A Fresh View on the Outcome-Testing Approach versus Price-Setting Approach: Discussion and Recommendations Regarding the Timing of Benchmarking Studies

The authors discuss the practical implications of the outcome-testing approach and the price-setting approach, and, based on an example, offer some suggestions and recommendations regarding the wording of paragraphs 3.67 to 3.71 of the OECD Guidelines (2010) for setting and testing transfer prices in practice, with the aim of minimizing the risk of double taxation.

1. Introduction

The issue of outcome testing versus price setting is certainly very topical, but what is this really about? In its letter dated 6 June 2012,1 OECD Working Group No. 6 requested that companies comment on the proposed changes regarding the following topics, especially with reference to their practicability (feasibility of practical implementation):

- the point in time of the review of the appropriateness of transfer prices;
- the point in time of the review of the appropriateness of transfer prices.

Many businesses and other stakeholders responded to this request. The collection of their comments comprises 135 pages and was released on 29 October 2012.2 A major finding of the OECD was that member countries have different approaches with respect to the timing of the application of the arm’s length principle, namely:

- **ex ante approach** (price-setting approach). In this case, the arm’s length test is applied at the moment the transfer prices are determined. Such transfer prices may not be adjusted retrospectively. Therefore, the formation of transfer prices, in this case, is based on information and data which are available at the moment of carrying out the transaction (historical and/or forecast). The arm’s length data always derive from earlier periods than the year in which the transaction actually takes place; and

- **ex post approach** (outcome-testing approach). The arm’s length test is carried out ex post (in retrospect). In this case, the result of the intercompany transactions is tested with respect to available arm’s length data. According to the OECD this should be done at the end of the financial year or when filing the tax return. The arm’s length data are more recent than that of the ex ante approach, and ideally comes from the period actually being tested.

This article will provide a brief discussion of the practical implications of these diverging approaches and, based on an example, offer some suggestions and recommendations regarding the wording of paragraphs 3.67 to 3.71 of the OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations (OECD Guidelines) (2010) for setting and testing transfer prices in practice, with the aim of minimizing the risk of double taxation.

2. Consequences of the Differing Approaches

The existence of these two approaches leads to a number of difficulties or questions in practice. These include:

- which is the "right" moment to collect arm’s length data so that both companies and tax authorities can rely on a uniform, reliable and verifiable data basis;
- which is the "right" moment for companies to adjust transfer prices, in other words during or at the end of the financial year, when filing the tax return, etc. Will such (positive or negative) adjustments to transfer prices be accepted by all of the member countries;
- is it allowed or appropriate to rely on ex post data, i.e. information and data that are available solely after the end of the financial year or after the point in time of the formation of transfer prices, in order to check the appropriateness of the "old" periods, transactions, budgets or forecasts;
- how does each approach influence business-steering purposes? Which enterprise resource planning (ERP) requirements and investments would be necessary for each approach? Which internal process costs are related to the implementation of each approach;
- as IP valuation is always subject to uncertainty, the particular interest of tax authorities to conduct subsequent, retrospective price adjustments becomes visible (for example in Germany, unfortunately, only in a one-sided manner, to the detriment of the taxpayer); and

---

ultimately, different interpretations of the member countries regarding the time reference lead to double taxation and, subsequently, to mutual agreement procedures or arbitration proceedings.

In this regard, it has been proposed to change the paragraphs of the OECD Guidelines (2010) as follows:

3.67 **There are timing issues that arise in comparability** with respect to the time of origin, collection and production of information on comparability factors and comparable uncontrolled transactions that are used in a comparability analysis. See paragraphs 5.3, 5.4, 5.5, 5.9 and 5.14 of Chapter V for indications with respect to timing issues in the context of transfer pricing documentation requirements.

3.68 **In principle, information relating to the conditions of comparable uncontrolled transactions undertaken or carried out at the same time** during the same period of time as the controlled transaction (‘comparative uncontrolled transactions’) is expected to be the most reliable information to use in a comparability analysis, because such information reflects how the prices and other conditions that independent parties have established in an economic environment that is the same as the economic environment of the taxpayer’s controlled transaction. It is therefore important in conducting a transfer pricing analysis to identify and use information that is as contemporaneous as possible to the controlled transaction, taking into account practical constraints created by the availability and collection of data. Availability of information on contemporaneous uncontrolled transactions may however be limited in practice, depending on the timing of collection.

3.69 In some cases, taxpayers establish transfer pricing documentation to demonstrate that they have made reasonable efforts to comply with the arm’s length principle at the time the transactions were undertaken, i.e. on an *ex ante* basis (hereinafter “the arm’s length price-setting approach”), based on information that was reasonably available to them at that time. Where such an approach is followed, pricing determinations should be based on information that was known or reasonably foreseeable by the associated enterprises at the time the transaction was entered into. Such information includes not only information on comparable transactions from previous years, but also information on economic and market changes that may have occurred between those previous years and the time at which the controlled transaction was undertaken, as well as information on economic and market changes likely to occur after the time the transaction was undertaken that could have reasonably been anticipated at the time the transaction was undertaken and that would have affected the pricing that would have been agreed between independent enterprises in similar circumstances. In effect, independent parties in comparable circumstances would not base their pricing decisions on historical data alone.

3.70 In other instances, taxpayers might test the actual outcome of their controlled transactions to demonstrate that the conditions of these transactions were consistent with the arm’s length principle, i.e. on an *ex post* basis (hereinafter ‘the arm’s length outcome-testing approach’). Where such an approach is followed, pricing confirmations should be based on information available at the time the tax return is prepared, providing such information is related to the outcome of comparable uncontrolled transactions undertaken at the same time as the controlled transaction, or prior to the controlled transaction, where suitable comparability adjustments are made to reflect economic changes that occurred between the time that the comparable transaction was undertaken and that at which the tested transaction was undertaken. Such a test typically takes place as part of the process for establishing the tax return at year-end.

3.71 Both the arm’s length price-setting and the arm’s length outcome-testing approaches, as well as combinations of these two approaches, are found among OECD member countries. The issue of double taxation may arise where a controlled transaction takes place between two associated enterprises where different approaches have been applied and lead to different outcomes, for instance because of a discrepancy between market expectations taken into account in the arm’s length price-setting approach and actual outcomes observed in the arm’s length outcome-testing approach. See paragraphs 4.38 and 4.39. Competent authorities are encouraged to use their best efforts to resolve any double taxation issues that may arise from different country approaches to year-end adjustments and that may be submitted to them under a mutual agreement procedure (Article 25 of the OECD Model Tax Convention).³

Even after taking into account the proposed changes to the OECD Guidelines (2010), one can conclude that neither these changes, nor the current OECD Guidelines, nor e.g. section I of the German Foreign Tax Act (Aussensteuergesetz) fully clarify at what point in time the transfer pricing analysis should be performed and on the basis of what information.

Also, the very extensive, published comments⁵ by enterprises and consulting firms regarding the above-mentioned request do not show a clear preference for the one or the other approach. This is not really surprising because the interests of individual companies, as well as those of member countries, diverge. In other words, the question as to the most appropriate concept depends on several parameters, such as industry specifics; actual and individual arm’s length behaviour of corporate groups; restrictions and possibilities of the IT landscape; available human and systemic resources; number of products; and volatility of commodity prices, production costs or retail prices.

### 3. An Example

Despite the diverse interests and general uncertainty, in practice taxpayers need to choose and apply an approach in their day-to-day business. Accordingly, this article will address the timing of a transfer pricing analysis by reference to an example in order to formulate a recommendation that is both practicable and mitigates the risk of double taxation.

**Figure 1**

- **US Inc.**
- **DE GmbH (LRD)**

<table>
<thead>
<tr>
<th>Results DE GmbH (before transfer pricing adjustments)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiscal Year</td>
</tr>
<tr>
<td>Actual EBIT margin</td>
</tr>
</tbody>
</table>

3. OECD Ctr. for Tax Policy and Admin., supra n. 1.
4. OECD, *Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations (OECD Guidelines)* (OECD 2010), paras. 3.67 to 3.71; *International Organizations’ Documentation* IBFD.
5. OECD Ctr. for Tax Policy and Admin., supra n. 2.
The following points and the illustration in Figure 1 provide an overview of the facts regarding possible timings in the example.6

Since 2002, DE GmbH (a limited liability corporation) has acted as a low-risk distributor in the German market. The distribution agreement between the US-based entrepreneur, US Inc., and DE GmbH states that prices are determined based on the transactional net margin method (TNMM). The following events took place in the specified quarters between 2011 and 2015:

– Q4/2011: creation of benchmarking study;
– Q4/2012: update of benchmarking study;
– Q1/2013: preparation of the financial statement 2012;
– Q4/2013: filing of tax return 2012; and
– Q1/2015: tax audit for 2012.

Figure 2 presents chronologically the conceivable points in time for the arm’s length test with respect to 2012 and the interquartile ranges derived from database analyses (i.e. a benchmarking study) that are available at each of these times. In this example, a benchmarking study was performed in Q4/2011 resulting in an interquartile range of 2% (lower quartile) up to 4% (upper quartile), with a median of 3%. Subsequently, the transfer prices were set in Q4/2011 for fiscal year 2012 based on a target EBIT7 margin of 3%.

Consideration of this example leads to the following key questions:

(1) Must the taxpayer adjust its transfer prices 2012?

(2) In Case B (actual EBIT margin = 6%), the US tax auditor has imposed a transfer pricing adjustment equal to the difference between 3% and 6% based on the benchmarking study which was prepared in Q4/2011. Are there corresponding adjustments available to DE GmbH in Germany?

(3) How precisely should the transfer pricing method, the target EBIT margin, the point in time of the preparation of the underlying benchmarking study and a possible transfer pricing adjustment mechanism be regulated in the intercompany contract?

(4) Is it necessary (mandatory) to regulate contractually whether the outcome-testing approach or the price-setting approach should be applied?

Accordingly, the discussion below addresses each of these questions.

As a general remark, the TNMM, as well as the use of databases, to find arm’s length margins is accepted by the OECD and most non-OECD-member countries and is very often agreed upon in advance pricing agreements (APAs). A price-setting approach requires reliable forecasting. Otherwise, it is unrealistic to believe that prospective price adjustments during the fiscal year would lead to actual margins of a fiscal year which fall within the target range of arm’s length margins.

With regard to Issue (1), whether a transfer pricing adjustment is required in this example is dependent upon which benchmarking study should be applied when making this assessment. As outlined in the illustration above, there are a number of options. If one applies Case A (actual EBIT margin is equal to 2%), the application of any of the later benchmarking studies would mean that the profit actually achieved was outside the arm’s length range. Similarly with Case B (actual EBIT margin is equal to 6%), this actual result is not within all ranges, even though, in both cases, the target EBIT margin was within the range available at the time the target was set. Accordingly, the following paragraphs consider each of the possible timings for the benchmarking study in turn.

– Q1/2015: At the date of the tax audit, at least the use of comparable data from the year in question (fiscal year 2012) would be possible. However, such a use of hindsight ought to be rejected because otherwise there would be unequal situations for tax authorities on the one hand and taxpayers on the other. Such an approach effectively means that the tax authorities are
evaluating transfer prices based on information that was quite simply not available to the taxpayer at the time of the transaction itself.

Q4/2013: Referring to the date of the tax return filing also ought to be rejected, as the preparation and filing deadlines differ wildly throughout the world. Therefore, deadlines are different from country to country. Accordingly, it is quite possible that different countries would apply different benchmarks based on different time periods. This in turn leads to a risk of double taxation when the positions of the tax authorities are not consistent.

Q1/2013: Focussing on the date of the final statutory financial statements may be possible, as this is the last point in time where a company could perform a (retrospective) transfer pricing adjustment entry. On the one hand, the comparable data would be newer by roughly a year compared to the following options, but, on the other hand, there would still be no comparable data available from the actual year in question (here, fiscal year 2012). Further, this approach would require defining whether the date of the final group (later) or local statutory financial statements (earlier) of the involved related parties counts. Overall, similar argumentation applies if the time of the preliminary statutory financial statements in Q4 2012 were considered. Accordingly, this approach seems to be possible, although second best in comparison to the point in time discussed immediately below.

Q4/2011: Using the date when the transfer prices were determined (i.e. prior to the start of the financial year) appears to be the most logical and reasonable option. The definition of target EBIT margins is technically possible on the basis of ex ante data, it is arm’s length (in the sense that it mimics the behaviour of third parties that typically set prices and let them run) and it is verifiable retrospectively for all participants (taxpayers and tax authorities) by means of databases.

Following the above analysis, in Case A there should not be a necessity for any adjustment because the actual margin of 2% falls inside the range of the benchmarking study performed in Q4/2011 (2% to 4%). In Case B, an adjustment of the actual margin of 6% would be necessary in order to achieve an actual value that falls inside the range of 2% to 4%. In theory, such adjustments could be made prospectively during the year (ex ante) or with the preparation of the financial statements (retrospective transfer price adjustment). However, in practice – at least from a German or European perspective, the prospective approach may meet with a more ready acceptance by the tax authorities. Still, both options are equally available.

With regard to Issue (2): Subject to the wording of the distribution contract, following the logic above, the point in time of creation of the benchmarking study should be Q4/2011. This produces a situation where the actual profitability of DE GmbH is above the determinative arm’s length range. Therefore, the US tax auditor could impose a transfer pricing adjustment, for example, equal to the difference between 3% and 6%. Subsequently, DE GmbH would apply for a mutual agreement procedure in order to achieve a corresponding correction of the target EBIT margin in Germany. Because of the fact that the income tax treaty requires that the two countries come to an agreement, double taxation would be eliminated (with a bit of patience).

With regard to Issues (3) and (4): In general, there is no requirement that contracts be in written form in Germany. Nevertheless, from a practical perspective it is strongly advisable to set forth the essential parameters of the remuneration in written form. The more wiggle room allowed by the wording of the contract, the larger the risk of double taxation in the end. In this regard, the following remuneration clause would surely no longer be sufficient: “The parties agree to transfer prices which are consistent with the internationally agreed arm’s length principle”. According to the view taken here, the remuneration clause should point out clearly and consistently which transfer pricing method is applied and, especially, whether and exactly how a transfer price adjustment mechanism works. It is advisable to include a detailed definition of sales, costs, margins and accounting standards. Moreover, it should be clarified at which point in time the arm’s length test should be applied, i.e. at which point in time a benchmarking study should be created for the margin test. It should not be necessary to state the ‘outcome-testing approach’ or the ‘price-setting approach’, as both terms are too vague and therefore prone to interpretation.

4. Conclusion

The outcome-testing approach means that the actual margin (exaggerated, independently of how it was achieved) is validated ex post with arm’s length data. The draft of paragraph 3.70 of the OECD Guidelines proposes the use of arm’s length data available at the moment of filing the tax return. Following the opinion represented here, this would still lead to an unnecessarily high risk of double taxation. This approach should be rather based on arm’s length data which were available before the beginning of the financial year in consideration.

The price-setting approach assumes that the transfer price has been determined before the beginning of the financial year in consideration (based on benchmarking data available prior to the beginning of the fiscal year) and (perhaps) has been adjusted ex ante during the year. For this reason, the actual margins of routine functions may fluctuate significantly. A consequence of this fact could be that tax authorities in both the state of the principal and

---

8. PwC, Worldwide Tax Summaries: Corporate Taxes 2013/2014 (for the 150 most important countries). In Germany, for example, a fiscal year 2012 corporation tax return should be filed by 31 December, 2013 (if a tax advisor is involved); although sometimes further extensions are possible. In China, the submission deadline would be 31 March 2013, whereas in the United States the deadline would typically be 15 September 2013.
of the routine entity may be likely to challenge these results. When following this approach, taxpayers should rather be able to trust in the tax authorities accepting actual margins (even if they fall outside the interquartile range), as long as the ex ante transfer price determination was demonstrably arm’s length. In any case, ex ante transfer price adjustments during the year require very resilient and accurate forecasting. Finally, the price-setting approach seems to be much more complex to implement compared to the outcome-testing approach.

In general, a price-setting approach should not per se be preferable to an outcome-testing approach.

It is strongly advisable to include a precise remuneration clause in a written contract. The OECD should clarify in the above-mentioned paragraphs that the tax authorities of both states must accept the point in time for the arm’s length test which has been chosen by the taxpayer and has been explicitly written down in advance (this should also be valid for the transfer pricing method, transfer price adjustment mechanisms, etc.). In this manner, it is very likely that double taxation could be avoided effectively.

In the end, there are valid arguments to take the point in time when the transfer prices are set (i.e. prior to the start of the financial year – Q4/2011 in the example) as the most appropriate time for the arm’s length test. This is independent from the decision regarding whether to choose the price-setting approach or the outcome-testing approach, and provides a certain flexibility. Alternatively, preparing benchmarks immediately prior to the date of the final statutory accounts (group or local entity) would still encompass a group-wide consistent timing for the preparation of benchmarking studies. However, the question is whether this timing is more arm’s length than taking a benchmark study prepared at the start of the financial year.

Following the view taken here, the arm’s length test should be performed on the basis of benchmarking studies available prior to the start of the financial year. This is independent from the question regarding how the transfer price adjustment process works during the year (outcome-testing vs. price-setting approach). As transfer pricing is not an exact science, following the OECD view, it should be acceptable that the arm’s length data are a bit “older” and not available for the tested year.

A precise convention (by the OECD) without alternatives as defined above would have the great advantage that neither companies nor tax authorities would have an advantage or a disadvantage. Moreover, the risk of double taxation could be reduced significantly. Also, from a more pragmatic perspective, it should be acceptable to use data that are a little bit “older”, as seasonal fluctuations will be softened by the application of multiple-year data anyway. Apart from this, only a “timing” effect of roughly two years remains. In any case, it is questionable whether, in practice, two benchmarking studies for routine functions that are dated two years apart really would lead to significantly different interquartile ranges. However, this postulate would need to be analysed and assessed empirically.

As an alternative to the above-mentioned convention and in order to achieve the elimination of double taxation, the OECD Guidelines could stipulate that tax authorities of all member countries must accept completely the point in time (i.e. stating a certain quarter or month of a year) for the arm’s length test, as well as the procedure for determining transfer prices and the method for transfer price adjustments, which have been chosen by the taxpayer and have been written down in advance unambiguously and explicitly (it is, of course, assumed that the underlying transfer pricing method is adequate). This approach would allow flexibility with respect to both (i) different business models/industry specifics regarding price setting versus outcome testing and (ii) the definition of a group-wide timing to perform benchmark studies. An even more flexible situation where individual entities or tax authorities can determine individual timing of benchmarking studies does not seem to be advisable from the perspective of avoiding disagreements between taxpayers and tax authorities, and subsequently double taxation.

The authors would very much appreciate subsequent discussion of this topic.