ERASMUS UNIVERSITY ROTTERDAM Reproduction prohibited

Erasmus School of Economics

Master thesis

Remuneration of similar routine functions:

A practical approach for the determination of an aggregated arm's length interquartile range.

Name student: M.A.A. Oosterbaan

Student number: 330289

Mentor: dr. H. Vrijburg

Co-reader: Mr. R.B.N. van Ovost

Rotterdam, 27/08/2014

Acknowledgements

The past five years have been a great adventure. Now the time has come to finish what I started: the study 'Fiscal Economics'. Hereby I want to thank dr. Vrijburg for the support and great sparring sessions. Furthermore, I would like to thank mr. van Ovost for the additional support in regards to the fiscal aspect of this thesis.

I would also like to thank my colleagues whom allowed me to share some of my ideas regarding this thesis. In addition, I want to express my appreciation for the support my parents, brothers and friends have given me over the past five years.

And last but certainly not least, I want to dedicate this thesis to my late grandfather J.B.H.M. Beks. He has been, and still is, the greatest source of inspiration and motivation for my love of economics and fiscal law.

List of Abbreviations

CITA	-	Corporate Income Tax Act 1969
CoCG	-	Code of Conduct Group
CUP	-	Comparable Uncontrolled Price method
DTA	-	Dutch Tax Authorities
EBIT	-	Earnings Before Interest and Tax
EU	-	European Union
GIE	-	Globally Integrated Enterprise
IQ range	-	Interquartile range
IRS US	-	Internal Revenue Service of the United States
LRD	-	Limited Risk Distributor
MAD	-	Median Absolute Deviation
MNE	-	Multinational Enterprise
OECD	-	Organization for Economic Cooperation and Development
OECD Guidelines	-	OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations
OLS	-	Ordinary Least Square
OPM	-	Operating margin
PLI	-	Profit Level Indicator
RoS	-	Return on Sales
SME	-	Small and Medium-sized Enterprises
TNMM	-	Transactional Net Margin Method
TP	-	Transfer Pricing
TP Decree	-	Decree written by the Dutch Secretary of State with respect to application of the arm's length principle and the Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations

Chapter 1   Introduction

1.1   Introduction and research question

This thesis examines whether tax compliance costs of multinational enterprises (MNE's) could be substantially reduced by standardizing the transfer pricing (TP) practice. Through the current economic downturn, MNE's seem increasingly strained by their "moral responsibility".[1] Their strategic policy, sustainability policies and fiscal policies are increasingly scrutinized. The thriving days, in which aggressive tax structures could be set up without having to worry about any "substance", are over, as a growing international initiative is taken by governments in order to reduce tax avoidance.[2] In turn, this leads to an increasing compliance burden for MNE's.

Although the New York Times[3] thinks differently, they claim that TP is used to shift profits; TP is inter alia incorporated in the law to allocate profits to the jurisdiction where it has arisen. The Organization for Economic Cooperation and Development (OECD) has issued TP guidelines for MNE's and tax administrations.[4] On the one hand the OECD Guidelines require effort from MNE's, but the tax authorities should also meet requirements on the other hand. These guidelines are an internationally recognized standard for the assessment of the propriety of the intercompany transactions. The arm's length principle states, in short, that the terms, conditions and pricing of intercompany transactions should be the same between third parties undertaking similar transactions.[5] This ensures associated enterprises to act 'at arm's length'. The methodology relates to intercompany trade of goods and services.

Within the OECD Guidelines, several methods are described to determine the arm's length price, from which the comparable uncontrolled price method (CUP) is preferred.

---

[1] For example, more attention is devoted by the media to alleged tax evasion and the ethics which are associated with it. See, for example: Murphy, Richard. "Amazon, Google and Starbucks are struggling to defend their tax avoidance." *theguardian.com*. Guardian News and Media, 13 Nov. 2012.

[2] See for example; Choi, Wooje . "Tax benefits from structured financial transaction denied for lack of economic substance." ibfd.org. IBFD North America, 4 Apr. 2014; Schellekens, Marnix . "New substance requirements in effect." ibfd.org. IBFD Research, 5 Feb. 2014.

[3] Norris, Floyd. "Apple's Move Keeps Profit Out of Reach of Taxes.", The New York Times, 2 May 2013.

[4] OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations, July 2010 (OECD Guidelines)

[5] Article 9 OECD Model Tax Convention

This is in fact a price that is based on a (exactly) similar uncontrolled transaction, which has been established between third parties. Because it is normally not possible to find an exactly identical uncontrolled transaction, an inter quartile ranges IQ range (IQ range) is determined by means of a data survey (i.e. a benchmark). This survey is typically composed of publicly available financial statements, which have unspecified gross margins. The IQ range will therefore generally not see on identical uncontrolled transactions, but on the operating profit margin that is ultimately achieved at EBIT-level[6] by uncontrolled transactions between independent enterprises. The benchmark examines, in short, a group of independent comparable companies that substantially perform the same functions. Everything within the IQ range is usually considered to be at arm's length.

Up to the publication of the Primarolo report in 1999[7], the Netherlands had a practice where pre-determined IQ range were already known for certain common activities and was thereby part of the ruling practice. It is an agreement between the Dutch Tax Authorities (DTA) and an enterprise about, for example, the profits that should be generated with the Dutch operations. Within the (fixed) IQ range set by the tax authorities a specific margin, for the entity that is subject, was determined(i.e. a "safe harbour[8]"). Primarolo did research on tax systems in various States and its impact on cross-border level, whereby potentially harmful tax measurements where identified. After publication of the foregoing report, the ruling practice (legally) had to change on several levels because they allegedly led to a significant lower effective tax rate. The burden of searching for comparable companies was shifted to the business since a pre-determined IQ-range was no longer available.

When establishing an IQ profit margin, one tries to establish an economic reality that is not there. On one hand, the tax authorities have more specific information available than the enterprises doing the search[9], and on the other hand, a benchmark does not take the

---

[6] EBIT is the abbreviation of earnings before interest and tax
[7] "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999.
[8] A safe harbour is a statutory provision that applies for a predetermined group of taxpayers or transactions and relieves them from specific obligations that would otherwise be imposed.
[9] See in this context also article 2.65 OECD Guidelines

specific advantages of globally integrated enterprises. For example, the presence of knowledge, (money) resources, economies of scale and synergies. In principle, enterprises can manipulate the IQ-range that they establish with their own market knowledge and thus have the opportunity to shift profits across borders. In order to reduce the related compliance costs, governments could reform to the pre-Primarolo practice using a fixed IQ range set by, for example, the tax authorities, if it appears that the IQ range, resulting from the benchmarked functions, does not significantly differ between industries. If it becomes apparent that, for the determination of the IQ range for a distributor, it does not matter if one looks at a chemicals industry or an electronics industry, than inter alia cost saving could be realized by the annual determination of an IQ range, as occurred before the Primarolo report.

The above has resulted in the following research question:

"Do the profit margins for similar routine wholesale distributers differ significantly between industries?"

To answer the above-mentioned research question, a thorough statistical analysis needs to be conducted. In order to show a significant correlation, I will examine if the IQ ranges found per industry statistically differ from each other.[10] I will also consider whether or not each industry has a significant influence on the IQ range and mean. A Chi-square test will determine if the level of the IQ ranges are equal, as well as the total distribution. Finally, the filters that will be applied on the data found will lead to a comparable set, which could be used in the TP practice. The initial group that is considered to perform similar routine wholesale functions comprises of the Appearance, Chemicals, Electronics, Industrial and IT Products industry.

1.2   Scope

This thesis will only focus on the distribution function, for eight different and diverse industries. I will refrain from other non-standard TP methods discussed in the OECD Guidelines and literature. Finally, this thesis is primarily designed to test if a practical application of the combined set of industries is possible. Consequently, it is explicitly not

---

[10] Bases on the assumption the full set of industries selected (i.e. aggregated IQ range)

intended to provide an explanation for the profit margins and the associated distribution thereof in various industries.

1.3   Method and Approach

First of all the reason of the existence of TP will briefly be examined. Subsequenlty, the changes over time and the state of play at this time will be discussed briefly. This will include the relevant methods with their limitations.

Secondly, I will look by means of financial data of wholesale distribution companies whether an equal operational profit margin over time and within periods can readily be identified, in order to provide a possible foundation for my statements. This will be done on the basis of statistical tests, which will be discussed in detail.

This thesis will be concluded with a conclusion and discussion of the found insights and an answer on my research question.

Chapter 2   Transfer Pricing

2.1   Introduction

In order to demonstrate the relevance of this study, it is important to know the TP concepts that are applicable, their origin and changes over time. Firstly, a short introduction will be given about the origin, after which the main TP terms and concepts are discussed. Secondly, different TP methods will pass in review. In addition, the Primarolo report will be discussed because it demonstrates, inter alia, changes that had to be made over time. Finally, a brief summary will be given of the Dutch TP provisions and regulations to show the relevance for the DTA and local (multinational) enterprises.

2.2   History

A transfer price is the price that a division of a multinational operating enterprise charges for the provision of goods or services to another division of the same group. Normally, market forces are setting prices and conditions.[11] This is not the case when one looks at a MNE's. The affiliated divisions that are engaged in transactions with each other are able to influence prices and conditions. This could lead to problems when fiscal accounts are determined at year-end.

Tax authorities must reconcile with the right to only tax those profits attributable to the activities that arise within their territory. This will prevent economical double taxation. The risk of economical double taxation occurs when one or more tax authorities are not satisfied with the allocation of profits to their territory. On a more practical level, the correct allocation of income and expenses hampers when difficulties occur associated with obtaining financial data outside their own jurisdiction.

In order to avoid the possible disruptive double taxation on the international trade market, the OECD member countries have tried to set up general guidance on TP (OECD Guidelines).[12] Following the Pickle Hearings[13], the focus on TP intensified. The Internal

---

[11] §1.2 OECD Guidelines
[12] The first guideline is published in 1979; see foreword OECD Guidelines
[13] Eden, Lorraine. Taxing multinationals: transfer pricing and corporate income taxation in North America. Toronto: University of Toronto Press, 1998, p. 343-345

Revenue Service of the United States (IRS US) investigated if North American multinationals used TP to shift profits. Evidence of this phenomenon was found in 1990 and was further elaborated in Pickle Hearings. The IRS US started to challenge and penalize more and more enterprises, but was restricted due to lack of information.[14] Therefore the OECD Guidelines were updated in 1995.[15] It is important to note that tax authorities still recognize that MNE's are able to 'use' TP in their advantage to shift profits to other jurisdictions.[16]

2.3   The arm's length principle

For years the arm's length principle has been, and still is, the international standard in the determination of transfer prices.[17] Both the OECD and the United Nations (UN) define the arm's length principle in their model Tax Convention and is defined as follows:[18]

" *Where an enterprise of a Contracting State participates directly or indirectly in the management, control or capital of an enterprise of the other Contracting State, (…) and conditions are made or imposed between the two enterprises in their commercial or financial relations which differ from those which would be made between independent enterprises, then any profits which would, but for those conditions, have accrued to one of the enterprises, but, by reason of those conditions, have not so accrued,*

may be included in the profits of that enterprise and taxed accordingly."

This principle works both ways. The shareholders are interested in the consolidated value of the enterprises (i.e. the values of shares) and are less interested in the origin of the profits. In turn, the tax authorities are interested in a (higher) share of the profit that has arisen on their territory. Therefore an erroneous allocation of profit on the one hand, and double taxation on the other hand, lies just around the corner. In order to empower tax authorities to address profit shifting and decrease economical undesirable double (non)taxation, the arm's length principle has been an internationally accepted as

---

[14] Eden, Lorraine. Taxing multinationals: transfer pricing and corporate income taxation in North America. Toronto: University of Toronto Press, 1998, p. 346

[15] And later in 2010, see footnote nr. 4

[16] See for example: "Camp Opening Statement: Hearing on Transfer Pricing Issues." Committee on Ways and Means, 22 July 2010.

[17] For a comprehensive historical overview, see: Hamaekers, H. "Arm's length - how long?." International Transfer Pricing Journal 8.2 (2001): IBFD Tax Research Platform.

[18] Article 9 OECD Model Tax Convention; Article 9 UN Model Tax Convention

the method to allocate profits. However, it is important to note that TP does not cover the problems that arise from tax fraud or tax avoidance, regardless of whether TP policies are applied for that purpose or not.[19]

It may be difficult to determine the correct (total) profit level in a jurisdiction when you are dealing with a MNE. Think for example of non-meaningful consolidated financial statements and a possible financial organisation of one specific division that is divided cross border over multiple (sub-) divisions. Tax authorities therefore focus on statutory legal entities in order to determine the profit/losses that have accrued. This follows from the 'separate entity approach' as mentioned in the OECD Guidelines.[20] The OECD has provided further guidance for the application of arm's length principle. For example, the applicable functions, risks and assets as well as contractual terms need to be considered.

2.4 Separate entity approach and tested party

The OECD has defined the separate entity approach as follows:[21]

"(…) the arm's length principle follows the approach of treating the members of an MNE group as operating as separate entities rather than as inseparable parts of a single unified business."

The OECD Guidelines stipulate that one must not look at associated entities on a combined level, because profitability of reciprocal transactions per entity should not be affected by the fact that they are part of a MNE. It is therefore important that MNE's seek references from independent enterprises that are of comparable nature when applying the arm' length principle.[22]

The separate entity approach is selected to create a level playing field between associated MNE's and independent enterprises.[23] The OECD Guidelines endeavours to ensure that no tax advantage/disadvantage could be created between associated MNE's

---

[19] §1.2 OECD Guidelines
[20] §1.6 OECD Guidelines, to be discussed in chapter 2.4
[21] §1.6 OECD Guidelines
[22] This should consist of comparable transactions as well as comparable conditions (these are i.e. comparable uncontrolled transactions)
[23] §1.8 OECD Guidelines

and independent enterprises. This approach does not take economies of scale and synergies into account, whilst MNE's are driven by these advantages. Because the separate entity approach is used, only one of the affiliated enterprises involved in the intercompany transaction is included in the analysis.[24] The affiliated enterprise that is being compared to an unrelated party is classified as the 'tested party'. Generally the tested party is that affiliated enterprise that had the least complex functional analysis.[25] This is because the most reliable comparable companies could often be found for the enterprise with the least complex functional analysis.

The above can be explained when one considers two parties that enter into transactions with each other, and where one party owns intangibles and the other party carries out merely routine functions[26]. This assumption makes sense, because it is practically impossible for enterprises that own intangibles to find similar enterprises in the public domain since uniqueness of intangibles is assumed. However, there is a risk that whilst performing such a unilateral analysis, the aggregated profitability of the group is not considered.[27] Consequently, in this thesis companies that are assumed to be the least complex entity will be examined. This will manifest itself during the selection of the data.

2.5 Selection of the TP method

The comparability analysis is of great importance for all the available TP methods when one tries to determine an arm's length transfer price. It forms the foundation for every substantiation of the transfer price applied. This analysis includes a functional analysis, the contractual terms, economic circumstances, business strategies, and so on. It is important to note that all factors of the comparability analysis need to be examined, because (small) differences could lead to material deviations in the remuneration.[28]

---

[24] This approach is not applied at the Profit Split Method, which is discussed later on.
[25] §3.18 OECD Guidelines
[26] These are basic/simple functions. For instance, a company whose core-function is the distribution of goods. They do not perform marketing activities etc.
[27] Heijenrath, M.J.M., and P.W. van Meeuwen. "Introductie tot de praktische aspecten van de OECD-transfer-pricingmethodieken." Tijdschrift voor Fiscaal Ondernemingsrecht 2006/19
[28] §1.35 OECD Guidelines

In order to find the at arm's length transfer price, MNE's conduct a comparability analysis consisting out of a functional and economic analysis. The functional analysis determines, in short, that in order to identify and compare economically significant activities and responsibilities undertaken, one must look at the functions performed, assets used and risks assumed in the transaction between parties.[29] For the economic analysis one is, based on the functional analysis, looking for comparable transactions (and i.e. the level of their associated operating profit margins on EBIT level).

2.5.1   Functional analysis

In order to identify which party contributes to a transaction and to what extent, a functional analysis needs to be performed. This assessment will ensure that a clear overview is formed about who is responsible for, for example, manufacturing, sales, distribution, and so on. From here, the tested party can be characterised. The characterisation comprises, for example, of entities that use a manufacturing model, distribution model, perform contract R&D, etc. In order to identify a certain intercompany transaction, one gathers the relevant facts and circumstances and how these subsequently are allocated, by looking at the functions, risks and assets. Altogether, the functional analysis determines the key value drivers of the value chain and the allocation thereof between subsidiaries.[30]

---

[29] §1.42 OECD Guidelines
[30] Vollebregt, H.A. . "Verrekenprijzen: fiscus en functionele analyse." Weekblad voor Fiscaal Recht 2004/1009

The figure below shows a schematic overview of possible functions in a value chain / within a MNE:

Figure 1: Different profit-drivers

Decentralised					Centralised

Board of Directors

Corporate center

Shareholder activities	Finance function	Services/provision function	R&D function	Sales function	Manufacturing function

Functional analysis by function
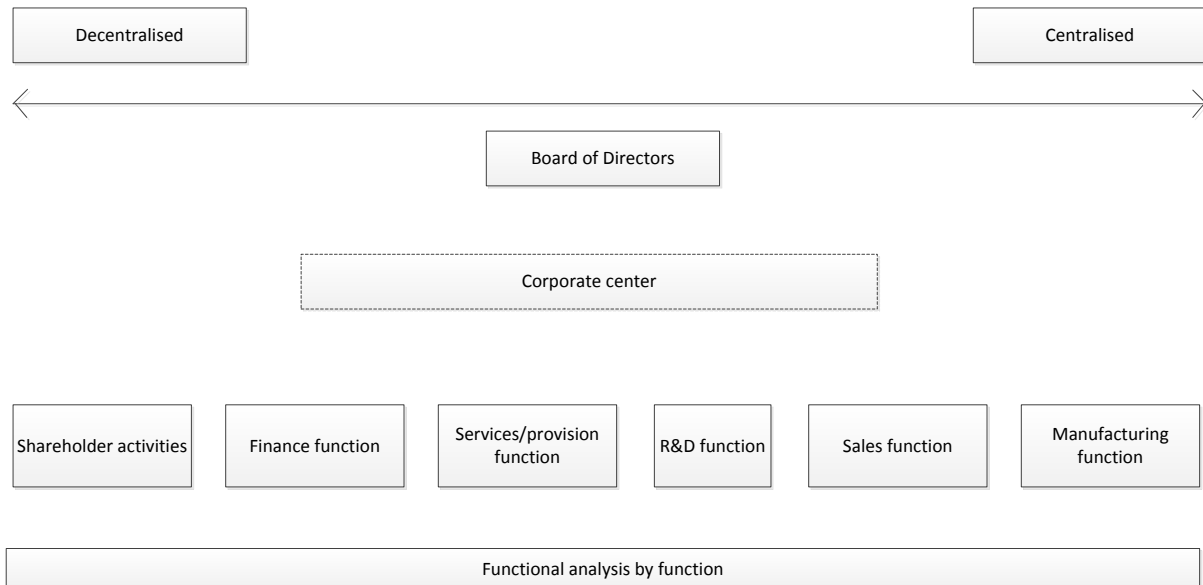
Source: Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.45

2.5.1.1   Functions

There are several functions to consider when one looks at the value chain of a MNE. Apple Inc. for example has divisions that perform R&D, production, design, assemblage, services, public relations, personnel and marketing.[31] It will be apparent that a routine function, that does not strongly depends on the performance of the entire business, should receive a relative low remuneration. And vice versa, a strongly complex core-function receives a relative high remuneration. The marketing function of Apple's products will be of greater importance than for instance the assembly function. It is more likely that the profitability, positive or negative, depends more on the marketing function because it attributes to Apple's business worldwide. It adds more value to the value drivers. In addition, it is for example relevant who purchases the raw materials in the production division. The remuneration for a subsidiary will be valued lower if the principal is responsible for the procurement of raw materials.[32]

---

[31] In §1.43 OECD Guidelines, some possible functions are suggested
[32] Functions/activities are derived from the TP decree

Consequently, a routine function is considered to be 'routine' because of limited responsibility and autonomy. Most risks are borne by the entrepreneurial function.

As previously indicated in figure 1, there are several functions that need to be considered (*see Appendix A for a detailed discussion)*. In this thesis the focus is on the sales function only, because that is the function under review (i.e. the wholesalers). Below this function is discussed in detail.

2.5.1.2 Risks

The TP-method is generally determined on the basis of the functions performed. However, in order to achieve a proper analysis, adjustments need to be made when one signals significant differences in the risks assumed. For instance, a higher remuneration will be appropriate if more (material) risks are allocated to an enterprise. The open market will take these risks into account.[33] Thus, the rate of return that is earned by an enterprise partly stems from the level of risks they bear (*see Appendix A for a more detailed discussion)*. This thesis will implicitly control for this by means of the data selection procedure.

2.5.2 Characterisation sales function

The sales function is often considered to be the last function in the value chain of an enterprise. The sales force is in contact with the final consumer.[34] This function varies from solely logistical/distribution activities, to activities related to marketing activities. The following characterisations can be discerned and are all based on the nature of the activities performed by the sales entity/distributor:[35]

- Agent
- Commissionaire
- Limited risk distributor
- Full-fledged distributor

---

[33] §1.45 OECD Guidelines

[34] Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.43

[35] The entity performing the sales function does not have to be (solely) a sales entity. Another core-function may be applicable. These characterisations will be discussed in more detail further on.

The sales function offers the products or services of an enterprise to the end-consumer. Every single sales characterisation differs on the basis of the functions that they perform, the risks that are assumed and the assets that are used. The level of functions performed, risks assumed and assets employed all contribute to the remuneration. A wholesale distributor will usually receive remuneration on the basis of a return on sales.[36] This is simply a percentage on top of their overall turnover/sales. The comparison of operational profits margins should consequently be done with companies that have the same characterisation.[37] The figure below shows the preceding clearly:

Figure 2: Allocation of profit
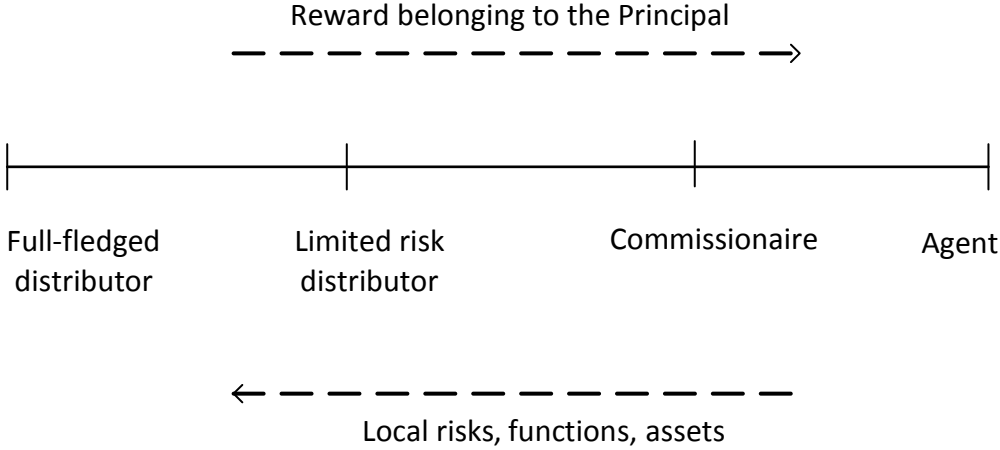
Source: Corlaciu, Alexandra. "Business models for tax and transfer pricing purposes." The Journal of the Faculty of Economics, University of Oradea 1 July 2013: 1182-1189.

For example, the agent receives a low remuneration and the reward belongs to the principal since the local risks are low.

Distinguishing various characterisations is important, because it can be used to select and examine a group of wholesale companies that are assumed to have, inter alia, similar risk levels and perform similar functions.

---

[36] NB: It relates to the remuneration earned by dependent companies/subsidiaries.
[37] NB: The percentage (mark-up) on top of the turnover is should ultimately be equal to the operational profit of an independent company divided to the turnover.

2.5.2.1 Agent

Firstly, the figure below shows the difference between an agent and a commissionaire:

Figure 3: Differences commissionaire and agent
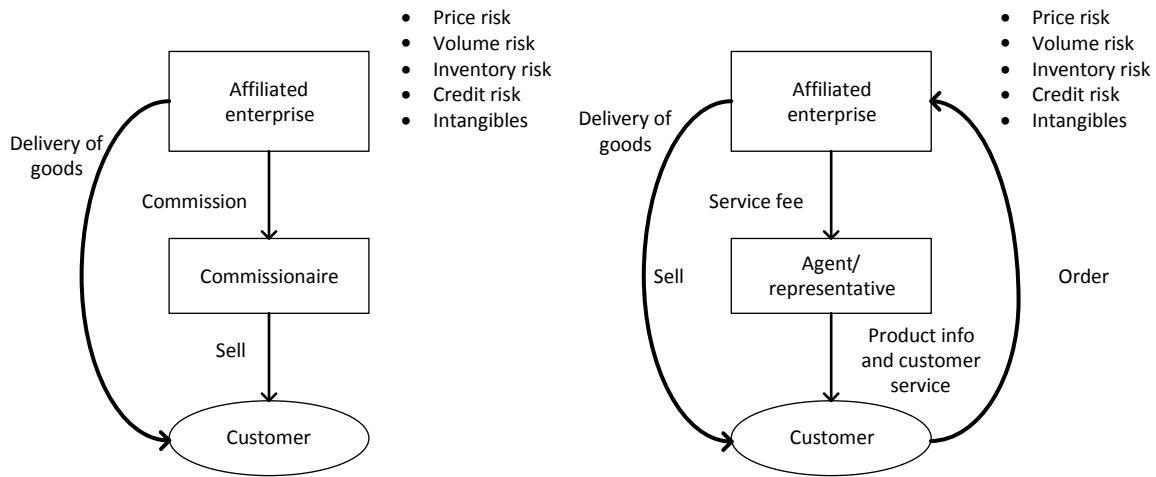
An agent functions as a representative by means of the sales company. The agent governs the sale of products to consumers by providing product information and several customer services. The agent needs to develop a sales team, identify and qualify new customers and maintain the sales relationships. In addition, the agent acts on behalf and on the name of the principal. The agent does not sign contracts nor has products in stock and therefore bears no inventory risk. The remuneration could be based on the CUP, cost-plus or a return on sales (i.e. a service fee).[38]

2.5.2.2 Commissionaire

The commissionaire-model is a civil law concept similar to the agent-model. However, although the commissionaire sells the goods on behalf of the principal, he does it in its own name. The risk associated with the goods as well as the benefits of the risk belong to the principal since the commissionaire does not become the owner of the goods. The remuneration could be based on the CUP, resale-minus or the TNMM (i.e. a commission).

2.5.2.3 Limited risk distributor

The limited risk distributor (LRD) (re)sells goods on its own behalf and acts in its account. The goods are typically purchased from a master distributor who then directly resells. The LRD owns the goods that are sold shortly prior to the moment of sale. As a

---

[38] See 2.6 for a better understanding of the applied methods of remuneration.

result, limited inventory risk is incurred. The remuneration could be based on the CUP, resale-minus or the TNMM with a profit level indicator return on sales.

2.5.2.4 *Full-fledged distributor*

This could be considered to be a conventional distributor model. The full-fledged distributor will typically procure the goods directly from a manufacturer. It therefore owns the inventory and bears the related inventory risk. The full-fledged distributor will act on its own behalf and could therefore be regarded as the principal. In addition, all the activities performed are decentralised with a limited amount of centralised control. It thus bears market risk. All the risks and functions performed should be reflected in the remuneration.

In short, this thesis will use data that is retrieved from financial public statements of independent companies. It therefore focuses on companies that are expected to have similar characteristics as a full-fledged distributor (i.e. independent companies always act on their own behalf).

2.5.3 Business strategies

The OECD Guidelines recognise that business strategies could influence the comparability between controlled and uncontrolled transactions and enterprises.[39] Enterprises may have different motives for their daily actions and thereby differ from the found uncontrolled transactions and enterprises. In order to achieve an at arm's length price, one needs to adjust for differences like, for example, the level of innovation, the degree of diversification and the absence or presence of product development. Common and well-known business strategies are market penetration and actions that expand market share (*see Appendix A for a more detailed discussion*).

2.6 Transfer pricing methods

The majority of TP models are binary. This means that an entity either exerts a routine function or an entrepreneurial function. It is an important distinction because the entity performing the entrepreneurial function is entitled to the residual profit (or loss). A routine entity merely receives remuneration on the basis of an operational (profit)

---

[39] §1.59 OECD Guidelines

margin, where the remuneration of the entrepreneurial thus consists of the residual profit[40].

Enterprises should adopt the most appropriate method in order to achieve an at arm's length transfer price. The selection is almost entirely driven by the functional analysis.[41] The applied method is only a means and not an end in itself in finding an at arm's length price. The OECD Guidelines describes three traditional transaction methods, such as the comparable uncontrolled price method, resale price method and the cost plus method, as well as two transactional profit methods, such as the profit split method and the transactional net margin method (TNMM).[42] However, enterprises are not bound by these methods and may choose other methods.[43] The diagram below indicates which level of profit the method relates to:[44]

Turnover	100	CUP	
Costs of goods sold	80		
Gross margin	20	Resale price/cost plus	
Operational expenses	10		
Net margin	10	TNMM/profit split	
Deprecation and interest	5		
Profit	5		

All methods are explained later on, this gives a comprehensive overview in advance regarding the different profit levels that could be identified

One should take the nature of the analysed transaction, the availability of reliable data from third parties, the degree of comparability and the extent to which adjustments are possible into account.[45] Only the most direct (CUP) and relevant methods will be discussed in detail (*see Appendix A for a discussion about the other methods*).

---

[40] This could also be a loss.
[41] The 'best method rule' was adopted making the specific facts and circumstances (i.e. the functional analysis) top priority. Abdallah, Wagdy M., Critical concerns in transfer pricing and practice. Westport, Conn.: Praeger, 2004. p.166
[42] §2.1 OECD Guidelines
[43] §2.9 OECD Guidelines
[44] Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.74
[45] §2.2 OECD Guidelines

2.6.1   Comparable uncontrolled price method

The comparable uncontrolled price method, also known as CUP, is the most direct method when it comes to determining an at arm's length intercompany price. It is the method that meets the actual transactions undertaken in the best possible manner. I.e. it compares the price charged for goods or services that are transferred in a controlled transaction to the price of goods or services that are transferred in a similar free-market (or uncontrolled) transaction under comparable circumstances.[46]

There are two different types of CUP's that can be distinguished. The internal CUP is the result of a comparable transaction that an affiliated enterprise had entered into with an unrelated enterprise. The external CUP comprises of a comparable transaction made between unrelated parties.

If none of the differences between the compared transactions result in a material effect on the price, then the CUP can be considered to be a comparable uncontrolled transaction. It can also be considered to be a comparable uncontrolled transaction when it is possible to make reasonably accurate adjustments to eliminate the material effects of such differences. For example, adjustments based on differences in sales volume or credit terms are easy to make, whereas adjustments based on differences in quality or the geographical market are difficult to make.[47]

2.6.2   Transactional net margin method

In this method the net operational profit margin relative to an appropriate base, which is obtained by a controlled transaction, is being compared to similar uncontrolled transactions. For example, an appropriate tax base could be the level of cost incurred, sales made or assets used. The calculated ratio of profit relative to the costs, sales or assets is known as the profit level indicator (PLI). It serves as an indicator for the comparability analysis. The most common PLI's relate to costs or sales. These PLI's are respectively a cost plus method and a resale minus method based on the net profit level instead of the gross margin. The advantage of the TNMM is that differences in the nature

---

[46] §2.13 OECD Guidelines et seq.
[47] NB: differences in volume, etc. can be quantified, whereas differences in quality, etc. are hard to quantify.

of products and services are less relevant at net profit level than at gross profit level. Operating margins are, in comparison to the gross margins, also less prone to small functional differences. This is because variations in the functions performed between enterprises often lead to differences in operating expenses. In a comparison on operational level, both higher remunerations and higher costs are reflected in net profit level. Similar companies may therefore have very different gross profit margins, but still achieve net profits that are at a similar level and may be comparable based on a functional analysis.[48]

In principle, the comparison of the net profits must take place before deducting interest and extraordinary gains and losses. This will only be different if, on the basis of the functional analysis, the before mentioned aspects have a significant influence on the operational (net) profit. For instance, a high interest burden could occur if an enterprise allows her clients pay at a relative late point in time.

2.6.3   Profit split method

This two-sided approach method divides the profit of a MNE in order to establish an arm's length transfer price *(see Appendix A for a more detailed discussion).* The profit split method ensures that (operational) profit is divided in such a way that would be expected of independent enterprises that find themselves in a joint-venture relationship.[49] Where the above-mentioned methods are all based on the entity that performs the least complex function(s), it can also occur that *both* enterprises add a unique and valuable contribution to the transaction.

When applying this method, an enterprise needs to determine the total profit level. The profits should subsequently be allocated to its affiliates. Just as with the TNMM, the allocation of the profits should be based on a transfer price derived from uncontrolled transactions. In the allocation of the total profit resulting from economies of scale or other benefits resulting from more efficiency, a two-sided analysis appears to result in a better outcome. A one-way analysis like the TNMM, where the focus is more oriented

---

[48] Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.92
[49] In some specific transactions the use of the gross margin is allowed as well; §2.131 OECD Guidelines

towards one of the parties involved, may cause that the economies of scale obtained actually end up with just one of the parties.

Depending on the specific facts and circumstances, parties must search for an appropriate allocation key that reflects the relative added value of the contribution by the parties involved.

The OECD Guidelines distinguish two main types of profit split methods, namely the contribution analysis and residual analysis. The *residual* profit split analysis is a two-stage method. In the first step, each affiliate receives an appropriate remuneration for its routine functions. This remuneration will usually be determined by means of the application of one of all the above-mentioned methods.

Thus, the results of this thesis, which are based on the application of the TNMM approach, could be applied in this method as well. The enterprises that enter into the controlled transaction will be deemed 'tested party'. This means that intangibles are not taken into account. The residual profit or loss will be allocated to the more complex functions on the basis of the particular facts and circumstances.[50]

2.7 Ruling practice

Usually taxpayers will receive certainty about their tax return years after filing. It simply takes a while before tax authorities assess the tax return and make it final. However, taxpayers might want to receive certainty about their tax position in advance, before filing, to acquire certainty about their tax liabilities. For example, the Dutch legislator has incorporated this possibility for those taxpayers, both legal and natural persons. It states that every taxpayer has the right to receive the (binding) opinion of the Dutch Tax Inspector, provided that the boundaries of the tax system are not being exploited. This is also known as a 'ruling'.

The consequences are the same for taxpayers with or without a ruling since Dutch Tax Law, Jurisprudence and Administrative Policies, such as the TP decree, bind the ruling. If deviations from the presented facts and circumstances occur, ruling arrangements will

---

[50] §2.21 OECD Guidelines

not be applicable anymore.[51] This practice exists in many jurisdictions, meaning that the foundation for a possible safe harbour IQ range is already in place.

Finally, a ruling is usually entered into for a four-year period, is not applicable after modifications in Tax Law and is based on the same corporate tax rate.

2.8 Primarolo report and consequences

The ruling practices of many countries were under review by Primarolo. The findings of the Primarolo report where sent to the ECOFIN Council on 29 November 1999. This report is the result of research done by the Code of Conduct Group (CoCG), section business taxation, and comprises of harmful features identified in European jurisdictions. It looked at particular tax systems in various States and its impact on international cross-border level *(see Appendix A for a more detailed discussion)*.

Predetermined IQ ranges (i.e. safe harbours) were especially present in the Netherlands and could therefore be regarded as a 'good' example for the purpose of this thesis. It was concluded that the Netherlands should revise their tax system and administrative practice on thirteen points. All the cross border measurements that need to be taken were listed and are all regarded as harmful because they all indicate a significant lower (effective) tax rate.

For example, the Netherlands had a pre-set cost-plus ruling, which relates to the allocation and determination of profits. It was applicable in those situations where the inter-company activities are of preparatory and supporting nature. One of the requirements is that a comparable uncontrolled price cannot be found for the services rendered. This led to a lower compliance burden and lower associated costs for eligible tax payers *(see Appendix A for a more detailed discussion)*.

Although the DTA still have a ruling practice in place, they no longer have pre-set rulings. Therefore, enterprises with simple routine functions need to comply with the high administrative burden as well.

---

[51] Müller, Johann. "Ruling Practice." The Netherlands in international tax planning. 2nd ed. Amsterdam, Netherlands: IBFD, 2007. 359-367.

2.9   Dutch Corporate Income Tax Act

The Dutch legislator endorses the arm's length principle as defined in the OECD Guidelines, codified this in article 8b Corporate Income Tax Act 1969 (CITA[52]) and has further elaborated the principle in several decrees[53] *(see Appendix A for a more detailed discussion)*. It is a codification of the existing practice as well as a confirmation of the arm's length principle. Double (non-)taxation due to different Dutch TP provisions is thus largely avoided. The incorporation has led to an extension of the at arm's length principle, since documentation requirements were included. This is done by the legislature to avoid unjustified profit shifting from the Netherlands to other jurisdictions and is therefore a defensive measure.[54] Many jurisdictions had already strict documentation requirements in place, causing the fear that profits would be unjustifiably shift to those jurisdictions, because the DTA could bring little to the table against it. The same applies to jurisdictions with a relative low tax rate.

The legislator takes a further stance by means of the TP decree. It discusses the relevant methods and other concerns in regards to TP. The position as well as the starting point of the DTA is displayed in the TP decree.

The arm's length principle applies to transactions between affiliated companies. The Dutch CITA does not define precisely when entities are affiliated to each other. It could be considered an 'open norm' because affiliation can originate from capital, management or supervision.[55] There must be sufficient control in regards to the entity under review, making the shareholder interest a non-decisive influence.[56]

Affiliated enterprises are required to document data in their administration that shows how the transfer prices have been established. It ensures that tax authorities have

[52] CITA is an abbreviation of Corporate Income Tax Act

[53] A decree is a written decision by an administrative authority constituting a public legal act. It provides the view taken by the Secretary of State with respect to a specific topic. For example: The State Secretary for Finance. "International Tax Law. Transfer pricing method, application of the arm's length principle and the Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations." Rijksoverheid.nl. Version IFZ2013-184M,. Directorate-General for Tax and Customs Policy and Legislation, International Tax Policy and Legislation Directorate, 14 Nov. 2013 (hereinafter: TP Decree)

[54] Nota n.a.v. het Verslag, *Kamerstukken II*, 2001-2002, 28 034, nr. 5, p. 33

[55] Peerbooms, J., and M. van der Breggen. "Transfer pricing: verplichte kost voor iedere fiscalist!." Forfaitair 173 (2007)

[56] MvT, *Kamerstukken II*, 2001-2002, 28 034, nr. 3

access to sufficient information with regards to the transfer prices set. It allows them to audit enterprises in a proper manner. The burden of proof lies initially on taxpayers whom are also obliged to provide the administration to the tax authorities.[57]

---

[57] Article 52 resp. article 47 State Taxes Act (Dutch: Algemene Wet Inzake Rijksbelastingen)

Chapter 3   Economic Literature & Theory

The above-mentioned theory will, inter alia, be offset against several economic studies in order to show the possible relation and impact of certain factors. For instance, can the European market be considered to be economically integrated? And can we find evidence that results can be influenced by an industry business cycle? Subsequently the use of pan-European comparable companies for each industry specific dataset will be substantiated.

3.1   Resource-based view

The economical rational behind the research question stems from the resource-based view. This theory states that companies can earn above normal returns if they own superior resources. This can be achieved by means of the following resources, it must be:[58]

- Scarce;
- Valuable;
- Inimitably; and
- There may be no substitutes available.

Thus, it is important to note that the competitive advantage can primarily be achieved by means of the combination of resources that are at their disposal.[59] Following the resource-based view, similar routine function and risk profiles/levels should thus receive a similar remuneration. For example, if a limited risk distributor assumes similar inventory risks in the durable goods industry and the non-durable goods industry, than they should receive a similar remuneration.

3.2   Safe harbour

The determination of an arm's length range is an intensive process with a lot of uncertainty and heavy administrative burdens for both the taxpayer and tax authorities. The use of safe harbours might ensure that the overall burden is reduced. A safe harbour

---

[58] Grant, Robert M. "The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation."california management review 33.3 (1991): 114-135
[59] Wernerfelt, Birger. "A Resource-based View Of The Firm." Strategic Management Journal 5.2 (1984): 171-180

is a statutory provision that applies for a predetermined group of taxpayers or transactions and relieves them from specific obligations that would otherwise be imposed.[60] This often means that there are simplified obligations applicable, if safe harbour provisions are followed. The OECD Guidelines distinguish two types of safe harbours, namely:

- Those who exclude certain transactions from the application of (local) TP provisions, often by means of setting thresholds; or
- Those who allow taxpayers to apply simplified rules, often by means of a pre-set range that is deemed to be at arm's length.

The use of safe harbours has, obviously, both advantages as disadvantages *(see Appendix A for a more detailed discussion)*. For example, a simplification of the compliance regulations can lead to a significant reduction of costs.[61] However, the inherent risk arises that one deviates from the arm's length principle. It should be noted that many countries already have introduced safe harbours for small and medium-sized enterprises (SMEs), because the benefits outweigh the problems identified in such businesses or non-complex transactions.[62]

It might be possible to identify an exogenous arm's length range by the examination of various industries, after which an arm's length remuneration could be determined. This could be done by examining the distributions of various industries and thereby look at an overlap. If an overlap can be identified in which arm's length ranges normally fall, then we can characterize that as a safe harbour.

3.3 Purchasing power parity

It will be clear that the prices of (similar) goods and services differ across marketplaces. The price levels could differ due to, inter alia, cheap labour in country 'A' and a high amount of capital in country 'B'. For example, prices levels between different countries

---

[60] Section E OECD Guidelines
[61] § 4.97 OECD Guidelines
[62] "Multi-Country Analysis of Existing Transfer Pricing Simplification Measures." OECD 6 June (2012)

are often positively related to income.[63] That is because of the high prices of non-tradables, such as services, relative to prices of tradables found in rich regions/countries, then prices found in poor regions.[64] This phenomenon could be explained, inter alia, by the productivity differences between rich and poor regions. For example, because the non-tradables sector often is more labour intensive than the tradables sector, differences arise due to the fact that labour is relatively cheaper in poor regions.[65] The market of an intercompany transaction should to be comparable to the market of an unrelated party in order to define an at arm's length price. For deviations in market circumstances should thus be corrected.

The OECD Guidelines cites several points that should be taken into consideration.[66] As previously indicated during the discussion about the possible risks that can be identified, the presence of close substitutes plays a major role in the determination of an at arm's length transfer price. Consequently, these close substitutes determine, among other things, the size of the market and the level of competition. The geographical location of a market and the purchasing power in the market may also play a role. The purchase power parity (PPP) is a well-known method to determine the relative values of various currencies. The PPP states that the national price levels across countries should be equal once adjusted to the same currency.[67] It might be possible to use this method to adjust intercompany transactions on the basis of transactions found in different markets. However, recent studies have indicated that the real exchange rates only tend toward a PPP in the very long run.[68]

Thus, differences in price levels between the EU countries are not to be expected, because they all reside in a similar rich area, with the same currency. However, if small

---

[63] For example, the 'Big Mac Index' gives an insightful presentation hereof. It determines to what extent exchange rates lead to similar costs of goods in various countries, by calculating the amount of time an average worker needs to work in order to purchase a Big Mac.

[64] Kravis, Irving B., and Robert E. Lipsey. National price levels and the prices of tradables and nontradables. Cambridge, MA (1050 Massachusetts Ave., Cambridge 02138): National Bureau of Economic Research, 1988.

[65] Kravis, Irving B., and Robert E. Lipsey. "Toward an Explanation of National Price Levels." Princeton Studies in International Finance, Princeton University no. 52 (1983)

[66] §1.55 OECD Guidelines, this is a none limitative list

[67] Rogoff, Kenneth. "The Purchasing Power Parity Puzzle." Journal of Economic Literature 34.2 (1996): 647-668. Jstor

[68] Idem.

differences arise, this might be due to the fact that real exchange rates only tend toward a PPP in the very long run.

3.4 Market integration

In addition, market integration should lead to a convergence of economies within the EU, because economic theory suggests that interaction causes more cross-border activities.[69] Inherent to more cross-border activity is an increasing degree of competition. This implies that the market size expands and profit margins for similar routine functions converge because companies (i.e. people) want to capture the above normal rent. Consequently, in a large market the profit margins for routine functions will converge for every industry because it will be saturated on the long run.

First of all it is important to note that there are little, if any, market barriers left within the European market due to several treaties and the European monetary union.[70] These treaties are entered into by some of the current EU member states in order to stimulate free trade between member states. This was deemed good for the local and European economic market. In addition, particular discrepancies in policies were eliminated. As result of the Maastricht Treaty, markets were further integrated by the introduction of the European Union and the subsequent implementation of the euro.[71]

Although the variability of profits between certain geographical regions may be prompted or influenced by different customer preferences, the market size, etc., this might not be the case when the market is integrated. Several studies have been performed in regards to the possible price and profit convergence as indicator for market integration in the EU. Haskel and Wolf concluded, for example, that difference in prices within countries seem, in the majority of cases, equal to or even higher than the

---

[69] Several studies have researched this theory. For example: Emerson, Michael. *The Economics of 1992: an assessment of the potential economic effects of completing the internal market of the European Community*. Luxembourg: Office of Official Publications of the European Communities, 1988 (whom made predictions). Or: Kleimeier, Stefanie, and Harald Sander. *European financial market integration: evidence on the emergence of a single eurozone retail banking market*. Maastricht: METEOR, Maastricht research school of Economics of TEchnology and ORganizations, 2002.

[70] See for example: Single European Act, June 1985; or European monetary system (regulation 3181/78), December 1978; or Maastricht Treaty, February 1992

[71] Kleimeier, Stefanie, and Harald Sander. *European financial market integration: evidence on the emergence of a single eurozone retail banking market*. Maastricht: METEOR, Maastricht research school of Economics of TEchnology and ORganizations, 2002.

differences found between countries.[72] These results are supported by Nielsen, who concluded that no significant price differences exist for a number of tested goods between EU countries, because the results demonstrated that the spread found did not differ from the countries on a stand-alone basis.[73]

Cross-border differences between different industries are thus harmonised by the market integration, making it reasonable to test this market.

3.5   Variance

Statistical research has been performed by means of the comparison of, inter alia, operating profit margins (i.e. RoS) in regards to the profitability between three integrated markets, being the EU, Japan and US.[74] Firstly, they found that the measured profit level indicators always show a lower variance of profitability in the EU than in Japan, but not compared to the US. The aforementioned three markets could be deemed to be economically integrated because, inter alia, trade barriers were removed, cross-border trade increased, free competition was made possible, etc. However, the authors found that although these markets are integrated, the financial characteristics of particular integrated markets could differ significantly.[75] Factors such as profitability and sales growth were significantly higher in the US than for example in Japan. Similar results were found in a more recent study that compared Mexican, Canadian and US manufacturing firms.[76] They conclude that manufacturing firms in the US have better means to meet their obligations and tend to have a greater liquidity. Differences can also be explained by the fact that US firms use less fixed assets when producing goods.

---

[72] Haskel, Jonathan, and Holger C. Wolf. "From Big Macs to iMacs : what do international price comparisons tell us?." World Economics 1.2 (2000): 167-178

[73] AC Nielsen. "New ACNielsen survey points to price convergence in Europe since the introduction of the Euro." Web. 25 June 2014. <http://nl.nielsen.com/site/documents/breakingnews_europe.pdf>

[74] These markets are Japan, the EU and US; Meric, Ilhan, Stephanie M. Weidman, Carol N Welsh, and Gulser Meric. "A Comparison of the Financial Characteristics of U.S., E.U., and Japanese Manufacturing Firms." American Business Review 20.2 (2002): 119

[75] Only manufacturing firms were taken into account.

[76] Meric, Ilhan, Herbert E. Gishlick, Leonore S. Taga, and Gulser Meric. "A Comparison of the Financial Characteristics of U.S., Canadian, and Mexican Manufacturing Firms."International Journal of Economics and Finance 5.7 (2013): 1-8

Furthermore, although many studies indicate that the variability of profits within the EU is similar to other integrated markets, no statistical studies have been performed to support these statements.[77]

Consequently, one may conclude that comparable companies from multiple pan-regions cannot be used simultaneously to determine an at arm's length range for a given industry, in a given pan-region. Therefore only the European market will be reviewed. Secondly, one could conclude that the profitability in the EU is not more dispersed than the other markets.

3.6   Economic circumstances

Furthermore, the OECD Guidelines mention several economic circumstances that might influence the comparability of the market/industry that is being benchmarked and thus influence profit margins.[78] They, inter alia, say that the place in the market can play a role when the market is being evaluated. For example, the market size and the extent of competition can influence the profit margins between different limited risk distribution wholesalers. In addition, the degree of substitution goods might even lower the overall profit margin for a particular limited risk wholesaler.[79] On this basis, any deviations may be explained by possible core-differences that come with those particular goods. For instance, the food industry might show significantly lower margins due to size of the market, the absolute simplicity of the business and (high) degree of competition. Conversely, the wholesale of pharmaceutical and medical equipment is accompanied by medical knowledge and legal advice. Therefore it is to be expected that wholesalers with a similar risk *level* have comparable margins (i.e. the wholesale of hardware, software, clothing, sanitary products etc.).[80] These industries all have in common that the core-function is wholesale and have little added-value, but do have some market knowledge.

---

[77] See for example: Monatsbericht. "Zur Unternehmensrentabilität im internationalen Vergleich." Bundesbank, 15 Oct. 1997: 33-44
[78] D.1.2.4. OECD Guidelines
[79] This is based on Porter's five forces model. The five forces determine the profitability of an industry. The degree of substitutions goods is one of the forces described: Hill, Charles W. L., and Gareth R. Jones. *Strategic management theory: an integrated approach*. 9th ed. Mason, OH: South-Western/Cengage Learning, 2010.
[80] NB: Although the risks might differ between industries, it is expected that the overall degree of risk exposure is (almost) the same.

3.7　　　　　　　Business cycle

Wholesalers that are situated in various industries might show different profit margins due to an industry business cycle. Differences in profit margins might occur although similar risks and functions are performed, simply because that industry had a 'good or bad' year. Consequently, fixed costs will lead to lower resp. higher margins. The possible existence of business cycles might thus blur the comparability of the result. The possible differences between industries might influence the relative profitability of companies because there are both high and low cyclical industries that could be distinguished.[81] For example, a wholesaler of food products is generally found in low-cyclical industry, where a wholesaler of specialised electronics might find himself in a more cyclical sector. Consequently, different industries are influenced by other market factors. Where the food industry might only be impacted by a change in the size of the population, the electronics industry might be influenced also by the economic climate.

Research has already been done in the manufacturing sector with regards to the profit margins and related business cycle.[82] They analysed a twenty-year period, starting from 1970. It could be concluded that the operational profit margins decreased heavily in times of recession. Consequently, this is in line with models that predict a similar pro-cyclical response of operational profit margins. The research holds for multiple sectors because producer goods as well as both durable and non-durable goods where included.

However, the moment of impact of the aggregated shocks seem to vary. Lima and Resende have performed a similar study for the Brazilian industry over a nine-year period by means of a data panel model, which applied a conjectural variation framework.[83] Comparable pro-cyclical results were found in regards to the aggregated cycle. However, the sector-specific variables did not show an obvious pro-cyclical behaviour. After a robustness check several variables showed a strong respectively

---

[81] Research has demonstrated that profits are influenced by the business cycle: Bikker, JA, and H Hu. " Cyclical patterns in profits, provisioning and lending of banks and procyclicality of the new Basel capital requirements." PSL Quarterly Review 55.221 (2012). Subsequently, business cycles exist: Schumpeter, Joseph Alois. "The Contours of Economic Evolution ."Business cycles : a theoretical, historical, and statistical analysis of the capitalist process. 1. ed. New York, NY McGraw-Hill, 1939. 179.
[82] Machin, Stephen, and John Van Reenen. "Profit Margins and the Business Cycle: Evidence from UK Manufacturing Firms."The Journal of Industrial Economics 41.1 (1993): 29-50
[83] Lima, Marcos, and Marcelo Resende. "Profit Margins and Business Cycle in the Brazilian Industry: a Panel Data Study."Applied Economics 36.9 (2004): 923-930

weak explanatory value. Both the lagged profitability and the import intensity showed a strong value, where union density had little impact.

Consequently, if differences in profit margins for similar routine wholesale industries occur, this might be explained by the difference in lagged profitability of that specific industry. It should be noted that the differences are not directly due to the lagged profitability. For example, after a (economical) shock the associated effect is not only present in year one, but is often persistent over time. Differences thus indicate that an asymmetrical shock occurred, impacting industries at another moment in time or did not impact some industries at all. Because the lagged profitability is persistent, it may take some time before such a shock is filtered from the profit margins.

Finally, in 2008 an economic crisis struck the European market.[84] This could influence the different industries significantly because they might be affected at a different point in time. Although no business cycle can be identified with a six-year dataset, a trend can be distinguished. It is to be expected that the profit margins of the different industries with similar "core-wholesale" distribution functions decline in a similar fashion.

3.8 European comparable companies

Further, the Deloitte White paper has compared country-specific data to a pan-European dataset.[85] More specifically they reviewed arm's length results of four different sets of comparable companies. Both the manufacturing and distribution function were analysed. NB: The papers mentioned in, for example, section 3.4 focussed on differences *between* continents (or pan-regions) were this paper is an addition which focuses on differences found *within* a pan-region.

The datasets were generated on the basis of statistical methods as well as TP principles. Several different, but representative datasets were made by making some conservative assumptions. These assumptions led to more comparable companies within each dataset. For example, only companies with an operating margin between minus five and plus fifteen percent were selected, in order to eliminated outliers and solely select

---

[84] Kotz, David M. "The Financial and Economic Crisis of 2008: A Systemic Crisis of Neoliberal Capitalism." *Review of Radical Political Economics* 41.3 (2009): 305-17
[85] Dawid, Roman, Peter Meenan, and Jörg Hülshorst. "Is Europe One Market? A Transfer Pricing Economic Analysis of Pan- European Comparables Sets." European Commision DOC: JTPF/007/BACK/2004/EN (2004): 1-54.

similar routine companies. All the filters that were applied are used in the daily practice of Deloitte.

Two different series of tests have been performed in order to increase the robustness of the results. This is done for two different datasets. The normal set was the result of a more detailed comparability analyses, where the broad set had less stringent selection criteria.

They found that the broad as well as the normal[86] dataset did not show significantly different arm's length ranges between the country-specific comparable companies and pan-European comparable companies.[87] In addition, the country-specific ranges that did differ from pan-European ranges showed no clear bias or patterns. They ultimately conclude that the pan-European comparable companies can be used to determine country specific arm's length ranges. Consequently, the pan-European comparable companies could also be used to compare different industries with similar routine functions across border. Furthermore, it can be concluded that the operational profit margins do not differ as result of a different national price levels.[88]

Finally, the use of a pan-European database is based on both the existence of an integrated European market and practical reasons. This reasoning is also followed by the EU Joint Transfer Pricing Forum whom states: "*comparables found in pan-European databases should not be rejected automatically*".[89] This suggests that non-domestic countries can be regarded as comparable.

---

[86] The Deloitte White Paper refers to the normal set as 'close' set.
[87] Based on the statistical equality of both the lower and upper quartile, under a 95% confidence level.
[88] In my opinion the Deloitte paper (indirectly) demonstrates that the relative cost and profits ratio are similar under different prices levels (i.e. under a different purchasing power parity).
[89] "Report on the Activities of the EU Joint Transfer Pricing Forum in the Field of Documentation Requirements"; EU Joint Transfer Pricing Forum JTPF/020/REV4/2004/EN (2005): §141

Chapter 4   Methodology

This section introduces the various statistical tests that will be executed in order to examine the data and demonstrate whether the profit margins of the different industries are equal.[90] The first test will examine the equality of the means, as it is the first moment of a distribution. Subsequently, the IQ range is being examined in more detail, as it is the second moment of a distribution: the variance of a distribution. Both the median, lower and upper quartiles are examined, in order to determine if the level of all quartiles are statistically equal and can therefore be used when establishing an at arm's length IQ range.[91] Finally, the total distribution of each industry is separately tested relative to the distribution of all industries. This should substantiate and 'group' the findings of the aforementioned tests. Finally, the data-selection process is part of the methodology, but will be discussed later in detail.

4.1   First-test: Comparing means

This thesis examines whether different industries have different profit margins. A first test is whether different industries have different mean profit margins. Thus, in order to test if an aggregated wholesale distribution IQ range could be applied for the comparability analysis, several statistical tests will be performed. The mean of each industry is compared to the aggregated mean. In order to test the equality of the means on profit level an ordinary least squares regression test (OLS) will be executed. Significantly different mean operating profit levels between industries imply different distributions of profit margins. The profit level (i.e. operating margin or OPM) will be denoted as the dependent variable, where the industry dummy variables will be denoted as the independent variables. The specification of the OLS variables can be considered to be an equivalent to an analysis of variance test.[92]

$$OPM_{it} = \propto\ + \beta_j \sum_{j=1}^{I} D_j +\ \varepsilon_{ijt} \qquad (1)$$

---

[90] The distribution of the profit margins is inherent to this.
[91] More specifically, the quartiles indicate the levels of the profit margins. It therefore is designed to test the level of the upper and lower quartiles.
[92] Also known as ANOVA test.

The OPM depicts the operating margin of a particular wholesale industry for a particular point in time. Further, the α stands for a constant. The 'I' depicts the total number of industries considered and the 'j' is a characteristic of the firm observation, the industry. The β_j shows the effect that different industries have on the operating margin. The D denotes all the dummy variables that are being tested. The dummy variables are binary: it equals one for the respective industry and zero otherwise. For example, the OPM obtained by a specific industry ('one') is thus being offset against the remaining industries (zero). In addition, each regression omits one of the four industries, which subsequently forms the intercept; this makes sure that there is no interference between dummies (i.e. the omitted industry always depicts the value zero).[93] The four regressions will each omit a different industry. Moreover, this test is a practical approach for the determination of an aggregated arm's length IQ range, since the distribution of the tested industry is included in the aggregated set.

The hypothesis (1) that is being tested could be defined as follows:

H_0 : *the mean of the TNMM range derived from the selected companies is equal to the mean of the TNMM range derived from respective comparable companies based on the aggregated dataset, in regards to the full-fledged distribution function.*

In order to test if the distribution is fundamentally different or not, an additional OLS test will be performed that compares the OPM of one industry to the remaining industries.[94] For example, in order to test whether a specific industry is significantly different, I add one dummy at a time, rather than all the four dummies at once and omit one (the intercept). More specifically, the dummy that is tested is being compared to the mean of the 'rest'. The 'rest' subsequently forms the intercept.

---

[93] Please note that the dataset is biased with the tested industry. This will be discussed in more detail in section 5.1.2

[94] NB The first regression tests the total aggregated set to the tested industry, including the tested industry, whereas this test demonstrates if the unbiased aggregated set (a set without the distribution of the tested industry itself) is different from the distribution of the tested industry.

The hypothesis (2) that is being tested could be defined as follows:

H_0 : *the mean of the TNMM range derived from one of the selected companies is equal to the mean of the TNMM range derived from respective comparable companies based on the remaining dataset, in regards to the full-fledged distribution function.*

This test will be performed on a year-by-year basis for a six year period in order to increase the strength of the model. For example, the null-hypothesis might be accepted when the analysis is solely done for one year, but rejected in subsequent years.[95]

The collected data is based on the assumption that functions performed and risks assumed are comparable. It is therefore to be expected that no significant differences appear on mean profit level since this formula only tests industries that perform similar routine wholesale functions and assume comparable risk-levels. N.B. these tests are thus solely focussed on wholesalers' distribution between different industries. This will implicitly control for sources that might affect profit margins, as was discussed in more detail in section 2.5.3. Thus, it controls for risks and differences in functions performed.

In addition, in line with the resource-based view, comparable profit margins are to be expected if similar risk levels are assumed for companies that have a similar function, namely wholesale distribution. Therefore no coefficients are included to correct for different risk-profiles.

However, if the results show a significant difference in operating margin, further analysis need to be done in respect to the risk-profiles. Divergent results might originate from the overall market size of that particular industry, liquidity of goods, but also the degree of risk that is assumed.

4.2   Second-test: comparing IQ ranges

As discussed in chapter two, both the profit split method and the TNMM uses IQ ranges for the determination of an at arm's length remuneration. More specifically, the range is defined as the distance between the lower (25th) and upper (75th) percentile of a distribution. It indicates the relative ranking of individual observations for a given

---

[95] Given that the same confidence interval is being applied.

variable (i.e. the variation), in this case the profit margin.[96] However, the interpretation and implementation of the ranges may differ between tax authorities.

For example, the DTA take the median as a starting point, while other tax authorities might choose a different point in the range because they want to deal in favour of the taxpayer. Because these different treatments exist, both the level of the lower and the upper IQ range will be statistically tested. Subsequently, statistical test will also be performed on the basis of the medians found. Consequently, test are performed to check if differences exist between the medians of industry-specific routine wholesale distribution functions and the aggregated median of industry-specific routine wholesale distribution functions as well as IQ ranges. The same tests are performed for the median as well as the lower and upper quartile.[97]

The arm's length IQ ranges are separately tested by means of the Chi-square $-\chi^2-$ test. This test analysis the relationship between two categorical variables.[98] Statistical conclusions can be drawn with the Chi-square test by comparing the medians of two or more independent datasets/populations. The Chi-square test is a nonparametric test, and can therfore be used to calculate the arm's length quartiles without further adjustments with regard to the distribution of the underlying data.[99] In addition, this test can be applied when relative small sample sizes are available. However, the Chi-square test does not say much about the strength of the relationship.

The hypothesis (3) that is being tested could be defined as follows:

H_0 : *the median repectively the upper respectively the lower quartile of the TNMM range derived from the selected companies are equal to the median respectively the upper respectively the lower quartile of the TNMM range derived from respective comparable*

---

[96] Carlson, William L., and Betty Thorne. Applied statistical methods: for business, economics, and the social sciences. Upper Saddle River, N.J.: Prentice Hall, 1997. P. 53-54

[97] Please see below for a detailed discussion of the tests that will be performed at 1st (25%), 2nd (50%) and 3rd (75%) quartile level.

[98] Myatt, Glenn J.. Making sense of data: a practical guide to exploratory data analysis and data mining. Hoboken, N.J.: Wiley-Interscience, 2007. p.82-83

[99] No stringent assumptions are made with regard to the distribution of the data under nonparametric tests.

companies based on aggregated industry-specific data, in regards to the full-fledged distribution function.

This test will be performed on a year-by-year basis for a six year period in order to increase the strength of the model. For example, the null-hypothesis might be accepted when the analysis is solely done for one year, but rejected in subsequent years.

In addition, the use of multiple year data takes into account the cyclical nature of businesses. The OECD Guidelines recognize the importance of using multiple years of data when presenting an arm's length result.[100] Therefore the weighted average of two *three year periods* will be compared. The weighted average is determined as follows:

$$Weighted\ average = \frac{\sum_{t=1}^{3} Operating\ Profit}{\sum_{t=1}^{3} Turnover} \qquad (2)$$

4.2.1 Standard Chi-square test

The chi-square test makes use of contingency tables, which is simply an array of numbers in a matrix. For example, a researcher may state that he observed 204 cars on the main road, or he can state that he observed:

BMW	Volkswagen	Others	Total
66	46	92	204

and thus uses a 1x3 contingency table. The numbers represent frequencies or observations found within a population. The more rows he includes, the more specific he can be. For example, he could make a distinction between old and new cars.

---

[100] The OECD 'Revision of chapters I-III of the Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations' (The Guidelines), paragraph 3.75-3.79.

This will lead to an 2x3 contingency table and so on (i.e. the r x c contingency table). For example:

	BMW	Volkswagen	Others	Totals
New car	40	21	15	76
Old car	26	25	77	138
Totals	66	46	92	204

The observed values are than offset against the expected values. The use of the word "expected value" has the following meaning in the context of that what is tested here: namely, whether the three types of cars have a similar distribution across the "old" and the "new" group. For example, the expected value for a new BMW car can be calculated as follows:

$$E_{New\ BMW(i,j)} = \frac{n_i C_j}{N} = \frac{66 \cdot 76}{204} = 25 \qquad (3)$$

Subsequently, the χ^2 is then calculated by means of the function provided below:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (4)$$

Hereby c depicts each random sample and r the observed number of observations of a subgroup in a random sample. The O_{ij} respectively E_{ij} stands for the specific observed respectively expected value in cell (i, j). "H_0 is true if there are small, if any, differences between the observed and expected value. That means that the number of observations in cell(i, j) should approximate the expected value found. That is the ith sample size n_i multiplied by the proportion of all the observations in the same category (j)."[101]

The χ^2 is than compared to the test statistic T (i.e. its critical value) with (r1)(c-1) degrees of freedom.[102] NB: Since Excel is being used, only the p-values are given/returned. Consequently, these values are used to determine the level of significance of the results found.

---

[101] Conover, W. J.. Practical nonparametric statistics. 2d ed. New York: Wiley, 1980: 155
[102] Conover, W. J.. Practical nonparametric statistics. 2d ed. New York: Wiley, 1980: 156

Finally, in order to test if the IQ range is fundamentally different or not, additional Chi-sqaure tests will be performed that compares the profit margins of one industry to the remaining industries.

More specifically, this method can tackle two statistical problems:[103, 104]

- The median test; and
- The goodness-of-fit test

4.2.2 Medians Test to test the IQ range

The Median's test examines whether or not various samples of populations have the same median and can be considered to be a special application of the standard Chi-square test.[105] It gathers all the observations and determines the overall 'grand' median of the aggregated set (N). Subsequently, each c group (i.e. industry) is compared to the grand median. For example:

Sample	1	2	...	c	Totals
> Grand median	O_{11}	O_{12}	...	O_{1c}	a
≤ Grand median	O_{21}	O_{22}	...	O_{2c}	b
Totals	n_1	n_2	...	n_c	N

In order to apply the Median test on the lower respectively upper quartiles, the dataset needs to be modified to medians of the specific distribution. Firstly, each individual string of industry data will be divided in two subsets of equal observations (i.e. at the level of the median). Subsequently, the lower resp. upper subsets are tested for the equality of the lower resp. upper quartile. This is done by counting the number of observations greater than the grand median and the number of observations that are equal or smaller than the grand median. For example, all 2007 data strings are divided in two subsets at the level of the median, a lower resp. upper subset. Subsequently, the median of the lower subset of the chemical industry (i.e. the old lower quartile of that

---

[103] Conover, W. J.. Practical nonparametric statistics. 2d ed. New York: Wiley, 1980.
[104] These methods will be discussed in more detail further on.
[105] It is solely a special application of the before described Chi-square test.

industry) is compared to the grand median. The grand median is the median that stems from the aggregated lower subset.

In addition, the median of each separate industry is tested on equality with respect to the aggregated industry median by means of the Chi-square Medians test. The test supports the research question if it appears that the medians do not significantly differ from the aggregated median.

4.2.3 Third-test: Comparing total distribution (goodness-of-fit)

The goodness-of-fit test compares random samples with a sample that would be expected from a hypothesized distribution (i.e. "if the hypothesized distribution function 'fits' the date of the sample"[106]). This is also a special application of the standard Chi-square test. It can be used to test random industry samples against a hypothesized expected frequency. The hypothesized distribution is based on the median, upper and lower quartile of the aggregated set of operating margins. Thus, instead of testing the median, upper and lower quartile on a standalone basis, the goodness-of-fit test compares the whole distribution of one industry to the aggregated distribution. More specifically, it tests the equality of the median, upper and lower quartile of a specific industry as a whole, to the median, upper and lower quartile of the aggregated distribution.

The hypothesis (4) that is being tested could be defined as follows:

H_0 : *the distribution of the TNMM range derived from the selected companies is equal to the distribution of the TNMM range derived from respective comparable companies based on the aggregated dataset, in regards to the full-fledged distribution function.*

---

[106] Conover, W. J.. Practical nonparametric statistics. 2d ed. New York: Wiley, 1980: 189

The observerd frequency is being compared to the expected frequency in a 1x4 contingency table. The same Chi-square formula applies with (c-1) degrees of freedom. For example:

	Subgroup 1 $(-\infty, 25^{th}]$	Subgroup 2 $(25^{th}, 50^{th}]$	Subgroup 3 $(50^{th}, 75^{th}]$	Subgroup 4 $(75^{th}, \infty)$	Total
Observed frequency	25	30	44	33	132
Expected frequency	33	33	33	33	

(The expected count is 25% of the total per subgroup)[107]

For example, all the operating margins of the year 2007 are included in one string of data, after which the median, upper and lower quartile are determined. Let 1% be the lower quartile, 2% the median and 3% the upper quartile of the aggregated distribution. The 1%, 2% and 3% are considered to be the hypothesized distribution, creating subgroup 1: $(-\infty, 1\%]$, subgroup 2: $(25^{th}, 2\%]$, etc. The 2007 dataset of the chemical industry is subsequently compared to the hypothesized distribution. The number of observations up to and including 1% are counted. Then the number of observations from 1% up to and including 2% are counted, and so on. The expected frequency is the total number of observations multiplied by a quartile.

---

[107] In this example: 25% * 132 = 33.

Chapter 5   Data and data selection

Firstly, this section will discuss the selection procedure that would normally be done by MNE's when performing the economical analysis (i.e. the comparability analyses). This partly explains the data-selection procedure performed in this thesis. Secondly, the selection procedure will be discussed in detail, after which the final dataset will be presented. The final dataset will be used to execute the aforementioned tests.

5.1   Comparability analysis in practice

In order to objectively quantify their intercompany transfer price, enterprises must search for comparable transactions once an appropriate transfer pricing method is chosen. However, the problem that presents itself is the collection of financial information of individual uncontrolled transactions on a similar functional level. Most unrelated enterprises perform more than one function. In addition, accounting standards do typically not require enterprises to split their profit and loss accounts into multiple functions, through which specific transactions cannot be traced back.

The OECD Guidelines prescribe a nine-step plan for a structured approach to a comparability analysis.[108] Once the tested party and method is chosen, a distinction should be made between internal and external comparable companies. It is not always possible to find similar comparable companies. Therefore, after appropriate comparable companies are selected, if any, there might have to be made adjustments in order to increase the comparability. In addition, multiple year data could be used to increase the comparability.

5.1.1   Types of comparable companies

If comparable companies can be derived from similar third party transactions, they then can be used to demonstrate the validity. The comparable companies can be obtained by a variety of sources and could be divided into two categories:

- Internal comparable companies; and
- External comparable companies.

---

[108] §3.4 OECD Guidelines and further elaborated between A2-B5 of chapter 3

5.1.1.1 Internal comparable companies

An internal comparable is a similar transaction that the enterprise under review, or an affiliated enterprise, has with a third party (i.e. an internal CUP). These comparable companies are more likely to match the transaction that is being assessed since they are part of the same group and often carry out comparable functions. In addition, more information regarding the compared situation will be available, allowing more certainty towards the comparability.

5.1.1.2 External comparable companies

External comparable companies are similar (uncontrolled) transactions entered into between unrelated enterprises. Information is typically obtained through existing data in the public domain by means of, for example, commercial databases, industry organisations or the knowledge of employees. Employees could assist in the search of comparable companies since they are able to provide valuable information about, for instance, competitors. In turn, industry organisations publish several documents and studies that could assist in, for instance, specific industry information. However, the actual external comparable companies are typically found by (commercial) online databases. These databases differentiate both potential comparable companies and associated financial information. The financial information is often derived from published financial statements.

However, there are disadvantages to acknowledge when information is solely derived from financial statements. First of all, information obtained from these databases may be difficult to compare due to different accounting standards across border. For example, some countries only have a limited disclosure requirement. Secondly, comparability factors of unaffiliated enterprises are difficult to be deduced from the financial statements as well as specific transaction levels. Thirdly, it is not always evident whether or not a part of the data found contains intercompany transactions. In addition, it is not always evident to what extent intangibles are used within the transactions found.

These disadvantages can be partially alleviated by a portfolio approach or a package deal.[109] Transactions that are closely intertwined and related with each other can be evaluated in conjunction under the application of the correct TP method. The portfolio approach states that enterprises could bundle certain transactions as part of a business strategy giving them an appropriate return across a bundle of products. The package approach bundles several, usually intangible, benefits. These benefits comprise, for example, of patents, know-how, the provision of headquarter services and the lease of distribution facilities. Foregoing, together with qualitative research of the facts and circumstances ensure that the use of databases can lead to an arm's length remuneration. The quality of the data found can be refined with publicly available information.[110] The reliability of the comparable companies found can, for example, be verified through an analysis of annual reports or by information found on the Internet.

5.1.2   Dataset

The profit margins used in this thesis are similar to those that will be derived when applying the TNMM with the profit level indicator 'return on sales'. This means that the data retrieved looks at EBIT level and not gross margin level. Furthermore, the dataset will consist of a comparable dataset, which is solely derived from independent parties. It is important to note that the aggregated industry IQ range needs to include the tested industry within the aggregated dataset because the tests look at mutually exclusive options (i.e. the tests demonstrate if significantly different arm's length IQ ranges occur when the aggregated wholesale industry data or the routine wholesale industry-specific data is used).[111] For example, the distribution of the electronics industry is compared to the aggregated set of the distributions of all the industries (including the electronics industry). As a result, the aggregated industry dataset will include the tested routine industry if it turns out that the IQ ranges do not significantly differ from one another (and can subsequently be used in the day-to day practice).

---

[109] §3.10 and §3.11 OECD Guidelines
[110] §3.33 OECD Guidelines
[111] The same approach is used in: Dawid, Roman, Peter Meenan, and Jörg Hülshorst. "Is Europe One Market? A Transfer Pricing Economic Analysis of Pan- European Comparables Sets." European Commision DOC: JTPF/007/BACK/2004/EN (2004): 1-54.

However, since the underlying question would be whether the distributions are different or not, a additional tests will be performed. These tests will exclude the industry that is being tested from the aggregated dataset. For example, the distribution of the electronics industry is compared to the aggregated set of distributions of the remaining industries (one versus the rest). Consequently, the remaining set is not biased with the tested industry.

5.2   Data selection procedure for exploration safe harbours

This study uses data obtained from Amadeus, a European database with financial information of roughly twenty million companies.[112] Amadeus gathers financial data from all the companies across Europe, if available. For example, some countries do not require entities to share their financial statements with the tax authorities and public, making it difficult to obtain country-specific data. The pan-European comparable companies come from the following countries: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom plus Iceland, Norway and Switzerland (pan-European + 3).

The Amadeus database makes use of the NACE-code[113] classification system, which allows enterprises to further select potential comparable entities. Both the relevant functions and product characteristics (i.e. industries) can be selected because the NACE-codes classify all the productive economic activities that take place in the EU.[114] NACE makes use of a hierarchical structure that comprises out of four levels. For example:
- section G describes the wholesale and retail trade;
- division 46 describes the wholesale trade, motor vehicles excluded;
- group 46.4 describes the wholesale of household goods; and
- the class level narrows it further down to wholesale of textile goods (46.41), etc.

---

[112] Amadeus is published by Bureau van Dijk; version 221 consist of 19.772.536 companies in total.
[113] NACE is an acronym of "Nomenclature statistique des activités économiques dans la Communauté européenne".
[114] NACE Rev. 2: statistical classification of economic activities in the European Community. Luxembourg: Office for Official Publications of the European Communities, 2008. P. 15

In order to test the hypothesis, several wholesale distribution industries where selected on the basis that they are expected to have a similar degree of competition. In addition, a similar response in regards to the economical circumstances is expected.[115] At first view, these industries all have in common that the core-function is wholesale and have little added value, but do have some market knowledge. Please note that this is a preliminary selection of industries. Various NACE-codes were selected that represent the following industries[116]:

- Appearance products
- Chemicals
- Electronics products
- Industrial products
- IT Products
- Automotive

This will give a comprehensive overview of different industries with different NACE-codes, making it possible to extensively test the hypothesis'.

Additional NACE-codes were selected of industries that are expected to have lower, resp. higher risk profiles, namely the food and pharmaceutical industry. N.B. this is based on the idea that Amadeus can only identify and characterise companies to a certain level. This is a valid assumption since the Amadeus database is used on a global scale, for which no better alternative for the day-to day TP practice exists.

Comparability adjustments may be appropriate if potential comparable companies found are initially not comparable. The type of adjustment that needs to be made depends on the facts and circumstances *(see Appendix B for a detailed discussion)*. Therefore, adjustments need to be made in order to get a *comparable dataset* out of Amadeus.

Consequently, some initial filters will be applied to obtain comparable companies that are potentially more in line with the industry and function under review. For example, companies that are recorded by Amadeus to be active or have an unknown status were selected. Since this study focuses on MNE's/globally integrated enterprises (GIE),

---

[115] This refers to the preliminary expectation that they have similar business cycles.
[116] See Appendix C for all the selected NACE-Codes

subsidiaries need to be eliminated from the dataset.[117] Subsequently, companies that hold shares of 50-100 percent in subsidiaries are rejected. Subsidiaries can be industrial companies, hedge funds, bank and insurance companies, trustees etc. This way, underlying TP issues taking place in those entities, will be out of the dataset. In addition, entities with a lack of financial information from preceding years as well as start-up businesses are eliminated from the dataset. With respect to the former, the data can be perceived to be unreliable. The same reasoning applies to start-up businesses because they typically show start-up losses etc. and thus blur the results.

Normally companies that show an absolute loss over a three-year period are eliminated under the assumption that they are not comparable to the tested party. Please note that this thesis assumes a tested party that represents a routine wholesale company and is full-fledged.[118] This will result in a uniform set of comparable wholesale companies. Consequently, a filter has been applied to eliminate these loss making companies. This is done for two periods, namely 2007-2009 and 2010-2012. For example, a company is eliminated when he has an absolute negative average operating profit for the period 2007-2009.

Furthermore, an intellectual property filter is applied in order to increase comparability (i.e. create a comparable set of companies). Companies that have a significant different (/high) level of intangibles to total asset ratio are not comparable. Firstly because of their ability to earn above normal returns.[119] And secondly, simply because differences in the level of intangibles indicate dissimilar functions, where this thesis is designed to test a similar group of companies.

Finally, no filters where applied that could influence differences in reported profit margins due to, for example, divergent accounting standards or different levels of foreign exchange exposure.

---

[117] Subsidiaries are by definition dependent and can therefore not be used as reliable data. NB: it is not possible to check whether or not they have at arm's length prices themselves.
[118] Regardless of what industry he is in.
[119] Please see the discussion in regards to the resource-based view, section 3.1.

In summary, firms satisfying the following quantitative[120] criteria are **not** selected:[121]

- Non-active companies

- Shareholders with more then 49,9 percent ownership

- Companies that were not incorporated in the preceding five years

- Companies with no financial information over the last three years from t_1 / failed to report

- Companies that have a high intangible asset : total asset ratio

- Companies that show an absolute negative operational profit over a three year average

In order to further enhance the dataset, a qualitative search will be performed. The qualitative search compares the potentially comparable enterprises to the tested party. Recall that the tested party is a routine wholesale company and full-fledged. This assumption is necessary because the results retrieved from the Amadeus database can be to dispersed and unreliable. Therefore, quantitative search need to be performed in order to end up with a uniform dataset that represents a set that is used in the day-to day practice.

This can initially be done by the exclusion and/or inclusion of specific word stems. For example, the exclusion of the word stem 'consul*' will lead to the exclusion of derived words therof, such as consultants, consulting, consult, etc. The functions performed and risks assumed are further compared by means of, for example, a website check and trade description check. Amadeus provides short trade descriptions of the companies in the dataset, if publicly available. Based on this check, companies can be eliminated from the dataset if the description differs from the tested party. Normally, a website check, or any other publicly available source will be consulted. This obviously is a labour-intensive process and out of scope for this thesis. It is compensated for by the removal of outliers. These comparable companies have extraordinarily high resp. low profits and are thus assumed to have a functional difference or abnormal risk-level of which it is not possible to identify them with Amadeus.

---

[120] This refers to criteria that could be quantified, meaning that Amadeus has identified objectified criteria that can easily be identified and subsequently removed without further research.
[121] N.B.: The companies that fall within the scope of the screens will be eliminated.

Outliers have also been removed because they can strongly influence the OLS regression results. Outliers can significantly influence the value of the mean and thus give less reliable results. An outlier can be defined as an observation, which greatly deviates from the other observations.[122] Companies were removed that fell out of the range that is based on three times the Median Absolute Deviation (MAD).[123, 124] Unlike a standard deviation based on the mean, the MAD is not sensitive to outliers when the standard deviation is determined. For example, all the observations that had an operating margin greater than the median plus three times the MAD were eliminated.

A six year period will be analysed in order to avoid corrupted results. For example, the test might indeed show a similar remuneration between different industries, but is influenced by a business cycle in year one.

Finally, it is important to note that this thesis is not designed to fully explain the profit margins. We will implicitly control for other deteminants (i.e. missing variables) by means of these filters.

---

[122] Carlson, William L., and Betty Thorne. Applied statistical methods: for business, economics, and the social sciences. Upper Saddle River, N.J.: Prentice Hall, 1997. p.40

[123] Ellison, S. L. R., and Trevor J. Farrant.Practical statistics for the analytical scientist a bench guide. 2nd ed. Cambridge, UK: RSC Publishing, 2009: 55-56

[124] The MAD can be used as a 'classical' standard deviation under the assumption that the distribution is approximately normal. The found MAD is subsequently multiplied by 1,483.

5.3 Dataset selection

The above has resulted in the following dataset:[125]

Figure 4: NACE-codes, inclusions and exclusion of the Amadeus dataset

Wholesale Distribution Industry	Results dataset (six year average)							
	Appearance Products	Automotive	Chemicals	Electronica Products*	Food Products	Industrial Products	IT Products**	Pharmaceuticals
Geograpical Area	pan-European + 3							
NACE Rev. 2 Codes	46.45: Wholesale of perfume and cosmetics 46.48: Wholesale of watches and jewellery 46.42: Wholesale of clothing and footwear	45.11: Sale of cars and light motor vehicles 45.31: Wholesale trade of motor vehicle parts and accessories	46.12: Agents involved in the sale of fuels, ores, metals and industrial chemicals 46.75: Wholesale of chemical products	46.43: Wholesale of electrical household appliances 26.11: Manufacture of electronic components	46.31: Wholesale of fruit and vegetables 46.32: Wholesale of meat and meat products 46.33: Wholesale of dairy products, eggs, edible oils and fats	46.64: Wholesale of machinery for the textile industry 46.66: Wholesale of other machinery and equipment	46.5: Wholesale of information and communication equipment	46.46: Wholesale of pharmaceutical goods 47.74: Retail sale of medical and othopaedic goods in specialised
Industry description (inclusions)***	Trade*, Distrib*, Wholesale*, Sale*, Cloth*, Parf*, Appear*	Trade*, Distrib*, Wholesale*, Sale*, Auto*, Car*, Motor*	Trade*, Distrib*, Wholesale*, Sale*, Chemi*, Fluor*, Lac*, Wax*, Adhesiv*	Trade*, Distrib*, Wholesale*, Sale*, Appliance*, Electr*	Trade*, Distrib*, Wholesale*, Sale*, Food*, Beverage*	Trade*, Distrib*, Wholesale*, Sale*, Industr*, Machin*, Construct*	Trade*, Distrib*, Wholesale*, Sale*, Software, Hardware, Conductor*, Compu*, Tele*, Peripher*	Trade*, Distrib*, Wholesale*, Sale*, Medic*, Pharma*
Industry description (exclusions)***	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*, Food*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*, Food*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*
Average size of dataset / number of observations****	915	1276	469	423	1479	135	494	595

* Manufacturing codes can be applied due to the exclusion and inclusions words. In addition, every independent companie performs a sales function, because it can not be done by a principal / headquarter.

** Contrary to the electronica products, which are lamps, vacuums, toasters, etc., the IT products comprise of software and hardware used in the ICT sector.

*** The asterisk implies that both those exact words and derivatives thereof are taken into account.

**** Data is gathered from 2007 till 2012, the six year average is shown.

N.B. the number of companies found (i.e. profit margins) per industry is not distributed homogeneous across country and over time. This is partially the result of different levels of compliance across Europe because not all companies have to make their financial statements public. For example, in Germany companies do not have to publish their financial statements. When the Dutch companies are obliged to make their financial statements public, which they are, relatively more Dutch companies are expected to be included in the dataset.

The Amadeus database can only apply several filters on the preliminary dataset, namely the specific NACE-codes, level of independency, geographical region, non-active companies and the inclusion resp. exclusion words. Therefore, all the companies that

---

[125] Please see Appendix C for a more detailed view.

failed to report twice over the last three years were manually eliminated, including companies that were not incorporated (i.e. failed to report trice).

Several intellectual property filters were examined in order to increase comparability. These filters eliminated companies that own more than a particular percentage of intangibles with respect to their total assets. Several levels were examined:

- One percent;
- Three percent;
- Five percent; and
- Ten percent.

None of the manually applied intangible filters led to a significant elimination or change in profit margins. These filters are therefore not taken into account, since insignificant alterations downgrade the quality of the dataset and the subsequent findings.

Normally companies that show an absolute loss over a three-year period are eliminated under the assumption that they are not comparable to the tested party. More specifically, loss making companies are assumed to go bankrupt over time. Consequently, a filter has been applied to eliminate these loss making companies. This is done for two periods, namely 2007-2009 and 2010-2012. For example, a company is eliminated when he has an absolute negative average operating profit for the period 2007-2009.

This resulted in the following timeline:

Figure 5: Medians per industry over time without loss makers
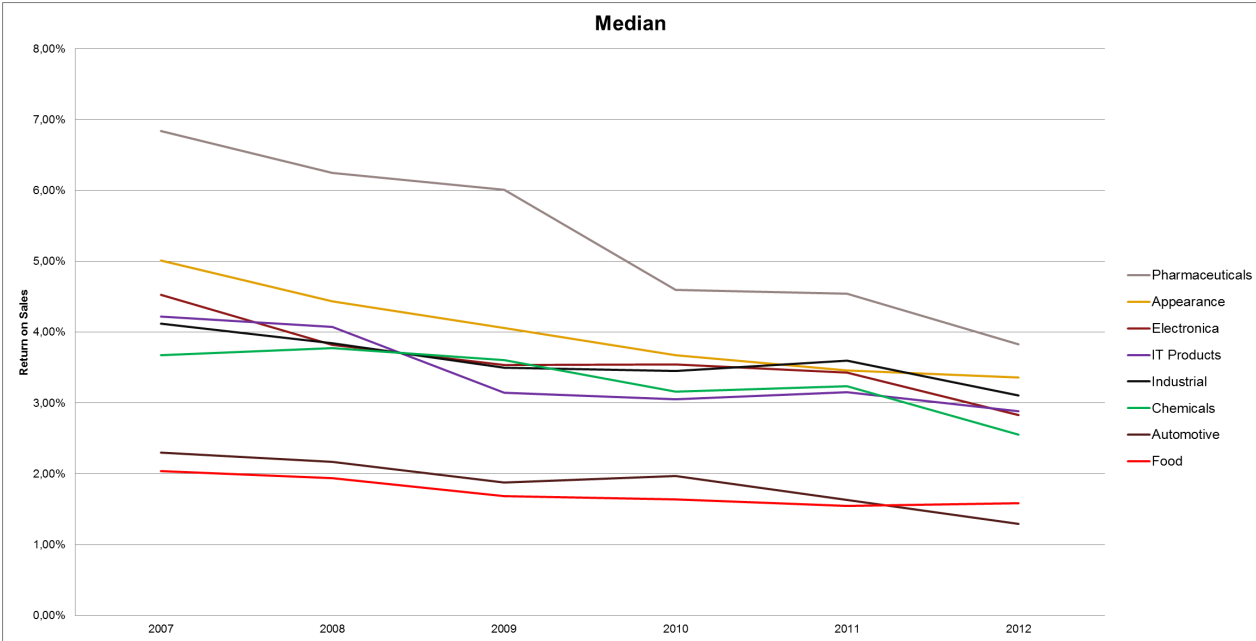
NB: Only the median is depicted in this chart. Appendix D shows the median as well as the upper and lower quartile in more detail.

Even without loss making companies a clear downwards trend emerges. This is in line with the expectation that the economic crisis impacts all the industries. The group of companies in the middle show a similar level of decline (i.e. this is the group that we are testing). The pharmaceutical group differs significantly from the middle group. This is also in line with expectations, since their risk-profile is expected to be higher, inter alia, due to more specific advice needed for the distribution of goods. Therefore, the pharmaceutical companies will not be taken into account when the aggregated set is formed. Consequently, the industry is only included for reference.

In addition, a lower group can be identified that includes the automotive and food industry. As previously stated, the wholesale distribution food companies are expected to have a lower risk-level due to a different business cycle, if any, due to the size of the market, the absolute simplicity of the business and (high) degree of competition. Foregoing could also be deduced from the timeline since the food industry responds relatively less intense on the impact of the economic crisis. Consequently, the food industry is also only included for reference.

50

Finally, the automotive industry has a remarkably low median, with an absolute downfall in the lower quartile (2012). This can be explained by the fact that the 'hard-core' automotive *wholesale* industry is limited to a handful of multinational enterprises. For example, BMW, Volkswagen and Peugeot are all dependent enterprises and therefore by definition excluded from the dataset. More specifically, these enterprises were already eliminated from the dataset since they are dependent companies. The remaining ''ordinary'' independent companies do not find their existence in (whole)selling cars, but from the additional services: repair. Consequenly, their core-function is not considered to be wholesale. In addition, buying a car or the maintenance thereof is usually expensive and thus often postponed in times of crisis. This is less the case when buying, for example, a personal computer or chemicals that allows a consumer to maintain his own car. Consequently, relative low profit margins arise in times of crisis and vice versa. It can thus considered to be a more cyclical industry and will therefore not be included in the dataset.

However, the finding of a clear downward trend led to the conclusion that loss making companies cannot simply be excluded from the dataset anymore, since it might well be at arm's length in times of crisis to consider the profit margins of these loss making companies. The inclusion of loss making companies will thus be more in line with the current economic climate and reality. Finally, from this moment on the outliers were removed.

As shown below, an overall drop in margins is the result from the inclusion of loss making companies:

Figure 6: Medians per industry over time with loss makers
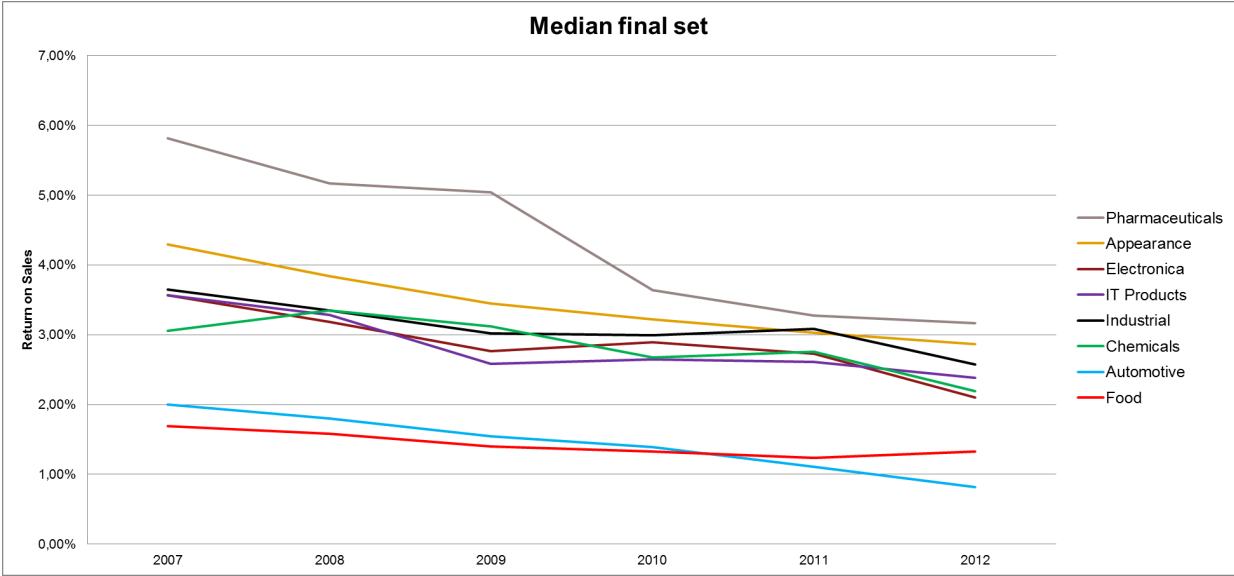
NB: Only the median is depicted in this chart. Appendix E shows the median as well as the upper and lower quartiles more detailed.

All the above has resulted in the following final dataset and will be further statistically analysed:

Figure 7: Final results after qualitative and quantitative filters

Refined results dataset (six year average)*								
Wholesale Distribution Industry	Appearance Products	Automotive	Chemicals	Electronica Products	Food Products	Industrial Products	IT Products	Pharmaceuticals
Size of dataset / number of observations	787	1064	375	332	1249	113	407	501
Mean Sales (x1000)	€ 3.208	€ 10.121	€ 8.394	€ 8.633	€ 8.536	€ 4.010	€ 7.613	€ 9.243
Profit Level Indicator (PLI)	Return on Sales							
Mean	3,97%	1,43%	3,57%	3,29%	1,74%	3,64%	3,54%	5,63%
Lower Quartile	1,43%	0,36%	1,43%	1,02%	0,56%	1,58%	1,22%	1,85%
Median	3,27%	1,37%	2,91%	2,83%	1,46%	3,15%	2,85%	4,36%
Upper Quartile	6,07%	2,60%	5,43%	5,36%	2,71%	5,40%	5,44%	8,71%

 * Corrected for outliers (median + 3x median absolute deviation)

Chapter 6   Results

This section will explain the results found with regard to the empirical analysis. Do the results indicate that it is indeed possible to use the aggregated wholesale distribution IQ range of multiple industries on a stand-alone basis when determining an at arm's length IQ range? The analysis is done by means of several regressions and Chi-square tests, which are displayed in the following tables.

6.1   Presentation of results

Because Excel was used when performing both the Median's test and the goodness-of-fit test, no χ^2 coefficients are presented. Instead, only the found probability values (p. value) were returned in Excel and therefore implemented in the table. Further, the table shows the results for all the years and periods tested. In addition, both the tests on lower and upper subsets are given as well as the test on median level of the total set. For example, table 3 shows that the Median's test on lower quartile subset returned a p. value of 0,905 in the year 2008, an insignificant result.

Where most studies try to demonstrate significant results, this research is interested in insignificant results in order to confirm the null hypothesis. A significant result would reject the null hypothesis. This means that the quartile could not have been used to state that the aggregated IQ range could be applied when establishing an arm's length remuneration. Due to this difference, the significant results are presented in a similar way, but need to be interpreted differently. For example, note that a three stars (***) represent a significance level of $\propto = 1\%$ and two stars (**) represent a significance level of $\propto = 5\%$. Finally, a single star (*) will represent a significance level of $\propto = 10\%$. Please note that significant results are perpendicular to the research question.

In addition, the regression output is similar to what one would expect, meaning that it shows the coefficients and standard errors. The significance level will be presented in a similar way as done with the Chi-square tables. NB It uses the same stars and the underlying meaning. Furthermore, the regression results are multiplied by 100 in order to give a clear presentation of the results. The coefficients and standard errors showed

low values due to the fact that they originate from percentages. For example, the actual standard error of the appearance industry in 2007 is 0,00313 (Table 1).

Finally, if there are more significant results to be found in the tables, than that indicates that there are differences between industries and the hypotheses should be rejected.

6.2 Preliminary statements

This section will demonstrate that the first important results indicate that the appearance industry is significantly different compared to both the aggregated set and the 'rest'. Firstly, the following table depicts the regression results whereby the aggregated set is tested:

Table 1: Summary of regression results aggregated set (all five industries)

	2007	2008	2009	2010	2011	2012	07-09	10-12
Excluding electronics industry								
Appearance	1.038***	0.884***	0.758**	0.286	0.514***	0.909***	0.790***	0.423**
	(0.313)	(0.3190	(0.325)	(0.194)	(0.195)	(0.219)	(0.278)	(0.177)
IT Products	0.456	0.513	-0.348	-0.252	0.009	0.306	0.195	-0.038
	(0.324)	(0.333)	(0.340)	(0.238)	(0.239)	(0.271)	(0.289)	(0.216)
Chemicals	-0.179	0.108	0.146	0.097	0.114	0.317	-0.044	0.120
	(0.344)	(0.353)	(0.359)	(0.227)	(0.229)	(0.255)	(0.308)	(0.207)
Industrial	0.386	0.884	0.239	-0.002	0.722**	0.480	0.196	0.437
	(0.495)	(0.319)	(0.520)	(0.335	(0.340)	(0.368)	(0.443)	(0.306)
Intercept	4,005	3,629	3,459	3,457	3,119	2,500	3,903	3,229
	(0,262)	(0,269)	(0,273)	(0,162)	(0,163)	(0,185)	(0,232)	(0,147)
R square	0.013	0.007	0.011	0.003	0.005	0.010	0.008	0.004
N (total panel observations)	1728	1764	1626	2423	2441	2153	1880	2589

This regression tested the equality of the aggregated mean. The constant of the regression (intercept) is the mean of the omitted industry[126]. On the basis of this regression, it is possible to state whether the means of the industries differ or not. More specifically, an insignificant result implies that one cannot state that the industries are different.[127] These first results clearly show that the appearance industry differs significantly from the aggregated mean. For example, highly significant results in the

---

[126] That industry whose dummy is always zero, which is the electronics industry in this regression.
[127] Please note that this regression cannot state whether or not the omitted industry differs from the average, because the omitted industry is part of the aggregated average.

years 2007, 2008, 2011, 2012 and even the time period 2007-2009 lead to the rejection of the null hypothesis (1). Consequently, the aggregated mean seems to be influenced significantly by the appearance industry. The low r-square is due to the fact that no control variables were used in the regression, because the selection procedure implicitly controlled for this. The r-square value depicts the level of the variation of the operating margins that is explained by the OLS of the margins on the dummies.[128]

Subsequently, multiple regressions were performed in order to determine if an industry differs from the rest on a stand-alone basis. For example, the mean of the appearance industry is compared to the mean of the remaining industries (first row). See the following table for the five stand-alone regressions for multiple years:

Table 2: Summary of regression results 'one versus the rest'

	2007	2008	2009	2010	2011	2012	07-09	10-12
Appearance	0.874***	0.642***	0.832***	0.322**	0.407***	0.674***	0.708***	0.353***
	(0.210)	(0.212)	(0.219)	(0.140)	(0.141)	(0.155)	(0.188)	(0.127)
Intercept	4,170	3,871	3,385	3,421	3,226	2,735	3,984	3,299
	(0,121)	(0,125)	(0,128)	(0,091)	(0,091)	(0,101)	(0,109)	(0,082)
R square	0.010	0.005	0.009	0.002	0.003	0.009	0.008	0.003
IT Products	-0.001	0.066	-0.756***	-0.415**	-0.319*	-0.254	-0.173	-0.303*
	(0.224)	(0.230)	(0.237)	(0.191)	(0.191)	(0.216)	(0.202)	(0.173)
Intercept	4,462	4,077	3,867	3,621	3,447	3,060	4,271	3,494
	(0,117)	(0,118)	(0,121)	(0,075)	(0,076)	(0,084)	(0,104)	(0,069)
R square	0.000	0.000	0.006	0.002	0.001	0.000	0.000	0.001
Chemicals	-0.792***	-0.443*	-0.080	-0.002	-0.201	-0.253	-0.454**	-0.119
	(0.250)	(0.256)	(0.261)	(0.177)	(0.179)	(0.196)	(0.225)	(0.161)
Intercept	4,618	4,180	3,685	3,557	3,434	3,070	4,313	3,469
	(0,111)	(0,113)	(0,116)	(0,077)	(0,077)	(0,086)	(0,099)	(0,070)
R square	0.006	0.002	0.000	0.000	0.001	0.000	0.002	0.000
Industrial	-0.075	-0.444	0.030	-0.106	0.470	-0.044	-0.134	0.232
	(0.435)	(0.442)	(0.458)	(0.303)	(0.308)	(0.329)	(0.390)	(0.277)
Intercept	4,466	4,119	3,668	3,562	3,372	3.024	4,232	3,434
	(0,102)	(0,104)	(0,107)	(0,071)	(0,072)	(0,080)	(0,092)	(0,065)
R square	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000

---

[128] Moore, David S.. *The practice of business statistics: using data for decisions*. 2nd ed. New York: W.H. Freeman and Co., 2009. p. 124-125

Electronics	-0.532* (0.285)	-0.541* (0.291)	-0.246 (0.296)	-0.121 (0.179)	-0.340* (0.180)	-0.631*** (0.204)	-0.378 (0.252)	-0.266 (0.163)
Intercept	4,538 (0,107)	4,170 (0,109)	3,705 (0,113)	3,578 (0,077)	3,459 (0,077)	3,131 (0,085)	4,280 (0,096)	3,495 (0,070)
R square	0.002	0.002	0.000	0.000	0.001	0.004	0.001	0.001
N (total number of observations)	1728	1764	1626	2423	2441	2153	1880	2589

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

Table 2 clearly shows that the appearance industry significantly differs from the rest (i.e. the mean of the appearance industry differs from the mean of the remaining four industries). Although (almost) each industry has a significant result over time, they do not show an obvious pattern.

Based on these two results, one may conclude that the appearance industry strongly influences the aggregated set. In addition, it seems to create an aggregated mean that cannot be used in the TP practice (Table 1). In order to confirm the above results, and thus eliminate the appearance industry, a chi-square Median's test was performed. This might show that, for example, the upper quartile shows significant results, meaning that the (individual) industries indeed have significantly different upper quartiles when compared to the aggregated set.

The Chi-square Median's test focused on the median of the total dataset[129] and both the median of the upper and lower subsets (i.e. the upper quartile and lower quartile). This was done for all five industries on a combined level, which means that, for example, the lower quartile of each industry was tested on equality in regards to the aggregated lower quartile. The Chi-square Median's test also showed that the appearance industry led to a large amount of highly significant results (Table 3). A significant result rejects the null hypothesis (2). For example, of the 24 tests that were performed, 7 tests showed

---

[129] The total dataset is the initial dataset of each industry before it was splitted into a lower resp upper subset.

highly significant results ($\propto = 1\%$). This means that only few results could have been used to state that the aggregated IQ range could be applied when establishing an arm's length remuneration.

Table 3: Summary of Median's test aggregated set (all five industries).

Chi-square Median's test - Appearance, Chemicals, Electronics, Industrial & IT Products

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,201	0,306	0,234	0,181	0,368	0,030**	0,061*	0,440
Upper Quartile	0,017**	0,177	0,470	0,001***	0,000***	0,412	0,000***	0,192
Lower Quartile	0,167	0,905	0,002***	0,004***	0,000***	0,000***	0,210	0,000***
Total								

The given values are all p.values

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

After further manual tests and a closer, more detailed view at the both the dataset and timeline, the appearance industry was deemed not to be comparable. The manual tests evaluated the impact that each industries had on the results. A more detailed view at the timeline revealed that the appearance industry has an overall significantly higher RoS. For example, from the year 2009 up to and including 2012, four tests on the lower quartile returned a highly significant result. This means that the lower quartile cannot be used for the determination of an arm's length remuneration on the basis the aggregated average based on these five industries.

More specifically, it appears that the appearance industry has a statistically significant different distribution, because both the upper and lower quartile of the aggregated distribution is unequal relative to the distributions of the individual industries. Since a safe harbour range can never contain this industry, the exploration for the establishment of a safe harbour range will therefore be resumed without the appearance industry.

6.3   Comparing means

After the exclusion of the appearance industry, various regressions were preformed. On the one hand, the aggregated set was tested extensively, and on the other hand each industry was tested on a stand-alone basis to the aggregated set of the remaining industries.

6.3.1   Testing aggregated set

Table 4 shows the results of the four regressions of the aggregated set. This test is a practical approach for the determination of an aggregated arm's length IQ range, since the distribution of the tested industry is included in the aggregated set.

Table 4: Summary of regression results aggregated set (all four industries)

	2007	2008	2009	2010	2011	2012	07-09	10-12
Excluding electronics industry								
IT Products	0.456	0.513*	-0.348	-0.252	0.009	0.306	0.195	-0.038
	(0.285)	(0.288)	(0.288)	(0.222)	(0.217)	(0.246)	(0.253)	(0.202)
Chemicals	-0.179	0.108	0.146	0.097	0.114	0.317	-0.044	0.120
	(0.303)	(0.306)	(0.304)	(0.211)	(0.208)	(0.231)	(0.269)	(0.193)
Industrial	0.386	0.045	0.239	-0.002	0.722**	0.480	0.196	0.437
	(0.435)	(0.437)	(0.441)	(0.312)	(0.309)	(0.334)	(0.387)	(0.286)
Intercept	4,005	3,629	3,459	3,457	3,119	2,500	3,903	3,229
	(0,231)	(0,232)	(0,231)	(0,150)	(0,148)	(0,168)	(0,203)	(0,137)
R square	0.006	0.004	0.004	0.002	0.004	0.003	0.001	0.002
Excluding industrial industry								
IT Products	0.070	0.468	-0.587	-0.250	-0.713**	-0.174	-0.001	-0.475
	(0.405)	(0.408)	(0.413)	(0.318)	(0.315)	(0.340)	(0.363)	(0.291)
Chemicals	-0.565	0.063	-0.093	0.098	-0.608**	-0.163	-0.239	-0.317
	(0.418)	(0.420)	(0.425)	(0.311)	(0.308)	(0.329)	(0.374)	(0.285)
Electronics	-0.386	-0.045	-0.239	0.002	-0.722**	-0.480	-0.196	-0.437
	(0.435)	(0.437)	(0.441)	(0.312)	(0.309)	(0.334)	(0.387)	(0.286)
Intercept	4,391	3,647	3,698	3,456	3,841	2,980	4,098	3,666
	(0,369)	(0,370)	(0,376)	(0,273)	(0,271)	(0,288)	(0,330)	(0,250)
R square	0.006	0.004	0.004	0.002	0.004	0.003	0.001	0.002
Excluding chemicals industry								
IT Products	0.635**	0.405	-0.495*	-0.348	-0.105	-0.011	0.239	-0.158
	(0.258)	(0.261)	(0.262)	(0.220)	(0.216)	(0.240)	(0.232)	(0.201)
Industrial	0.565	-0.063	0.093	-0.098	0.608**	0.163	0.239	0.317
	(0.418)	(0.420)	(0.425)	(0.311)	(0.308)	(0.329)	(0.374)	(0.285)
Electronics	0.179	-0.108	-0.146	-0.097	-0.114	-0.317	0.044	-0.120
	(0.303)	(0.306)	(0.304)	(0.211)	(0.208)	(0.231)	(0.269)	(0.193)
Intercept	3,827	3,737	3,605	3,554	3,233	2,817	3,859	3,349
	(0,196)	(0,198)	(0,198)	(0,148)	(0,146)	(0,159)	(0,176)	(0,135)
R square	0.006	0.004	0.004	0.002	0.004	0.003	0.001	0.002

Excluding IT products industry								
Chemicals	-0.635**	-0.405	0.495*	0.348	0.105	0.011	-0.239	0.158
	(0.258)	(0.261)	(0.262)	(0.220)	(0.216)	(0.240)	(0.232)	(0.201)
Industrial	-0.070	-0.468	0.587	0.250	0.713**	0.174	0.001	0.475
	(0.405)	(0.408)	(0.413)	(0.318)	(0.315)	(0.340)	(0.363)	(0.291)
Electronics	-0.456	-0.513*	0.348	0.252	-0.009	-0.306	-0.195	0.038
	(0.285)	(0.288)	(0.288)	(0.222)	(0.217)	(0.246)	(0.253)	(0.202)
Intercept	4,461	4,143	3,111	3,206	3,128	2,806	4,098	3,191
	(0,167)	(0,170)	(0,172)	(0,163)	(0,159)	(0,180)	(0,151)	(0,148)
R square	0.006	0.004	0.004	0.002	0.004	0.003	0.001	0.002
N (total number of observations)	1150	1152	1070	1409	1415	1238	1242	1507

*** Significant at α = 1%

 ** Significant at α = 5%

 * Significant at α = 10%

For example, the top section (of rows) depict the results of the comparison of the aggregated set, which omitted the electronics industry. It shows that the industrial industry seems to differ from the electronics industry in 2011, based on the total aggregated set. However, this is not a strong result. The year 2011 is the only year that indicates that the mean of the TNMM range derived from the selected companies is *unequal* to the mean of the TNMM range derived from respective comparable companies based on the aggregated dataset (Table 4). However, this result is not highly significant ($\alpha = 1\%$) over time and only holds when the industrial industry is omitted from the regression. For example, aside from the year 2011, only 6 of the remaining 84 results were significant at a significance level between $\alpha = 5 - 10\%$.

Finally, the two periods that were tested all showed insignificant results. Thus, although the annual results show varying results due to, for example, business cycles, the three-year periods are robust.

6.3.2   Testing one industry against the rest

The regression, as show below in Table 5, might show additional information with regards to the different industries that influence the mean significantly. More specifically, these regressions test the underlying and more fundamental question whether the distribution of one industry is equal to the distribution of the rest.

Table 5: Summary of regression results 'one versus the rest'

	2007	2008	2009	2010	2011	2012	07-09	10-12
IT Products	0.491**	0.453**	-0.454**	-0.294	-0.134	0.096	0.189	-0.148
	(0.217)	(0.220)	(0.221)	(0.190)	(0.187)	(0.209)	(0.195)	(0.173)
Intercept	3,970	3,689	3,565	3,500	3,262	2,710	3,909	3,339
	(0,138)	(0,140)	(0,139)	(0,098)	(0,097)	(0,107)	(0,123)	(0,090)
R square	0.004	0.004	0.004	0.002	0.000	0.000	0.000	0.000
Chemicals	-0.488**	-0.191	0.315	0.196	0.011	0.123	-0.177	0.075
	(0.234)	(0.237)	(0.236)	(0.180)	(0.178)	(0.195)	(0.201)	(0.165)
Intercept	4,315	3,928	3,290	3,358	3,222	2,694	4,036	3,275
	(0,127)	(0,129)	(0,130)	(0,102)	(0,101)	(0,113)	(0,114)	(0,093)
R square	0.004	0.001	0.002	0.001	0.000	0.000	0.001	0.001
Industrial	0.242	-0.215	0.342	0.038	0.679**	0.272	0.124	0.405
	(0.386)	(0.387)	(0.393)	(0.287)	(0.285)	(0.304)	(0.345)	(0.263)
Intercept	4,150	3,890	3,357	3,418	3,162	2,707	3,974	3,261
	(0,112)	(0,113)	(0,113)	(0,089)	(0,087)	(0,097)	(0,100)	(0,081)
R square	0.001	0.000	0.000	0.000	0.004	0.001	0.000	0.002
Electronics	-0.209	-0.309	0.095	0.053	-0.155	-0.336*	-0.105	-0.101
	(0.261)	(0.263)	(0.262)	(0.181)	(0.179)	(0.200)	(0.230)	(0.166)
Intercept	4,215	3,938	3,364	3,405	3,275	2,837	4,008	3,331
	(0,121)	(0,122)	(0,123)	(0,102)	(0,100)	(0,110)	(0,108)	(0,093)
R square	0.001	0.001	0.000	0.000	0.000	0.002	0.000	0.000
N (total number of observations)	1150	1152	1070	1409	1415	1238	1242	1507

Table 5 shows that the mean of IT Products industry differs from the mean of the remaining industries. This result only applies for the years 2007, 2008 and 2009 on a year-by-year basis, because the weighted average in that period showed statistically insignificant results. Consequently, the shocks/business cycle in which the industry seems to find itself in, is being filtered out. The means of the remaining three industries do not differ from the mean of the rest. For example, the industrial industry was

compared to the aggregated set of the IT products, chemicals and electronics industry. With only significant results in, again, the year 2011, the null hypothesis (2) is not rejected. This means that the mean of the TNMM range derived from one of the selected companies is equal to the mean of the TNMM range derived from respective comparable companies based on the remaining dataset, in regards to the full-fledged distribution function. Please note that the three-year periods all show insignificant results, which implies that these results are more robust and distributions are equal.

Finally, the even lower r-squared can be explained by the fact that one of the original five industries was excluded, leaving one extra explanatory dummy out of the regression.

6.4   Comparing the interquartile ranges of the distribution

The first set of Chi-square tests focuses on the median of the total dataset and the median of both the upper and lower subsets (i.e. the upper quartile and lower quartile), as discussed in section 4.2.2. This was done for all four industries on a combined level, which means that, for example, the lower quartile of each industry was tested on equality with regards to the aggregated lower quartile.[130] The second set of tests compared one industry in relation to the remaining industries. For example, the Median's test was performed with two groups: the chemicals industry versus the remaining industries. Consequently, this Chi-square Median's test is testing the underlying and more fundamental question whether the IQ range of one industry is equal to the IQ range of the rest.

---

[130] This is i.e. a practical approach, because the results can be used in the determination of an arm's length IQ range.

6.4.1   Testing aggregated set

The results of the first test are presented below:

Tabel 6: Summary of Median's test aggregated set

Chi-square Median's test - Chemicals, Electronics, Industrial & IT Products

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,452	0,872	0,112	0,637	0,336	0,298	0,730	0,270
Upper Quartile	0,050*	0,372	0,468	0,050*	0,022**	0,114	0,713	0,674
Lower Quartile	0,800	0,763	0,001***	0,08*	0,009***	0,001***	0,060*	0,001***
Total								

The given values are all p.values

   *** Significant at α = 1%

   ** Significant at α = 5%

   * Significant at α = 10%

Table 6 clearly shows that all the individual medians combined can be used when determining an at arm's length median, because there was no test statistically significant. This means that there was no industry that significantly deviated from the expected count. In addition, there is no industry that significantly influenced the expected count, which would in turn lead to a Grand Median that differs from the median of the other industries. The null hypothesis (3) is thus accepted. Furthermore, the absolute difference in the number of observations per industry had no significant effect. In addition, both the year-by-year results as the weighted results indicate that the medians are equal. This is in line with expectations as it is statistically insignificant each year.

The above-mentioned is also applicable to the upper quartile of the combined industries since, taken as a whole, no statistically significant results were found (Table 6). However, under a less strict critical level[131] the results indicate that the upper quartile of the combined group deviated from the aggregate upper quartile (i.e. the median of the upper subset). Although the years 2010 and 2011 tend to go towards a more critical value, the results can be used for transfer pricing purposes because the weighted (period) results are highly insignificant.

---

[131] When ∝ is considered to be 5-10%.

Less clear results are shown in regards to the lower quartile, as was expected from the results presented in the trend line (*see Appendix E*). The Median's test shows that four out of the eight tests performed have statically significant results in regard to the aggregated lower quartile (i.e. the median of the lower subset). This implies that one or more industries deviate significantly from the expected number of observations and thus have widely spread lower quartiles in regards to aggregated set. Furthermore, no at arm's length lower quartile could be determined under the assumption that a weighted average filters some, if any, cyclical differences. A shift in the ratio of *total* observations is apparent between the IT products and the other industries, which may indicate that the IT products suffer relatively more from the crisis.

However, a stronger impact was the changing, and thereby deviating, ratio *between* the number of industrial industry observations per year. It fluctuated greatly over time. More specifically, a change in the ratio above and below the Grand Median, can strongly influence the results. The small number of observations and large deviations from the aggregated set indicate that the industrial industry not only does not influence the level of the aggregated lower quartile, but also deviates significantly from the aggregated lower quartile.[132] Combined, this causes an unrepresentative aggregated lower quartile for three of the six years reviewed. Consequently, the weighted average in the second period cannot be applied on the basis of *this* test.

6.4.2   Testing one industry against the rest

The second set of tests compared, for example, the median of one industry to the aggregated median of the remaining industries. This test was performed in order to check if the results hold under more strict conditions. More specifically, the four industries were all tested for the equality of the median, upper and lower quartile in relation to the median, upper and lower quartile of the remaining three industries.

---

[132] I.e. due to the inability to influence/bias the aggregated set

The chemicals, IT products and electronics industry all showed similar insignificant results, with all one highly significant result on lower quartile level (please see Appendix F). All three industries have highly insignificant results on weighted average level, except for the lower quartile of the electronics industry versus the rest in the period 2010-2012.

However, the industrial industry was less comparable in relation to the remaining industries, as shown in the table below:

Tabel 7: Summary of Median's test Industrial versus the rest

Chi-square Median's test - Industrial versus rest

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,443	0,242	0,019**	0,369	0,001***	0,058*	0,784	0,051*
Upper Quartile	0,773	0,334	0,783	0,154	0,008***	0,627	0,262	0,599
Lower Quartile	0,443	0,242	0,019**	0,369	0,001***	0,058*	0,059*	0,001***
Total								

The given values are all p.values

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

Although the results are highly significant for the median, upper and lower quartile in the year 2011, they are in fact similar to the results of the regression (Table 5). As with the regression, the results of the year 2011 indicate that the distribution of the industrial industry was unequal to distribution of the remaining industries. In addition, even Table 4 showed that the industrial industry differed from the mean of the aggregated set (year 2011). Consequently, it can be stated that the year 2011 strongly influenced the industrial industry.

6.5   Comparing the distribution

The Goodness-of-fit Chi-square tests focuses on the total IQ range of each industry in regards to the aggregated IQ range (i.e. the total distribution). This was done for all four industries on a standalone basis, which means that, for example, the ratio of observations of the chemicals industry was tested on equality with regards to the aggregated ratio of observations. The results are presented below:

Tabel 8: Summary of Goodness-of-fit test on a standalone basis

Chi-square Goodness-of-fit test - Chemicals, Electronics, Industrial & IT Products

	2007	2008	2009	2010	2011	2012	07-09	10-12
Chemicals	0,607	0,763	0,227	0,029**	0,130	0,908	0,809	0,688
Electronics	0,175	0,891	0,375	0,162	0,446	0,140	0,298	0,179
Industrial	0,937	0,434	0,339	0,265	0,010**	0,485	0,270	0,020**
IT Products	0,312	0,724	0,285	0,603	0,543	0,608	0,165	0,874
Total								

The given values are all p.values

   *** Significant at α = 1%

  ** Significant at α = 5%

 * Significant at α = 10%

The electronics industry and the IT products industry can use the aggregated IQ range when determining an at arm's length remuneration (Table 8). Both show highly statistical insignificant results over time, meaning that their IQ range does not differ from the aggregated IQ range on a standalone basis. In addition, the weighted tests confirm these results, inter alia, after the mitigation of several varying business cycles.

Table 8 shows that the chemicals industry has a similar IQ range with respect to the aggregated IQ range. All results were insignificant over time. Although the results are significant at a significance level of 5% in the year 2010, the results can be used for transfer pricing purposes. This can be explained by the fact that the year-by-year results are apparently influenced by a business cycle, because the weighted average corrected the relative low results found in 2010 and 2011.

Finally, one may conclude that the industrial industry can also apply the aggregated set when determining an at arm's length IQ range. However, the results are less statistically insignificant, especially when one looks at the weighted average. The industrial industry seems to, as with the Chi-square Medians test, lead to small discrepancies.

Chapter 7   Conclusion

Profit margins for similar routine wholesale distributors do statistically significantly differ between the initial group of industries selected. However, it is possible to identify a group of similar routine wholesale distributors that do not statistically significantly differ. The initial set comprised out of five industries that were deemed to be comparable, where the final set consisted out of four. This means that the hypothesis needs to be slightly altered. It should state that a specific group has similar profit margins. Subsequently, one can state that some industries seem to have a different operating margin, in respect of which a different group needs to be identified. Thus, the results only hold for a fixed set of industries. Consequently, it is possible to construct a safe harbour for certain industries when, inter alia, the same characterization and risk level is assumed.

The five industries that were not comparable due to differences in, inter alia, the level of risks assumed. For example, the risk-profile of the pharmaceuticals industry is expected to be higher, inter alia, due to more specific advice needed for the distribution of goods, such as medical knowledge and legal advice. Furthermore, the food industry is expected to have a lower risk-level due to a different business cycle, if any, due to the size of the market, the absolute simplicity of the business and (high) degree of competition. In addition, the 'hard-core' automotive wholesale industry is limited to a handful of multinational enterprises and was therefore eliminated. The appearance industry was eliminated based on the fact that the distribution was statistically different from the other industries. The selected industries all have in common that the core-function is wholesale and have little added-value, but do have some market knowledge.

In addition, caution should be taken when one establishes the lower quartile for the aggregated sample, because less robust results were found. For example, industries with a relative small number of companies can fall out of the aggregated IQ range due to, for example, their inability to influence the IQ range. For example, small deviations tend to lead to significant results while they appear to carry out similar core functions, as with the wholesale of industrial products. Furthermore, the industrial industry showed deviating results in the year 2011, which appear to be caused by exogenous shocks in

the business cycle, since the weighted average of the period 2010-2012 did not deviate in general.

Consequently, the aggregated IQ range should be based on the weighted average because results are more robust when these average margins are used. For example, it corrects for influences, if any, which are the result of shocks in the business cycle. In addition, this approach is in line with the common TP practice.

Finally, no strong/obvious differences were found between the tests on aggregated level and the tests on the level of remaining industries. The IQ range based on the (biased) aggregated set could thus be applied in practice without prejudice to the fundamental and underlying question whether the profit margins of one industry is equal to the profit margins of the rest.

7.1 Discussion

As aforementioned, the TP practice makes use of databases that consists out of financial data, which can be used for the determination of an at arm's length remuneration. It is used to calculate, for example, the return on assets, return on sales (i.e. the operating margin) of the routine operations. Although these databases and this method is widely used and accepted within western countries, there is an emerging debate with regards to the use and specifically the quality of the databases.[133] Tax authorities indicate that there is a lack of data and the quality of the data provided is often not sufficient. This applies especially for developing countries. Those countries have less comparable companies available, due to fewer organised taxpayers. Furthermore, information can be incomplete or impossible to analyze due to a lack of local knowledge. The OECD, and others, conclude that the current databases provide an inaccurate pool of financial data for the use of comparability analyses.[134] These problems might be tackled partially by expanding the current commercial databases. However, that does not mean that the quality of the financial data is optimal.

---

[133] OECD discussion paper on Transfer Pricing Comparability Data and Developing Countries, March 11, 2014

[134] §9 OECD discussion paper on Transfer Pricing Comparability Data and Developing Countries, March 11, 2014

The establishment of an aggregated IQ range might partly resolve the problem of insufficient data. This can be done based on the results found, by grouping the available data of industries that were found to perform similar wholesale functions (i.e. the chemical, electronics, industrial and IT product industries). That way, tax authorities and MNE's only have to look, on the basis of the functional analysis, at which point in the range the entity belongs. In turn, this might safe time or diminish the administrative burden and compliance costs.

Furthermore, enterprises are required to document all details of the transactions and the facts and circumstances therewith. However, the business communities are not interested in the creation of the margins, but attach value to management accounting, the vision, product strategy, business unit, etc. They do not look at the statutory legal entity for accounting purposes because it is often consolidated. The business will usually solely look at the consolidated operating profit. Based on this profit level, decisions are taken or assessed for their value. They only look at the statutory legal entity for tax purposes because they have to know exactly what is attributable to each separate legal entity. The different set of requirements between countries could lead to a heavier burden for MNE's, and thus result in higher costs to meet the requirements, than a business that operates only in a single jurisdiction. The use of an aggregated IQ range for certain industries might enable tax authorities to reduce the administrative burden of these companies.

In addition the current binary TP models do not take specific advantages of globally integrated enterprises (GIE) into account, when determining the profit levels of routine functions.[135] That is because the remuneration for those functions is the result of a benchmark that is (partly) based on companies that do not have these benefits (i.e. independent companies). Palmisano describes the GIE as:

"(…) a company that fashions its strategy, its management, and its operations in pursuit of a new goal: the integration of production and value delivery worldwide."[136]

---

[135] Please see Appendix G for a more detailed introduction.

[136] Palmisano, Samuel J., "The Globally Integrated Enterprise." Foreign Affairs 85.3 (2006): 127-136.

A controlled transaction is being compared to profit margins accrued by comparable uncontrolled transactions without correcting for benefits that arise like, for example, economies of scale or synergies. The current TP databases are more valued by the tax authorities and companies for the provision of a common platform in order to avoid double taxation. It can therefore be regarded as a safe harbour.[137] However, it is not considered to be a reliable source for an arm's length profit level for the benchmarked functions. A safe harbour IQ range can reduce the documentation requirements and creates opportunities to discuss and focus more on the functions performed, assets employed and risks assumed by the GIE's. This can be done by annually establishing a three-year average IQ range for the predetermined set of industries. The appropriate return on sales remuneration for each specific tested party should subsequently be based on the functional analysis. As a result, the compliance burden and associated costs are reduced.

In addition to the above, a recent study demonstrated that managerial transfer pricing could be used to influence the value creation.[138] One of the value drivers mentioned is the administrative transfer pricing system. More specifically, Steens states that TP can directly influence the value creation by administrative costs. For example, some TP methods are easier to support because there are different (administrative) systems in place or the procedures are more efficient. The defined safe harbour could subsequently be considered as an additional 'tool' to minimise the administrative burden and at the same time create more value for the MNE.

7.2   Recommendations

In order to be able to say more about the robustness of the results and the underlying explanatory variables, further analyses is suggested. These recommendations go hand in hand with the limitations of this study.

---

[137] Vroemen, Eric, and Hans Geluk. "Tax follows business - profit split to the rescue!." International Transfer Pricing Journal 13.3 (2012): 4-9.
[138] Steens, B. "Impact of Transfer Pricing on Value Creation." Dissertation Vrije Universiteit Amsterdam, 14 November 2008

Firstly, a follow-up analysis across multiple pan-regions and between more specific industries can be performed. This should show whether or not significant differences are visible within one industry and over multiple regions.

Secondly, volatility-analyses needs to be performed in order to further test the differences that may exist as result of business cycles. Thus, the financial performance of various industries needs to be analysed in order to be able to increase comparability.

Finally, an analysis on RoS level relative to the inventory-to-turnover ratio might allow for arm's length prices to be made by means of more precise working capital adjustments. This could especially be done when bins could be identified that indicate the effect of a certain ratio. In addition, closer comparability could also be reached after the examination of the RoS level relative to the operating expense-to-turnover ratio.

Bibliography

Works Cited

The State Secretary for Finance. "International Tax Law. Transfer pricing method,
 application of the arm's length principle and the Transfer Pricing Guidelines for
 Multinational Enterprises and Tax Administrations." Rijksoverheid.nl. Version
 IFZ2013-184M,. Directorate-General for Tax and Customs Policy and Legislation,
 International Tax Policy and Legislation Directorate, 14 Nov. 2013.

Abdallah, Wagdy M. Critical concerns in transfer pricing and practice. Westport, Conn.:
 Praeger, 2004.

Ahmad, Nadim. "Measuring Trade in Value Added, and Beyond." Upjohn Institute . OECD,
 <http://www.upjohn.org/MEG/papers/Ahmad_Trade_in_VA.pdf>.

Bikker, JA, and H Hu. " Cyclical patterns in profits, provisioning and lending of banks and
 procyclicality of the new Basel capital requirements." PSL Quarterly Review
 55.221 (2012)

Bilaney, Sunny Kishore . "Supply chain management using alternative manufacturing
 models ." International Transfer Pricing Journal 12 Mar. 2014: 85-95

Bogetoft, Peter, and Henrik Ballebye Olesen. Design of production contracts: lessons
 from theory and agriculture. Copenhagen: Copenhagen Business School Press,
 2004.

"Camp Opening Statement: Hearing on Transfer Pricing Issues." Committee on Ways and
 Means. Committee on Ways and Means, 22 July 2010.
 <http://www.waysandmeans.house.gov/News/DocumentSingle.aspx?Documen
 tID=199859>.

Carlson, William L., and Betty Thorne. Applied statistical methods: for business,
 economics, and the social sciences. Upper Saddle River, N.J.: Prentice Hall, 1997.

Choi, Wooje . "Tax benefits from structured financial transaction denied for lack of
 economic substance." IBFD North America, 4 Apr. 2014.
 <http://online.ibfd.org/kbase/#topic=doc&url=/highlight/data/tns/docs/html
 /tns_2013-04-
 04_us_1.html&q=Tax+benefits+from+structured+financial+transaction+benefit+
 taxes+transactions&WT.z_nav=Navigation&colid=4913>.

"Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999. <http://ec.europa.eu/taxation_customs/resources/documents/primarolo_en.pdf>.

Conover, W. J.. Practical nonparametric statistics. 2d ed. New York: Wiley, 1980.

Corlaciu , Alexandra. "Business models for tax and transfer pricing purposes ." The Journal of the Faculty of Economics, University of Oradea, 1 July 2013: 1182-1189.

Dawid, Roman, Peter Meenan, and Jörg Hülshorst. "Is Europe One Market? A Transfer Pricing Economic Analysis of Pan- European Comparables Sets." European Commision DOC: JTPF/007/BACK/2004/EN (2004): 1-54.

Eden, Lorraine. Taxing multinationals: transfer pricing and corporate income taxation in North America. Toronto: University of Toronto Press, 1998.

Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011.

Ellison, S. L. R., and Trevor J. Farrant. Practical statistics for the analytical scientist a bench guide.. 2nd ed. Cambridge, UK: RSC Publishing, 2009.

Emerson, Michael. *The Economics of 1992: an assessment of the potential economic effects of completing the internal market of the European Community*. Luxembourg: Office of Official Publications of the European Communities, 1988.

Governors Staff, Federal Reserve System Board . The Code of Federal Regulations of the United States of America: Title 12 - Banks and Banking. Washington : Office of the Federal Register National Archives and Records Administration, 2004.

Grant, Robert M.. "The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation." California management review 33.3 (1991): 114-135.

Hamaekers, H. "Arm's length - how long?." International Transfer Pricing Journal 8.2 (2001): IBFD Tax Research Platform.

Haskel, Jonathan, and Holger C. Wolf. "From Big Macs to iMacs : what do international price comparisons tell us?." World Economics 1.2 (2000): 167-178.

Heijenrath, M.J.M., and P.W. van Meeuwen. "Introductie tot de praktische aspecten van de OECD-transfer-pricingmethodieken." Tijdschrift voor Fiscaal Ondernemingsrecht 1 Jan. 2006

Hill, Charles W. L., and Gareth R. Jones. *Strategic management theory: an integrated*

approach. 9th ed. Mason, OH: South-Western/Cengage Learning, 2010.

Kleimeier, Stefanie, and Harald Sander. *European financial market integration: evidence on the emergence of a single eurozone retail banking market*. Maastricht: METEOR, Maastricht research school of Economics of TEchnology and ORganizations, 2002.

Kotz, David M. "The Financial and Economic Crisis of 2008: A Systemic Crisis of Neoliberal Capitalism." *Review of Radical Political Economics* 41.3 (2009): 305-17

Kravis, Irving B., and Robert E. Lipsey. National price levels and the prices of tradables and nontradables. Cambridge, MA (1050 Massachusetts Ave., Cambridge 02138): National Bureau of Economic Research, 1988. Print.

Kravis, Irving B., and Robert E. Lipsey. "Toward an Explanation of National Price Levels." Princeton Studies in International Finance, Princeton University no. 52 (1983)

Kroppen, Heinz-Klaus. "Is Europe One Market? A Transfer Pricing Economic Analysis of Pan-European Comparables Sets." EU Joint Transfer Pricing Forum. Version JTPF/007/BACK/2004/EN. European Commission, 24 Feb. 2004.

Lima, Marcos, and Marcelo Resende. "Profit Margins and Business Cycle in the Brazilian Industry: a Panel Data Study." Applied Economics 36.9 (2004): 923-930.

Loots, Jozua Johannes. A comparability adjustment transfer pricing model. Potchefstroom: Potchefstroom Business School University of North West, 2006.

Machin, Stephen, and John Van Reenen. "Profit Margins and the Business Cycle: Evidence from UK Manufacturing Firms." The Journal of Industrial Economics 41.1 (1993): 29-50.

Meric, Ilhan, Stephanie M. Weidman, Carol N Welsh, and Gulser Meric. "A Comparison of the Financial Characteristics of U.S., E.U., and Japanese Manufacturing Firms." American Business Review 20.2 (2002): 119.

Meric, Ilhan, Herbert E. Gishlick, Leonore S. Taga, and Gulser Meric. "A Comparison of the Financial Characteristics of U.S., Canadian, and Mexican Manufacturing Firms." International Journal of Economics and Finance 5.7 (2013): 1-8.

Monatsbericht. "Zur Unternehmensrentabilitat im internationalen Vergleich." Bundesbank 15 Oct. 1997: 33-44.

Moore, David S.. *The practice of business statistics: using data for decisions*. 2nd ed. New York: W.H. Freeman and Co., 2009.

Müller, Johann. "Ruling Practice." The Netherlands in international tax planning. 2nd ed. Amsterdam, Netherlands: IBFD, 2007. 359-367.

"Multi-Country Analysis of Existing Transfer Pricing Simplification Measures." OECD 6 June (2012): 155.

Murphy, Richard. "Amazon, Google and Starbucks are struggling to defend their tax avoidance." theguardian.com. Guardian News and Media, 13 Nov. 2012.

Myatt, Glenn J.. Making sense of data: a practical guide to exploratory data analysis and data mining. Hoboken, N.J.: Wiley-Interscience, 2007.

NACE Rev. 2: statistical classification of economic activities in the European Community. Luxembourg: Office for Official Publications of the European Communities, 2008.

Nielsen, AC . "New ACNielsen survey points to price convergence in Europe since the introduction of the Euro." .
<http://nl.nielsen.com/site/documents/breakingnews_europe.pdf>.

Norris, Floyd. "Apple's Move Keeps Profit Out of Reach of Taxes." The New York Times. The New York Times, 2 May 2013.
<http://www.nytimes.com/2013/05/03/business/how-apple-and-other-corporations-move-profit-to-avoid-taxes.html?pagewanted=all&_r=0>.

Owens, Jeffrey. "The Taxation of Multinational Enterprises: An Elusive Balance." Bulletin for International Taxation 67.8 (2013): IBFD Tax Research Platform.

Palmisano, Samuel J. . "The Globally Integrated Enterprise." foreign affairs 85.3 (2006): 127-136.

Peerbooms, J., and M. van der Breggen. "Transfer pricing: verplichte kost voor iedere fiscalist!." Forfaitair 173 (2007)

"Report on Compensating Adjustments." EU JOINT TRANSFER PRICING FORUM JTPF/009/FINAL/2013/EN (2013): 5.

"Report on the Activities of the EU Joint Transfer Pricing Forum in the Field of Documentation Requirements"; EU Joint Transfer Pricing Forum JTPF/020/REV4/2004/EN (2005)

"Revised Section E on Safe Harbours in Chapter IV of the Transfer Pricing Guidelines" Committee on Fiscal Affairs and the OECD Council C(95)126/FINAL(2013)

Rogoff, Kenneth. "The Purchasing Power Parity Puzzle." Journal of Economic Literature 34.2 (1996): 647-668. Jstor.

Rubin, Allen. Statistics for evidence-based practice and evaluation. 3rd ed., student ed.

Belmont, CA: Brooks/Cole, Cengage Learning, 2013.

Schellekens, Marnix . "New substance requirements in effect." ibfd.org. IBFD Research, 5
Feb. 2014.
<http://online.ibfd.org/kbase/#topic=doc&url=/highlight/data/tns/docs/html
/tns_2014-02-
05_nl_1.html&q=New+substance+requirements+in+effect+effects+requirement
+substances&WT.z_nav=Navigation&colid=4913>.

Schumpeter, Joseph Alois. "The Contours of Economic Evolution ." Business cycles : a
theoretical, historical, and statistical analysis of the capitalist process. 1. ed. New
York, NY [u.a.: McGraw-Hill, 1939. 179.

Silberztein, Caroline . "Transfer pricing aspects of intangibles: the OECD project."
Transfer Pricing International Journal 12.8 (2011): 1-6.

Smith, Gordon V., and Russell L. Parr. Intellectual property: valuation, infringement, and
joint venture strategies. 4th ed. New York: Wiley, 2005.

Sodhi, ManMohan S., and Christopher S. Tang. A long view of research and practice in
operations research and management science the past and the future. New York:
Springer, 2010.

Steens, B. "Impact of Transfer Pricing on Value Creation." Dissertation Vrije Universiteit
Amsterdam, 14 November 2008

Vollebregt, H.A. "Verrekenprijzen: fiscus en functionele analyse." Weekblad voor Fiscaal
Recht 1 Jan. 2004: n. pag. Kluwer Portal . Web. 13 May 2014.

Vollebregt, H.A. "Transfer pricing: Aanbevelingen ter beperking van de administratieve
lasten voor de verrekenprijsadministratie van art. 8b Wet VPB 1969." Weekblad
Fiscaal Recht 16 (2005): Kluwer Portal.

Vroemen, Eric, and Hans Geluk. "Tax follows business - profit split to the rescue!."
International Transfer Pricing Journal 13.3 (2012): 4-9.

Wernerfelt, Birger. "A Resource-based View Of The Firm." Strategic Management Journal
5.2 (1984): 171-180.

<div align="center">Court Cases</div>

Supreme Court May 31, 1978, nr. 18 230, BNB 1978/252

Supreme Court 28 June 2002, nr. 36 446, BNB 2002/343

Amsterdam court of Appeal May 23, 2000, nr 99/1625, V-N 2000/39.10

Other

MvT, *Kamerstukken II*, 2001-2002, 28 034, nr. 3

Nota n.a.v. het Verslag, *Kamerstukken II*, 2001-2002, 28 034, nr. 5, p. 33

Single European Act, June 1985

European monetary system (regulation 3181/78), December 1978

Maastricht Treaty, February 1992

Appendix A – TP General

This appendix provides a more detailed discussion about the discussed topics as well as other relevant topics from a TP perspective.

Selection of TP methods

There are several functions to consider when one looks at the value chain of a MNE. Apple Inc. for example has divisions that perform R&D, production, design, assemblage, services, public relations, personnel and marketing.[139] It will be apparent that a routine function, that does not strongly depends on the performance of the entire business, should receive a relative low remuneration. And vice versa, a strongly complex core-function receives a relative high remuneration. The marketing function of Apple's products will be of greater importance than for instance the assembly function. It is more likely that the profitability, positive or negative, depends more on the marketing function because it attributes to Apple's business worldwide. It adds more value to the value drivers. In addition, it is, for example, in the production division relevant who purchases the raw materials. The remuneration for a subsidiary will be valued lower if the parent is responsible for the procurement of raw materials.[140]

The key value drivers of MNE's can be determined in various ways. The importance of a specific function performed could, for example, be examined by means of interviews with personnel or by job descriptions. However, the value of the functions performed may also be derived from the management control-system. For example, the incentive scheme of an enterprise can say a lot about the value of specific functions performed.

Consequently, the MNE needs to register all the functions performed because the documentation requirements set by the OECD Guidelines[141], among other things, depends on the complexity of the enterprise. Hence, the difference between a central and a decentralized policy is of great importance for the analysis.[142] The degree of

---

[139] In §1.43 OECD Guidelines, some possible functions are suggested
[140] Functions/activities are derived from the TP decree
[141] §5.6 OECD Guidelines
[142] §7.4 OECD Guidelines

autonomy of an entity will therefore be decisive for the starting point of the functional analysis.

As previously indicated in figure 1, there are several functions that need to be considered.

The shareholder activities consist primarily of the management of subsidiaries. The related costs are therefore (partly) known as shareholder costs. An enterprise tries to optimize their financial ratios, minimize their costs of capital and manage the financial risks by means of the finance function. In addition, the service function provides support services in a specialised and harmonised manner to, for example, the whole group. Typically, this is not the core function of the group and is often classified as a shared service centre. This way support activities can thus be centralised and thereby increase efficiency. Research and development (of intangibles) is usually considered to be a core function of an enterprise. Enterprises could distinguish themselves with the developed intangibles, allowing them to set, for instance, higher prices. These functions will not be discussed any further because they are not in scope of this thesis.

The manufacturing function relates to the entities inter alia that manufacture products. It is important to characterise the contribution of each entity involved in the manufacturing function of related parties in order to compare it with the uncontrolled/independent entities. The following characterisations can be discerned and are all based on the nature of the activities performed by the manufacturer:[143]

- Toll manufacturer
- Contract manufacturer
- Licensed manufacturer
- Full-fledged manufacturer

---

[143] Corlaciu , Alexandra. "Business models for tax and transfer pricing purposes ." The Journal of the Faculty of Economics, University of Oradea 1 July 2013: 1185-1187

The sales function is often considered to be the last function in the value chain of an enterprise. The sales force is in contact with the final consumer.[144] This function varies from solely logistical/distribution activities, to activities related to marketing activities. The following characterisations can be discerned and are all based on the nature of the activities performed by the sales entity/distributor:[145]

- Agent
- Commissionaire
- Limited risk distributor
- Full-fledged distributor

Risks

The TP-method is generally determined on the basis of the functions performed. However, in order to achieve a proper analysis, adjustments need to be made when one signals significant differences in the risks assumed. For instance, a higher remuneration will be appropriate if more (material) risks are allocated to an enterprise. The open market will take these risks into account.[146] Thus, the rate of return that is earned by an enterprise partly stems from the level of risks they bear.

The purported allocation of risks should be in line with the economic substance of the transaction.[147] It is, for instance, not possible to determine an at arm's length price when transactions lack any economical substance or assumptions, simply because uncontrolled transactional cannot be found as unrelated parties would not agree on those terms and conditions. It makes sense that one looks at the materiality of transactions, because that party actually controls and bears the financial consequences of the risks. Thus, even though the contractual agreement should be the starting point, one must look at the economic substance as well. This is also endorsed by the economic

---

[144] Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.43
[145] The entity performing the sales function does not have to be (solely) a sales entity. Another core-function may be applicable.
[146] §1.45 OECD Guidelines
[147] §1.48 OECD Guidelines

theory of profit maximisation of Bogetoft and Olesen.[148] They state, in short, that any rational party strives to maximize their profit by entering into contracts. In turn, this leads to a similar thought when it comes to the division of risks and related incentives, because with an increasing (overall) profit level, there is more to divide.[149] Therewith is it in the best interest of both parties to divide the risks in such a way, that they minimise the risks and consequently the costs. Therefore risks would generally be contractually allocated to the party with the lowest risk premium.

There are several risks that can be considered. For example, one recognizes, market risks, inventory risks, credit risks, product liability risks, foreign exchange risks and so on.[150] The risks should be allocated to the party that has significant control over the risks. However, the OECD Guidelines recognise that not all risks could be allocated to the party that has significant control. Enterprises usually do not have (significant) control over a general business cycle and the risk that accompanies it.[151]

Market risk consists for example of losses resulting from selling at an precarious marketplace, the degree of competition or the general economic conditions. Severe competition on the marketplace will lead to a higher risk premium compared to a market where they are not troubled by substitutes. In addition, the market risk-exposure could differ between industries. General market risks are reflected in the change of market value as result of movements in the market prices.[152] An aerospace industry may perform extremely different over the business cycle. They will suffer for instance more of the current economic downturn then the food industry. Simply because people have to eat, whether the economy finds itself in an expansion or recession. Hence, a safeguard of market risks by a principal to a subsidiary may be more valuable in the aerospace industry. It thus depends on the kind of marketplace and the specific demand in that industry.

---

[148] Bogetoft, Peter, and Henrik Ballebye Olesen. Design of production contracts: lessons from theory and agriculture. Copenhagen: Copenhagen Business School Press, 2004

[149] Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.48

[150] §1.46 and §5.24 OECD Guidelines

[151] §1.49 OECD Guidelines

[152] Governors Staff, Federal Reserve System Board. The Code of Federal Regulations of the United States of America: Title 12 - Banks and Banking. Washington : Office of the Federal Register National Archives and Records Administration, 2004, p. 274

Assets

The OECD Guidelines indicate that the assets used and risks assumed are part of the functions performed.[153] However, they do not provide a rule-of-thumb, nor further guidelines. When the assets of the affiliated enterprise are identified, one should value the assets accordingly. The value should be based on the risks that accompany the assets. For example, a sales entity that does not bear the inventory risk should receive a relative low remuneration.

Characterisation manufacturing function[154]

MNE's are always exploring multiple ways to increase the quality of their products, lower their total costs and limit the time needed to respond between the demand and supply in order to maintain or even expand their market share or market value.[155] This led, in turn, to the separation of manufacturing functions across different jurisdictions among different members of the MNE (*see Appendix H)*. This resulted in a cross border fragmentation of production. Furthermore, a manufacturer will usually receive a remuneration on the basis of a mark-up on their total costs.
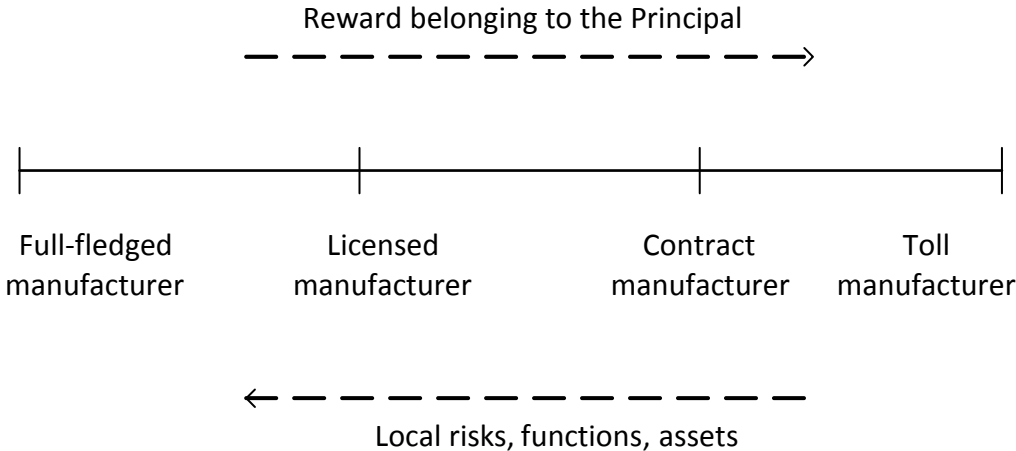
---

[153] See for example: §1.36, 2.21, 5.23 OECD Guidelines

[154] The characterisation is, inter alia, based on: Ahmad, Nadim. "Measuring Trade in Value Added, and Beyond." Upjohn Institute. OECD **and** Corlaciu, Alexandra. "Business models for tax and transfer pricing purposes." The Journal of the Faculty of Economics, University of Oradea 1 July 2013: 1182-1189 **and** Egdom, J.T. van. Verrekenprijzen; de verdeling van de winst van een multinational. Deventer: Kluwer, 2011. p.42-43

[155] Ahmad, Nadim. "Measuring Trade in Value Added, and Beyond." Upjohn Institute. OECD

The figure below shows the preceding clearly:[156]

Figure 8: Allocation of profits

Reward belonging to the Principal

— — — — — — — — — — — →

Full-fledged
manufacturer

Licensed
manufacturer

Contract
manufacturer

Toll
manufacturer

← — — — — — — — — — — —

Local risks, functions, assets

Toll manufacturer

A toll manufacturers main function is simply the production of (finished) goods for, and under supervision and behalf of, a principal and can therefore be regarded as a service provider that does not bear selling risk. The manufacturer usually produces a predetermined volume of business and is thus protected from those market risks. The raw or sub-assembled materials, necessary for the production, are fully owned by the principal. Although the toll manufacturer is not involved in the procurement of, for example, the raw materials, they do receive the physical flow of goods 'directly' of the supplier.[157] It is characteristic that it does never take title on anything produced or to be produced. Therefore no inventory risk is born by the toll manufacturer. No intangible property is owned considering it does never take title nor adds any value on its own behalf. At the same time, no responsibility is taken by the toll manufacturer for the production scheduling and quality control. They do however provide feedback on the scheduling plans and confirm invoices for the materials received. The toll manufacturer typically owns the fixed assets. This includes a production plant, machinery as well as

---

[156] Corlaciu, Alexandra. "Business models for tax and transfer pricing purposes." The Journal of the Faculty of Economics, University of Oradea 1 July 2013: 1182-1189.
[157] Bilaney, Sunny Kishore . "Supply chain management using alternative manufacturing models ." International Transfer Pricing Journal 12 Mar. 2014: p.87

equipment. The remuneration is generally a mark-up on total costs. The cost base consist solely out of operating expenses because raw materials are not being kept.

Contract manufacturer

A contract manufacturer is, in essence, a service provider and thus follows the specifications given by the principal. They procure and own the raw materials after which they are engaged in the production of the goods and the coherent quality control. This means that the contract manufacturer bears the inventory risk that is associated with the raw materials. However, since it produces a predetermined quantity, they do not bear the risk of inventory associated with the finished goods. A contract manufacturer sells all the finished goods directly to the principal, provided that it meets the standards specified by the principal. Consequently, the contract manufacturer is not responsible for any type of warranty.

The contract manufacturer does typically not own any significant intangible property. All the intangibles that are related to the product are supplied and owned by the principal. This includes, for example, patents, specific formulas and know-how. The contract manufacturer may however develop its own intangibles during the manufacturing process leading to an increase in efficiency. It owns fixed assets as well as working capital.

Licensed manufacturer

A manufacturer under a licensing model performs manufacturing as well as assembly functions. It not only procures the raw materials, but it also owns and bears the risk of the raw materials as well as the work-in progress, inventory and the fixed assets, including the warranty risk etc. The market risk and production risks are also borne by the licensed manufacturer since they are responsible for both the raw materials and the finished goods that remain in their inventory. They are also responsible for scheduling, the product quality control and logistics. However, a licensed manufacturer is generally not concerned with the research and development of the product and does not own any intangible property. Other parties own the intangibles associated with the production and manufacturing process. This can be both dependent as independent enterprises whom licence the intangibles to the manufacturer. They can utilize the intangibles by

paying an at arm's length royalty.

Full-fledged manufacturer

This could be considered to be a conventional model. It not only performs the same functions as a licensed manufacturer, but also owns and develops the intangible property. The developed intangibles may include patents, industrial know-how as well as secret formulas and designs.[158] They create more intangible property by research and development programs in order to create more value or decrease production costs. They engage for example in process development or engineering.

Considering the fact that the full-fledged manufacturer owns and develops the intangibles and bears the majority of the risks, if not all, a remuneration based on the residual profit (or loss) should be deemed at arm's length, provided that the distributor bears very limited risks. The distributor would be compensated on an arm's length basis, after which the 'entrepreneurial' function is compensated by means of the remaining manufacturing profit or loss.

Contractual terms

The contractual terms of a transaction state in what manner the risks, responsibilities and benefits are to be divided between parties.[159] The economic substance is addressed as well. One must look for example at the conduct of an enterprise in absence of a written contract. The contractual relationship shall be derived on the basis of the facts and circumstances in combination with the manner an unrelated party would act. The economic substance becomes even more evident in the situation where a contract is in place between affiliated enterprises. Even though a contract is in place, they must also behave in accordance with those contractual terms. The actual/factual division of risks, responsibilities and benefits will be taken as a starting point for the comparability analysis when they do not behave in accordance.

---

[158] Bilaney, Sunny Kishore . "Supply chain management using alternative manufacturing models ." International Transfer Pricing Journal 12 Mar. 2014: p.86
[159] §1.52 OECD Guidelines

Business strategies

Finally, the OECD Guidelines recognise that business strategies could influence the comparability of controlled and uncontrolled transactions and enterprises.[160] Enterprises may have different motives for their daily actions and thereby differ from the found uncontrolled transactions and enterprises. One needs to adjust for differences like, for example, the level of innovation, the degree of diversification and the absence or presence of product development, in order to achieve an at arm's length price. Common and well-known business strategies are market penetration and actions that expand market share. The strategies all have in common that they want to cause an impact by means of a temporary increase in the cost or a reduction of the market price. The comparatively lower profits do not necessarily result in an erroneous arm's length price. The underlying idea is that the reduced profits in the present can be offset by higher profits in the future. The tax authorities may demand a more detailed explanation concerning the reduced profits, because higher profits in the future do not always materialize and tax authorities are often constraint by national law to re-examine earlier tax years.[161] However, they may not ignore a business strategy for TP purposes when the expected results are substantiated and credible.[162] When the parties act in accordance to the purported business strategy, there conduct will generally be deemed at arm's length. The purported business strategy must comply with economic reality. For example, a reduced sales price would generally be accompanied by an increase in marketing expenses. Whereby the costs need to be divided at arm's length. An unaffiliated party will bear no costs if the nature of the purported business strategy does not support it, because when risks are incurred remuneration is appropriate.[163]

---

[160] §1.59 OECD Guidelines
[161] §1.61 OECD Guidelines
[162] §1.63 OECD Guidelines
[163] §1.62 OECD Guidelines

TP Methods

Resale price method

The resale price method is based upon the price at which a product is being (re)sold to an independent enterprise and analyses the gross margin (i.e. resale price margin, RPM) that is to be derived thereof. The price that is being charged to an independent enterprise is by definition at arm's length. The sales price will be reduced by a RPM in order to cover the expenses of the vendor, including an appropriate profit margin. The balance will serve as an arm's length intercompany price for the original transfer.[164]

This method will primarily serve industries that are located at the end of the value chain because the reseller typically adds less value. There are two different types of RPM's that can be distinguished. The internal RPM is based on the RPM that is achieved by the sale and resale of products from and to third parties. The external RPM comprises of the margin that is achieved by independent enterprises in comparable transactions. In the determination of the remuneration, products sold do not have to be similar, provided that the function performed is similar.[165] However, differences in business operations may be important because differences in functions and activities can influence the (external) RPM. For this reason, the resale price method is difficult to be applied to resellers that contribute to the development and preservation of the intangible property.

Cost plus method

This method is based on the costs that are borne by the supplier of goods or services between associated enterprises. An arm's length price is achieved by an appropriate mark-up on total costs. The mark-up is a percentage of the costs incurred and is based on the functions performed, assets used and risks assumed.[166] It is important that, for the sake of the comparability analysis, that attention is devoted to the differences in the nature and extent of the costs related to the functions performed. Thus, the cost can reflect, for example, a difference in function or may be evoked by additional functions

---

[164] §2.21 OECD Guidelines et seq.
[165] The resale price method focuses' mainly on the functions performed. Sales entities that sell apples or pears exert the same function allowing the remuneration of the sales function for both products to be comparable
[166] §2.39 OECD Guidelines et seq.

for which additional remuneration is appropriate. In addition, differences in the cost base may arise from efficiency or inefficiency differences in operations performed by the parties that are being compared. Thus (production) capacity adjustments need to be made if occurring. One can distinguish three different types of costs and expenses:

- Direct costs of production such as the cost of raw materials
- Indirect costs of production, these costs are closely related to the production process but could be allocated to more than one division, such as costs of electricity
- (General) operating expenses such as management and administrative costs.

Important to note is that the operating expenses are not to be included in the cost base as well as financing costs and disbursement costs. It is also possible that not all direct costs are taken into account as a result of the contractual terms and conditions or the factual practice, as is the case with wage producers[167].

The cost plus method is based on budgeted costs, since sales prices and costs are usually determined in advance. This is also known as 'price-setting'. Differences in actual costs are often due to, for example, inefficiencies or the economic circumstances. 'Price evaluating' is less common, because when one retrospectively determines the actual costs, inefficiencies are not taken into account. NB: the price/the remuneration is determined afterwards on the basis of the actual costs. There is no incentive to produce more efficiently.

---

[167] These producers work for a fixed fee per hour (i.e. a wage)

Profit Split Method

This two-sided approach method divides the profit of a MNE in order to establish an arm's length transfer price. The profit split method ensures that (operational) profit is divided in such a way that would be expected of independent enterprises that find themselves in a joint-venture relationship.[168] It is a preferred method when one deals with highly integrated operations.[169] A one-sided approach would not be appropriate in such a relationship. Where the above-mentioned methods are all based on the entity that performs the least complex function(s), it can also occur that both enterprises add a unique and valuable contribution to the transaction. In such case, this two-sided approach method might provide an appropriate transfer price. This means that the two parties involved in the transaction are assessed within the functional analysis. Affiliated parties may also be intertwined in such a way and therefore seek a method that takes the relative value of their contributed intangibles into account.[170]

When applying this method, an enterprise needs to determine the total profit level. The profits should subsequently be allocated to its affiliates. Just as with the TNMM, the allocation of the profits should be based on a transfer price derived from uncontrolled transactions. In the allocation of the total profit resulting from economies of scale or other benefits resulting from more efficiency, a two-sided analysis appears to result in a better outcome. A one-way analysis like the TNMM, where the focus is more oriented towards one of the parties involved, may cause that the economies of scale obtained actually end up with just one of the parties.

Depending on the specific facts and circumstances, parties must search for an appropriate allocation key that reflects the relative added value of the contribution by the parties involved. If similar enterprises in an uncontrolled transaction allocate the profit based on a combination of weighted salaries and number of employees, this will be the base for the profit allocation of the controlled transaction. A great deal of prudence should be taken into account when selecting the apportionment.

---

[168] In some specific transactions the use of the gross margin is allowed as well; §2.131 OECD Guidelines
[169] Owens, Jeffrey. "The Taxation of Multinational Enterprises: An Elusive Balance." Bulletin for International Taxation 67.8 (2013), IBFD Tax Research Platform.
[170] §2.109 OECD Guidelines et seq.

In complex transactions, the moment between the costs incurred and the time the revenues are obtained are often widely separated. For example, a key that is based on the costs incurred by both parties does not have to say anything about the value of the contribution in the partnership. In the development of unique intangibles, costs incurred hardly say anything about the added value. An arm's length transfer price will be obtained when the above is taken into account.

However, the external market data used ensures that the transfer price is less related to the transactions as such. In addition, it is sometimes difficult to obtain sufficient information of the foreign related party participating in the transaction. And even if the profit is divided on the basis of an apportionment, problems may arise when one implements the apportionment due to different tax and accounting systems.

The OECD Guidelines distinguish two main types of profit split methods, namely the contribution analysis and residual analysis. Taking the conduct of independent enterprises into consideration, the contribution analysis divides the combined profits between the associated enterprises on the basis of the relative value of the functions performed. External market data that demonstrates how third parties would have divided these combined profits could be used to support the division. The residual profit split analysis is a two-stage method. In the first step, each affiliate receives an appropriate remuneration for its routine functions. This remuneration will usually be determined by means of the application of one of all the above-mentioned methods. The enterprises that enter into the controlled transaction will be deemed 'tested party'. This means that intangibles are not taken into account. The residual profit or loss will be allocated to the more complex functions on the basis of the particular facts and circumstances.[171]

---

[171] §2.21 OECD Guidelines

Primarolo report and consequences

The findings of the Primarolo report where sent to the ECOFIN Council on 29 November 1999. This report is the result of research done by the Code of Conduct Group (CoCG), section business taxation, and comprises of harmful features indentified in European jurisdictions. It looks at the particular tax systems in various States and its impact on international cross-border level. The CoCG examined provision that relate to, for example, business taxation and withholding taxes on both cross-border interest and royalty payments between entities.[172] The scope that is set up in the report is agreed upon by the Member States and the Council. It comprises of criteria associated with tax measures/adjustments that countries need to make. The criteria are subdivided in groups from A till F, where A covers (broad) tax measures that need to be made on both the tax system and administrative practice and F discusses more specific tax measurements.

If the criteria are met and countries thus need to adjust their policies, they can defend themselves when they do not agree with the outcome.[173] It was concluded that the Netherlands should revise their tax system and administrative practice on thirteen points, ranging from more critical changes in group A till less critical changes in group Z. All the measurements that need to be taken across border where listed and are all regarded as harmful because they all indicate a significant lower (effective) tax rate.

Given that the research question, inter alia, originates from the old tax ruling practice with regard to cost-plus and resale minus rulings, these will be discussed in more detail. Subsequently, other less relevant rulings will be discussed briefly, in order to set a scene.

Cost-plus and resale minus calculations

Tax liabilities are especially present within companies that carry out cross-border transactions with affiliated companies. It should be clear that those taxpayers have a

---

[172] "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999. §1
[173] "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999. §17

valid interest in knowing in advance whether the intercompany prices charged are considered to be at arm's length.

The DTA differentiate, inter alia, the cost plus and resale minus the ruling.[174] The cost-plus ruling, which relates to the allocation and determination of profits, is applicable in those situations where the inter-company activities are of preparatory and supporting nature. One of the requirements is that a comparable uncontrolled price cannot be found for the services rendered. The remuneration for simple activities cannot be lower than five percent, whereas the functional analysis may lead to a higher remuneration. Typically, after it is compared to third parties, the cost-plus remuneration ranges between five and fifteen percent. However, no ruling will be issued if the activity under review is deemed to be the core-business. The taxpayer will ultimately be deemed to deal at arm's length if they have a profit that comprises of a fixed percentage of all operating expenses. Disbursements are the only type of costs that are not included in the cost base. The ruling is usually concluded for a period of four years.

In addition, specific cost plus rulings are possible when a company from the USA carries on sales activities in the Netherlands.[175] This involves inter-company activities of preparatory and supporting nature as well. Local US companies are allowed to allocate a part of the income earned to these sales entities, also known as Foreign Sales Company (FSC). This will lead to an exemption in the US if conditions are met. The ruling determines the minimum acceptable inter-company price level. Consequently, a mark-up based on the total operating expenses and can vary between five and fifteen percent. The taxpayer will ultimately be deemed to deal at arm's length if they have a profit that comprises of the confirmed fixed percentage of all operating expenses, or more. If it appears that the commercial profit calculated under US rules and regulations exceeds the total remuneration based on the cost-plus ruling, then that amount is considered to be informal capital for the FSC. Informal capital is i.e. equity and is therefore not taxable under the CITA.

---

[174] The cost plus resp. resale minus ruling can be found under A008 resp. A009 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999.
[175] The cost plus US ruling can be found under E003 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999

Contrary to the cost-plus ruling, the resale minus ruling is applicable in those situations where the inter-company activities are of preparatory and supporting nature with a selling function. Activities that qualify for such determination of profit are not the selling activities as such, but the underlying activities like, for instance, marketing activities. One of the requirements is that a comparable uncontrolled price cannot be found for the services rendered. Typically, after the entity is compared to third parties, the resale minus remuneration ranges between one and three percent. However, again no ruling will be issued if the activity under review is deemed to be the core-business. The taxpayer will ultimately be deemed to deal at arm's length if the price charged for inter-company services is (at least) based on a percentage of sales revenue. The ruling is usually concluded for a period of four years.

Miscellaneous

A royalty-ruling is possible when a Dutch entity functions as a licensee as well as a sub-licensor for at least one affiliated entity and can thus be considered to act as an intermediate license-holding.[176] A certain spread will be applied to calculate the profits over the royalties that are received.

Similar profit determination is applied for intra-group finance activities. A spread is thus calculated for an affiliated finance entity that re-lends funds.[177] In addition, a risk-reserve need to be formed by these financing entities.[178] Firstly, this will limit finance entities to set up artificial loan schemes with affiliated entities that allows them to deduct (more) interest and thus counters artificial base erosion. Secondly, Dutch enterprises that operate international are allowed to protect themselves against substantial risks by a similar risk-reserve. Subsequently, certainty in advance will also be provided to those finance entities that carry out their core-business in a (foreign)

---

[176] The royalty-ruling can be found under A015 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999
[177] The intra-group finance ruling can be found under A010 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999
[178] The risk-reserve can be found under B004 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999

permanent establishment and want certainty about their profit allocation.[179]

Furthermore, the Netherlands has a participation exemption for holding entities located in the Netherlands and thus subject to the CITA. This provision exempts certain domestic or foreign profits and losses from the corporate income tax base. On the one hand, dividends and capital gains received from subsidiaries are exempt from the tax base, and on the other hand capital losses and costs of acquisition are not deductible. The participation exemption ruling does not only confirm the applicability of the exemption, but also states what remuneration is deemed to be at arm's length for managing subsidiaries.[180]

In addition, it is possible to deduct certain expenses from the tax base, even though these are not actually incurred.[181] The Dutch tax system does not tax formal contributions to the equity of an enterprise. The same applies for withdrawals. As result of the aforementioned court decision, interest free loans need to be corrected to an arm's length interest level. The unpaid correction arises from shareholder motives and need thus to be considered (in)formal capital.

Finally, enterprises are allowed to deduct a percentage of the amount invested in fixed assets, under the provision that the purchased fixes assets are not sold in any way within five years.[182] For example, specific energy investment allowances are applicable for legal and natural persons.[183]

---

[179] Certainty with regards to a permanent establishment can be found under B005 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999
[180] Participation exemption is discussed under A014 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999 and can be derived from article 13 CITA
[181] This is based on a Dutch Court decision; Supreme Court May 31, 1978, nr. 18 230, BNB 1978/252
[182] The investment allowance can be found under E018 in: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999
[183] See for example: Implementation regulation of the Energy Investment Deduction of 2001 in conjunction with article 3.42 of the second, fifth and seventh paragraph of the Income Tax Act 2001 (Dutch: "Uitvoeringsregeling energie-investeringsaftrek 2001 juncto 3.42, tweede, vijfde en zevende lid, van de Wet inkomstenbelasting 2001") *and* remarks made in C015of the "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999

Non-standard rulings

Besides all the above, it is also possible to conclude non-standard rulings.[184] Although the facts and circumstances of these rulings differ, the same regulations, conditions and exclusions are applicable. There are a few rulings that apply for a non-standard approach;

- Cost-plus ruling;
- Intra-group finance ruling;
- Participation exemption ruling;
- Informal capital ruling; and
- Other. Enterprises may request certainty about other tax positions that are taken.

Finally, although most rulings have a maturity of four years, some cases can apply for a longer period. This will be the case if it is already clear in advance that the nature of the activities or investments are long-term.

---

[184] This can be deducted from Z003: "Code of Conduct (Business Taxation)." ec.europa.eu. Version SN 4901/99. Council of the European Union, 23 Nov. 1999

Dutch TP regulations

The arm's length principle is, as mentioned earlier, incorporated in article 8b CITA in 2002. It is a codification of the existing practice as well as a confirmation of the arm's length principle. TP was already incorporated by a clause in the Dutch Income Tax Act 1964 and 2001, but was never mentioned explicitly.[185] The wording is knowingly done in a similar way as compared to the OECD Tax Convention.[186] Double (non-)taxation due to different Dutch TP provisions is thus largely avoided. It is important to note that the OECD Guidelines are solely a directive and is therefore no leading legislation.[187] However, the incorporation has led to an extension of the at arm's length principle, since documentation requirements were included. This is done by the legislature to avoid unjustified profit shifting from the Netherlands to other jurisdictions and is therefore a defensive measure.[188] Many jurisdictions had already strict documentation requirements in place, causing the fear that profits would be unjustifiably shift to those jurisdictions, because the Dutch tax authorities could bring little to the table against it. The same applies to jurisdictions with a relative low tax rate.

The legislator takes a further stance by means of the TP decree. It discusses the relevant methods and other concerns in regards to TP. The position as well as the starting point of the Dutch tax authorities is displayed in the TP decree. Finally it discusses the bilateral perspective and refers to Dutch case law.

Affiliation

The at arm's length principle applies to transactions between affiliated companies. The Dutch CITA does not define precisely when entities are affiliated to each other. It could be considered an 'open norm' because affiliation can originate from through capital, management or supervision.[189] There must be sufficient control in regards to the entity under review, making the shareholder interest thus of a non-decisive influence.[190] Following the advice from the Council of State, the legislature chose not to define a fixed

---

[185] Article 3.8 Dutch Income Tax Act 2001 (after 2000); and Article 7 Dutch Income Tax Act 1964 (before 2000)
[186] MvT, *Kamerstukken II*, 2001-2002, 28 034, nr.3, p. 20
[187] The Dutch Supreme Court supports this, Hoge Raad June 28, 2002, nr. 36 446, BNB 2002/343: §3.2.2
[188] Nota n.a.v. het Verslag, *Kamerstukken II*, 2001-2002, 28 034, nr. 5, p. 33
[189] Peerbooms, J., and M. van der Breggen. "Transfer pricing: verplichte kost voor iedere fiscalist!." Forfaitair 173 (2007)
[190] MvT, *Kamerstukken II*, 2001-2002, 28 034, nr. 3

percentage of the capital (i.e. shares) because this allows for tax planning structures to be set-up.[191] Efforts are thus made to avoid manipulation to, for example, evade the documentation requirements.

The figure below demonstrates the types of affiliation mentioned and will be discussed below:

Figure 9: Types of affiliation
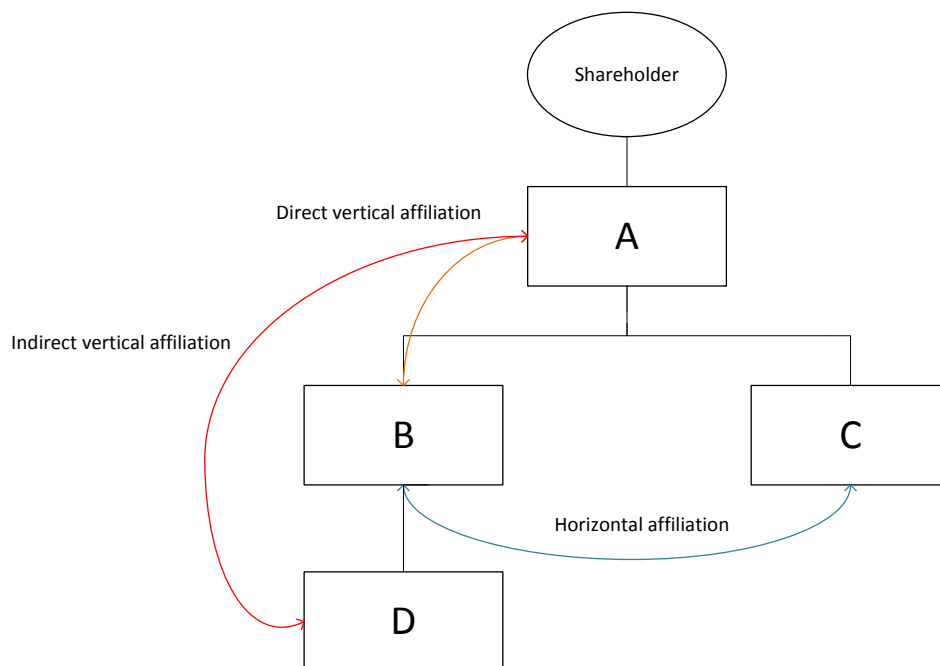
In the first paragraph of article 8b Dutch CITA the vertical affiliation is defined. This connection arises through direct or indirect participation of an enterprise in the before-mentioned management, control or capital of other enterprises. The second paragraph discusses the horizontal affiliation and its application to paragraph one. Horizontal affiliation arises when two enterprises are under the supervision or control of the same person or have the same ultimate shareholder. This could be either a natural person or legal person.

---

[191] Council of State, *Kamerstukken II*, 2001-2002, 28 034, nr A, p. 7

Documentation and audits

The third paragraph determines that affiliated enterprises are required to document data in their administration that shows how the transfer prices have been established. It ensures that tax authorities have access to sufficient information with regards to the transfer prices set and allows them to audit enterprises in a proper manner. Foregoing information cannot be obtained by the tax authorities without contribution of the taxpayer.[192] Taxpayers are required to document conform OECD TP practices.[193] It is the intention that the administrative burden therewith is as low as possible.[194] No 'best practice' is given, in order to prevent that individual enterprises are obliged to collect information that is not relevant for their particular case.

The burden of proof lies initially on taxpayers whom are also obliged to provide the administration to the tax authorities.[195] When a taxpayer remains in default after the tax authorities have requested for information, the tax authorities may then issue an information disposal.[196] The objective of the information disposal is to provide legal protection in situations where the tax authorities requests information of taxpayers that doubt the legitimacy of the request. The information disposal is therefore subject to objection. One has to clarify that the transfer price is set correctly by means of a plausible set of information and therefore no compelling evidence is required. However, if the information disposal has become irrevocable, then the burden to proof lies with the taxpayer when the tax authority takes up a TP position.

The tax authorities need to demonstrate why the used TP system would be inappropriate when a taxpayer has met documentation requirements. If the conclusion is that the transfer prices applied were in fact not at arm's length, then this will result in one or more corrections. However, tax authorities are not allowed to change the TP method applied by the taxpayer without reasonable substantiation. The State Secretary

---

[192] Nota n.a.v. het Verslag, *Kamerstukken II*, 2001-2002, 28 034, nr. 5, p. 35
[193] The OECD Guidelines are considered to be the starting point; Nota n.a.v. het Verslag, *Kamerstukken II*, 2001-2002, 28 034, nr. 5, p. 34-36
[194] Vollebregt, H.A.. "Transfer pricing: Aanbevelingen ter beperking van de administratieve lasten voor de verrekenprijsadministratie van art. 8b Wet VPB 1969." Weekblad Fiscaal Recht 16 (2005): Kluwer Portal.
[195] Article 52 resp. article 47 State Taxes Act (Dutch: Algemene Wet Inzake Rijksbelastingen)
[196] Article 47 State Taxes Act

of Finance indicates in the TP decree that an audit should start from the perspective of the taxpayer and method applied by them at the time of the transaction.

Market penetration

As mentioned earlier, a market penetration strategy can be regarded to be at arm's length when expected results are substantiated and credible. The Dutch tax authorities have already unsuccessfully tried to make adjustment in a market penetration case.[197] As could be expected, it became apparent that the tax authorities should accept a market penetration strategy if it is substantiated in a reasonable manner. The tax authorities need to make a compelling case if they wish to make adjustments. In this case the comparative material of the inspector was not sufficiently as he used transactions based on raw coffee instead of roasted coffee. In addition, the strategy was established in line with a market penetration strategy because the owner has, inter alia, trained local people to also conquer the market.

Corrections can thus not simply be imposed by the Dutch tax authorities on the basis of a chosen strategy. It can therefore be concluded that IQ range set by a dependent entity may differ from an at arm's length range if there is market (penetration) strategy. This will thus be part of results found by enterprises to substantiate their inter-company transfer prices.

---

[197] Amsterdam Court of Appeal May 23, 2000, nr 99/1625, V-N 2000/39.10

Safe Harbour

Benefits

The simplification of compliance regulations and reduction of related costs for a particular group of taxpayers is a key element of a safe harbour, because the documentation and collection of data can be both difficult and costly. This is especially true if compliance burdens are disproportioned to the size of the taxpayer.[198] The functions that some SMEs perform and the TP risks that are inherent can be, for instance, non-complex. Strict compliance regulations are therefore often regarded as disproportioned for those companies.

Furthermore, safe harbours will provide certainty in that jurisdiction where the taxpayer is compliant with the (local) provisions. Tax authorities can provide certain parameters to the eligible taxpayers. These given parameters will set a transfer price that is deemed at arm's length for the transactions that qualify. The transfer prices will be accepted with no or limited examination when the eligible taxpayer is complies. However, a functional analysis still needs to take place in order to establish the appropriate characterization and TP method.

Since tax authorities themselves have administrative advantages, safe harbours are not only beneficial for taxpayers. The information that needs to be provided by the eligible taxpayers will be less complex and/or will contain lower TP risks. Consequently, minimal examination will be required by the tax authorities, whom could even shift the administrative burden from auditors with TP experience to less qualified auditors.

Subsequently, the tax authorities can put more resources towards the examination of more complex or higher risk transactions and thus kill two birds with one stone. On the one hand, tax revenues are secured in 'low risk' cases with a minimal amount of resources, and are on the other hand able to devote more attention towards the, for example, more complex transactions.[199]

---

[198] "Revised Section E on Safe Harbours in Chapter IV of the Transfer Pricing Guidelines" Committee on Fiscal Affairs and the OECD Council C(95)126/FINAL(2013): 4.104
[199] "Revised Section E on Safe Harbours in Chapter IV of the Transfer Pricing Guidelines" Committee on Fiscal Affairs and the OECD Council C(95)126/FINAL(2013): 4.107

Concerns

Although the tax authorities try to establish a correct transfer price using parameters, there nevertheless remains a risk that one deviates from the arm's length principle because that is inherent to a simplification. Thus it is possible that the use of a safe harbour leads to the application of, for example, the cost plus method, where a resale minus would be more appropriate under the general TP provisions. This is undesirable because small deviations in facts and circumstances could thus result in an inconsistency with regards to the arm's length principle, as the most appropriate TP method should be applicable.[200] This problem could be avoided when eligible taxpayers can opt for both the application of the safe harbour or general TP requirements. Eligible taxpayers who think that the amount of income they have to report, on the basis of a safe harbour, exceeds the arm's length amount can then apply the general TP provisions. The downfall of this approach is a possible loss of tax revenue due to 'cherry picking'.

In addition, there is an increased risk of double (non-)taxation when safe harbour provisions are introduced unilateral. A jurisdiction that introduces a safe harbour on its own account might define parameters that cannot be considered at arm's length in order to increase the reported profits in its jurisdiction. Enterprises may be inclined to adjust their transfer prices in the safe harbour jurisdiction to avoid scrutiny. They would thus report a different amount paid or charged for the transaction under review in, for example, country A and country B. Subsequently, the risk of double taxation or double non-taxation arises.[201] It may occur that SME's, who can elect both the application of a safe harbour or general TP provisions, accept the double taxation therewith. This will be the case if the costs of double taxation incurred under the application of a safe harbour outweigh the costs of compliance under the general TP provisions. Double non-taxation will generally not lead to arbitration[202], but is unavoidable when taxpayers are allowed to report profits that are below an arm's length remuneration. In turn, this will result in distortions of investment as well as trade. A solution will be to only adopt safe harbours

---

[200] This will lead, inter alia, to profit shifting, more challenges and debates from and between tax authorities, double (non-)taxation because the transfer price is not in line with economic reality.

[201] "Revised Section E on Safe Harbours in Chapter IV of the Transfer Pricing Guidelines" Committee on Fiscal Affairs and the OECD Council C(95)126/FINAL(2013): 4.112

[202] In short, arbitration refers to the possibility of tax payers to request the competent tax authorities to remove the double (non-)taxation. Both tax authorities can, for example, enter into a binding or non-binding mutual agreement procedure.

on a bilateral or multilateral basis. This ensures that there are practically no divergences in reported transfer prices between jurisdictions.

Furthermore, safe harbours open the possibility to enter into tax planning schemes and possible tax avoidance therewith. It becomes possible to shift profits to other jurisdictions, simply by a modification of their TP policy. For instance, tax structures could be set up by the use of safe harbour jurisdictions, in order to reduce the overall compliance burden and/or the taxable income. A reduction of the effective tax rate could be the result. Abuse is possible when safe harbours are applicable for only 'simple' or 'small transactions.[203] Transactions can be presented in such a way that it appears to meet the parameters. It can be difficult for the tax authorities to obtain the actual facts and circumstances because the documentation requirements are often limited when using a safe harbour. In addition, tax authorities will not, or with limited scrutiny, audit these transactions. Cost efficient enterprises could also benefit from the application of a safe harbour by shifting the remainder of the actual arm's length mark-up to low tax jurisdictions.

Equity and uniformity issues can arise when similar taxpayers are treated differently through the safe harbour provisions. If the safe harbour criteria are not precise enough defined, then similar taxpayers could receive a contradictory treatment. In turn, this will lead to both discrimination and competitive distortions in the market place.

---

[203] "Revised Section E on Safe Harbours in Chapter IV of the Transfer Pricing Guidelines" Committee on Fiscal Affairs and the OECD Council C(95)126/FINAL(2013): 4.120

Appendix B - Comparability

This section provides a more detailed discussion in regards to the comparability adjustments that can be made, in order to get comparable companies. Subsequently, the timing possibilities are further examined, because they can lead to more robust arm's length prices.

Comparability adjustments

Comparability adjustments may be appropriate if potential comparable companies found are initially not comparable. The type of adjustment that needs to be made depends on the facts and circumstances. Several comparability adjustments can be distinguished. For instance, adjustments can be made due to different accounting standards across border. Differences in accounting standards may, for example, lead initially to a non-comparable gross margin. At the same time adjustments are appropriate when differences in functions performed, assets used and risks assumed occur. This is typically done by a working capital adjustment. These adjustments are aimed to recognise the comparable cost or price basis. The comparable cost basis will be adjusted on inventory level, whereas adjustments based on the comparable price basis are derived from accounts receivables and payables. The inventory level will be compared to the level of sales.[204]

Furthermore, it will generally be difficult to determine one price that is deemed reliable enough to be represented as 'the' arm's length price. Therefore, in order to improve comparability, less reliable comparable companies need to be eliminated after which statistical methods could be applied. A statistical analysis could be used to minimize the range and increase the comparability in the event the remaining comparable companies show deficiencies.[205] The IQ-range is a widely used and recognised statistical method in

---

[204] Loots, Jozua Johannes. A comparability adjustment transfer pricing model. Potchefstroom: Potchefstroom Business School University of North West, 2006, p.158
[205] §3.62 OECD Guidelines

which the upper and lower quarter of the total number of observations in the population found are eliminated.[206]

If, based on the comparable companies used, there are still doubts about the method applied; a so-called sanity check can be performed, whereby a different transfer pricing method is used for the analysis. This could be used to determine a more accurate arm's length range if, for example, a different range arises. In addition, if the relevant values of the controlled transactions (e.g. price or margin) fall within the computed arm's length IQ-range, no adjustments may be made and the transaction is considered to be at arm's length.[207]

Timing

The OECD Guidelines recognise that it could be useful to take more than just the year of the transaction into consideration. Such a consideration can reveal facts that may (have) influence(d) the determination of the transfer price. Think, for example, of a product life cycle or an economic crisis. It refers particularly to situations where the price or profitability is affected over time. The ups and downs in a cycle could be partially filtered out with data of multiple years and show if comparable transactions are influenced in a similar fashion due to a cycle. In addition, the contemplation of a single year may lead to a distorted view, because of the absence of a complete overview. It may also be clear that the profit margins are influenced by, for instance, an independent enterprise that is in a start-up phase or executes a market penetration strategy.

The above is also known as the ex-ante or price setting approach.[208] It means that a taxpayer needs to make reasonable efforts to establish an at arm's length transfer price with the information available at the moment intercompany transactions are undertaken. This contradicts with the ex-post or price evaluation approach, which allows taxpayers to test the actual outcome of the undertaken transactions.[209] They need to demonstrate that the conditions, of the transactions undertaken, are in line with

---

[206] Rubin, Allen. Statistics for evidence-based practice and evaluation. 3rd ed., student ed. Belmont, CA: Brooks/Cole, Cengage Learning, 2013, p.67
[207] §3.61 OECD Guidelines
[208] §3.69 OECD Guidelines
[209] §3.70 OECD Guidelines

the arm's length principle. A year-end adjustment will have to be made if it appears that the transfer price set is not at arm's length. The risk of double taxation or double non-taxation occurs when two jurisdictions have different approaches.[210]

Although the data obtained typically originates from earlier years, tax authorities are not allowed to simply make year-end adjustments and corrections with the use of 'hindsight'. This means that afterwards developed insights cannot (solely) be used to make corrections, despite the fact that this data can give a better understanding of the actual conditions and agreements between parties.

---

[210] "Report on Compensating Adjustments." EU Joint Transfer Pricing Forum JTPF/009/FINAL/2013/EN (2013)

Appendix C – Amadeus results

Wholesale Distribution Industry	Appearance Products	Automotive	Chemicals	Electronica Products*	Food Products	Industrial Products	IT Products**	Pharmaceuticals
Results dataset (six year average)								
Geographical Area	pan-European + 3							
NACE Rev. 2 Codes	46.45: Wholesale of perfume and cosmetics 46.48: Wholesale of watches and jewellery 46.42: Wholesale of clothing and footwear	45.11: Sale of cars and light motor vehicles 45.31: Wholesale trade of motor vehicle parts and accessories	46.12: Agents involved in the sale of fuels, ores, metals and industrial chemicals 46.75: Wholesale of chemical products	46.43: Wholesale of electrical household appliances 26.11: Manufacture of electronic components	46.31: Wholesale of fruit and vegetables 46.32: Wholesale of meat and meat products 46.33: Wholesale of dairy products, eggs, edible oils and fats	46.64: Wholesale of machinery for the textile industry 46.66: Wholesale of other machinery and equipment	46.5: Wholesale of information and communication equipment	46.46: Wholesale of pharmaceutical goods 47.74: Retail sale of medical and othopaedic goods in specialised
Industry description (inclusions)***	Trade*, Distrib*, Wholesale*, Sale*, Cloth*, Parf*, Appear*	Trade*, Distrib*, Wholesale*, Sale*, Auto*, Car*, Motor*	Trade*, Distrib*, Wholesale*, Sale*, Chemi*, Fluor*, Lac*, Wax*, Adhesiv*	Trade*, Distrib*, Wholesale*, Sale*, Appliance*, Electr*	Trade*, Distrib*, Wholesale*, Sale*, Food*, Beverage*,	Trade*, Distrib*, Wholesale*, Sale*, Industr*, Machin*, Construct*	Trade*, Distrib*, Wholesale*, Sale*, Software, Hardware, Conductor*, Compu*, Tele*, Peripher*	Trade*, Distrib*, Wholesale*, Sale*, Medic*, Pharma*
Industry description (exclusions)***	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*, Pharma*, Medic*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*, Food*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Pharma*, Medic*, Textile*, Car*, Auto*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Pharma*, Medic*, Textile*, Industrial*, Car*, Auto*, Food*	Retail*, Assembl*, Research*, Develop*, Design*, Consult*, Service*, Holding*, Software*, Semiconductor*, Tele*, Communic*, Textile*, Industrial*, Food*
Average size of dataset / number of observations****	915	1276	469	423	1479	135	494	595

* Manufacturing codes can be applied due to the exclusion and inclusions words. In addition, every independent companie performs a sales function, because it can not be done by a principal / headquarter.
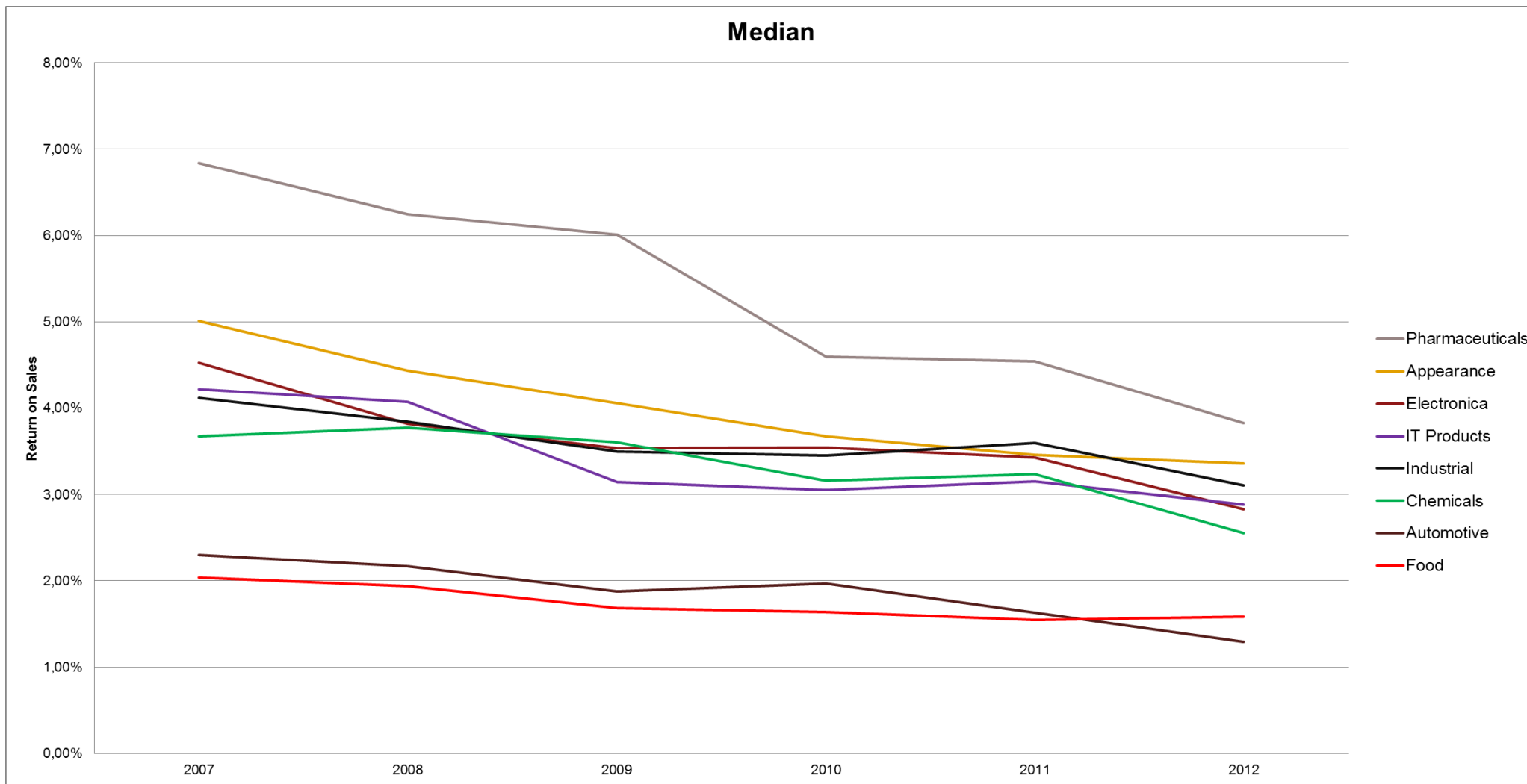
** Contrary to the electronica products, which are lamps, vacuums, toasters, etc., the IT products comprise of software and hardware used in the ICT sector.

*** The asterisk implies that both those exact words and derivatives thereof are taken into account.
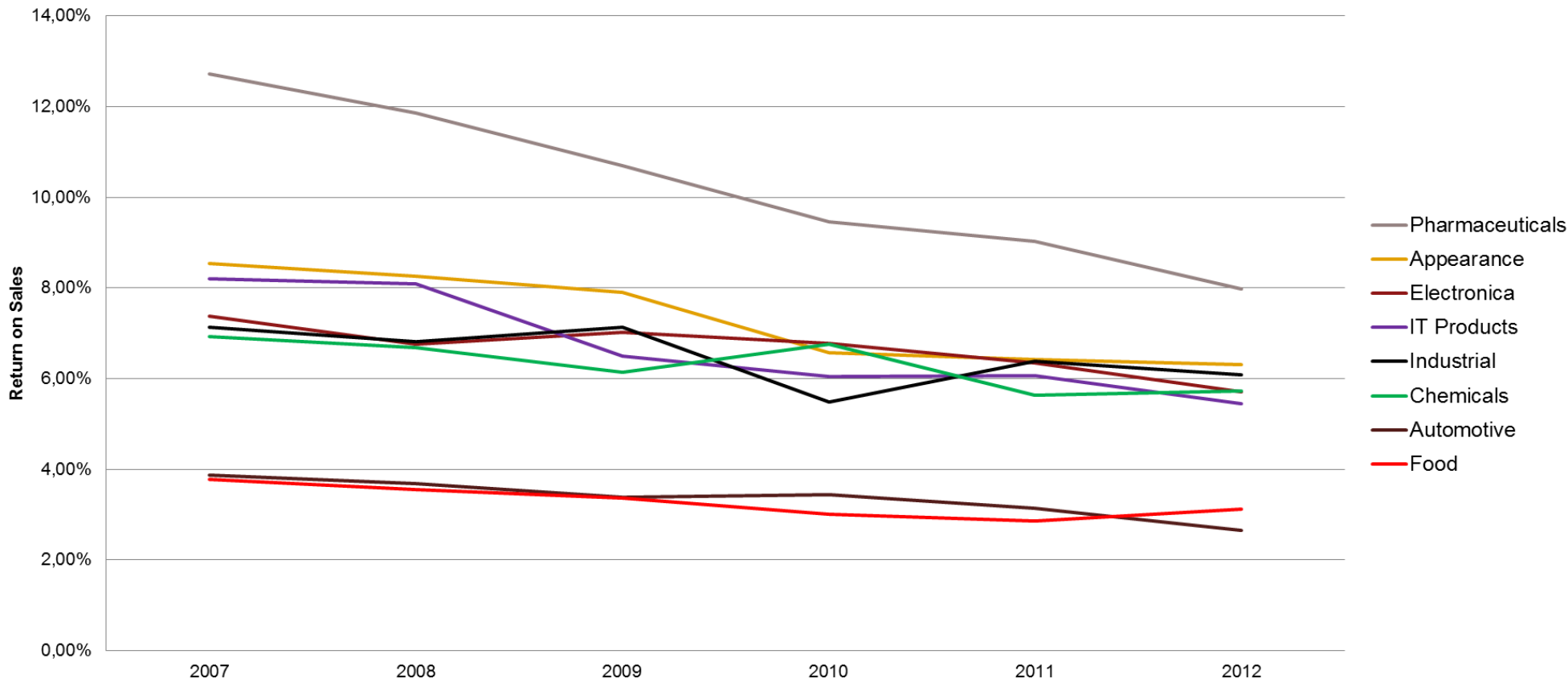
**** Data is gathered from 2007 till 2012, the six year average is shown.

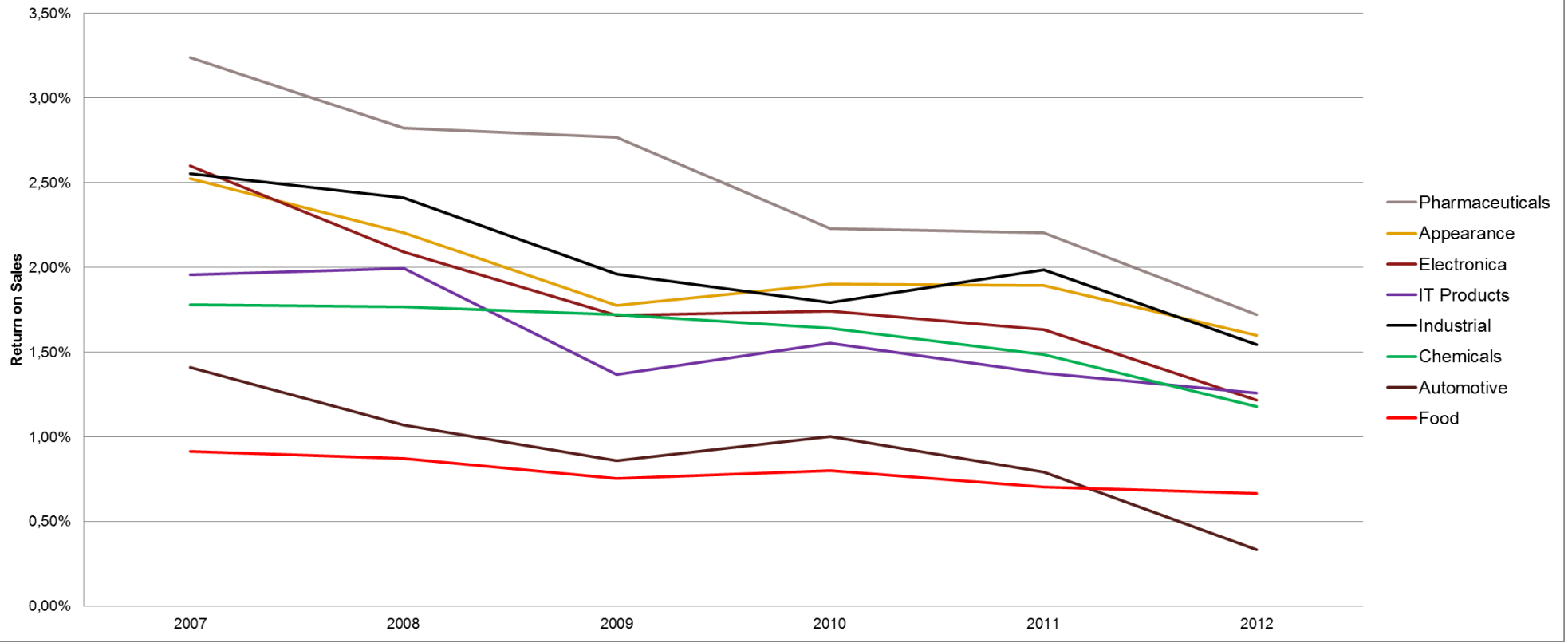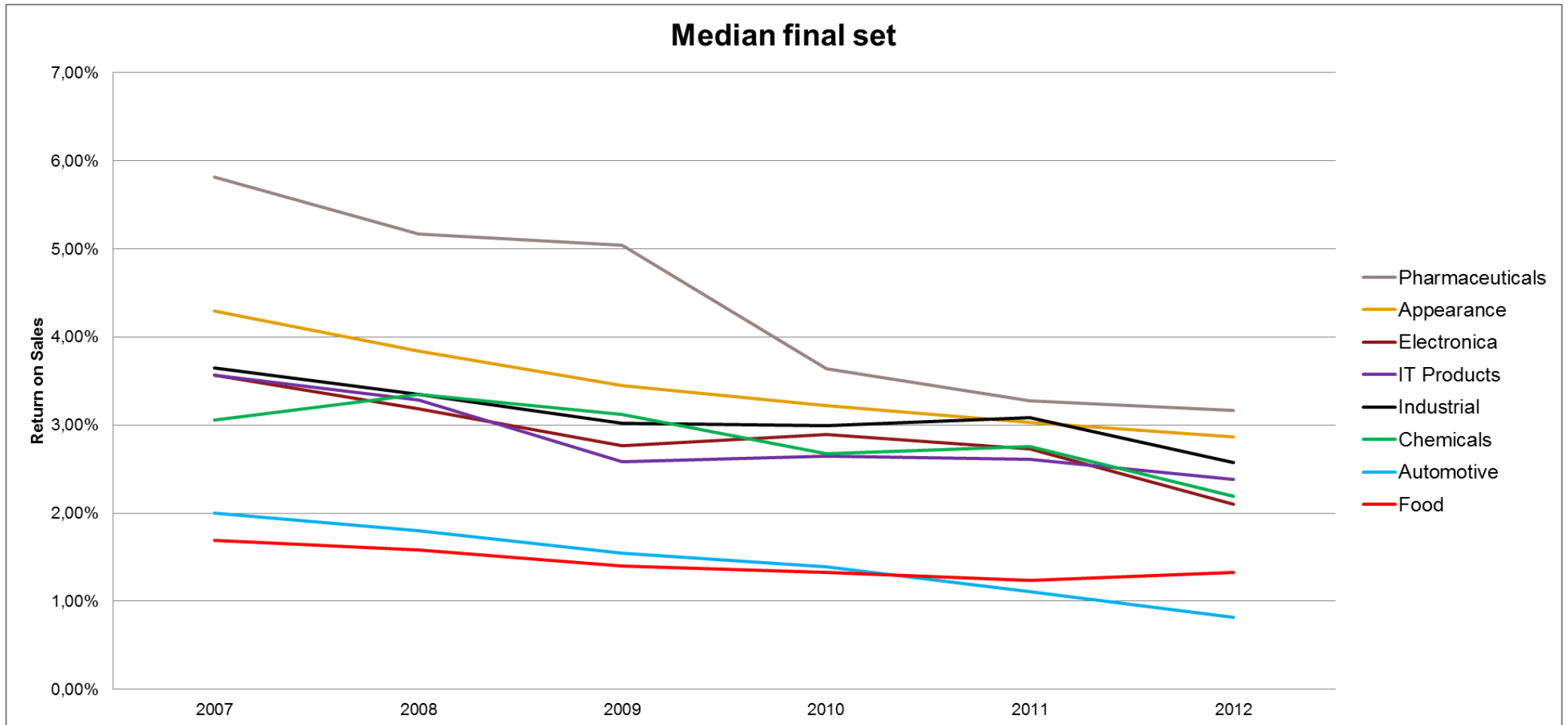# Appendix D – Dataset without loss making companies

Median chart showing Return on Sales (%) from 2007 to 2012 for: Pharmaceuticals, Appearance, Electronica, IT Products, Industrial, Chemicals, Automotive, Food.

Upper

Return on Sales

- Pharmaceuticals
- Appearance
- Electronica
- IT Products
- Industrial
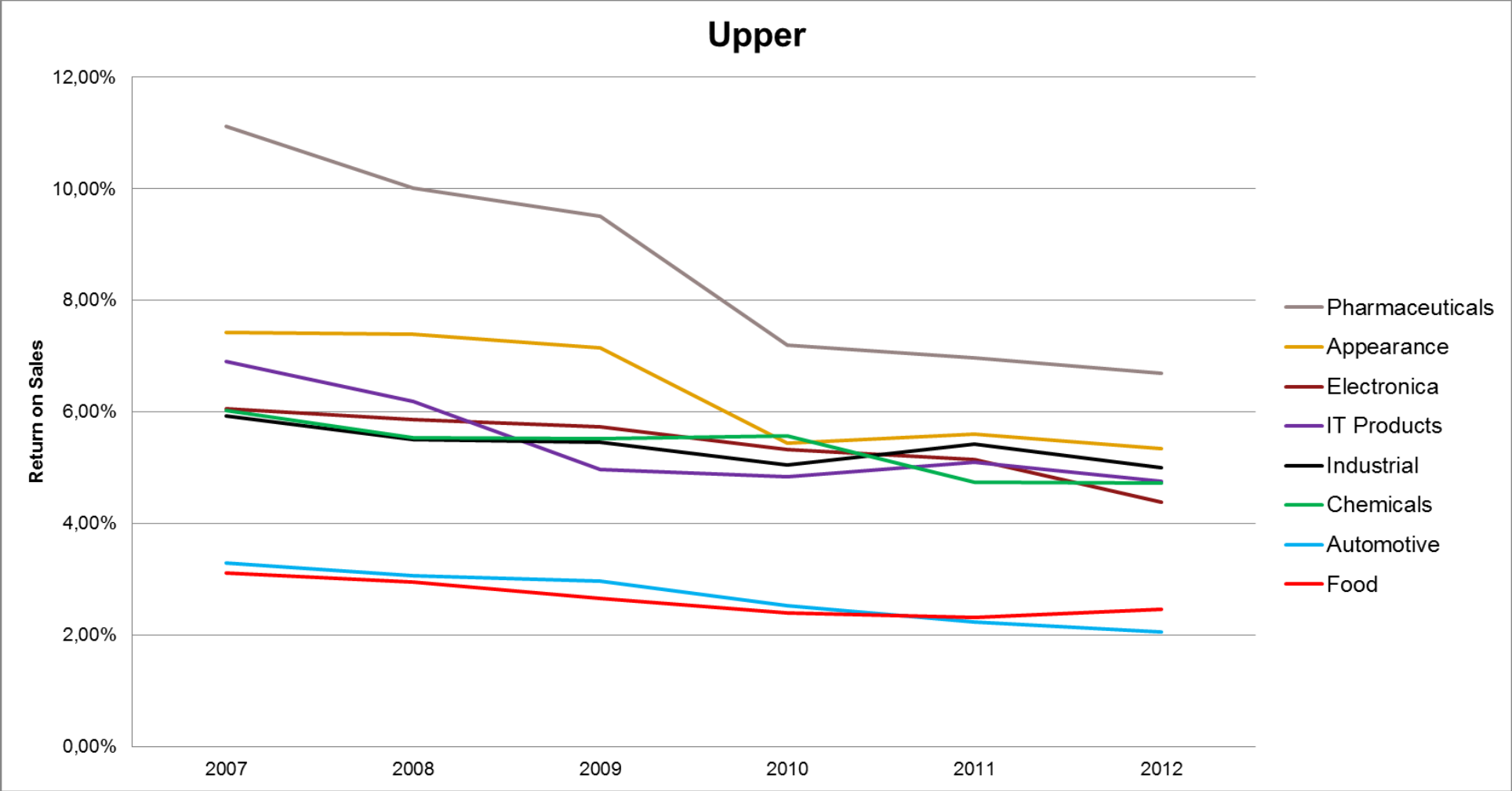- Chemicals
- Automotive
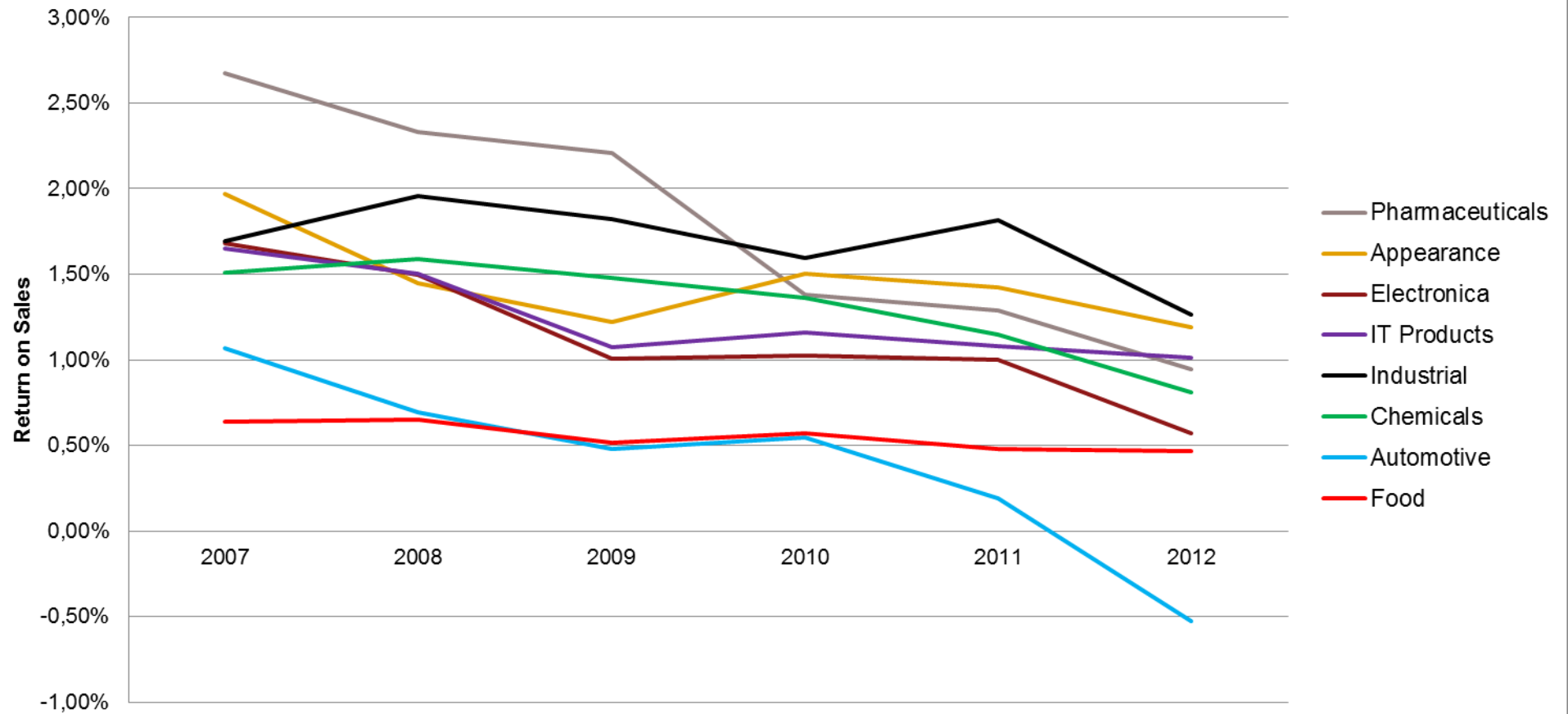- Food

Lower

Appendix E – Dataset final

Median final set

Upper

Appendix F – Chi-square one versus the rest

Chi-square Median's test - Chemicals versus rest

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,115	0,867	0,043**	0,543	0,837	0,602	0,971	0,754
Upper Quartile	0,276	0,273	0,630	0,031**	0,035**	0,901	0,993	0,659
Lower Quartile	0,377	0,898	0,004***	0,036**	0,807	0,537	0,395	0,120
Total								

The given values are all p.values

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

Chi-square Median's test - IT Products versus rest

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,556	0,991	0,054*	0,444	0,356	0,244	0,118	0,874
Upper Quartile	0,007***	0,148	0,128	0,058*	0,556	0,050*	0,557	0,649
Lower Quartile	0,838	0,684	0,005***	0,444	0,612	0,006***	0,099	0,101
Total								

The given values are all p.values

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

Chi-square Median's test - Electronics versus rest

	2007	2008	2009	2010	2011	2012	07-09	10-12
Median	0,371	0,540	0,593	0,380	0,698	0,177	0,485	0,447
Upper Quartile	0,073*	0,839	0,279	0,587	0,888	0,043**	0,949	0,227
Lower Quartile	0,744	0,729	0,151	0,052*	0,271	0,001***	0,107	0,033**
Total								

The given values are all p.values

*** Significant at α = 1%

** Significant at α = 5%

* Significant at α = 10%

Appendix G – History of Globally Integrated Enterprises

This appendix will introduce the rise of the globally integrated enterprises.

Due to a number of fundamental changes, MNE's were forced to reconsider their business model and approach. The trend of globalisation increased over the last twenty years because investment and trade barriers gradually declined. The rise of (information) technology has increased the speed of global communication and has also ensured a reduction in related costs. In addition, changes in the geopolitical climate allowed MNE's to explore both new markets and hidden talents.[211] Enterprises changed their business model from a supply chain that solely focuses on efficiency, to a model that uses the supply chain to increase total revenue as well as profit and overall consumer satisfaction.

A well-known trend is that European and American enterprises outsource their manufacturing needs to, for example, cheap labour countries. This led to a more sustainable business model, since the centralisation of all the functions is no longer efficient. Although foregoing centralisation of functions in one jurisdiction is still possible, globalisation thus allows MNE's to increase revenue and cut cost. In addition, diversification and decentralisation ensures that risks are being dispersed.

Ultimately, a GIE can enable itself to allocate their operations in the most-favourable locations due to the global network that it has established. Thus, they allow themselves to choose that particular location that either offers them the lowest cost level or best strategic value.[212]

---

[211] Sodhi, ManMohan S., and Christopher S. Tang. A long view of research and practice in operations research and management science the past and the future. New York: Springer, 2010. P. 195
[212] Sodhi, ManMohan S., and Christopher S. Tang. A long view of research and practice in operations research and management science the past and the future. New York: Springer, 2010. P. 196-197

Appendix H – Diversification

This figure shows the index of relative length of global value chains. It represents the world average for a number of selected industries.
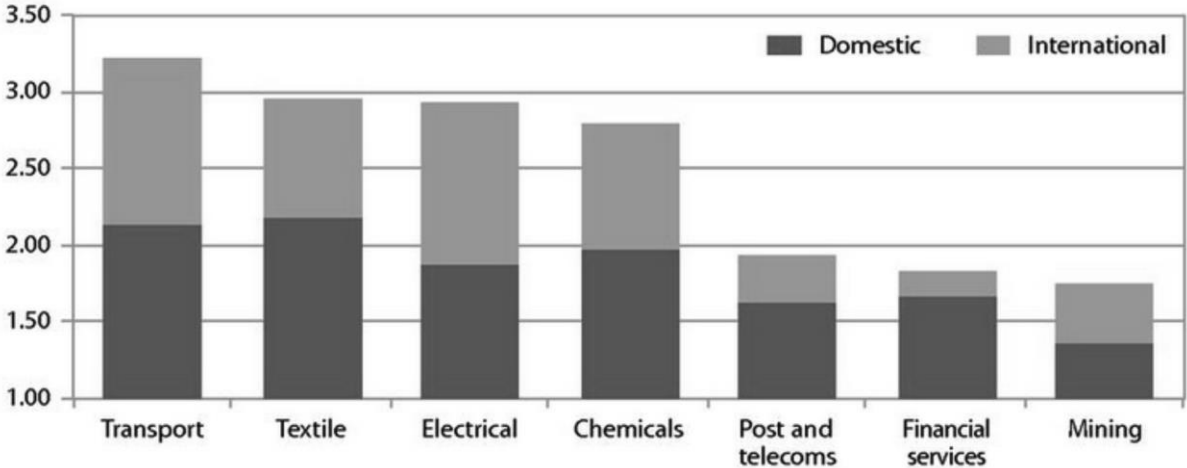
Source: OECD, Statistics Directorate, Global Value Chains: OECD Work on Measuring Trade in Value-Added and Beyond, internal working document, (OECD 2012).