BENCHMARKING AND INTERPOLATION (CALENDARISATION)

Using Annual Benchmark Data to Align or Derive Quarterly/Monthly Estimates

Fenella Maitland-Smith, OECD

The paper has the following structure:

Introduction

This paper discusses the processes involved in producing optimal monthly/quarterly estimates consistent with annual data (or monthly with quarterly). The general term for this is 'temporal disaggregation' and the common variants of this are benchmarking and interpolation. The terms calendarisation and quarterisation are also used. The paper will deal mainly with benchmarking in relation to quarterly and annual national accounts estimates.

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. quarterly and annual estimates of value-added from different sources. It is essential that quarterly national accounts (QNA) are consistent with the annual accounts (ANA) so that a clear view of economic developments is presented - differences in growth rates between quarterly and annual GDP would confuse and irritate users. Benchmarking is generally done retrospectively as annual benchmark data are available some time after quarterly data. Benchmarking does have a forward-looking element however, in that the relationship between benchmark and indicator data (benchmark:indicator ratio) is extrapolated forward to improve quarterly estimates for the most recent periods for which benchmark data are not yet available.

Interpolation refers to the case where no genuine quarterly (or monthly) measurements exist, and annual totals are distributed across quarters (months), using a proxy indicator for the quarterly pattern where possible, otherwise using a simple curve-fitting algorithm. A common example of this is the quarterly estimation of NPISH output from annual data. A quarterly pattern for interpolation may be derived :

- from previous (discontinued) survey data
- from proxy variables
- as a smooth mathematical function.

It is important that seasonal factors are not simply applied to annual totals as this will lead to steps between years. Correct benchmarking techniques must be used – in other words, individual years must not be treated in isolation from their neighbours, rather the time series of quarterly data over several years should be processed using benchmarking software. Also, where a seasonal indicator series is used, it should be pre-adjusted (using seasonal adjustment software) to remove trend and cyclical components, i.e. the seasonal factors only are used.

The need for benchmarking

For any year, it is important to have consistency between annual and quarterly (or monthly) estimates of levels of any variable, e.g. GDP:

$$\sum_{i=1}^{4} Q^{i}GDP = AGDP$$

or rates of change:

$$\frac{\sum_{i=1}^{4} Q^{i,t}GDP}{\sum_{i=1}^{4} Q^{i,t-1}GDP} = \frac{A^{t}GDP}{A^{t-1}GDP}$$

Differences are inevitable, however, as a result of the different data sources used for QNA and ANA, but they must be reconciled.

Commonly, comprehensive (high accuracy) and detailed information is only available at a fairly low frequency, e.g., an annual or ten-yearly enterprise census, due to the expense of collecting and processing large volumes of data. Also, the data are often available only after a considerable lag because of the time needed for collection and processing. These comprehensive <u>benchmark</u> data may form the basis for the ANA but it is usually necessary to use more timely, but less complete, data for QNA (and even for the most recent ANA years in some cases), e.g., sample surveys of enterprises. These more frequent, timely data are generally less accurate but they give an indication of how the comprehensive data would behave if they were available.

The common feature of such <u>indicators</u> is that their coverage is less complete than is the case for the less frequent benchmark data, i.e. coverage of establishments, variables, commodities, industries or geographical areas. Thus, indicators suffer from bias in comparison to the more comprehensive 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 quarterly and annual accounting methods used by respondents, and respondent error. Other causes of bias are weak assumptions about the relationships between proxy indicators and target variables (e.g. constant IO ratios), and failing to account for quality change or changes in product mixes.

As a result of indicator bias, quarterly source data are viewed as serving only to determine short-term movements (quarterly path), whereas annual data determine the overall levels and the medium- to long-term trends. Chart 1 shows differences in medium term trends shown by annual and quarterly data. Thus, quarterly data and QNA estimates are adjusted as necessary so that they correspond with the trends shown by annual estimates, i.e. once benchmark (annual) data become available, the indicators need to be brought into line (benchmarked) with the long-term trends shown by the benchmarks. Indicators will also need to be subsequently revised in line with any later revisions to the benchmark data.

Some of the causes of bias are likely to have systematic and thus predictable effects so that it is possible to make adjustments each quarter, i.e. "real-time" adjustments. Of course, it would be better to determine the causes of the bias and remedy them, but this type of analysis may be too expensive to carry out in practice. However, it will inevitably be the case that these "real-time" adjustments will not be sufficient, and that once the benchmark data are available retrospective adjustment of the quarterly data will also be needed.
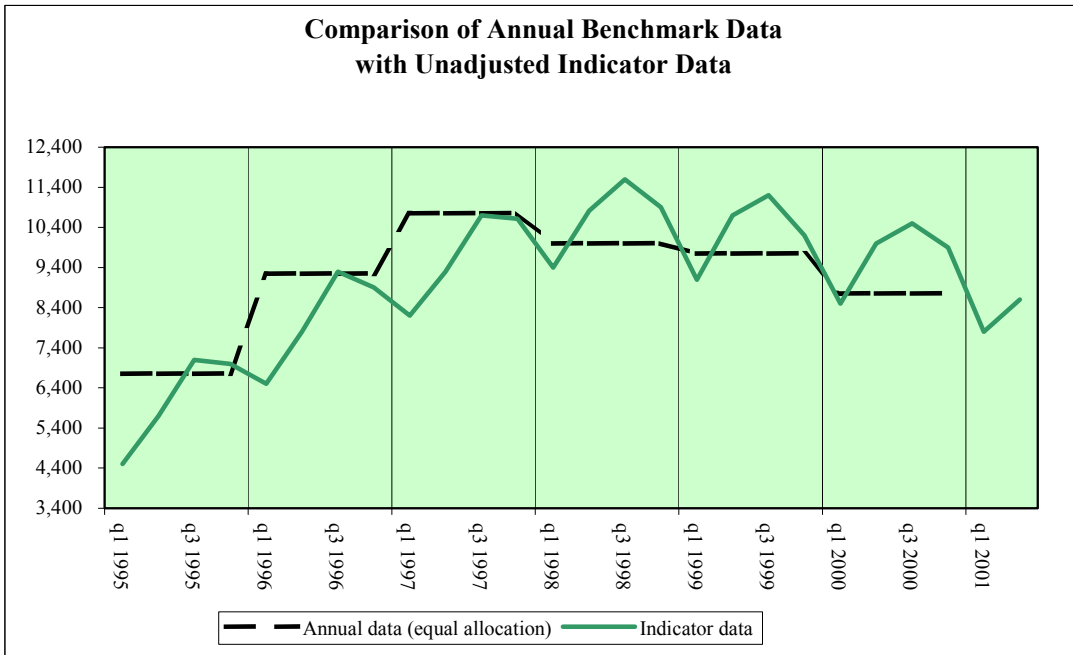
Chart 1.

The benchmark to indicator (BI) ratio

The benchmark to indicator (BI) ratio for any benchmark period is a measure of indicator bias, and adjustments are made so that the BI ratio becomes one.

$$Adjust \quad Q^i GDP \quad \Rightarrow \quad \frac{AGDP}{\sum_{i=1}^{4} Q^i GDP} = 1$$

The BI ratio usually changes from year to year, as shown in Chart 2 where the large change between 1997 and 98 should be noted for future reference.
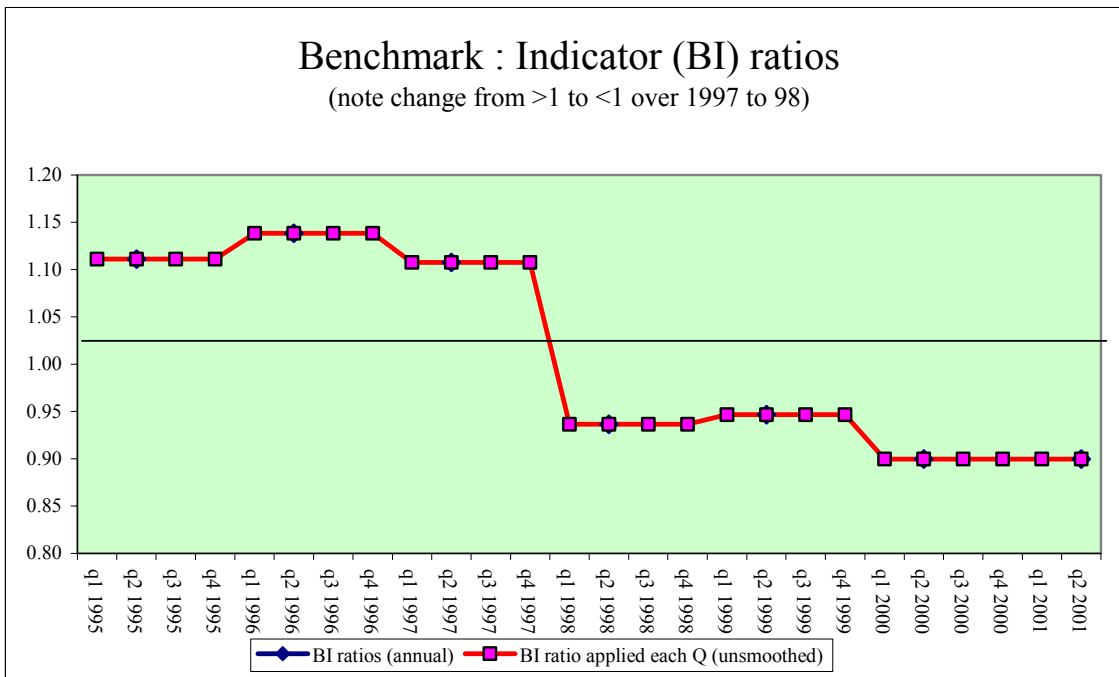

Chart 2.

The BI ratio may also be an expression of the relationship between sample survey levels and annual population levels. For example, a sample of 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 is therefore 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:

$$\sum_{i=1}^{12} SM^i \neq SA$$

$$\sum_{i=1}^{12} sm^i *(100/20) \neq SA$$

$$\frac{SA}{\sum_{i=1}^{12} sm^i} \neq (100/20)$$

So, although it was assumed at the start of the year that:

$$\frac{SA}{\sum_{i=1}^{12} sm^i} = (100/20)$$

In fact, benchmark data show that, for example:

$$\frac{SA}{\sum_{i=1}^{12} sm^i} = (100/15) = 6.67$$

So the BI ratio for the year in question is 6.67, and not 5.0

Pro-rata distribution and the 'step problem'

For any benchmark period (e.g. a year for which ANA are available) the BI ratio can be calculated. Since this can be viewed as a measure of the indicator bias it would be tempting to simply adjust for the bias by distributing the annual level data for each year according to the distribution of the quarterly indicator, i.e. some means of pro-rata distribution across quarters.

$$Adjusted \ Q^iGDP^* = \frac{Q^iGDP}{\sum_{i=1}^{4} Q^iGDP} \ x \ AGDP$$

This would be fine in cases where the BI ratio is constant from year to year. If, however, BI ratios for adjacent years are different, and pro-rata adjustments are used in each year, a discontinuity in the growth rate from the last quarter of one year to the first quarter of the next will be introduced. This is known as the "step problem". In other words, the growth between Q4 of one year and Q1 of the next will reflect the change in the BI ratio (adjustment) between the years as well as any genuine growth between Q4 and Q1.

This can be seen in Chart 3 and Table 1 where the distortion (step) of the quarterly path between Q4 97 and Q1 98 should be noted
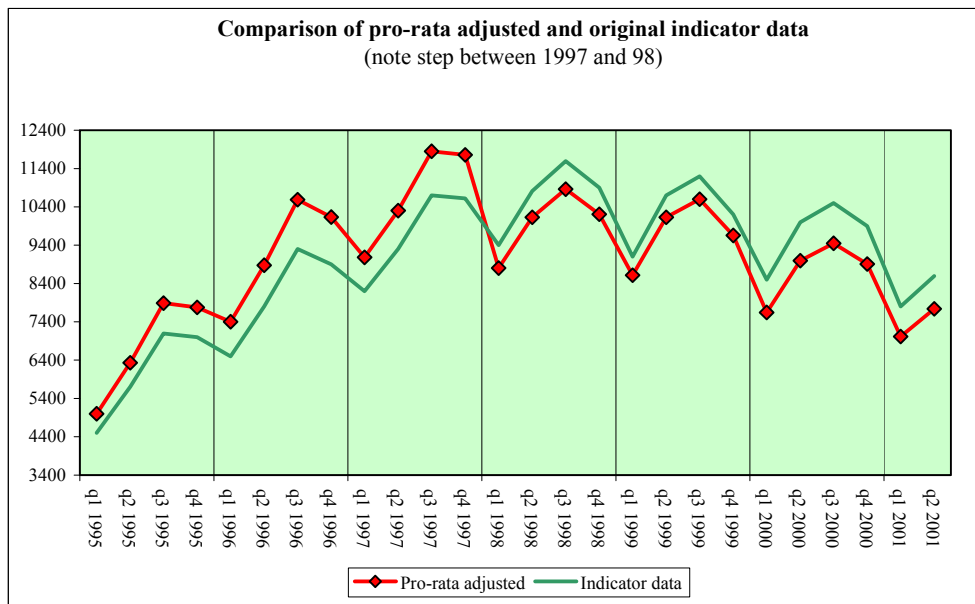

Chart 3.

Table 1 also highlights the fact that all the correction for bias in the growth rate of the indicator data is concentrated in the Q4-Q1 growth rate:

	Indicator data	Pro-rata adjusted	Difference in growth rate		Indicator data	Pro-rata adjusted	Difference in growth rate
q1 1995				**q1 1999**	-16.5	-15.6	*0.9*
q2 1995	26.7	26.7	*0.0*	q2 1999	17.6	17.6	*0.0*
q3 1995	24.6	24.6	*0.0*	q3 1999	4.7	4.7	*0.0*
q4 1995	-1.4	-1.4	*0.0*	q4 1999	-8.9	-8.9	*0.0*
q1 1996	**-7.1**	**-4.9**	**2.3**	**q1 2000**	**-16.7**	**-20.8**	***-4.1***
q2 1996	20.0	20.0	*0.0*	q2 2000	17.6	17.6	*0.0*
q3 1996	19.2	19.2	*0.0*	q3 2000	5.0	5.0	*0.0*
q4 1996	-4.3	-4.3	*0.0*	q4 2000	-5.7	-5.7	*0.0*
q1 1997	**-7.9**	**-10.3**	***-2.5***	q1 2001	-21.2	-21.2	*0.0*
q2 1997	13.4	13.4	*0.0*	q2 2001	10.3	10.3	*0.0*
q3 1997	15.1	15.1	*0.0*				
q4 1997	-0.8	-0.8	*0.0*				
q1 1998	**-11.5**	**-25.1**	***-13.7***				
q2 1998	15.0	15.0	*0.0*				
q3 1998	7.3	7.3	*0.0*				
q4 1998	-6.0	-6.0	*0.0*				

Table 1.

Benchmarking techniques to avoid the step problem

The step problem is solved by smoothing the changes in BI ratios. In other words, the BI ratios themselves are treated as a quarterly time series which initially appears as a series of steps with each year (step) consisting of four equal data points. This stepped series is smoothed, to give a series made up of BI ratios (see Chart 4) which are then applied to the indicator data to give a smoothly benchmarked series (Chart 5).
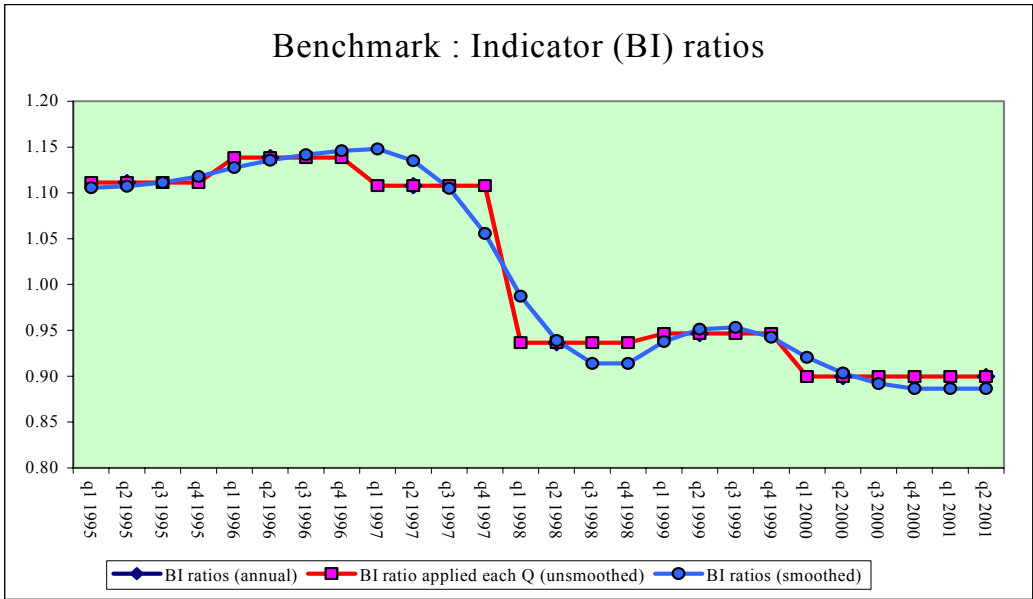
Chart 4.

So smoothing methods aim to produce a series of quarterly BI ratios that change smoothly from one period to the next, whilst averaging to the annual BI ratios within each year. They try as far as possible to maintain the original quarterly growth rates (quarterly path) under the constraints of a smooth transition between years (Q4-Q1 growth) and the need for the annual totals to reconcile (ÿQGDP=AGDP).

All quarterly growth rates will be adjusted by gradually changing but relatively similar amounts. This can be seen in Chart 5 and Table 2 where the series has been benchmarked using the Denton least squares approach. Note that the change in BI ratio between 1997 and 98 is accommodated throughout all quarters in 1997 and 98 rather than in a single step at Q4 97 – Q1 98 as in the pro-rata approach.
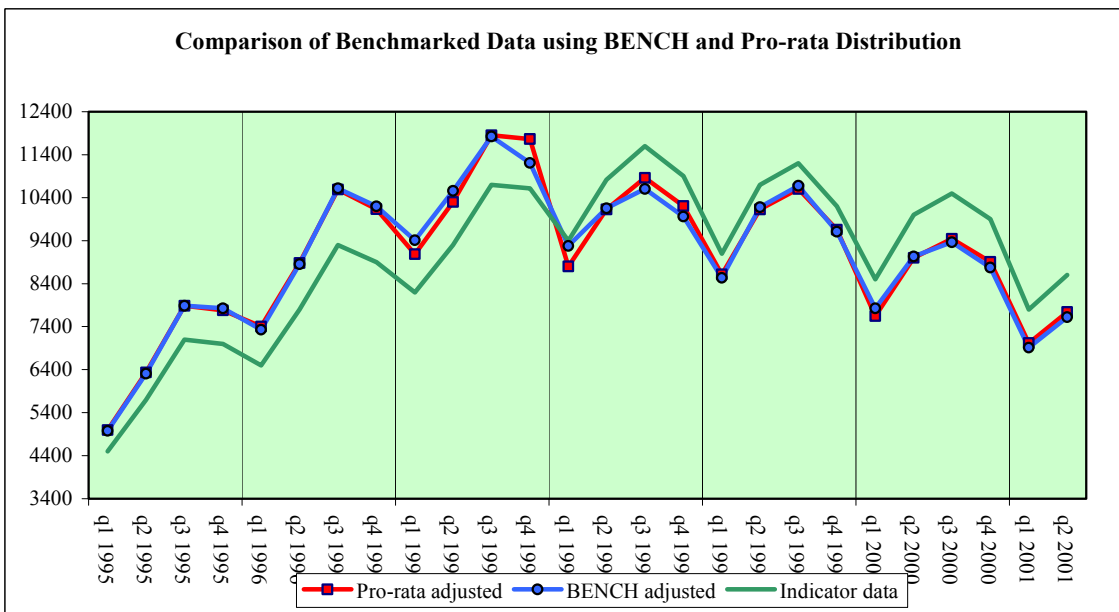

Chart 5.

	Indicator data	BENCH adjusted	*Difference in growth rate*		Indicator data	BENCH adjusted	*Difference in growth rate*
q1 1995				**q1 1999**	-16.5	-14.4	***2.2***
q2 1995	26.7	26.9	*0.2*	q2 1999	17.6	19.3	***1.7***
q3 1995	24.6	25.0	*0.4*	q3 1999	4.7	4.9	*0.2*
q4 1995	-1.4	-0.8	*0.6*	q4 1999	-8.9	-9.9	***-1.0***
q1 1996	-7.1	-6.4	*0.8*	**q1 2000**	-16.7	-18.6	***-1.9***
q2 1996	20.0	20.9	*0.9*	q2 2000	17.6	15.5	***-2.2***
q3 1996	19.2	19.9	*0.7*	q3 2000	5.0	3.7	***-1.3***
q4 1996	-4.3	-4.0	*0.3*	q4 2000	-5.7	-6.3	*-0.6*
q1 1997	-7.9	-7.7	*0.2*	**q1 2001**	-21.2	-21.2	*0.0*
q2 1997	13.4	12.1	***-1.3***	q2 2001	10.3	10.3	*0.0*
q3 1997	15.1	12.0	***-3.0***				
q4 1997	-0.8	-5.2	***-4.4***				
q1 1998	-11.5	-17.2	***-5.8***				
q2 1998	15.0	9.4	***-5.6***				
q3 1998	7.3	4.4	***-2.8***				
q4 1998	-6.0	-6.0	*0.0*				

Table 2.

Extrapolation

Benchmark data will not be available for current and recent periods. For example:

1. in May 2001, there are no benchmark data for months in 2000 and 2001 (annual enterprise census data for 2000 become available in June 2001),
2. so, BI ratios of years pre-2000 should be used in 2000 and 2001 monthly estimates, i.e. the relationships are extrapolated
3. then, in June 2001, the 2000 benchmark data become available
4. and the resulting BI ratio for 2000 is extrapolated and used for the months of 2001.

Extrapolation can be viewed as the linking of quarterly source data onto previous annual estimates, or, as constructing forward series by adjusting the last available benchmark level according to movements in an indicator.

For forwards extrapolation, the BI ratio from the last benchmarked quarter may be applied to all new quarterly data, i.e. the BI ratio is projected forwards 'flat' and applied to each new quarterly indicator estimate in "real time". This can have the undesirable result, however, that the annual BI ratio for this extrapolated year is different to the average annual BI ratio in the previous year. If there is a large amount of noise in the indicator's annual rate of change relative to systematic bias, there is danger of a "wagging tail effect".

So, in the more sophisticated models (enhanced Denton), the most recently available BI ratio is not necessarily extrapolated directly (flat), but a series of rising/falling BI ratios is forecast for the current and future years, based on the time series properties (error structure) of the BI series. For example, if the annual growth rate of the indicator is systematically biased, then the BI ratio series shows a trend, and the best forecast for the next year's BI ratio would be obtained using trend extrapolation of the BI ratio series. Thus the BI ratio is not projected flat during the operational phase, but rising or falling in line with its behaviour in previous years. Chart 6 gives an extremely simplified example:
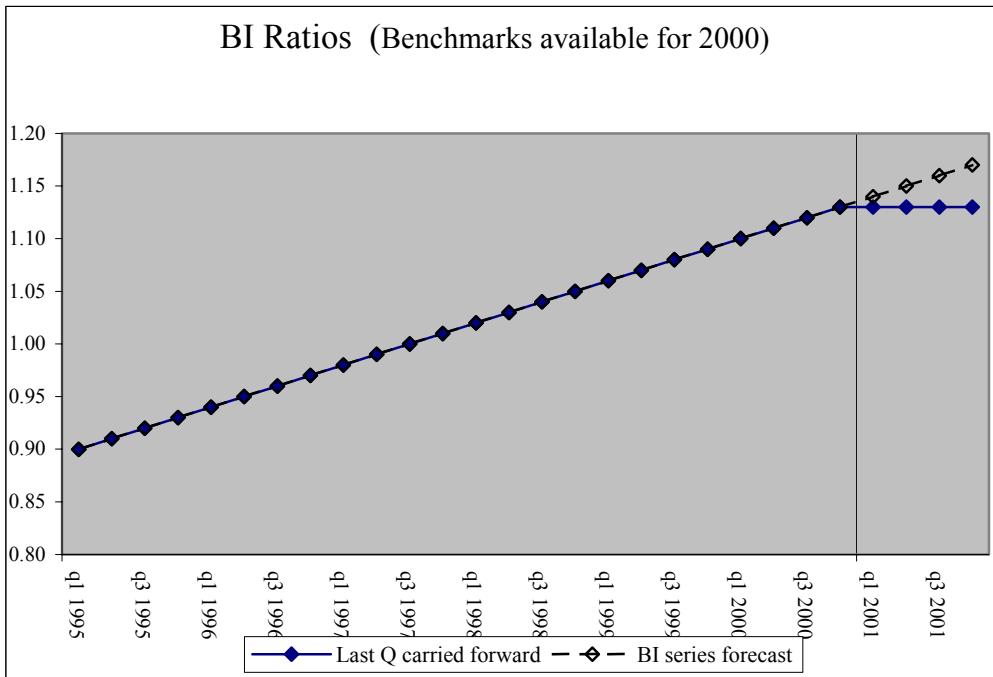
Chart 6.

So, in fact, effective benchmarking has both a retrospective and a forward-looking dimension, derived from common principles, and combined in the more sophisticated QNA software applications. In other words, extrapolated series generally need to be retrospectively benchmarked at a later date. Thus, two phases can be envisaged:

1. In the operational phase, there will be no annual benchmarks for the most recent quarters. So, the challenge is to extend the series beyond the last benchmark period, anticipating future ANA estimates so that future revisions are minimised, whist preserving, as far as possible, the short-term movements in the quarterly source data. Most extrapolation techniques are based on the idea that the last BI ratio based on actual data is projected forwards (flat) for each quarter and applied to each new quarterly estimate in real time. Further real-time adjustments are made to quarterly data if and when particular bias issues arise.

2. Following the operational phase, quarterly data will be treated retrospectively and subjected to continuing cycles of revisions. These revisions will arise from
    ▪ the arrival of annual benchmarks for the most recent year
    ▪ revisions to the quarterly source data
    ▪ revisions to the annual benchmarks of the previous year
Any new information needs to be incorporated into the QNA estimates as quickly as possible.

In order to satisfy the need for constant yet smoothed, updating and reconciliation, sophisticated models, such as the proportional Denton technique are available. These provide an integrated means of dealing with extrapolation, alignment and updating and give superior results to methods that treat these phases separately.

Denton, and other smoothing methods

This smoothing of the BI ratios can be achieved to some extent using straightforward techniques such as a Henderson moving average. Another alternative is to use a least squares method to minimise the difference in the first difference between the pro-rata adjusted quarterly series and the original quarterly

series, subject to the constraint that the sum of the quarters equal the annual estimate [1]. This method is applied to estimates for two years each time new annual estimates become available, so each year is adjusted twice.

The more sophisticated approaches to benchmarking may be classified as purely numerical, or statistical modelling, approaches. The numerical approach does not specify a time series model for the series, but instead uses least squares minimisation, e.g. the methods proposed by:

- Denton and others (1971) – least squares methods to minimise the difference between the pro-rata adjusted quarterly series and the original quarterly series, subject to the constraint that the sum of the quarters equal the annual estimate, The additive Denton version (D1) minimises (least squares) the absolute differences in absolute adjustments to neighbouring quarters subject to the constraints of the benchmarks (binding or non-binding). The proportional Denton version (D4) minimises the absolute differences in relative adjustments to neighbouring quarters (better when multiplicative seasonality present). D4 requires positive values only (so, pre-adjustment may be needed),
- Bassie (1958) – a stepwise benchmarking method, working on two consecutive years at a time (otherwise steps are introduced). It does not allow extrapolation.

The statistical modelling approaches include:

- ARIMA model based methods (Hillmer and Trabelsi, 1987),
- state space models (Durbin and Quenneville, 1997),
- series of generalised least squares regression models, e.g. BENCH (Statistics Canada). These take account of information about the stochastic properties of the error-generating process (ARMA),
- Chow and Lin (1971) have proposed a multivariate additive generalised least squares regression approach with binding benchmarks, i.e. several indicators are related to a single benchmark. The result is not benchmarking in the strict sense, but rather an approach for interpolation, distribution and extrapolation (Eurostat's ECOTRIM follows the same approach).

For good indicators, the results are fairly insensitive to the choice of technique, but where significant bias exists different techniques are better suited to differently behaved biases, i.e. it is important to understand the nature of the bias.

A detailed explanation of the theory behind the Denton technique and other methods can be found in the IMF's draft *Handbook on Quarterly National Accounts Compilation*[2].

Revisions to published data

When benchmark data for the most recent year become available, that year will be retrospectively benchmarked, and quarters in the year before (and maybe for several years before that) will also be revised. This is because smoothing the series of BI ratios may require trend adjustments stretching back more than four quarters. In other words, when new annual (year t-1) data become available, quarters in t-1 are benchmarked, and quarters in t-2, t-3 are also revised as required by revision of the smoothing function. In general, the best results will be obtained if the entire time series is revised each time new annual benchmarks become available. However, with the recommended proportional Denton technique, the impact on data for proceeding years will gradually become smaller and smaller, and will normally become insignificant after three to four years.

---

[1] This is the method used by Statistics Sweden.

[2] The draft can currently be viewed at: http://www.imf.org/external/pubs/ft/qna/2000/Textbook/index.htm

For the aligned back series, quarter to quarter growth rates will differ from those of the source data, and in extreme cases, new or different turning points may be introduced. This should be viewed as a necessary consequence of the reconciliation process, and and users should be educated to accept this. For the forward series, quarter to quarter growth rates will be the same as those of the source data, but the annual rate will be different, as expected.

Thus, the level and movements of final QNA estimates will depend on:

- Movements in short-term indicators, which determine the quarterly path/pattern of QNA,
- Levels of ANA for the current year, which, through retrospective benchmarking, determine the sum of the quarterly levels,
- Levels of ANA estimates for several preceding and following years, which through retrospective benchmarking, determine the trend.

It should be emphasised that in the case of incorporation of revised or new benchmarks, the calculations should be based on the original non-seasonally adjusted quarterly indicator, not on the preliminary QNA estimates that have already had adjustments made to them. Otherwise, the compilation process risks deteriorating into data hashing, in which the compilers lose track of the original data, the effects of benchmarking, and the effects of other adjustments.

Conclusions and recommendations
- QNA estimates should be aligned with ANA estimates, as soon as ANA data become available.
- QNA estimates should be revised in line with subsequent revisions to ANA data.
- Benchmark all source data before NA compilation, to reduce later benchmarking of NA aggregates.
- Benchmarking is best done in a supply-use framework to preserve balance between GDP O, E and I.
- It is better to seasonally adjust before benchmarking to avoid benchmarking distortions to seasonal patterns.
- Except in the extraordinary case where BI ratios are constant from year to year, it is important to use a system which smoothes the transition between years. Do not simply distribute annual data according to quarterly patterns – when BI ratios change from year to year, such pro-rata adjustment introduces steps between years.
- Smooth BI ratio series using standard time series techniques or integrated models.
- Various benchmarking applications are available which provide an integrated approach to retrospective alignment and forward extrapolation. It is worth spending time selecting and setting up a system correctly, as different models are better suited to data with different time series characteristics.
- If possible, forecast BI ratios for current quarters based on recent trends in the ratios.
- The importance of good benchmarking methods increases as quarterly indicators show more divergence in movements from annual data.