

# A GUIDELINE TO A COST BENCHMARK OF MARKET DATA

How to obtain reasonable prices of market data

DANISH SECURITIES DEALERS  
ASSOCIATION  
11 JULY 2019

AUTHORS

Sigurd Næss-Schmidt, Partner  
Malte Lisberg Buhr, Senior Economist  
Jonas Bjarke Jensen, Economist

## PREFACE

---

As documented in the recent report by Copenhagen Economics “*Pricing of market data*”, the cost of market data and the complexity in data access has increased significantly following the implementation of the first MiFID in 2007. We argued that this impairs financial market efficiency, which in turn has negative effects for market participants and eventually end-investors. Therefore, we suggested capping revenues from raw market data using a so-called Long Run Incremental Costs (LRIC+) model.

As a follow-up to this report, the Danish Securities Dealers Association has asked Copenhagen Economics to prepare a guideline on how concretely to implement regulation to ensure reasonably priced market data. We have specifically been asked to design an operational guideline that is consistent with the regulatory framework as prescribed by MIFIDII/MiFIR. This includes data collection, establishing a cost benchmark and a framework for a true LRIC+ model.

Establishing a Consolidated Tape Provider (CTP) has recently been put forward as a solution to contain costs and improve the functioning of the market for market data. This report will not directly address the CTP but cost benchmarking of input data can in turn be an important prerequisite for a well-functioning CTP.

# TABLE OF CONTENTS

---

<b>Preface</b>	<b>0</b>
<b>Executive summary</b>	<b>2</b>
<b>1 A new regulatory framework for market data</b>	<b>6</b>
1.1 Market data activities in trading venues should be regulated	6
1.2 How to implement the revenue cap: a dialogue-based cost benchmark	9
1.3 What is possible within the current regulation?	11
<b>2 Operational guideline on how to implement cost benchmarking</b>	<b>13</b>
2.1 A FIVE step regulatory implementation process	13
2.2 How to initiate the first cost benchmark	27
2.3 Timeline for the first cost benchmark	28
<b>References</b>	<b>34</b>

## EXECUTIVE SUMMARY

---

Over more than a decade, following the introduction of MiFID I in 2007, the overall costs of purchasing market data have increased significantly. The cost increases are not due only to increased pricing of existing products, but to a large extent also additional fees that have added to overall costs. The size and nature of these increases have been thoroughly documented in several reports, and they are also briefly referenced in this report.

In particular, it is clear that the underlying change in the costs of *producing* and *disseminating* market data do not justify the increase in the costs of *purchasing* market data. In other words, the trading venues are setting their prices of market data far above the costs of production, in conflict with the MiFID stipulation (reinforced in the revised MiFID II/MiFIR), which states that market data should be sold at “*reasonable commercial prices*” with a “*reasonable relationship to the cost of producing and disseminating that data*”.

Copenhagen Economics have in earlier reports suggested that a so-called Long Run Incremental Cost model (LRIC+) should be implemented as the long-term solution to ensure alignment between the costs of producing market data and the market data fees. The model is used extensively in other industries, characterised by a similar monopoly market position. Essentially, it defines a cost benchmark for a given trading venue, i.e. the overall costs the exchange would have – if efficiently operated – of providing raw market data to its customers.

We are proposing a *revenue cap* as a first-best solution – *not a price cap* on individual products. Indeed, the recent decade has demonstrated the futility of working with price caps; the overall prices charged for market data have increased not only as a consequence of higher fees of “old” products/services, but also because trading venues have added new fees as a precondition for getting access to the data. A regulation based on ESMA being involved in assessing individual costs of hundreds of products across the different trading venues – that are being changed on an on-going basis – is neither doable nor meaningful.

Instead, provided that an overall revenue cap is in place, trading venues could be given some freedom in structuring prices and products in ways that improve market efficiency in on-going dialogue with users across the industry.

### *Five steps to implement a revenue cap*

Concretely, we suggest a five step approach, stated in sequential order (cf. also Figure 1):

- 1. Collecting and verifying data on production costs:** This is a highly critical phase, which should give answers to the following questions: (1) Which market data products should be included under the regulation? The focus should be on the raw data integral to execution of trades, not derived data nor value-added services. (2) Which costs at the trading venues can be linked to production of these data? (3) Which procedures should be in place

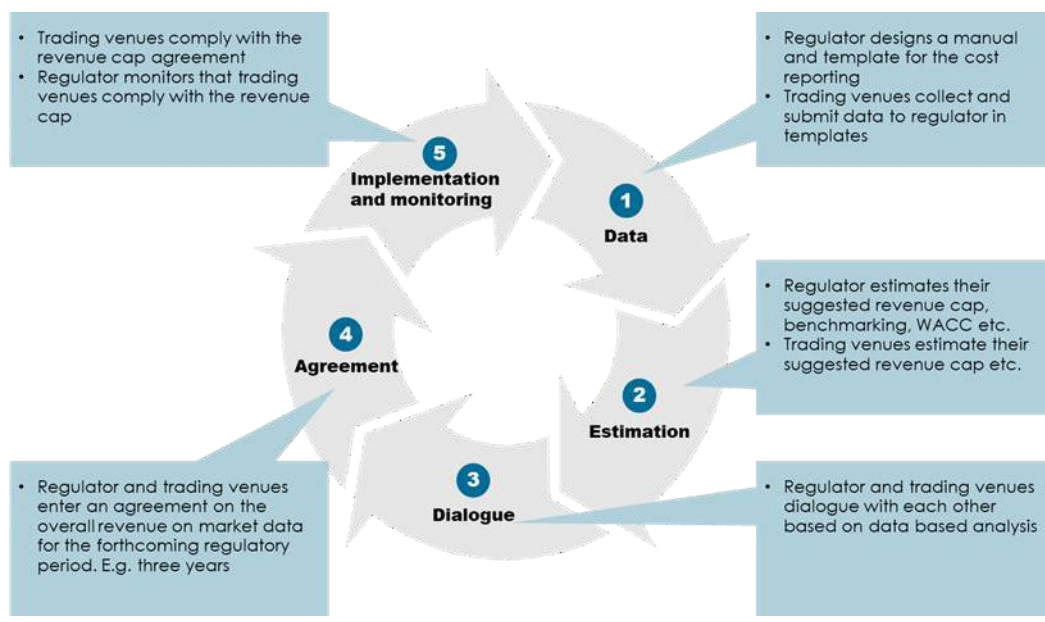
within the trading venues to ensure a uniform application of the cost assignment estimates? This step is important because the directive does not contain a precise definition of what data and what types of costs that should be used to measure whether the pricing of data is excessive. As described, we argue the costs should be defined in line with the LRIC+ model, possibly in the upcoming review of MiFID/MiFIR.

2. **Design of cost benchmark models:** Once the cost data is available, cost benchmarks can be developed. We recognise that the origination of market data is strongly linked to the execution of trade. The costs that we propose are those linked to the distribution of market data, i.e. are incremental to execution of trades. We outline how this can be done in practice and how sensitivity tests are crucial to get it right and ensure efficient regulation.
3. **Dialogue/negotiation:** Experience from other regulated industries suggests that cost benchmarks in the context of defining revenue caps should not be used in a mechanistic manner. There should be room for dialogue between parties to discuss whether identified factors should allow revenue caps to deviate from the benchmark models. Indeed, good benchmarking may imply several competing indicators.
4. **Agreement:** Optimally, the negotiations lead to an agreement and common understanding of the efficient cost level. Experiences from other regulated industries propose that it is important that both parties honour the agreement for the future collaboration.
5. **Implementation/enforcement:** Once an agreement is reached, each trading venue should implement a pricing of market data that complies with the agreed-upon cost level. If the parties do not reach an agreement, it can be necessary for the regulator to enforce a revenue cap based on regulator's estimation. However, the present MiFID II/MiFIR directives do not provide the regulator with tough means of enforcement as described below.

We recommend that every third year, the regulator follows the five-step regulatory wheel, cf. Figure 1. After the first round, step 1 and 2 will be significantly less demanding, as the benchmarking model should only be updated and not developed from scratch.

We strongly recommend carrying out step 1 and 2 as an open process. Experience from application of LRIC+ models in other industries shows that a common understanding of how data is being collected and benchmarked is needed to deliver high quality and obtain acceptance among the regulator, trading venues and users.

**Figure 1**  
**A three-year regulatory wheel**



Note: Illustration  
Source: Copenhagen Economics

*How much can be implemented under the current regulation?*

**Steps 1 and 2:** Are clearly fully within the remit of the existing regulation. Indeed, without data collection and some measurement of whether total costs and revenue grossly match, there is no way to ensure that market data is provided on a “reasonable commercial” basis.

**Step 3 and 4:** Also, nothing precludes ESMA from engaging in dialogue with the trading venues on their costs of producing market data, and *potentially* agree on a reasonable level of market data fees. Although the incentives for the trading venues to find an agreement is diminished as there is no threat of a hard fall-back with a binding revenue cap, in case of no agreement, cf. step 5 below.

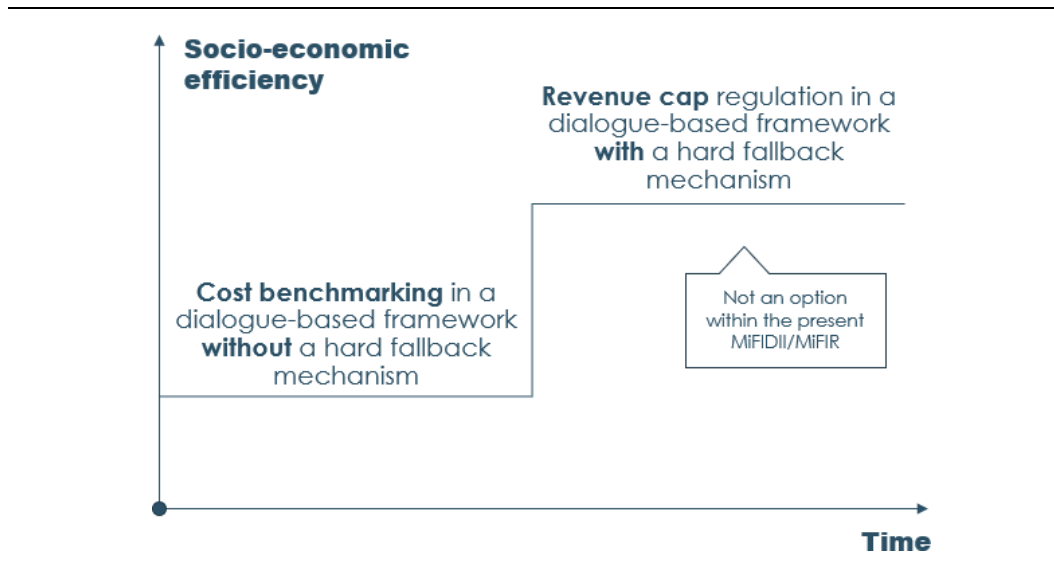
**Step 5:** Under the current regulatory structure, ESMA does not have the remit to implement and enforce a binding revenue cap for the trading venues, in case step 4 does not lead to an agreement.

As such, we can split our proposed regulatory model into two sequential phases, cf. also Figure 2:

- Only step 1-4:** Constitute in itself a **second-best regulatory model** – a dialogue-based cost benchmark model without any strict revenue cap, but with a reasonable obligation for regulators to sanction trading venues if prices for market data are not set according to existing regulation.
- Including step 5:** Once the regulation has been amended to allow for a dialogue-based framework with a hard fallback, providing a **first-best regulatory model**. Moreover, in this regulation the regulator can adjust future revenue caps if the trading venue does not comply with the agreement.



**Figure 2**  
**The regulatory progression model**



Note: Illustration  
Source: Copenhagen Economics

### *Link to Consolidated Tape Provider (CTP)*

The construction of a consolidated tape has been put forward as a possible solution to a number of the identified issues of market data access and excessive pricing. As described in our report “*Pricing of market data*”, the extent to which a CTP can solve the issues depends very much on the exact design of it, although we are sceptical that a CTP can be complete solution.

This guideline is not directly addressing a CTP. However, if a consolidated tape is established, it will be necessary to estimate how much the trading venue should be compensated for distributing the market data to the consolidated tape, and how much the consolidated tape provider should be remunerated, i.e. pricing of the consolidated tape. In doing so, the above described framework could be used.

### *Structure of the report*

The rest of the report is structured as follows: *In chapter 1*, we argue why the market structures call for a regulation, we provide an overview of the suggested regulatory model and we analyse whether any regulatory changes are needed to implement the regulatory model. *In chapter 2*, we provide a concrete five step technical guideline on how to implement the regulation and describe how and when the regulation can be initiated.



## CHAPTER 1

**A NEW REGULATORY FRAMEWORK FOR  
MARKET DATA**

In this chapter, we set the stage for regulating market data activities in a more socio-economically optimal manner to secure the essential role that access to market data has for the functioning of the financial market. In *section 1.1*, we describe why trading venues' sale of market data must be regulated. In *section 1.2*, we provide an overview for the guideline to ensure a uniform understanding and usage of cost benchmarking. In *section 1.3*, we discuss what will be possible to implement within the current MiFID II/MiFIR, and what requires regulatory changes.

**1.1 MARKET DATA ACTIVITIES IN TRADING VENUES  
SHOULD BE REGULATED**

The primary functions of trading venues are listing as well as matching buyers and sellers of securities at fair prices based on transparent rules, as also described in the recent report by Copenhagen Economics "*Pricing of market data*", as well as in reports by Oxera "*Pricing of market data services: an economic analysis*" and "*The design of equity trading markets in Europe*". These primary functions generate operational expenditures (OPEX) due to activities within sales and marketing, software, infrastructure operations, quantitative research, compliance and R&D. Moreover, the primary functions generate capital expenditures (CAPEX) due to depreciations and maintenance of the physical assets such as software, databases and infrastructure.<sup>1</sup>

Besides these primary functions, trading venues offer secondary products and services, which have become a very profitable business line<sup>2</sup>, namely the distribution of market data offered to security dealers, asset managers and in extension also to end-customers. Raw market data consists of blended bids and asks from all participants and executed trades on the specific venue. The trading venue has a natural monopoly of these raw market data as the data is initially only available to the trading venue, as described in detail in the Copenhagen Economics report "*Pricing of market data*". The "willingness" to pay for this market data is very high, partly since market data contains fundamental knowledge that is indispensable when participating in the market for trading, and partly for regulatory reasons of "best execution" and other requirements.

This implies that the trading venues produce two connected services:

- A commercial service (matching buyers and sellers)
- A monopoly service (selling market data)

Concretely, trading of securities generates market data that is published as a data feed, which is divided into two general categories, cf. Box 1:

- Pre-trade market data
- Post-trade market data

---

<sup>1</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 2, <https://iextrading.com/docs/The%20Cost%20of%20Exchange%20Services.pdf>

<sup>2</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 2

### **Box 1 Raw market data definition**

We define raw data as: 1) Pre-trade data: all bids and asks at the venue and 2) Post-trade data: data on executed trades.

Both pre- and post-trade data should include the following information:

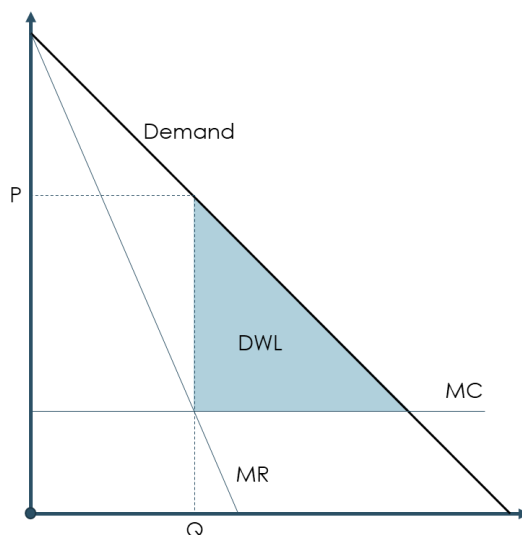
- Price
- Volume
- Identification of the traded security (ISIN, etc.)
- Timestamp
- Counterparty information

Covering all types of latency.

In an unregulated monopolistic market, a profit-maximising trading venue's pricing strategy is to set prices to maximise net revenues. Hereby, market data fees are determined by the customers' willingness to pay (demand) in contrast to a competitive market where prices are determined by the marginal costs of generating and distributing market data. In this context, customers sometimes – due to regulation – are obligated to buy market data. The lack of competition in this market results in an outcome with a suboptimal low level of market data purchases, creating a socio-economic loss, cf. Figure 3. On top of that, this pricing strategy does not comply with the MiFID II/MiFIR directive as described in the Copenhagen Economics report "*Pricing of market data*".

**Figure 3**  
**Profit-maximising monopoly**

Y axis: Price and costs



Note: DWL: Dead Weight Loss, which is the loss of economic efficiency resulting from restricted output and too high prices; MR: Marginal Revenue; MC: Marginal Cost; P: Price and Q: Quantity

Source: Perloff, Jeffrey M. (2016), *Microeconomics* – seventh edition, page 396

### *Monopolistic markets are often regulated through a revenue cap*

Many markets in the EU and other countries use regulatory models specifically designed to handle similar monopoly situations, especially when it comes to private ownership companies managing infrastructure critical to the overall economy, as in the present case. Often, these models involve revenue cap regulation, cost benchmarking and rules on allowed rate of return. The overall purpose is to imitate the outcome of perfect competition, e.g. an overall price setting that equals marginal costs.

Natural monopolies such as distribution of telecommunications, broadcasting, electricity, gas, water, waste water and district heating as well as airports are examples of industries subject to full-scale revenue caps, benchmarking and rules on allowable rate of return. Other network economies are regulated in a similar fashion.<sup>3</sup>

By and large, comparable industry regulation focuses on total revenue caps, not price regulation of individual services. The reason is that the pricing of individual services would be subject to uncertainty and it will limit the trading venue's commercial maneuvering room. Moreover, a regulator does

<sup>3</sup> There are also discussions on whether social media – e.g. Facebook – is a reasonable comparable industry, given that their platform also includes i) registration of new users, ii) users' interaction on the platform and iii) sale of information and commercials to external parties – where the latter may be comparable to the sale of market data from venues. However, there are several crucial differences between the business model of social media and the business model of venues. Firstly, it is statutory that security dealers collect data from different exchanges in order to obtain the best execution on behalf of end-customers, meaning that their demand for market data is mandatory. Additionally, security dealers must document that they obtained the best execution at the specific time. This is a significant difference to Facebook, which also experiences increasing competition from other social networks, as Facebook data is not unique and unsubstitutable which is the case for market data from venues. Finally, the functioning of social media platforms does not have the same crucial role for real economic productivity.

not want to micromanage the trading venue, but create the overall framework, which would guide the industry in the right direction.

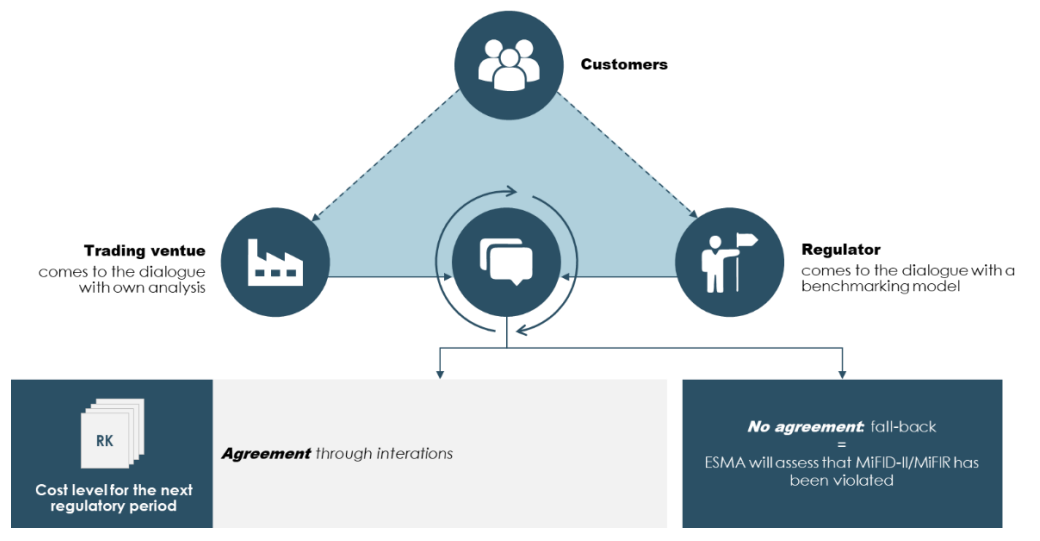
## 1.2 HOW TO IMPLEMENT THE REVENUE CAP: A DIALOGUE-BASED COST BENCHMARK

In this section, we introduce the overall regulatory framework. In the next chapter, we provide details of the regulatory model and how to carry it out step by step.

*We recommend a dialogue-based cost benchmark*

Often regulators of monopolistic sectors have determined revenue caps, price caps, benchmarking estimates and other remedies without significant involvement of the regulated entity. This is suboptimal regulation, and there is a new trend going toward what we call *dialogue-based regulation*, which minimises the information asymmetry between the regulated entity (agent) and the regulator (principal), cf. Figure 4. In addition, this type of regulation is significantly more flexible, forward- and customer-oriented than a traditional revenue cap estimated and determined by an independent regulator. Thus, we recommend that the forthcoming Reasonable Commercial Basis (RCB) guideline implements a dialogue-based cost benchmarking.

**Figure 4**  
**Dialogue-based cost benchmarking model**



Note: Illustration  
Source: Copenhagen Economics

### *How it works*

In our dialogue-based regulation model, both parties – being the regulated entity and the regulator – build a cost benchmark model to estimate the expected costs for the next three years. Both parties form their own expectations about the customer’s demand, willingness to pay, inflation, rate of return etc., which are important elements to estimate the efficient costs of providing market data to market participants. The trading venue submits their documentation and the two cost benchmarks can be compared.

If the parties reach an agreement on the cost-based prices of market data in the next regulatory period, both parties should comply with this agreement. The trading venue should price accordingly, and the regulator should monitor the compliance according to this level. In case the parties cannot reach an agreement, the parties end up in the fallback mechanism, where a hard revenue cap is implemented, based on the cost benchmark estimation of the regulators, cf. also example in Box 2.

### **Box 2 Example of a fallback mechanism**

In the Danish aviation industry, the fallback mechanism is quite loosely defined, and the agreement process works because the fallback is a harsh threat, which none of the parties (airlines negotiate with Copenhagen Airports) wants to face. In this case, the Ministry of Transport steps in and dictates the future revenue cap. The UK regulation of electricity and gas transmission and the US Federal Energy Regulation Commission also apply a negotiated settlement approach (dialogue) in the determination of future prices. If the parties cannot reach an agreement, the Federal Energy Regulation Commission settles the pricing with traditional remedies.

### *Coverage*

We suggest, that the following trading venues should be covered by this regulatory setup:

- Regulated Markets
- Multilateral Trading Facilities
- Organised Trading Facilities

However, as both Approved Publications Arrangements (APA) and Consolidated Tape Providers (CTP) are also required to provide market data on a “reasonable commercial basis”<sup>4</sup>, it should be considered how these entities fulfil this requirement. In this way, the above described cost benchmark model can be useful in pricing a consolidated tape and to identify how the trading venues should be compensated for distributing market data to the consolidated tape. Indeed, without a price or revenue cap on a CTP, it can hardly solve the issues of limited access to market data for users.

### *A three-year regulation period*

Optimally, the cost benchmark should be evaluated every three years. This will keep administrative costs at a reasonable level both for trading venues and the regulator. Furthermore, a regulation period of three years corresponds to the normal depreciation period of IT systems within the financial industry<sup>5</sup>, cf. Box 3.

We suggest that the exercise should be carried out simultaneously for all trading venues in order to have as many observations for the benchmarking as possible.

---

<sup>4</sup> MiFID II, art. 64.1, and MiFID II, art. 65, 1, Delegated Regulation 2017/565, art. 84 - 89

<sup>5</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 3

**Box 3 Pros and cons on the length of the regulatory period**

The shorter the regulatory period, the costlier the regulation will be in terms of administration, also increasing the stress that trading venues are exposed to as the model replicates competitive markets. Moreover, a shorter regulatory period increases difficulties in "planning ahead" over several years. On the other hand, errors and methodological updates to the benchmarking model is corrected earlier.

**1.3 WHAT IS POSSIBLE WITHIN THE CURRENT REGULATION?**

The current MiFID II/MiFIR provides the regulator with the legal right to collect data as well as to analyse the underlying costs of generating market data. Moreover, the regulator has the legal right to engage in information and data exchanges with the trading venues as well as to engage in dialogue on the underlying costs of generating market data. This means that regulators already today have the right to collect data, analyse costs and develop cost benchmarking to estimate trading venues' efficient cost levels as well as spreading best practise.

We also consider it perfectly consistent with the current regulation that trading venues and the regulator agree on actual costs via a benchmarking approach, but the regulator do not have legal means to implement a given revenue cap in the absence of the trading venues' consent.

**Table 1**  
**Which parts of revenue cap regulation can be implemented under MiFID II**

<b>ELEMENTS</b>	<b>ALREADY POSSIBLE DUE TO MIFID LEGISLATION</b>
Defining RCB and costs of "producing and disseminating market data"	Yes
Data collection	Yes
Cost benchmarking across trading venues	Yes
Compare venues' estimates of their cost of producing market data against own data and benchmarks	Yes
Engaging in dialogue with trading venues on pricing of market data	Yes
Asking for changes of prices to comply with RCB	Yes
Agreement on a three-year revenue cap between regulator and trading venue	Yes
Implementing a fallback mechanism if an agreement is not reached	No

Note: Illustration  
Source: Copenhagen Economics

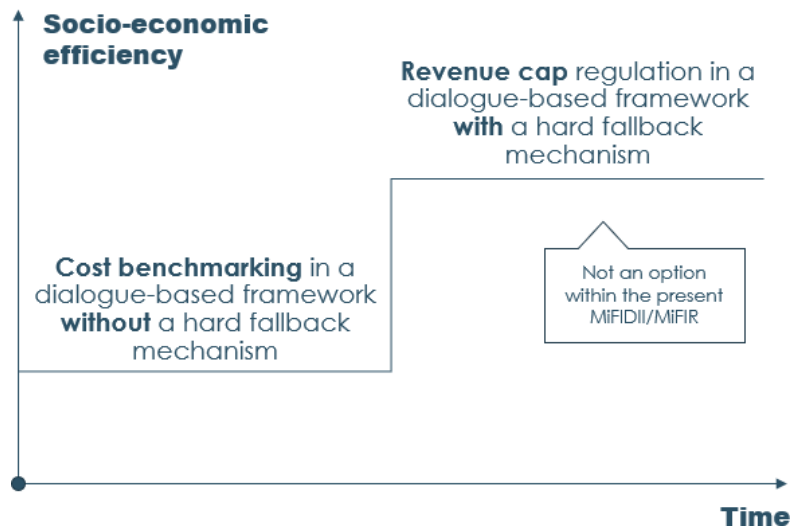
This implies that step 1-4 (which will be outlined in next chapter) are within the current regulation, but not step 5.

*An implementation in two steps.*

As such, we can split our proposed regulatory model into two sequential phases, cf. Figure 5:

1. **Step 1-4:** Currently, it is possible to implement a **second-best regulatory model** – a dialogue-based cost benchmark model without any strict revenue cap, but with a reasonable obligation for regulators to sanction trading venues if prices for market data are not set according to existing regulation.
2. **Step 5:** Once the regulation has been amended to allow for a dialogue-based framework with a tough fallback, providing a **first-best regulatory model**. Moreover, in this regulation the regulator can adjust future revenue caps if the trading venue does not comply with the agreement.

**Figure 5**  
**The regulatory progression model**



Note: Illustration  
Source: Copenhagen Economics



## CHAPTER 2

**OPERATIONAL GUIDELINE ON HOW TO  
IMPLEMENT COST BENCHMARKING**

In this chapter, we describe the regulatory process and how to implement a dialogue-based cost benchmarking in five steps. The essence of this guideline is to demonstrate how to build and implement a framework, which implies reasonably priced market data based on the actual costs of generating and distributing market data.

To do so, it is crucial that the regulator has the right tools to measure the trading venue's idiosyncratic costs of providing market data to security dealers. In this context, the IEX-report *The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange* forms a good starting point with respect to the costs and information components to be collected by the regulator from trading venues in order to set up cost benchmarking.

In addition, several trading venues – e.g. Deutsche Börse – have in the past been quite detailed in their annual reports on how employees were allocated to different activities. This could also give some indications of the cost structures. At the same time, this also underlines the fact that trading venues are fully capable of allocating their costs to the production of the different

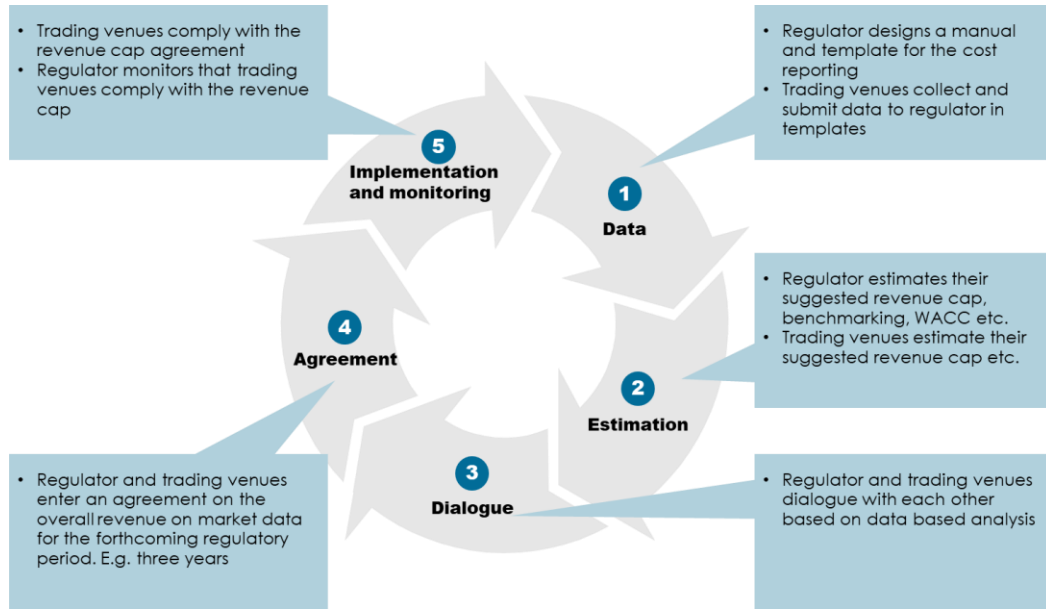
The rest of the chapter is structured as follows. In *section 2.1*, we present the five main steps in the regulatory implementation process. Moreover, a thorough description of the expectations for the regulator and the trading venues in each step. In *section 2.2*, we describe how the regulation of the five-step model can be initiated. In *section 2.3*, we present a possible implementation timeline.

**2.1 A FIVE STEP REGULATORY IMPLEMENTATION  
PROCESS**

To determine the true costs via a cost benchmarking model for a regulatory period, the following five steps must be taken, cf. also Figure 6:

1. Data collection
2. Design of cost benchmarking model
3. Dialogue/negotiation
4. Agreement
5. Implementation and monitoring

**Figure 6**  
**Regulatory implementation process**



Note: Illustration  
Source: Copenhagen Economics

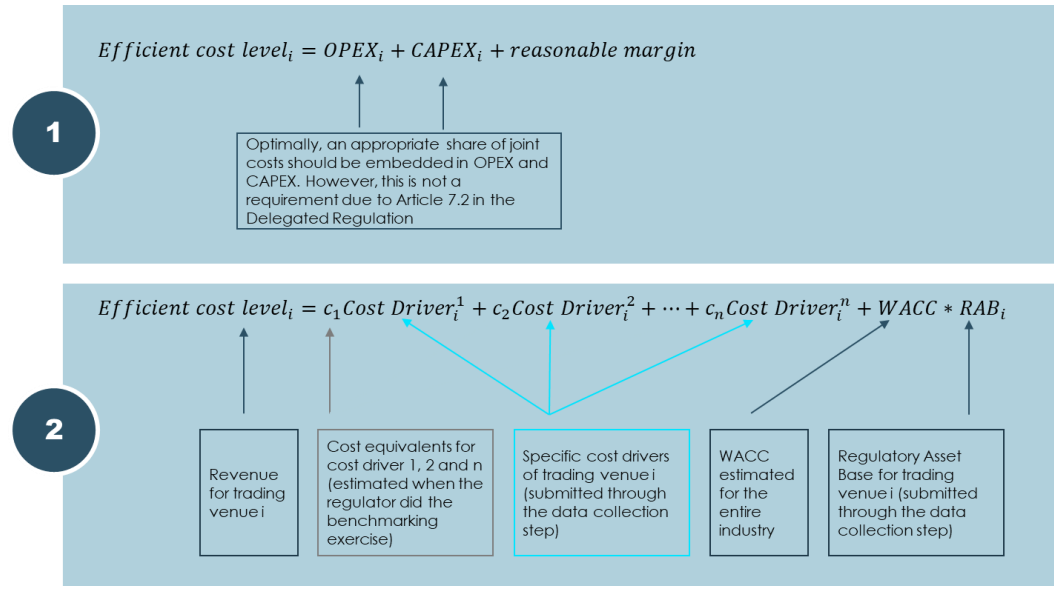
While the cost benchmark should reflect the costs of an efficient trading venue, the underlying cost structures of venues may differ, e.g., due to differences in the number of consumers. If these cost differences between venues are large, this would require the benchmarking model to be calibrated to the different venues and maintained over time.<sup>6</sup>

From a helicopter view, the regulator and the trading venue should negotiate and agree on what they believe is the efficient cost level for producing and disseminating market data for the next three years, cf. Figure 7. This is expressed in equation 1, which states that the efficient cost level is a sum of the efficient OPEX, CAPEX and a reasonable margin in each trading venue. As the regulator cannot observe the efficient cost levels in each trading venue, the regulator needs to estimate this cost level, which is the purpose of equation 2.

The five steps in the implementation process is a description of how to estimate the parameters in equation 2 in order to avoid trading venues ‘gaming’ the regulator due to opportunistic behaviour. This means that cost driver 1 to n and parameter value 1 to n in equation 2 should be designed to replicate the efficient level of OPEX and CAPEX in equation 1. Moreover, the Weighted Average Cost of Capital and the Regulatory Asset Base in equation 2 should replicate the reasonable margin in equation 1.

<sup>6</sup> Copenhagen Economics (2014), How to ensure reasonable prices of market data, page 25

**Figure 7**  
**Efficient cost level equation**



Note: Illustration  
Source: Copenhagen Economics

### 2.1.1 Step 1: Data collection

The regulator initiates the regulatory process by collecting relevant data to estimate the trading venue's cost function and thereby identify the true costs of producing and disseminating market data for the individual trading venues. In this context, the regulator must have a thorough understanding of what drives costs (OPEX and CAPEX)<sup>7</sup> and therefore needs an in-depth understanding of the physical infrastructure that comprises the core trading system as well as the operational tasks. Moreover, the regulator needs to understand the market data process to be able to guide trading venues and ensure that they submit the correct and true costs. It is crucial that the cost benchmarking ensures equal treatment of each trading venue, implying that the benchmarking must be conducted on the same cost types defined in the same way. The benchmarking must also take account of scale, scope and idiosyncratic conditions, which are specific to each trading venue.

The level of maintenance and recalibration depends on the underlying cost volatility, which tends to differ significantly across sectors. In general, CAPEX tends to be less volatile, especially in industries with a relatively stable production apparatus. OPEX on the other hand varies more as it depends on the produced amount of goods and/or number of customers. As OPEX changes over time, the cost benchmark should be recalibrated to reflect these changes.

The production and distribution of market data is characterised by a relatively stable production apparatus and – more importantly – quite invariable OPEX. This reason is that the cost of disseminating market data does not depend significantly on the number of consumers who buy the data. Once the production and distribution setup are in place, serving the marginal consumer is not associated

<sup>7</sup> OPEX is the operating expenditure necessary to distribute market data. CAPEX is capital expenditure.

with many additional costs. This service provision is in stark contrast to the supply of standard manufactured goods or many other services, where the cost of input is strongly correlated with the amount of output: e.g. you need to buy more coal to produce more energy.<sup>8</sup>

Overall, the regulator needs to collect the following four types of data:

- Actual costs of generating and providing market data including costs for software development, maintenance, patching etc.
- Underlying drivers of the cost of producing and disseminating market data, e.g. Market Data Process Servers and Market Data Feeds Switches.
- Underlying variables for the estimation of correlations between costs and cost drivers.
- Output measures, e.g. the type of market data which the trading venue delivers as well as the amount.

For the data collection, the regulator must design a written manual<sup>9</sup> on which costs, cost drivers and underlying variables trading venues have to submit about their physical infrastructure and activities.

The process flow of producing and disseminating market data is a complex system. Below we provide a simplified overview divided into three main parts, cf. also Figure 8:

- Part I – physical connectivity: This part connects to the members of the exchange. In the first place, it allows the exchange’s members to submit a new order (bid/ask) to the exchange, also known as a user message. See Part I on the left side of Figure 8.
- Part II – logical connectivity: The orders are then passed to this part, which is used to validate and transform each user message into IEX’s internal message protocol. See Part II in Figure 8. These internal messages are then passed on to the core trading system for trade execution.
- Part III – producing and transmitting market data: After trade execution, part III is used to produce the market data in accordance with the specified format, which then can be send back to market data users. See Part III in Figure 8.
- Part I: Part I is then used again to distribute the market data back to the subscriber. See Part I on the right side of Figure 8.

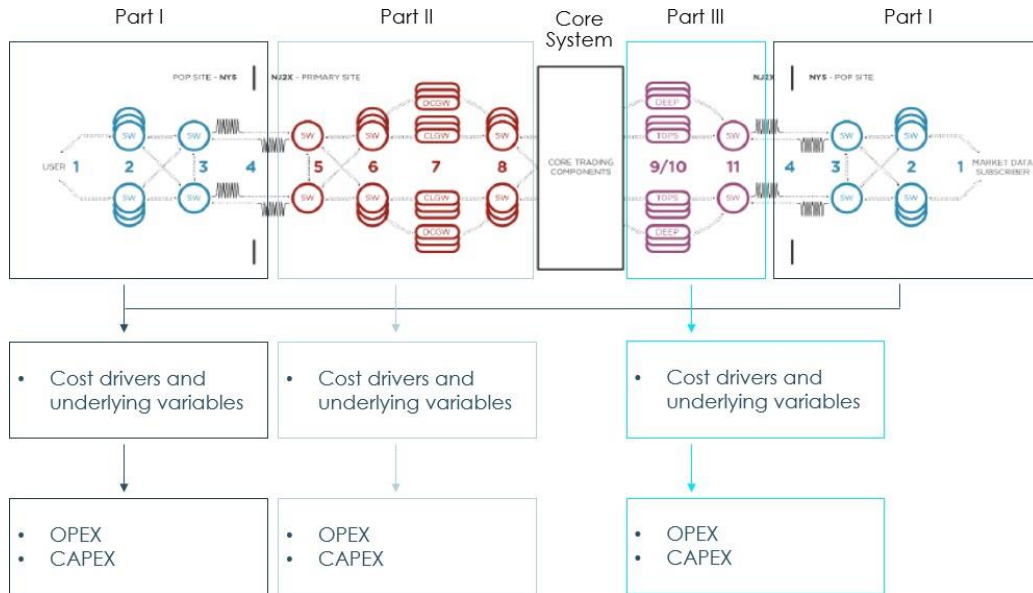
In the subsequent sections, we describe which costs, cost drivers and underlying variables trading venues must submit within each part, for the regulator to be able to design a benchmarking model.

---

<sup>8</sup> Copenhagen Economics (2014), How to ensure reasonable prices of market data, page 25

<sup>9</sup> Besides the manual, the regulator must develop a standardised template such that every trading venue submits their specific data in the same way, which is important for the benchmarking exercise. The manual must also describe the documentation that the trading venue must provide to the regulator.

**Figure 8**  
**The physical infrastructure that comprises the Core Trading System of the IEX**



Note: Illustration

Source: IEX (2019): The Cost of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 3, Figure 2

### *Part I – physical connectivity*

In this section, we describe the costs, cost drivers and underlying variables associated to physical connectivity. The physical connectivity is provided by a trading venue via a network switch and cabling infrastructure that allows participants to ‘plug into’ an exchange’s system for market data receipt, order entry, execution receipt and other messaging.<sup>10</sup> In this context, IEX has analysed their costs associated with physical connectivity, cf. Box 4.

<sup>10</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 21

#### **Box 4 IEX's costs of physical connectivity and associated personnel**

##### **Costs for physical connectivity used in Phase I**

- Cross Connect via Patch Panel: the physical point at which a participant plugs its cable into the IEX system
- Access Layer Switches (the 'leaf layer'): each connection at the patch panel from a participant has a corresponding connection at this initial array of switches
- Distribution Switches (the 'spine layer'): connections from access layer switches are aggregated at distribution switches
- Equidistant Cabling and POP Networking Equipment: IEX uses a fibre spool, optics equipment, and fibre lines between two data centres to implement the 350µs of designed latency experienced by each message to/from participants
- ITF System: specific internet-facing switches and firewalls to allow access for FIX certification via the Internet
- Space, Power and Security: physical space, electrical power and security at the data centres considered for the physical assets in scope
- Monitoring: servers, switches and software licences used to monitor the physical assets in scope as well as the health of the connectivity service provided by such assets

##### **Costs for associated personnel**

- Product Management: resources responsible for managing the exchange offering
- Infrastructure: resources responsible for installing and maintaining the physical infrastructure
- Operations: resources responsible for supporting participants and maintaining the IEX system
- Legal and Regulatory: resources responsible for the exchange's compliance with applicable laws and regulations
- Information Security: resources responsible for maintaining and monitoring information security

Source: IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 22

To estimate the true cost function and how efficient different trading venues are, the regulator needs cost driver data and underlying variable data that generate the above-mentioned costs, cf. Box 5.

#### **Box 5 Cost drivers and underlying variables**

Some of the key cost drivers on which the regulator wants to collect data are listed below. However, there might be other relevant cost drivers:

- Cost driver 1: Number and capacity of the Cross Connect via a Patch Panel
- Cost driver 2: Number and capacity of Access Layer Switches
- Cost driver 3: Number and capacity of Distribution Switches
- Cost driver 4: Number and capacity of the Equidistant Cabling and POP Networking Equipment

Source: Copenhagen Economics based on IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 22, Figure 5

*Part II – logical connectivity*

In this section, we describe the costs, cost drivers and underlying variables associated to logical connectivity for order entry. The logical connectivity for order entry is provided via the network switch and cabling infrastructure that deliver orders, execution and drop copy messages, as well as server infrastructure that runs software processes responsible for validating and formatting such messages for either internal or external use.<sup>11</sup> In this context, IEX has analysed their costs associated with logical connectivity, cf. Box 6.

**Box 6 IEX's costs of logical connectivity and associated personnel****Costs for logical connectivity used in Part II**

- Order Entry Distribution Switches: data from the distribution switches at the POP traverses the POP networking equipment to the order entry distribution switches in the primary data centre
- Order Entry Access Layer Switches: data from the distribution switches is distributed into a specific top-of-rack access layer switch
- Client Gateway & Drop Copy Servers: the servers running either a client gateway or drop copy process that either receive orders or send execution/drop copy messages
- System Access Layer Switches: data validated and formatted by a client gateway process is distributed via an initial array of top-of-rack switches to the IEX system
- ITF System: hardware and software to facilitate FIX certification and order entry testing
- Space, Power and Security: physical space, electrical power and security at the data centres considered for the physical assets in scope
- Administrative Access: command and control infrastructure used by operations teams to administer the exchange services
- Monitoring: servers, switches and software licences used to monitor the physical assets in scope as well as the health of the connectivity service provided by such assets

**Costs for associated personnel**

- Product Management: resources responsible for managing the exchange offering
- Legal and Regulatory: resources responsible for selling of market data
- Information Security: resources responsible for maintaining and monitoring information security
- Development and Quality Assurance: resources responsible for building and testing new exchange functionality
- Operations: resources responsible for supporting participants and maintaining the IEX system
- Infrastructure: resources responsible for installing and maintaining the physical infrastructure
- ITF Operations: resources responsible for supporting participants and maintaining the IEX Testing Facility

Source: IEX (2019): The Cost of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, pages 28-29

To estimate the true cost function and how efficient different trading venues are, the regulator needs cost driver data and underlying variable data that generate the above-mentioned costs, cf. Box 7.

<sup>11</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 28



**Box 7 Cost drivers and underlying variables**

Some of the key cost drivers on which the regulator wants to collect data are listed below. However, there might be other relevant cost drivers:

- Cost driver 1: Number and capacity of Order Entry Distribution Switches
- Cost driver 2: Number and capacity of Order Entry Access Layer Switches
- Cost driver 3: Number and capacity of Client Gateway & Drop Copy Process Servers
- Cost driver 4: Number and capacity of System Access Layer Switches

Source: Copenhagen Economics based on IEX (2019): The Cost of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 29, Figure 7

**Part III –producing and transmitting market data**

In this section, we describe the costs, cost drivers and underlying variables associated to producing and transmitting market data that must be submitted by trading venues, cf. Box 8.

**Box 8 IEX's costs of physical assets and associated personnel****Costs for physical assets used in Part III**

- Market Data Servers: the server running a market data process that either publishes data or responds to retransmission requests
- Market Data Feeds Switches: data formatted by a market data process is distributed to participants via an initial array of top-of-rack access layer switches
- ITF Infrastructure: hardware and software used to facilitate market data testing
- Space, Power and Security: physical space, electrical power and security at the data centres considered for the physical assets in scope
- Administrative Access: command and control infrastructure for operations teams to administer the exchange services
- Monitoring: servers, switches and software licences used to monitor the physical assets in scope as well as the resiliency of the market data product provided by such assets

**Costs for associated personnel**

- Product Management: resources responsible for managing the exchange offering
- Legal and Regulatory: resources responsible for the exchange's compliance with applicable laws and regulations
- Information Security: resources responsible for maintaining and monitoring information security
- Development and Quality Assurance: resources responsible for building and testing new exchange functionality
- Operations: resources responsible for supporting participants and maintaining the IEX system
- Infrastructure: resources responsible for installing and maintaining the physical infrastructure
- ITF Operations: resources responsible for supporting participants and maintaining the IEX Testing Facility

Source: IEX (2019): The Cost of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, pages 15 and 16

For benchmarking purposes, for instance, it is important to standardise CAPEX across the entities which are compared. In that context, IEX applies a depreciation rate of three years on its physical assets, which might be different in other companies. This means that the regulator might adjust for differences in depreciation rates across exchanges before conducting the benchmarking analysis.

The companies should submit personnel costs in the same way. As an example, IEX has analysed the activities of their personnel and identified that they use a full-time equivalent of 4.05 for market data generation relative to 33 employees in total.<sup>12</sup> Overall, IEX estimates their annual cost per Data Recipient to be \$11,943 associated with the costs in this part.<sup>13</sup>

To estimate the true cost function and how efficient different trading venues are, the regulator needs cost driver data and underlying variable data that generate the above-mentioned costs, cf. Box 9.

#### **Box 9 Cost drivers and underlying variables**

Some of the key cost drivers on which the regulator wants to collect data are listed below. However, there might be other relevant cost drivers:

- Cost driver 1: Number and capacity of Market Data Process Servers
- Cost driver 2: Number and capacity of Market Data Feeds Switches

Source: Copenhagen Economics based on IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 16, Figure 3

### **2.1.2 Step 2: Cost benchmarking**

In this step, the regulator tests, cleans and structures the collected data (costs and cost drivers) to estimate the true cost function of producing and disseminating market data. We recommend a benchmarking process of four steps, cf. Box 10. For a deep dive into the benchmarking analysis, see Appendix A.

---

<sup>12</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 17

<sup>13</sup> IEX (2019): The Cost Of Exchange Services: Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange, page 18, Figure 4

**Box 10 Benchmarking**

Typically, cost benchmarking consists of the following steps after the data has been collected, tested and verified:

1. Use descriptive statistics to understand what drives costs
2. Specify the cost drivers and the overall model:
  - Define inputs, outputs, exogenous variables
  - Use econometric analysis, expert information, economic theory
3. Estimate the cost function and efficient frontier:
  - Use methods such as Data Envelopment Analysis, Stochastic Frontier Analysis, Best-of-More Approach
4. Do second-stage analysis:
  - Do statistical analysis, sensitivity analysis, robustness analysis
  - Make sure all relevant variables are included and follow up on omitted variables

Source: Copenhagen Economics

Based on the descriptive and econometric analysis, the regulator defines a set of cost drivers, which should be used for the benchmarking analysis. This could be done through a multi-dimensional benchmarking where every cost driver and cost measure are reflected on individual dimensions. In addition, one could also add cost drivers together and design a smaller set of aggregated cost driver categories, e.g. a benchmarking model with five dimensions, where the output measures are included as a fifth dimension, cf. Table 2.

**Table 2**  
**Dimensions in the benchmarking model**

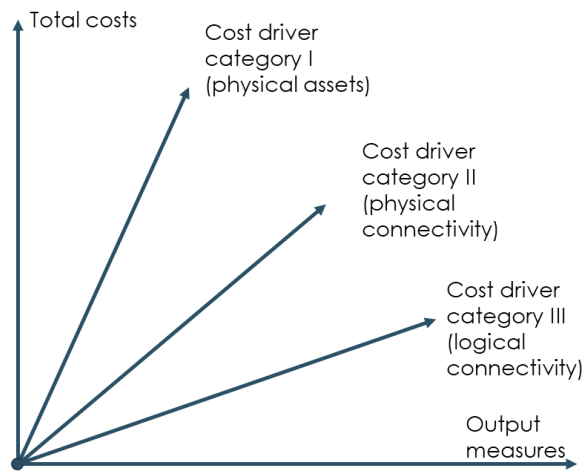
<b>DIMENSION</b>	<b>COST DRIVERS</b>
<b>Total costs</b>	OPEX and CAPEX on physical assets
<b>Cost driver category I</b>	OPEX and CAPEX on physical connectivity
	OPEX and CAPEX on logical connectivity
<b>Cost driver category II</b>	Number and capacity of Market Data Process Servers
	Number and capacity of Market Data Feeds Switches
<b>Cost driver category III</b>	Number and capacity of the Cross Connect via a Patch Panel
	Number and capacity of Access Layer Switches
	Number and capacity of Distribution Switches
	Number and capacity of the Equidistant Cabling and POP Networking Equipment
<b>Outputs</b>	Number and capacity of Order Entry Distribution Switches
	Number and capacity of Order Entry Access Layer Switches
	Number and capacity of Client Gateway & Drop Copy Process Servers
	Number and capacity of System Access Layer Switches
	Depth of Book
	Top of Book
	Last Sale
	Auction Imbalance

Note: One might have to convert the individual cost drivers into the same measure, e.g. a cost measure based on cost equivalents.

Source: Copenhagen Economics

Based on the five dimensions, the regulator can estimate the efficient frontier and relative performance of the trading venues, cf. Figure 9.

**Figure 9**  
**Multi-dimensional benchmarking model**



Note: One can use different methodologies, e.g. DEA, SFA, COLS or other simpler econometric methodologies.  
Source: Copenhagen Economics

With the benchmarking analysis, efficient frontier analysis and cost equivalents per cost driver, the regulator is equipped to estimate the efficient cost level for each specific trading venue based on their specific cost driver information and the benchmarking model's estimated parameter values for the next three years.

In this context, it is important to state that the above-described benchmarking is based on the current situation and historical data, which implies that the estimation is not forward-looking. Therefore, the regulator must do other analysis to form their expectations on how the market for market data will develop over the coming regulatory period.

Concretely, the regulator must form its expectations on every component of the suggested revenue cap for the next regulatory period before initiating the negotiation with the trading venue, cf. Table 3.

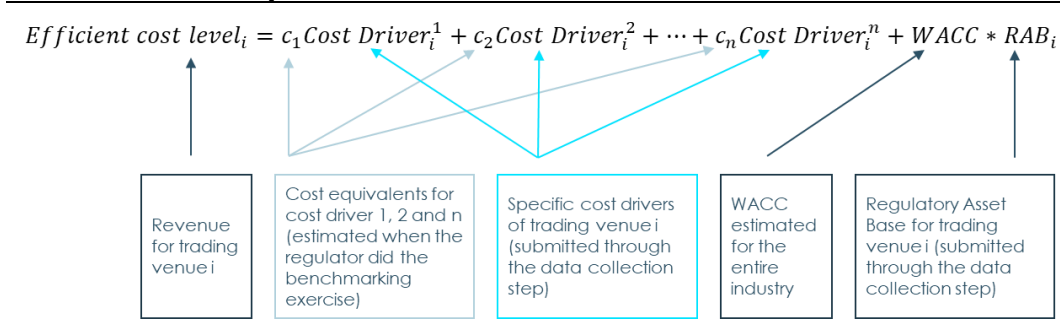
**Table 3**  
**Revenue is the sum of each cost driver's cost contribution plus WACC**

COST DRIVER	COST EQUIVALENT	EXPECTED COSTS
Number and capacity of Market Data Process Servers	1,000	2,000
Number and capacity of Market Data Feeds Switches	1,000	2,000
Number and capacity of the Cross Connect via a Patch Panel	1,000	2,000
Number and capacity of Access Layer Switches	1,000	2,000
Number and capacity of Distribution Switches	1,000	2,000
Number and capacity of the Equidistant Cabling and POP Networking Equipment	1,000	2,000
Number and capacity of Order Entry Distribution Switches	1,000	2,000
Number and capacity of Order Entry Access Layer Switches	1,000	2,000
Number and capacity of Client Gateway & Drop Copy Process Servers	1,000	2,000
Number and capacity of System Access Layer Switches	1,000	2,000
Rate of return (WACC x RAB)		2,000
<b>TOTAL</b>		<b>22,000</b>

Note: WACC: Weighted Average Cost of Capital. RAB: Regulatory Asset Base. The numbers are fictive.  
Source: Copenhagen Economics

Based on each individual trading venue's specific cost driver data, i.e. number and size of Market Data Process Servers, number and size of Market Data Feeds Switches etc. as well as the general cost equivalents (parameter values) – which were estimated in previous benchmarking analysis – and the efficient frontier, the regulator can estimate the efficient cost level of each specific trading venue, cf. Figure 10.

**Figure 10**  
**Efficient cost level equation**



Note: Illustration  
Source: Copenhagen Economics

Just like the regulator, the individual trading venues will prepare for the forthcoming negotiations. They will also form their own expectations on costs, expected activities and reasonable rate of return

relative to risk exposure. Most likely, they will not conduct a benchmarking exercise, as they are perfectly informed about their own costs. The information asymmetry between the trading venue and the regulator is one of the reasons why the regulators apply state-of-the-art benchmarking models.

### 2.1.3 Steps 3 & 4: Dialogue and Agreement

The regulator and the individual trading venue meet to negotiate on their individual perceptions of the efficient cost level for the next regulatory period. They will of course discuss the overall level as well as the underlying OPEX, CAPEX and rate of return elements that add up to total costs. Moreover, the parties will discuss expected developments in underlying parameters.

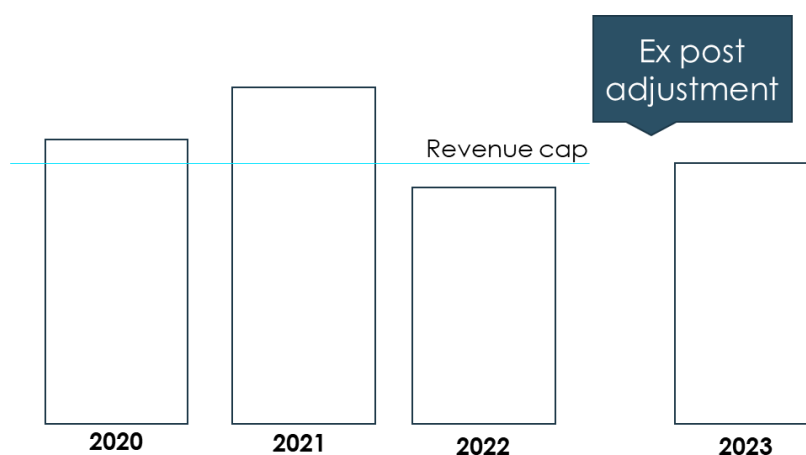
Based on the discussions, they will prepare a gross list of parameter values, assumptions and other relevant variables they disagree on. Moreover, they test each other's applied values and assumptions in their respective models and return to the negotiations with new arguments, tests and ultimately a new suggested cost level. This process continues until an agreement is reached and both parties are satisfied.

### 2.1.4 Step 5: Implementation

When an agreement has been reached, the trading venue must price their market data in line with the agreed efficient cost level for the next regulatory period. Additionally, the regulator will monitor the trading venue's compliance and engage in a dialogue with them if they do not comply.

The first-best solution is a binding agreement such as a revenue cap for the entire regulatory period. However, this would require significant changes to the existing regulatory framework, cf. MiFID II/MiFIR. In such a setup, a trading venue is only allowed to generate an overall revenue which complies with the above cost equation. We suggest that the trading venue may deviate from the yearly revenue cap as long as the overall cap over the three-year period is complied with, cf. Figure 11.

**Figure 11**  
**Revenue cap in a three-year regulatory period**

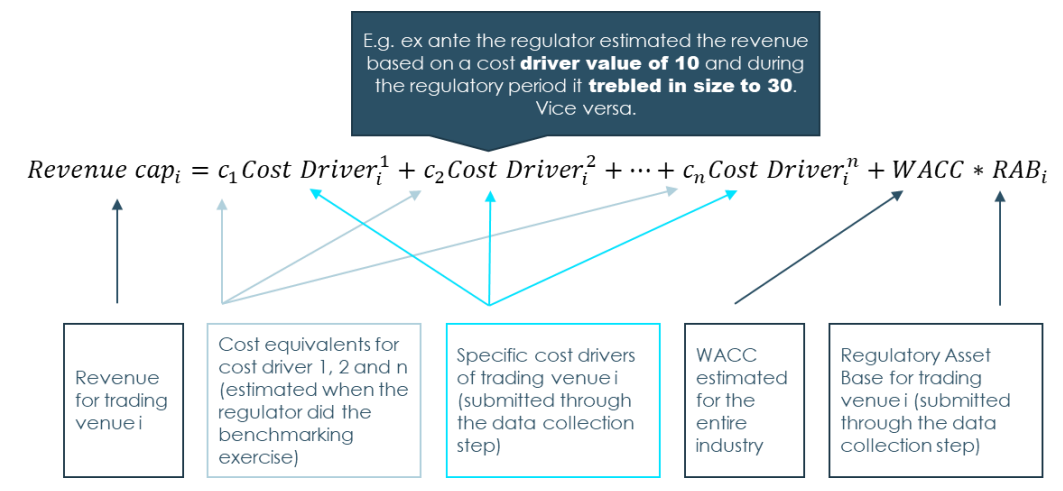


Note: Illustration  
Source: Copenhagen Economics



Coincident with the trading venue's implementation of the revenue cap, the regulator monitors the conduct of the trading venue. At the end of the regulatory period, the regulator will adjust the cost benchmarking and ideally the revenue cap based on any positive or negative discrepancy between actual revenue and the revenue cap. In this ex post adjustment, the regulator will take significant changes in the specific trading venue's cost driver composition, size, market development etc. into account before establishing a new benchmark and revenue cap for the next regulatory period, cf. Figure 12.

**Figure 12**  
**Ex post adjustment of the revenue equation**



Note: Illustration  
Source: Copenhagen Economics

## 2.2 HOW TO INITIATE THE FIRST COST BENCHMARK

Setting the first cost benchmark is more time-consuming than subsequent cost benchmarks, as the regulator must do several iterations before the correct cost level and benchmarking methodology is ascertained, cf. Box 11. Trading venues might have expected a stricter monopoly regulation to be implemented, which implies a risk of upward-biased costs, as rational agents will try to boost costs. Therefore, the regulator could test the cost level based on a 'neutral' period, meaning a period where trading venues do not expect cost benchmarking to come in place.

**Box 11 Steps to estimate the very first cost benchmark**

Concretely, we recommend the following steps when estimating the first benchmark:

- Involve the trading venue through workshops, bilateral meetings and other means to signal openness and responsiveness. It is crucial to start on the same page to get trading venue input on costs structures.
- Use workshops to get trading venue input on the theoretically correct cost function, cost drivers and other relevant variables before data collection and estimations.
- Collect data based on workshop input, own analysis and expert information. Use simple, user-friendly templates and manuals on how to submit data.
- Test, clean and structure the collected data and estimate the underlying cost function.

Source: Copenhagen Economics

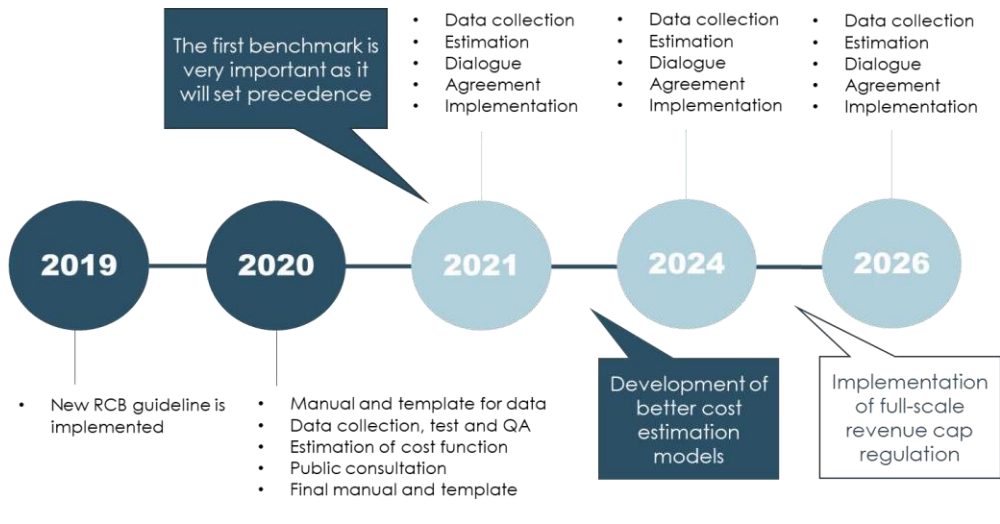
Overall, we recommend that the regulator publish their analysis continuously as well as letting the trading venues comment through public consultations when the first analysis on cost drivers, cost functions etc. is published. When the regulator is quite certain about the cost function, the regulator and trading venue start the formal negotiations with the purpose of reaching an agreement on the costs, which should form the foundation for market data pricing for subsequent years. This very first step is fully consistent with the current regulatory framework.

## **2.3 TIMELINE FOR THE FIRST COST BENCHMARK**

In this section, we discuss what it takes to implement a cost benchmarking framework which can be achieved within the next 18 months such that the new framework is in place on 1 January 2021.

In Figure 13 below, we provide an overview of the steps necessary to implement a cost benchmarking framework as the first regulatory step in this industry, which over time can develop into a full-scale revenue cap regulation after the MiFID II/MiFIR has been adjusted.

**Figure 13**  
**Implementation of a new regulatory framework**



Note: Illustration  
Source: Copenhagen Economics

Assuming the new RCB guideline is finalised in 2019, the regulator should be able to implement the new regulatory framework at the beginning of 2020. Concretely, the regulator will start collecting data, set up workshops and undertake several analyses throughout the first half of 2020, so the regulator and the trading venue can start the first negotiations in the autumn of 2020.

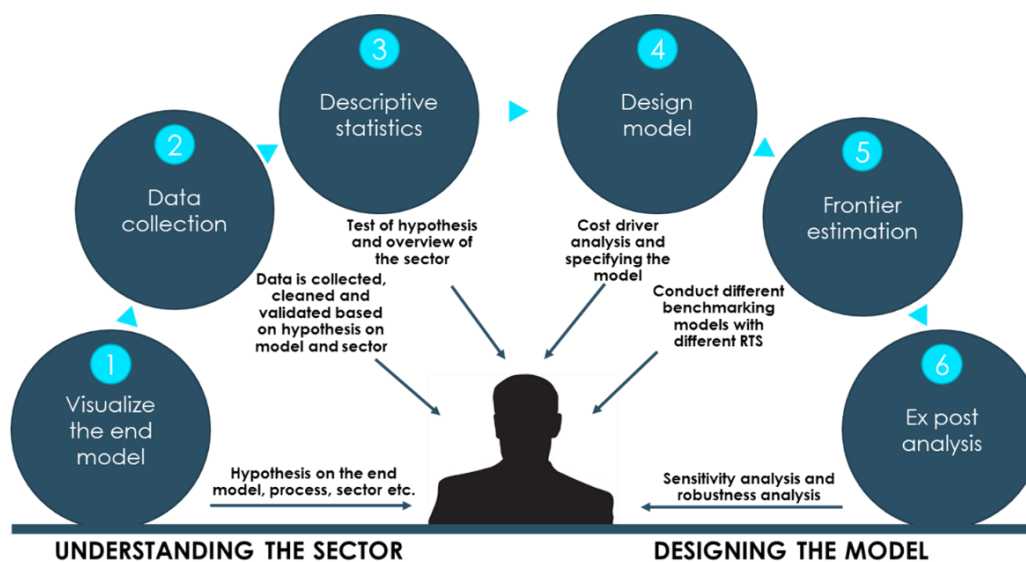
## APPENDIX A

# DEEP DIVE INTO THE SPECIFIC BENCHMARKING ANALYSIS

This appendix takes a deep dive into the four benchmarking steps described in chapter 2 as well as two steps of the data collection part, cf. Figure A.1:

1. Visualise the 'end model'
  - a. Minimise 'iterative feedback loops'
2. Data collection, cleaning and validation
  - a. Workshops with the trading venues
  - b. Standardisation of variables (also standard for the future)
  - c. Accounting standards, cost allocations, definitions of costs etc.
  - d. Use of manuals
3. Use descriptive statistics to understand what drives costs
4. Specify the cost drivers and the overall model
  - a. Define inputs, outputs, exogenous variables
  - b. Use econometric analysis, expert information, economic theory
5. Estimate the cost function and efficient frontier
  - a. Use methods such as Data Envelopment Analysis, Stochastic Frontier Analysis, Best-of-More Approach
6. Do second-stage analysis
  - a. Do statistical analysis, sensitivity analysis, robustness analysis
  - b. Make sure all relevant variables are included and follow up on omitted variables

**Figure A.1**  
**Benchmarking**

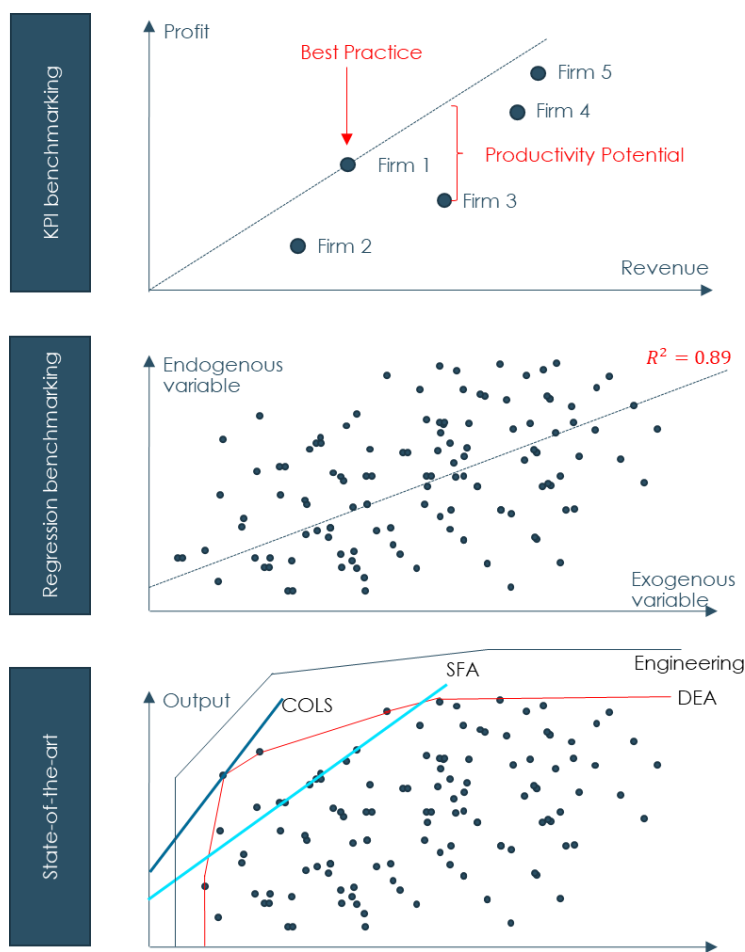


Note: FTS: Return to Scale  
Source: Copenhagen Economics

## STEP 1: VISUALISE ‘THE END MODEL’

It is crucial to design a hypothesis about the correct benchmarking model before data collection is initiated. In this context, it is also an advantage to design a hypothesis on how the benchmarking results should be presented in the end. And finally, one should design figures to get an overview and thorough understanding of the underlying structures of the trading venues to be benchmarked as well as the underlying cost structures, activities and external variables. Therefore, we recommend that the regulator forms a hypothesis on what kind of benchmarking methodology they expect to apply, cf. Figure A.2:

**Figure A.2**  
**Different benchmarking methodologies**



Note: Illustration  
Source: Copenhagen Economics

## STEP 2: DATA COLLECTION, CLEANING AND VALIDATION

The data should be collected having the ‘end model’ in mind. Overall, we recommend that the regulator visits different trading venues, arranges workshops and engages in expert meetings to get a thorough understanding of the underlying structures. Based on these interviews, the regulator creates a list of cost drivers and the costs to be analysed.

Once the regulator knows which data to collect, the regulator should:

- Use publicly available data (income distribution, customer density, geography etc.)
- Use data from other relevant regulatory authorities
- Collect data from the sector
- Prepare an Excel sheet for entering data
  - It must be easy to apply
- Design a manual to standardise and align data with the Excel sheet
  - It must be simple, concise and easy to understand to increase data quality
  - It must have clear definitions of OPEX, CAPEX, cost allocations, cost allocations on cost drivers as well as cost drivers
  - It should be able to guide trading venues on how to make a qualified estimate of costs and cost drivers in case the trading venue is lacking information or data
- It can be necessary to do a roadshow on 'How to use the Excel sheet'
- Create easy access for the trading venues to get support and submit data
  - Confidentiality is often important to the client

After the collection of raw data, the regulator needs to evaluate the data by performing the following steps:

- Clean data of errors
- Clean data of misunderstanding categories
- Clean data of misunderstanding applied units
- Test data using means, standard deviations, regression analysis, outlier analysis, confidential intervals, figures, filters, peer groups analysis etc.

### **STEP 3: USE DESCRIPTIVE TOOLS TO UNDERSTAND WHAT DRIVES COSTS**

The overall objective in this step is for the regulator to get a thorough understanding of the overall industry as well as the benchmarked trading venues. Therefore, the regulator needs to apply descriptive statistics and illustrations to get a complete overview. In this context, we suggest that the regulator applies different distributions, KPIs, ratios etc., which are easy to understand and get an overview from. It is quite normal that, after this step, the regulator realises that more data on some crucial dimensions needs to be collected.

### **STEP 4: SPECIFY THE COST DRIVERS AND THE OVERALL MODE**

The objective of this step is to identify the cost drivers that the regulator wants to implement in the benchmarking model and that explain the individual trading company's business model as closely as possible. Normally, the regulator wants to go through the following steps:

- The regulator wants to apply the initial hypothesis on the correct benchmarking model design
  - Identify relevant cost drivers using descriptive statistics from step 3 as well as KPIs
- First round of regression analysis
  - Look for signs, significance and functional form
  - Estimate correlations between costs and cost drivers as well as between cost drivers. Also identify a subset of cost drivers that the regulator will continue with

- Second round of regression analysis. Identify the optimal cost driver combination
- Third round of regression analysis based on final subset of cost drivers. Examine significance and fit. The regulator wants to make sure not to leave out important variables or over-specify the model.

## **STEP 5: ESTIMATE THE COST FUNCTION AND EFFICIENT FRONTIER**

In this step, the regulator will identify best practice, estimate the efficient frontier and the individual efficiency potential of each trading venue. Concretely, the regulator will have to go through the following steps and calculations:

- Decide on which models to apply. DEA, SFA or e.g. a Best-of-More approach.
- Identify outliers by using Banker's super efficiency criterion, Banker's F test and Cook's D.
- Decide on which observations and return to scale to apply by using scatter plots, tables and figures to get an overview.
- Estimate different models with different functional forms, e.g. DEA with different assumptions about return to scale (fdh, vrs, irs, crs) and adjustments for outliers, SFA with different function forms (linear, log linear) and adjustments for outliers. The regulator should also decide whether they will apply a Best-of-More approach and possibly a different cut-off level for the efficiency potential estimations.
- Estimate and present efficiency scores by using figures, plots, tables etc. to obtain an overview as well as a reality check of the results.

## **STEP 6: DO SECOND-STAGE ANALYSIS**

In the final benchmarking step, the regulator will test the benchmarking model and results with second-stage analysis. The purpose is to identify omitted variables and check if the regulator has reasons to include some of the previously excluded cost drivers.

The regulator will in this second-stage analysis estimate different models with some of the cost drivers which are most likely to be an omitted variable. Additionally, the regulator must compare results to get a thorough understanding of how robust the overall model is and how sensitive the specific estimates are to changes in the model.

## REFERENCES

- Copenhagen Economics (2013), Regulating access to and pricing of equity market data
- Copenhagen Economics (2014), How to ensure reasonable prices of market data
- Copenhagen Economics (2018), Pricing of market data
- European Commission (2004), Directive 2004/39/EC of 21 April 2004 on markets in financial instruments, Article 45. See also Article 44.
- European Commission (2014), Directive 2014/65/EU of 15 May 2014 on markets in financial instruments
- IEX (2019), The Cost Of Exchange Services – Disclosing the Cost of Offering Market Data and Connectivity as a National Securities Exchange
- Oxera (2019), The design of equity trading markets in Europe
- Perloff, Jeffrey M. (2016), Microeconomics – seventh edition
- Regulation (EU) No 600/2014 of 15 May 2014 on markets in financial instruments