types:

(a) comparison of activities of distributive trade units between two different points in time; and

(b) comparison within one reference period of two or more different sub-populations of units, such as between units in different geographical locations, or between trade units and other units classified in service activities, or between wholesalers and retailers, etc.

7.4. Different objectives and priorities of structural and short-term distributive trade statistics require from countries that appropriate statistical techniques are developed and implemented in order to combine these two sets of data. The main aim of the these techniques is the reconciliation of the statistical data derived from different data sources with different frequency, in order to obtain short-term data series that, while obeying the constraints imposed by the more reliable and accurate long-term information sources, preserve as much as possible the dynamic time-profile of the high frequency time series.

7.5. To facilitate the achievement of this aim, both sets of statistics should be based on the identical concepts and measurement principles, statistical units, classifications and definitions of data items. Short-term statistics should be built on a foundation of timely and accurate infra-annual data sources that cover an adequate proportion of units (size of the samples). Econometric methods and indirect estimation procedures should not be accepted as a substitute for data collection. They also should be made consistent with their annual equivalents, partly for the convenience of users and partly and more fundamentally because the benchmarking process incorporates the information content of the annual data into the monthly/quarterly estimates.

7.6. This chapter discusses some of the important aspects pertinent to short-term distributive trade statistics, such as the compilation of distributive trade indices, their time-series character and the necessity of seasonal adjustment and benchmarking.

B. Indices of distributive trade

1. Indices of distributive trade: An overview

7.7. *Types of distributive trade indices.* To analyze various aspects of distributive trade dynamics a number of indices can be constructed ranging from a rather simple indices of turnover changes in nominal terms (value index) to a more detailed and complex index of turnover volume and output of distributive trade sector (indices reflecting volume of production of retail and wholesale trade services). To obtain volume indices, the indices of retail and wholesale prices or appropriate volume indicators should be available. The development of international recommendations on compilation of such price indices is not a part of IRDTS, however some guidance on this topic can be found in paragraphs below.

7.8. *Purpose.* One of the main purposes of compilation of distributive trade indices is the description of the short-term changes in value and volume of turnover of wholesale and retail trade as well as in the output of distributive trade sector as a whole and of its components. If available on a monthly/quarterly basis, indices of volume of turnover complement indices of other economic activities in the short-term analysis of entire economy including the identification of the turning points in economic cycles. Indices of output of distributive trade sector, in addition to their importance for short-term analysis, provide a key input in the compilation of quarterly national accounts.

7.9. *Periodicity.* It is recommended that indices of turnover volume and output are compiled on a monthly basis as this better reveals the short-term fluctuations. Monthly indices are even more meaningful if produced without significant time lag that is within the month immediately following the reference period. Recognizing that national statistical offices may not have capacity to produce reliable monthly indices countries, in such cases, are advised to compile quarterly indices, as it gives sufficient flexibility in terms of time and resources. It should be noted, however that use of quarterly indices dilutes the effects of current market conditions, seasonal changes and other factors related to short term production.

7.10. *The choice of index formula and base year.* The detailed discussion of index types, their theoretical properties and comparative advantages and drawback is provided in various international sources[1] and is not reproduced in this publication. It is recommended that compilers of distributive trade indices use those manuals while developing their countries' distributive trade indices. Although, some policy guidelines in this respect are provided in this chapter, a more detailed discussion of good practices in this area is beyond the scope of these recommendations. That will be provided in *Distributive Trade Indices: A Handbook of Country practices,* which is to be issued as a follow-up publication to the current recommendations.

7.11. *General recommendations for the compilation of volume indices of distributive trade*. As a general guideline the chained-linked Laspeyres index with the weights being updated at least every five years is recommended as a preferred approach for the compilation of volume indices. This index formula satisfies most of the desirable criteria such as the monotony, homogeneity, up-to-date weighting structure, real comparison of volumes, cost efficiency and timeliness etc. Paasche formula does not have any strong advantages over the Laspeyres index and is more difficult to implement as it requires availability of current weights. Chained Fisher index where the current and the base period weights are used in its Laspeyres and Paasche components possesses several theoretical advantages such symmetry and time reversal but it looses on interpretability and is the most difficult one to implement..

7.12. Laspeyres volume index with the weights not been changed for a long period of time (more than five years) is used by countries with limited resources and persistent problems with obtaining updated weights. This approach has an advantageous property of producing data at constant prices which are additive (sum of components is equal to the total value) and, therefore, has a clear economic interpretation and is convenient for use. However, as the time gap between the base year and current period increases the quality of such an index deteriorates as it does not reflect dynamics of distributive trade. If a country uses Laspeyres volume index with fixed weights it is recommended that the periods between which weights are updated are as close to five years as possible. In the process of updating the weights countries are encouraged to make every effort and to chain-link the series with the new weights.

7.13. It is recommended that while choosing the index type, countries take into account the purpose of the index and practical considerations such as general policy of a given national statistical office in the area of price statistics, the availability and quality of data, resource constrains etc. As a further guidance for the compilation of distributive trade indices, countries are advised to use seasonally adjusted series when appropriate and available. Section 2 below provides some additional recommendations with respect to the indices of turnover volume and the output of distributive trade services.

2. Indices of wholesale and retail trade turnover

[1] See 2008 SNA, Chapter XVI; Compilation Manual for an Index of Service Production, OECD, 2007, Section 5; Handbook on price and volume measurement in national accounts, Eurostat, 2001; Manuals on Consumer and Producer Price Indices etc.

7.14. *Turnover value index.* The turnover value index is a direct index that compares the value of turnover in the current period (at current prices) with the value of turnover in the base year (at base year prices). This index can be calculated for both retail and wholesale trade and its components.

7.15. *Turnover volume index.* The turnover volume index, especially the volume of retail trade turnover, is one of the most closely monitored series. In order to eliminate the price effect on turnover, it has to be deflated. In principle, the deflator of turnover should be a price index representative of the particular distributive trade activity class and reflecting price changes in the goods sold rather than the trade services provided. The retail price indices (RPI) and wholesale price indices (WPI), or consumer price indices (CPI) and producer price indices (PPI), can be used as proxies for such deflators. The price deflator for a given activity should be calculated as a weighted average of the price indices for the relevant category of goods sold by that activity in the current period.

7.16. In order to compile the turnover volume indices at the higher levels of section G of ISIC, Rev.4, the indices at the lowest level have to be aggregated. This aggregation is done by using weights based on turnover share of each activity in the base year. For example, the index for Group 471 "Retail sale in non-specialized stores" is derived from all the indices of the lower level (i.e. classes included in 471). The index of Section G will be calculated by taking a weighted average of all the component divisions in the section.

7.17. *Alternative methods for measuring of turnover volume*. If appropriate price indices are not available to deflate turnover due to the difficulties in measurement of price changes or the complexity of data sources, its volume might be estimated using output volume indicators or input indicators.

(a) *Output volume indicators.* The output variables (e.g., physical quantity of goods sold) are accepted as the second best option if they represent well-defined products and are applied in sufficient level of details.

(b) *Input indicators.* Employment is considered as one of the main input indicators, which can be used as a proxy measure of production. Although, not recommended, there are many situations where information on input measures is the only readily available source. In this case, it is assumed that the changes in input and output are proportional to each other. Compilers should be very cautious regarding use of estimates based on input variables.

7.18. *Turnover volume index and index of output of wholesale and retail trade.* The objective of the turnover index is to show the evolution of the market for goods and services. It should be noted, in this connection, that there are significant conceptual differences between this index and the index of output of wholesale and retail trade activities (also called 'index of production of wholesale and retail services'). The main differences are:

(a) Turnover includes sales of goods bought for resale in the same condition as received which is not considered in the indices of output of wholesale and retail trade service;

(b) Goods produced (or purchased) and stocked before sale are included in both output and turnover, but are considered at different moments in time;

(c) Index of output of wholesale and retail trade services takes account of changes in the quality of the trade *service* supplied.

7.19. Both indices are important in their own rights. While the volume of turnover is recommended for compilation within the framework of short-term statistics, the indices of output of wholesale and retail trade services are meaningfully compiled only within the framework of national accounts, preferably within the framework of supply and use tables.

7.20. The indices of output measure changes in production of services by various distributive trade activities. One of the major reasons for compilation of these indices is their use as inputs in the quarterly national accounts compilation as an appropriate estimate of short-term changes in gross value added for the wholesale and retail trade services. Therefore, in principle, they should be calculated as weighted averages of the outputs of these activities using value added weights with the assumption that the ratio of value added to output is constant in the short-run. In practice, however, the required value added data might not be available at such a detailed level for the required periods. Therefore, in the absence of value added, alternative measures for producing these indices such as volume of turnover should be used.

C. Seasonal adjustments

7.21. *Need for seasonally adjusted distributive trade statistics*. Monthly and quarterly data on distributive trade statistics are an important tool for economic policy making, business cycle analysis, modelling and forecasting. However, they are often characterised by seasonal fluctuations and other calendar/trading-day effects, which are obstacles in the clear identification of important features of time series such as their short and long-term movements, turning points and consistency with other economic indicators. Seasonal adjustment is a process by which changes that are due to seasonal or calendar influences are removed from time series in order to achieve a better knowledge of the underlying behaviour. This section contains a brief overview of the basic concepts and recommendations for compilation of seasonally adjusted time series. The more detailed guidance on this issue will be provided in the forthcoming *Distributive Trade Statistics: Compilers Manual.* Seasonal adjustment issues of particular interest for distributive trade statistics like trading day and moving holidays effects are presented in section 3. Calendar effects.

7.22. As a general recommendation, countries should consider producing seasonally adjusted series as an integral part of their long term programme of quality enhancement of their distributive trade statistics. They are encouraged to begin production of seasonally adjusted series of distributive trade data items as a matter of priority. The seasonal adjustment method chosen once should not be changed often. If the changes are necessary, they should be thoroughly justified.

1. Basic concepts for use in compilation of seasonally adjusted data

7.23. *Time series.* When statistical data are collected at regular intervals of time they form a time series. Turnover of retail trade for each sub-period (week, month, quarter) of the year, in a given country is a good example of a time series. In contrast, data collected irregularly or only once do not represent a time series. There are two types of time series - stock and flow. Stock series are measures of activity *at a point* in time while the flow series measure the level of activity *over* a time interval.

7.24. *Components of time series.* A time series is generally considered to be made up of the following components:

(a) *The trend component* (T_t) which reflects long term movements lasting many years. It is generally associated with structural causes, for example, institutional events, demographic and technological changes, new ways of organization, general economic development, etc. In many series such as wholesale and retail sales, or production of goods and services, this may be termed the growth element.

(b) *The cycle component* (C_t) indicates the longer term irregular fluctuations, usually referred to as business cycle. In much analytical work, the trend and the cycle are combined because, for series covering a short period of time, the long-term trend cannot be estimated adequately. As such, the trend-cycle component is the underlying path or general direction reflected in the data, that is, the combined long-term trend and the business-cycle movements in the data.

(c) *The seasonal component* (S_t) is a movement within the year with a characteristic shape for each time series which represents the effect of climatic and institutional events that repeat more or less regularly each year. This component includes seasonal effects narrowly defined and calendar related systematic effects that are not stable in annual timing, such as trading day effects and moving holiday effects (see para. 7.31-7.37). The seasonal effect narrowly defined is an effect that is reasonably stable in terms of magnitude. Possible causes for this effect are natural factors, administrative or legal measures, social/cultural traditions, and calendar-related effects that are stable in annual timing (e.g., public holidays such as Christmas).

(d) *The irregular component* (I_t) represents unforeseeable movements related to events of all kinds. It is the residual variations due to developments or to momentous occurrences such as wars or national catastrophes, which affect a number of series simultaneously. In general, the irregular component has a stable random appearance and it captures effects that are unpredictable unless additional information is available, in terms of timing, impact, and duration. The irregular component includes the following: i) irregular effects narrowly defined; ii) outlier effects; iii) other regular effects such as the effects of unseasonable weather, natural disasters, strikes, irregular sales campaigns, etc. However, it should be noted that these effects can be estimated separately to the irregular component and that it is important to do this in order to ensure that the best quality seasonal adjustment is achieved.

7.25. *Seasonal adjustment*. The process of estimating and removing the seasonal component from a time series is known as seasonal adjustment. It removes all variations that are systematic (seasonal effects) and calendar related (institutional events which repeat more or less regularly every year).

2. Main principles and models of seasonal adjustment

7.26. As a general rule, the seasonal adjustment process should be performed at the end of a survey cycle when the survey has been designed and conducted; data has been collected, processed and edited; and estimates are produced. The seasonal adjustment process starts once the original estimates are available and the original time series of data are formed.

7.27. *Basic principles of seasonal adjustment.* In order to remove the seasonal component from a time series it should first be decomposed into its constituting components - the trend-cycle, the seasonal component and the irregular component; each of which may be made up of several subcomponents. The seasonal variations can be distinguished from the trend by their oscillatory character, from business cycle by having annual periodicity and from irregulars by being systematic. The four above-mentioned components can be combined in a number of ways. The most commonly found are two types of decomposition models: the additive decomposition model and the multiplicative decomposition model.

7.28. *Additive decomposition model.* The additive model assumes that the components of the time series behave independently of the other. In particular, the size of the seasonal oscillations is independent of the level of the series. For example, an increase in the trend-cycle will not cause an increase in the seasonal component. This model is used if the irregular and the seasonal effects are independent of the trend behaviour, i.e. if the seasonal effects are the same from year to year.

$$X_t = T_t + C_t + S_t + I_t$$

7.29. *Multiplicative decomposition model.* The multiplicative model is generally taken as the default model in seasonal adjustment software packages. This model assumes that the components of the series are interdependent and thus the seasonal variation's size increases and decreases with the level of series, a characteristic of most seasonal macroeconomic series. For example, an increase in the trend will cause an increase in the magnitude of the seasonal component.

$$X_t = T_t.C_t.S_t.I_t$$

7.30. *Quality of seasonal adjustment*. The most fundamental requirement of seasonal adjustment quality is that there is no estimable seasonal effect still present in the seasonally adjusted series. The presence of estimable seasonal effects in either the seasonally adjusted series or the de-trended seasonally adjusted series (i.e. the irregular component) is generally what is referred to as *residual seasonality*. To detect whether the seasonally adjusted time series contains residual seasonality and trading day effects a special "spectral diagnostic" should be carried out, for monthly data or for sufficiently long quarterly series. Depending upon the package used for seasonal adjustment there are other diagnostics that can be used to assess the presence of residual seasonality.

7.31. Other important requirements of a good seasonal adjustment are lack of bias in the level of the series and the stability of the estimates. A lack of bias in the level means, that the level of the series will be similar for both the original series and the seasonally adjusted series. Stability of the estimates means that as new data become available and are incorporated into the estimation procedure, the revisions to the past estimates are small. Large revisions can indicate that the estimates are misleading or even meaningless.

7.32. *Concept of direct and indirect seasonal adjustments*. Many of distributive trade data represent aggregates or residual items. For instance, the trade margin or value added is calculated as a difference between two components. In the first case these are values of turnover and goods bought for resale; and in the second case these are the output and intermediate consumption. A seasonally adjusted estimate of value added can be derived either as seasonally adjusting value added directly, or as a difference between the seasonally adjusted output and intermediate consumption.

7.33. Under most circumstances, the direct and indirect adjustments for an aggregate series are not identical. There are some very limited situations in which the two types of adjustment coincide, particularly if the adjustments are additive. Whether direct or indirect adjustment is more appropriate for a given set of series will to a greater extent depend on the set of series in question. Because neither theoretical nor empirical evidence uniformly favours one approach over the other, countries are advised to deal with this issue on a case by case manner, after a thorough analysis based on the characteristics of the series in question and on the aggregation constraints imposed by the context (national accounts, geographical breakdown etc.). The following is a practical guidance on how to deal with direct/indirect seasonal adjustment in some particular cases:

(a) Indirect seasonal adjustment should be preferred when the component series that make up the aggregate series have quite distinctively different seasonal patterns and have adjustments of good quality. The indirect seasonal adjustment in this case is of better quality than the direct adjustment.

(b) Direct seasonal adjustment should be preferred when the component series have similar seasonal patterns and summing the series may result in noise cancellation.

7.34. *Outliers in seasonal adjustment.* Outliers are abnormal values in the time series, usually caused by one-off economic or social event. Their detection and correction prior to implementation of the adjustment process is an important precondition for the quality of seasonal adjustment. It is essential to distinguish between different types of outliers because their treatment differs. Outliers are divided into two groups (i) errors in the data; and (ii) the "true" special events. The first step of any outlier analysis should be the detection and correction of plain data errors and after that, the detection and correction of "true" outliers. The correction of outliers aims at preventing the trend path from distortion. The trend path is intended to measure the long-term growth of time series and it is not desirable for it to respond to a one-off irregular movement. It should be noted that all seasonal adjustment packages have built-in option for the detection and the treatment of outliers, at least for the historical part of the series. For the most recent values, however, a sophisticated automatic correction is not possible.

7.35. *Concurrent versus extrapolated seasonal factors*. For the current year seasonally adjusted data can be computed either by running every month/quarter the seasonal adjustment procedure or by using extrapolated coefficients computed once a year. In the first case, data are revised every month/quarter. In the second one, data are not revised within the year but only once a year. In terms of accuracy of seasonally adjusted series the concurrent approach is the recommended one. It allows producing up-to-date seasonally adjusted data by recalculating the values every time new data become available. The use of extrapolated seasonal factors, although preferred by some users, can lead to biased results especially when unexpected events occur during the year. It is recommended that countries schedule the revisions of seasonally adjusted series in a regular manner, preferably in accordance with the established common revision policy.

3. Calendar effects

7.36. *Calendar effects.* Variations associated with the composition of the calendar play an important role in the analysis of distributive trade statistics. *Calendar effects* are regular effects that do not necessarily occur in the same month or quarter each year but can be identified and removed from the series. The most important of them are the moving holidays' effects and the "trading-day" variations which represent the "within-month effects". These variations are usually treated as seasonal in character and should be removed together with the other seasonal variations when producing a seasonally adjusted series.

7.37. *Moving holidays.* Moving holidays are holidays that occur at the same time each year based on the different calendars other than Gregorian calendar which is widely used as a world standard for statistical time series. Therefore, their exact timing shifts systematically each Gregorian calendar year. The influence of these moving holidays in economic and social behaviour can usually be country specific, making it difficult to build them into standards routines and practices. Examples of moving holidays include Easter, Chinese New year, Korean Thanksgiving day and Ramadan.

(a) *Easter* generally falls in April but can also fall in late March and can affect a variety of types of series, for example industrial production or retail trade sales especially in the western hemisphere. Easter effect is the variation due to the displacement from April to March of the volume of activity when Easter falls in March instead of the usual April occurrence.

(b) *Chinese New Year* affects in a similar way trade activities. It mostly occurs in February but can also occur in January. As for Easter, Chinese New Year's effects have a predictable magnitude and direction.

7.38. Two types of effects are generally associated with the moving holidays – (i) an immediate effect associated with the fact that some retail stores are closed during the holidays; and (ii) gradual effect associated with the fact that the level of trade activity is affected during several days preceding the holidays.

7.39. *Trading day.* Trading day is a common calendar related effect that is often found in economic time series, especially in distributive trade time series. This effect is due to the number of times each day of the week occurs in a given month/quarter and the length of the month/quarter. The number of trading days is also affected by the number of holidays in the in the given time period, that do not fall on weekends. The number of trading days may differ not only from period to period, but it may also vary between the same time periods in different years.

7.40. Trading day effect is present when the level of activity varies with the days of the week. Trading day variations imply the existence of an underlying daily pattern of activity defined over the week. This daily pattern states the relative importance of the days in the week. For example, five Sundays in a month impacts retail trade series because Sunday is not a business day and marks a low point in the economic activity. Also, for these series, the number of Fridays and Saturdays has a significant impact, as these days are those when people do much of their shopping activities. Trading day variations are associated also with the accounting and reporting practices of trade units. Stores that do their bookkeeping activities on Friday tend to report higher sales in months with five Fridays than in months with four Fridays. Trading day effects need to be accounted for because they lead to apparent changes in level of activity when the underlying level is in fact unchanged.

7.41. *Length of month effect.* Because different months of the year have different lengths - 28, 29, 30 or 31 days, one way to think of the trading day effect is to consider each month of the year as a block of 28 days (four days in each type of weekday) plus one, containing zero, two or three extra days. If the level of activity for each type of weekday is to be constant through the years, the only difference between the months in a given year will be due to the number of extra days (0, 1, 2, or 3). Hence, if June and July have the same levels of activity on the respective days of the week, the total level of activity for July may still be greater that that for June purely because July has an extra day. This effect is called a *length of month effect.* If a series does not have a trading day correction, then the length of month effect will be accounted for automatically in the seasonal factors. If the series have a trading day correction, the length of month can still be accounted for in the seasonal factors or alternatively in the trading day factors.

7.42. *Methods for trading day adjustment***.** Trading day adjustment can be carried out in either the proportional or regression methods for adjustment. Under the first approach, the effects of trading days is estimated by counting the proportion of them in the month/quarter while under the second the effects of trading days is estimated in a regression framework. In general, the regression based approach should be preferred by countries as a method of trading day adjustment. As for the other moving holidays effects, statistical packages have built-in options for the detection and treatment of trading day effects. They offer default calendars, however**,** it is recommended that in trading day adjustment countries use country-specific calendars as they ensure more accurate results.

4. Seasonal adjustments software packages

7.43. The most commonly used seasonal adjustment packages can be grouped into two main categories: (i) based on uni-variate time series decomposition, namely moving average techniques; and (ii) based on explicit models with a small number of parameters for each component. Choice of the countries between the two packages should be done on the basis of the thorough analysis of the time series subject to seasonal adjustment and/or on the past experience.

7.44. *Seasonal adjustment packages based on moving average methods*. The majority of seasonal adjustments methods used by statistical offices belong to the class of moving averages. The seasonal adjustment methods that belong to this category are mainly descriptive, non-parametric procedures in the sense that they lack explicit parametric models for each unobserved component. Major computational differences between the various approaches of this class are usually due to different techniques used at the ends of the time series. Some methods use asymmetric filters at the ends while others extend the series using ARIMA models and apply symmetric filters to the extended series. The general procedure in the class of moving average models follows an iterative estimation procedure, the core of which is based on a series of moving averages. Census X-11/X-12 and ARIMA belong to the seasonal adjustment techniques based on moving average methods.

7.45. *Seasonal adjustment packages based on model methods*. The model based approach requires the components of the original time series, such as the trend, seasonal and irregular to be modelled separately. This approach, assumes the irregular component is "white noise". Major computational differences between various methods in the model based approach are usually due to model specification. In some cases, the components are modelled directly. In other cases, the original series is modelled and the component models are derived from that model. Model based seasonal adjustment programs include, among others, TRAMO-SEATS, STAMP, and BV4.

7.46. *Seasonal Adjustment Diagnostics.* A set of diagnostics to assess the outcome, both from the modelling and the seasonal adjustment parts are provided in the programs and should be used. These diagnostics range from advanced tests targeted for the experts attempting to fine-tune the treatment of complex series to simple tests that as a minimum should be looked at by all users of the programs. While the programs sometimes are used as a black box without the diagnostics, they should not be used that way, because many tests can be readily understood.

7.47. *Minimum Length of the Time Series for Seasonal Adjustment.* Five years of data and relatively stable seasonality are required in general as minimum length to obtain properly seasonally adjusted estimates. For series that show particularly strong and stable seasonal movements, it may be possible to obtain seasonally adjusted estimates based on only three years of data. A longer time series, however, is required to identify more precisely the seasonal pattern and to adjust the series for calendar variations (i.e. trading days and moving holidays), breaks in the series, outliers, and particular events that may have affected the series and may cause difficulties in properly identifying the seasonal pattern of the series. If a country has gone through severe structural changes resulting in radical changes in the seasonal patterns, it may not be possible to seasonally adjust its data until several years after the break in the series. In such cases, it may be necessary to seasonally adjust the pre-break and post-break part of the series separately.

7.48. *Seasonal adjustment and consistency with annual data.* Annual totals based on the seasonally adjusted data will not automatically (and conceptually) be equal to the corresponding annual totals based on the original unadjusted data. The number of working days, the impact of moving holidays, and other calendar-related effects vary from year to year. Similarly, moving seasonality implies that the impact of the seasonal effect narrowly defined will vary from year to year. Thus, conceptually, for series with significant calendar-related effects or moving seasonality effects, the annual totals of a seasonally adjusted series *should differ* from the unadjusted series. In such cases, consistency with the annual series would be achieved at the expense of the quality of the seasonal adjustment and would be conceptually wrong.

7.49. However, in some particular cases like for national accounts or geographical breakdowns purposes, it may be necessary to maintain the additivity constraints in order to ensure consistency of data. In those cases, annual totals of the seasonally adjusted series must be "forced" to equal the annual total of the raw series. X-11-ARIMA and X-12-ARIMA provide options for forcing the annual totals from the seasonally adjusted data to be equal to the original totals.

7.50. *Revision policy and re-estimation of ARIMA models*. An important issue associated with model-based methods refers to how often the ARIMA models should be re-identified and re-estimated as new data become available. The stability of the models and their associated parameters depends on the nature of the series. In principle, the ARIMA models change slowly in time while their associated parameters are more sensible to new data. The recommended approach in such cases is to re-identify the models once per year and re-estimate the parameters every time seasonal adjustment is performed.

7.51. *Data dissemination and seasonal adjustment.* After removing seasonality and all calendar effects distributive trade data can be presented either in seasonally adjusted or trend-cycle form. The difference between them is the irregular component. In general, it is recommended that countries make available to users both the original and seasonally adjusted series. Dissemination of other series depends on users' interests and needs and country's capacity. Seasonally adjusted data, for example, are often considered more informative for uni-variate and multivariate purposes while trend-cycle data are in principle recommended for graphical representations and for series characterised by a high degree of volatility.

7.52. *Partially adjusted data*. Some countries publish as "non-seasonally adjusted data" data that have been adjusted for some seasonal effects, particularly the number of working days. It should be noted that partially adjusted data can be misleading and are of limited analytical usefulness. First, data presented as non-seasonally adjusted data should be fully unadjusted, showing what actually has happened, not partially adjusted for some seasonal effects. Working/trading-day effects are part of the overall seasonal variation in the series, and adjustment for these effects should be treated as an integral part of the seasonal adjustment process, not as a separate process. Second, working-day adjustments made outside the seasonal adjustment context are often conducted by simple methods such as using fixed coefficients based on the ratio of the number of working days in the month or quarter to the number of working days in a standard month or quarter which is not an appropriate procedure. Countries should make any effort to discontinue such practice. While still publishing the partially adjusted data, the appropriated explanatory note should be provided to warn users about limitations of such data. It is important that seasonal adjustment procedures used are appropriately documented and included in distributive trade statistics metadata.

D. Benchmarking

1. Need for benchmarking in distributive trade statistics

7.53. Common features of short-term statistics (monthly and quarterly) are their reduced scope, lower level of accuracy and details. For example, there may be differences in the coverage of units, degree of details of data items included, activity or geographical areas represented, etc. Thus, short-term statistics suffer more from bias in comparison to the more comprehensive annual data, arising from factors such as sampling error, differences in use of the business register (different versions of the register, grossing methods, reclassifications of establishments), different monthly/quarterly and annual accounting methods used by respondents. This section deals with the processes involved in producing optimal short-term distributive trade data consistent with annual ones. The general term for this is ‘temporal disaggregation’ and the common variants are benchmarking and interpolation

7.54. *Benchmarking* refers to the case where there are two sources of data for the same target variable, with different frequencies, and is concerned with correcting inconsistencies between the different estimates, e.g. differences between short-term and annual estimates of turnover of distributive trade sector derived from different sources. Benchmarking can be applied either to historical time series (distribution), where monthly/quarterly estimates have to add up at the end of the year to the new levels obtained by annual estimates or to forward series (extrapolation), where the time series are extended with the estimates for months/quarters for which no annual data are yet available.

7.55. *Interpolation* refers to the case where no genuine monthly (or quarterly) measurements exist, and annual totals are distributed across months (quarters), using a proxy indicator for the monthly (quarterly) pattern where possible, otherwise using a simple curve-fitting algorithm. A short-term pattern for interpolation may be derived from previous (discontinued) survey data, from proxy variables or as a smooth mathematical function.

7.56. *Objective of benchmarking.* The main objective of benchmarking is to combine the relative strengths of the low- and high-frequency data while preserving as much as possible the short-term movements in the source data under the constraints provided by the benchmarks. The low and high frequency data can be derived on the basis of conducted censuses or more accurate sample surveys or administrative data or even, some combination of these sources. It should be noted that the issue of benchmarking arises also with annual data, when a survey is only conducted every few years, and with monthly data, when they should be benchmarked to the quarterly estimates.

2. Main principles and methods of benchmarking

7.57. *Benchmark-to-indicator ratio*. For any year, it is important to have consistency between annual and infra-annual estimates of levels of any variable. To understand the relationship between the corresponding annual and monthly/quarterly data, it is useful to observe the ratio of the annual benchmark to the sum of the 12 months (or four quarters) of the indicator - the annual benchmark-to-indicator ratio (BI ratio). Usually, the value of BI ratio differs from 1 if estimates are obtained from data sources with different frequency. .

7.58. In the BI ratio framework the indicator determines the short-term movement of the estimates, while the annual data determine the overall level of the estimates and long-term movements. The BI ratio usually changes from year to year but the adjustments are necessary to be made so BI ratio becomes always 1.

7.59. The BI ratio may also be an expression of the relationship between sample survey levels and annual population levels. For example, a sample of trade enterprises is selected to cover 20% (*sm*) of population sales (*SM*) each month. So, for any month, population totals are estimated as *SM* = *sm* * 100 / 20. The BI ratio in this case is 100/20 or 5.0. But, the coefficient of 100/20 will become outdated as the sample becomes unrepresentative. So, when comprehensive data (*SA*) from an annual enterprise census become available it is likely that they will differ from the sum of twelve monthly sales (*SM*).

7.60. *Benchmarking methods*. There are two main approaches to benchmarking of time series - a purely numerical approach and a statistical modeling approach. The numerical approach differs from the statistical modeling approach by not specifying a statistical time series model that the series is assumed to follow. The numerical approach encompasses the prorating method and the family of least squares minimization methods, i.e. the family of proportional Denton methods. The statistical modeling approach encompasses ARIMA model-based methods and a set of various regression models[2]. The most commonly used numerical approach methods are briefly described below. Further guidance on the implementation of the benchmarking techniques will be provided in the forthcoming *Distributive Trade Statistics: Compilers Manual.*

7.61. *Pro-rata distribution method and the step problem*. For any benchmark year, for which annual estimates of a particular variable are available, the BI ratio can be calculated. Usually, it differs from 1 and to adjust the ratio for the bias, the annual level data are simply distributed according to the distribution of the monthly/quarterly values of the variable, i.e. pro-rata distribution across months/quarters. As a result, the pro-rating method preserves the proportional movement within each year. BI ratios for adjacent years, however, are different and pro-rata adjustment introduces a potentially large discontinuity between the last month (or quarter) of a year and the first month (or quarter) of the next year known as a "step problem". The pro-rata distribution technique is the simplest benchmarking method but because of the step problem, it is not recommended for the reconciliation of low- and high-frequency distributive trade data.

7.62. *Proportional Denton technique.* This technique is an integrated way of dealing with both aspects of benchmarking (distribution and extrapolation). The Denton family is based on the principle of movement preservation, which can be expressed as requiring that (i) the month-to-month (or quarter-to-quarter) growth in the adjusted monthly series and the original monthly (or quarterly) series should be as similar as possible; or (ii) the adjustment to neighbouring months should be as similar as possible. Basic requirement of the technique is that the calculation has to be based on the original monthly/quarterly indicator (not revised or seasonally adjusted). Usually the incorporation of new annual data for one year requires revision of previously published monthly/quarterly estimates because the adjustment for the bias in the indicator is spread over several periods, not just

[2] Detailed explanations on these methods, as well as an analysis of the available software for reconciliation can be found in the Eurostat *Handbook on Quarterly National Accounts* and the IMF *Quarterly National Accounts Manual – Concepts, Data Sources, and Compilation*

within the same year. In practice, the impact of the implementation of proportional Denton benchmarking technique becomes insignificant after 3 to 4 years.

7.63. The proportional Denton technique is relatively simple and well suited for large scale applications and as such is the recommended approach for dealing with benchmarking in distributive trade statistics. The key feature of this particular technique is that through the implementation of the least squares method it minimizes the month to month (or quarter to quarter) movements in the BI ratio between the benchmarked series and the indicator. This means that the method also smoothes the changes made to month/quarter to month/quarter growth in the indicator series and constructs a time series of monthly/quarterly benchmarked estimates-to-indicator ratios from annual observed BI ratios. The technique can particularly be used to avoid the step problem, i.e. the distortion in the monthly/quarterly time series associated with the implementation of pro-rata distribution method (see para. 7.57), mainly caused by the change of one BI ratio to another. The practical implementation of proportional Denton technique, however, requires application of special software.

3. Benchmarking and compilation of distributive trade statistics

7.64. Benchmarking should be considered by countries as an integral part of the compilation process of short-term distributive trade statistics and should be conducted at the most detailed compilation level. In practice, this may imply benchmarking different series in stages, where data for some series, which have already been benchmarked, are used to estimate other series, followed by a second or third round of benchmarking. The actual arrangements will vary depending on the particularities of each case. While undertaking the benchmarking of distributive trade data, countries may consider the following as a way of guidance:

- The estimates of one and the same variable produced with different frequencies should be consistent, so the users will not be confused;
- As soon as new annual data become available, the monthly/quarterly estimate should be aligned with them;
- The growth rates of the indicator series, should be preserved;
- The importance of good benchmarking methods increases in the cases when the quarterly indicators show considerable deviation from the annual data. In this relation, the consistency of estimates between the infra-annual and annual data sources should be reviewed, which may identify biases or other problems and lead to improved estimation and compilation practices for both sources;
- The benchmarking methods should be regularly reviewed;
- Mechanical methods for distributing the difference between the monthly/quarterly and annual estimates, such pro-rata distribution, should be avoided because they introduce steps between years;
- Improved accuracy for short-term statistics achieved through benchmarking may enable lower sample sizes and reduce costs and/or provide opportunities for improving timeliness.

7.65. *Benchmarking and revisions.* To avoid introducing distortions in the series, incorporation of new annual data for one year will generally require revisions of previously published data for several years in order to maximally preserve the short-term movements of the infra-annual series. This is a basic feature of all acceptable benchmarking methods. In practice however, with most benchmarking methods, the impact of new annual data will gradually diminishing to zero for sufficiently distant periods. As a practical recommendation, countries may allow at least, two to three preceding (and following) years to be revised each time new annual data become available.

7.66. *Benchmarking and quality*. A broader application of benchmarking techniques has a key role to play in improving the quality of distributive trade statistics. In fact, the fundamental characteristics of benchmarking closely relate to the dimensions of quality such as accuracy, timeliness and coherence. In the short-to-medium term, when resources are restricted and the capacity of statistical offices to expand data collection is limited, these techniques often succeed in filing the gaps of missing data and solving shortcomings. In the longer term where data quality to a large degree, depends on the availability and quality of basic data sources, benchmarking techniques can play an important role in optimizing the use of available data.

7.67. *Benchmarking and seasonal adjustment*. As it has been explain in the previous section benchmarking also occurs in the context of seasonal adjustment. Seasonally adjusting a monthly or quarterly time series can cause discrepancies between the yearly sums of the raw series and the corresponding yearly sums of the seasonally adjusted series especially, for series with significant calendar-related effects or moving seasonality. In order to fulfil some geographical or accounting constraints, such seasonally adjusted series should be benchmarked to the yearly sums of the raw series. As a general rule in this case, the benchmarking should be performed at the end of a survey cycle when data has been collected, processed and edited; and estimates are produced. The benchmarking process starts once the original estimates are available and the original time series of data are formed. In most cases, benchmarking is performed before the seasonal adjustment process, to “fine tune” the raw series that will be used as an input to the seasonal adjustment process. However, in some cases, benchmarking is performed on the seasonally adjusted data. For example, consistency with the annual data (for geographical or accounting reasons), can impose that benchmarking be performed within the seasonal adjustment process.