UNIVERSITÀ
DI SIENA
1240

DIPARTIMENTO DI ECONOMIA POLITICA E STATISTICA
SCUOLA DI ECONOMIA E MANAGEMENT

CORSO DI LAUREA IN ECONOMIA PER L'AMBIENTE E LA SOSTENIBILITA

CONSUMERS' ATTITUDES TOWARDS SUSTAINABILITY OF FINFISH PRODUCTS WITHIN THE EUROPEAN MARKET: A META-REGRESSION ANALYSIS OF DISCRETE CHOICE EXPERIMENTS

Relatrice
Prof.ssa Silvia Ferrini

Correlatore
Dott. Javier Cantillo

Tesi di laurea di
Mattia Galante

Anno accademico 2022/2023

Contents

Abstract

The current system of seafood production arises increasing concerns for its impact on the environment. Greenhouse gas emissions from transportation of raw or finished products, depletion of fish stocks (overfishing), pressure on resources (e.g., for feed), generation of marine litter and release of chemicals in the aquatic environment are only some of the issues afflicting both fisheries and aquaculture. However, if correctly managed, seafood production may represent a valid alternative to reduce the ecological footprint of food industry and assure nutrition security, especially thanks to the promising technological developments of fish farming. Within this framework, the detection of consumers' preferences and attitudes towards seafood products is essential: it allows for better informed policy making, which exploits the existing trends and creates new ones, in order to direct markets towards more sustainable pathways. Based on a literature review from Cantillo, Martín and Román (2020), in this thesis two meta-regression analyses are performed and, then, compared to synthesize and extract new information from 18 primary studies employing discrete choice experiments (DCE) and using willingness to pay (WTP) as measure, with the purpose of identifying and quantifying, in monetary terms, the preferences about, respectively, production features and sustainable attributes of finfish, within the context of the European market. The focus on finfish (which constitutes 76% of the species harvested every year) and on a continental market is justified in view of building more homogeneous samples and providing more robust and applicable results for decision making. The present research detects positive and relatively high WTPs for domestically harvested, unprocessed and fresh finfish, one one hand, and for sustainable products, on the other hand. A good predisposition to market innovation and the effectiveness of informed consumption also arise from the results, together with an aversion to aquaculture. Other aspects, like the differences in attitudes of consumers belonging to three European sub-regions and the impact of households' income on purchase behavior, are also explored. These are valuable findings to plan adequate policies, some of those are suggested in the final part of this work.

Chapter 1

Introduction

1.1 Food market and sustainability

1.1.1 Issues and trends

Food production is one of the most important activities of modern social metabolism, at a global level. It is essential for sustenance of our species, for our economy (it generated 3.6 trillion $ of value added in 2020, with a workforce of 866 million individuals as of 2021)[1] and for trade (1.42 trillion $ of export in 2020), and accounts for a substantial share of the flows and outflows that human system exchanges with the environment. The volume of world production of crops, livestock, fish and derived amounted to 9.518 billion tonnes in 2020, most of which consisting in cereals, fruit, vegetables and, in general, products from agriculture (94.75%). An economic phenomenon of such magnitude has an impact on the environment itself. The present food industry is unanimously considered as not sustainable and, among the several concerns arising, can be mentioned (Reisch, Eberle, and Lorek, 2013): large and complex supply chains generating high levels of greenhouse gases (GHGs) emissions, increasing demand of resources (this also causing severe problems of nutrition security and access to food),[2] diet-related health problems due to contemporary lifestyle, release of chemicals and other pollutants in atmosphere, soil and water bodies, which has as a consequence eutrophication events, loss of habitats and biodiversity. Some studies also point out the correlation of the spread of food habits based on highly processed meals (e.g., ready-to-cook) with the climate change (Saarinen et al., 2012). The

[1] These and the following data on food production come from the Food and Agriculture Organization (FAO, 2022b).

[2] El Bilali et al. (2019) highlight that there is a connection between sustainability and nutrition security in the broader process of ecological transition.

inventory of EDGAR-FOOD (Crippa et al., 2021) identifies 51 anthropogenic activities involved in food production and calculates a total of 18 gigatonnes of CO_2 equivalent they generated in 2015, about one third of the global GHGs emissions.

There are several possible solutions to make food sector more sustainable, some of which are already widely employed; these are production standards setting specific principles and requirements to be followed, organic production, development of new technologies, environmental management systems, impact assessment analyses (e.g., LCA), and so on. Very important are, also, the reduction of food waste, eco-design of packagings and improvements in resource and energy efficiency and in yields. Another interesting option would be shortening supply chains and support local markets; many studies agree on the positive effects these policies have on decreasing environmental footprint (Canfora, 2016; Jarzębowski, Bourlakis, and Bezat-Jarzębowska, 2020; B. Smith, 2008). Unfortunately, except for the last case, these measures are often voluntary or employed as marketing strategies and, however, not sufficient.

On the other hand, the activity of organizations like the already mentioned FAO and the increasing adoption of regulations both at a national and international level are successfully promoting sustainable consumption, although more specific interventions on the food industry are needed. The European Union, for instance, has a wide regulatory framework and ambitious policies for sustainability, which also drives its Common Agricultural Policy (CAP). Moreover, at the time of writing, a new regulation proposal "aimed to integrate sustainability into all EU food-related policies" results scheduled as part of the "Farm to fork" strategy, presented by the European Commission in 2020, and the European Green Deal. The enlisted topics also underlie some of the UN Sustainable Development Goals (SDGs).

1.1.2 Consumers' choices and policy making

The detection of consumers' preferences on food products and the factors behind their food choices plays a fundamental role in the design of policies and, in general, in decision making processes.

What is a food choice? Basically, it is any decision involving a consumer and a food product, in any purchase context. Lizin et al. (2022) describe it as driven by sometimes "past experience and satisfaction", sometimes low involvement, sometimes "active reasoning and deliberation"; it is made "in a multitude of choice contexts, combining different moments, occasions, situations, and types of company" by individuals "having heterogeneous sets of personal characteristics, knowledge, beliefs, perceptions, attitudes, and motivations", and "not only have an impact on a person's nutritional and health status and on his/her overall well-being", but also on the environment.

Understanding consumers' attitudes and behaviors means producing valuable information on the preferred commodities, their sustainability level, the heterogeneity in tastes and buying patterns across countries and generations, the presence of a possible aversion against innovation (e.g., new production techniques or ingredients), the reaction to price fluctuations, the trust in imported products and many other factors. These information lay the foundations for public and private entities to build effective policies. Economic valuation methods, like choice experiments or contingent valuations, are able to provide them and to quantify their magnitude, in monetary terms. The latter is further useful to decide on the appropriateness of certain policies or market strategies and when planning subsides or compensations. In this sense, consumers' choices drive markets. However, their knowledge can also be exploited to target them towards a set of goals, or towards more sustainable pathways.

1.2 Sustainability of seafood products

1.2.1 Seafood market: current situation and future perspectives

Worldwide, the seafood production of 2020 was 177.8 million tonnes, considering both fisheries and aquaculture products, or 213.8 million tonnes, also including algae (EUMOFA, 2022; FAO, 2022a). Most part of it was intended for human consumption, and a considerable share (33.7%) was actually exported from one country to another. Finfish[3] represented the main category of harvested species (76%). The major trend to be highlighted here is the constant growth of aquatic food consumption, with an average rate of almost 2.5% in 2010s; more than half of this growth can be explained by an increase in *per capita* consumption, rather than by simple increase in population.[4] This means that individual attitudes towards this type of food are changing. Seafood is essential for human nutrition, contributing to the intake of animal

[3]"Finfish" is defined by the International Seafood Sustainability Foundation (ISSF) glossary as the "term used to describe the strictly classified biological group of fishes, sometimes called true fishes to distinguish them from other aquatic life whose common names also end in -fish, including mollusks (e.g., cuttlefish), crustaceans (e.g., crayfish), echinoderms (e.g., starfish), and other animals (e.g., jellyfish); or any other aquatic life harvested in fisheries or aquaculture (e.g. shellfish)" (ISSF, 2023).

[4]Also in this case, where not otherwise indicated, data come from the Food and Agriculture Organization (FAO, 2022a).

proteins (with an higher proportion in middle-income countries) and fatty acids.

Concerning the balance between wild catch and aquaculture, the production from the latter had a steep increase within the last 30 years to reach the same volumes of fisheries. However, while fisheries production is stable, there is general agreement on projections stating that aquaculture will continue to grow (DNV, 2023).

Here resides one of the specificities of EU countries, where "only" 1.09 out of the total 4.96 million tonnes of seafood come from aquaculture; not too different is the situation in another European country which is also the largest producer of the continent: Norway. The most consumed species in EU are tuna, salmon and cod, although the *per capita* expenditure for seafood is still much lower than the other world countries (especially compared to Asia). Despite this, the European production (including EU and non-EU countries) is the third highest of the globe and, in general, it seems to follow the same common trends.[5]

1.2.2 Impacts on the environment

Seafood production also suffers of relevant criticalities with regard to sustainability, but environmental impacts of fishery and aquaculture are very specific to the sector.

For example, the related GHGs emissions are only 4% of global food production, although growing (Parker et al., 2018). Of course, this picture changes if indirect emissions from transportation are included. Ziegler et al. (2022) calculate that consuming one kilo of Norwegian fresh salmon shipped to Paris by truck emits 6.5 kg of CO_2 equivalent, whereas eating one kilo of Norwegian frozen cod transported to Paris by ship (through China) emits 2.5 kg of CO_2 equivalent. As a comparison, it would be necessary to drive a large car in urban traffic for 33 km to emit the same amount of CO_2 equivalent of one kilo of salmon.[6]

Thus, starting from fishery, one of the biggest environmental concerns derives from the state of fish stocks. According to FAO (2022a), the fraction of fish stocks within biologically sustainable levels decreased to 64.7% in 2019, meaning that one third of species is overfished. This threats biodiversity, ecosystem services and continuity of supply chain (thus, nutrition security).

[5]Data on European countries and the EU come from the European Market Observatory for Fisheries and Aquaculture Products (EUMOFA, 2022).

[6]Elaboration of data from the IEA (International Energy Agency) Mobility Model; data available at: `https://www.iea.org/data-and-statistics/charts/ghg-intensity-of-passenger-transport-modes-2019` (visited on 06/15/2023).

Marine litter, in particular plastic litter (Galgani et al., 2019), is also cause of concern, common to fishery and aquaculture; indeed, fishing gear, nets, installations and equipments, loss of items and packagings by crews are all sources of plastic and micro-plastic (Skirtun et al., 2022), which causes severe harm to marine biota (Werner et al., 2016) and human health (still to be confirmed). These kind of damages can potentially extend to non-target species.

Coming to aquaculture, also in this case a problem of resource management arises: modern fish feed is composed by a wide variety of ingredients (Winther et al., 2020), among which crops, whose production is already under pressure and subject to geo-political issues, fish meal, dependent on fishery which, as stated, is actually not sustainable and has not the same growing trend of aquaculture (Ankamah-Yeboah, Jacobsen, and Olsen, 2018), and soy.[7] Another relevant issue is related to the release of chemicals and other substances from farming installations to the aquatic environment, specifically phosphorus and nitrogen.[8] Also, impacts on human health are present, relatively to the intake of fatty acids (EPA/DHA): due to the substitution in fish feed of part of fish meal and fish oil with plant ingredients, the daily recommended intake of 500 mg cannot be granted (Hamilton, Newton, et al., 2020).

However, aquaculture potentially allows to better address sustainability in terms of fish stocks, adaptation to risks and, in general, in terms of assuring the continuity of supply chain (Subasinghe, Soto, and Jia, 2009) by, at the same time, making it more ecological. Indeed, solutions to the above enlisted issues of aquaculture exist: namely, pollution control, through the employment of technologies like the *closed-containement aquaculture* (CCA)[9] or land-based approaches based on *recirculating aquaculture systems* (RAS),[10], reduction of plastic marine litter,[11] innovative production methods like the *integrated multi-trophic aquaculture* (IMTA),[12] and the use of alternative

[7]Cultivation of soy involves deforestation, water consumption, use of pesticides and fertilizers (again, Ankamah-Yeboah, Jacobsen, and Olsen, 2018), and long-distance transportation since most of production is concentrated in Brazil.

[8]Phosphorus, originated by excretions and other losses, has eutrophication as a consequence (Hamilton, Brod, et al., 2016).

[9]It involves barrier technologies to reduce the release of waste and chemicals to the external aquatic environment.

[10]Which also reduce the release of pollutant to (and other dangerous interaction with) the aquatic environment. For a more detailed examination of aquaculture technologies, reference is made to the report realized by Hough (2022) for the FAO.

[11]Through new sea-proof installations and recycling programs (Skirtun et al., 2022).

[12]The IMTA provides a balanced aquaculture system in which different species, at different trophic levels, are farmed within the same environment; in this system, byproducts

feed. Moreover, simulations show how introducing stricter regulations in fishery would unavoidably result in a decrease in employment, whereas this is not necessarily true for aquaculture (McCausland et al., 2006).

1.2.3 Building a sustainable seafood market: product attributes

In addition to the interesting developments in aquaculture practices, the ways seafood market can become more sustainable are common with the broader food market. Thus, better environmental management of production systems and related resources, adherence to international standards, eco-design of packagings and industrial activities and structures, increase of efficiency, reduction of product processing and shortening of supply chains.

To be mentioned here, is that two specific groups of standards for seafood exist: those from the Marine Stewardship Council (MSC) and those from the Aquaculture Stewardship Council (ASC). The first are only applicable to fisheries compliant with the three principles of sustainability of the stock, ecosystem impacts and effective management which, in turn, are structured in targets and performance indicators.[13] The second, as can be guessed, are only applicable to aquaculture companies. Several standards exist for the most common species; for example, that for salmon provides to meet requirements related to fish feed, biodiversity, pollution, diseases and social responsibility.[14] In both cases, if the respect of the standards is assessed, companies are certified and can employ MSC and ASC eco-labels on their products. Moreover, specific standards for organic aquaculture also exist.[15]

All issues discussed so far are reflected in product characteristics on which build purchase choices of consumers and, in general, their preferences. Attributes like the origin, harvest method, production method, species, presence of an eco-label, presence of a nutrition, health or safety label, organic production, product form, product presentation, type of feed and so on are supposed to be essential in buying decisions. On these attributes, which will be further explored in section 2.4.4, focuses the present research.

from one species serve as an input for another, this having many advantages, among which a more natural nutrients cycle, the reduction of pressure on resources and, again, of interaction with (and pollution of) the outside environment.

[13]From the MSC website: https://www.msc.org/for-business/fisheries/fishery-certification-guide.

[14]From the ASC website: https://asc-aqua.org/producers/farm-standards/salmon/.

[15]More information on organic aquaculture basic are available on the website of IFOAM - Organics International: https://www.ifoam.bio/about-us/our-network/sector-platforms/ifoam-aquaculture.

1.3 Valuation methods

1.3.1 Review

The economic valuation methods employed to detect and quantify consumers' preferences and provide information about their attitudes are many. More generally, in environmental economics, different tools exist whose appropriateness depends on the specific field of study or the commodities under research; these can be environmental amenities in the case of non-market evaluation, or sustainable attributes of market products. Moreover, use of techniques like stated or revealed preferences also vary according to the research topic and the components of value to be included in the estimated measures, which can be based, respectively, on Marshallian or Hicksian demand curves, these identifying more or less broadly the *total economic value* (TEV).

The employment of whichever of these techniques has been growing in the last 20 years, in all fields of economics and environmental economics. For instance, studies analyzing the specific matter of attitudes towards sustainability of finfish and only based on DCEs have increased from 1 in 2000 to 25 in 2022, with a peak of 28 in 2021. **Figure 1.1** shows this trend with a bar chart. It represents the search output of the SCOPUS database using the boolean terms reported in section 3.2 and also used for data collection in the main analysis, and narrowing it to the year range 2000 - 2023.

Indeed, although some cases are present of application of *hedonic prices* (HP), the most common methods used when the research questions are related to the detection of consumers' preferences for sustainability of market products are *discrete choice experiments* (DCEs) and *contingent valuation* (CV), part of the stated preference methods. While in CVs the commodity/amenity is regarded as a whole, DCEs allows to "disassemble" it in its basic building characteristics[16] and to estimate the utility (and the monetary value) assigned by individuals to each of them (Pearce and Özdemiroğlu, 2002). Both of these methods are based on surveys delivered to respondents, which include more choices to make when DCE is employed, to valuate tradeoffs among product features, and give as output measure the *willingness-to-pay* (WTP) for the commodity or its attributes. More detailed information on the theoretical basis and the use of DCEs will be provided in section 2.4.2; here it is pointed out that, because of the complexity of the investigated issues and the need of information on consumers' attitudes towards every

[16]A technical term equivalent to *characteristics*, which will be widely employed in this thesis, is *attributes*. It refers to the set of elements used to describe any product or service according to the Lancaster's theory (Lancaster, 1966).

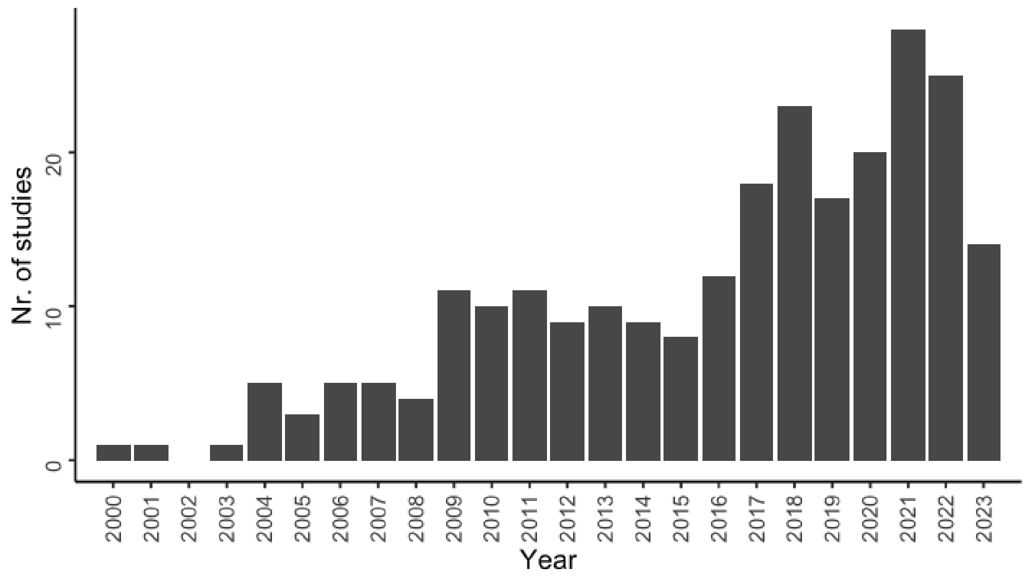


Figure 1.1: Search output for studies detecting consumers' preferences towards sustainability of finfish with DCEs, years 2000 - 2023, from the SCOPUS database. Since the search was made at mid-2023, a decrease can be observed for that year.

single aspect of them, the meta-regression analysis performed in this thesis, as anticipated above, takes into account only primary studies employing such valuation method.

1.3.2 Within the "jungle of findings"

But why meta-regression analysis? **Figure 1.1** indicates that several studies are available in literature, also if the scope of the analysis is narrowed down based on very specific criteria; this applies to almost any field of knowledge. However, single studies may not be totally reliable and their findings could be biased if generalized, because each of them is influenced by contextual factors, like country and region of study, sample size, socio-economic characteristics of respondents, specificities of the methodology employed, considered product features, types, attributes and trade-offs, publication selection and so on. Meta-regression analysis (MRA) and, in general, meta-analysis exploits the availability of data to control for such heterogeneity and build models as representative as possible of reality and universally valid, which also explain the impact of the enlisted differences and extract new informations contained in variability across studies.

Stanley and Doucouliagos, in their book "Meta-regression analysis in economics and business" (2012) refer to this incredible amount of data and findings, sometimes conflicting, and the increasingly easy access to them, as the "Tower of Research"; of course, the term recalls the "Tower of Babel", to highlight the concerns of some researchers[17] that this availability of information can threat genuine scientific developments and the related policy making, as a consequence of the loss of reference points useful to valuate reliability of results and a general confusion preventing the analysts to understand each other.

Here, we employ the term "jungle of findings" to indicate the amount and high heterogeneity of findings from a growing and essential field of study like that of sustainability of products. Likewise, we have the same trust of Stanley and Douculiagos in meta-regression analysis to make its way through this jungle by putting order and allowing to recognize authentic and reliable information. This methodology will be in-depth explored in the following chapter.

1.4 The analysis of this thesis: outlines, goals and structure

1.4.1 Finfish products, DCEs and MRA

The research conducted in this thesis will employ meta-regression analysis, exploiting data from previous DCE studies, to create models describing as accurately as possible consumers' attitudes, and influential factors on them, towards attributes of finfish products accounting for the above discussed sustainability issues, in the broader background of seafood and, therefore, food production, within the specific context of the European market. The reasons behind these choices and the relevance of detecting preferences, as well as of the considered environmental issues were already addressed, while those underlying the selection of European market as study context and finfish as study object will be addressed afterwards (section 2.4.7 and 3.1). The main goal is obtaining information about the *direction*, the *magnitude* (WTP), the *trends* and the *drivers* of consumers' attitudes which can be outlined to support decision making and design of policies, or to valuate the opportunity of those already existing and above enlisted. This will be done by performing two parallel MRAs which separately account for two main identified groups of attributes involved in buying choices, which activate different behavioral pat-

[17]E.g., the Nobel econometrician James Heckman.

terns, and by discussing policy options. The study will be complemented by other applications of meta-analysis tools and related statistical tests. Such aspects are also more exhaustively treated and explained in the following chapters.

1.4.2 Forerunning studies

It is important in this introduction to cite and briefly introduce the forerunning studies to which the present work is inspired and, first of all, the literature review "Discrete choice ex- periments in the analysis of consumers' preferences for finfish products: A systematic literature review" from Javier Cantillo, Juan Carlos Martín and Concepción Román (2020), whose first author is also the co-supervisor of this thesis. Their paper lays the groundwork for the research conducted and presented in these pages. It shares with the MRA performed here the just described research boundaries[18] and part of the primary studies. Indeed, the original idea triggering this analysis was applying *quantitative* methods on a topic only explored with *qualitative* methods.

Then, other previous MRAs concerning consumers' preferences within the context of food and seafood sector, although with different scope and focus, were taken as example for the planning of the analysis and to compare its results. These are the studies from Printezis, Grebitus and Hirsch (2019), Yeh and Hirsch (2023), Bastounis et al. (2021), Smetana, Melstrom and Malone (2022) and Li and Kallas (2021).

However, this is the first MRA conducted in the specific outlined field of study, the one encompassing the sustainability of finfish within the European market. Hopefully, further studies will follow.

1.4.3 Thesis structure

The thesis is structured as follows. Literature review and methodology is in chapter 2: the principles and basic issues of meta-analysis and meta-regression analysis are explored, as well as their specificities in environmental economics and when applied to DCEs, and set-ups and results from the previous studies are summarized. Chapter 3 presents more in detail data and methods, thus the data collection process, preliminary tests conducted, models and specifications. The results are presented in chapter 4 and discussed in chapter 5, which also contains some suggestions for policy making and

[18]Except that for the focus on European countries.

informs about the limitations of the research. Finally, conclusions are drawn in chapter 6.

Chapter 2

Literature review and methodology

2.1 Introduction to meta analysis

The last few decades have been characterized by a rapid increase in empirical research, the accumulation of scientific knowledge and the increasingly easier access to worldwide databases due to the diffusion of information technology. Such developments have raised the need for quantitative methodologies allowing to integrate or synthesize findings from similar studies, correct any biasing factor, explain their differences and improve reliability and validity of results. The aim is to better support, for instance, medical applications or policy making, in a objective, replicable and standardized way, by exploiting this new abundance of data. As many researchers argue (e.g., Stanley and Doucouliagos, 2012), variation in empirical findings in any scientific field of study, including economics, environmental economics and resource management, is the norm; meta-analysis aims to exploit the availability of terabytes of data to explain, with statistical foundations, this variation and, perhaps, to find values for such findings being as accurate as possible. Indeed, classical qualitative methods applied to summarize scientific knowledge on a certain topic, like literature reviews, are subject to reviewer's interpretation (Stanley, 2001) and do not provide clear quantitative values or models describing phenomena which can also be used for practical applications. These are the reasons driving the choice to employ meta-analysis for the present research.

Before deepening its basic concepts and the most common tool in economics, thus meta-regression analysis, a brief history of this methodology and its evolution over time and an overview of its key aspects and challenges is provided.

2.1.1 History of meta-analysis

The first attempts to develop statistical methods for quantitative synthesis of separate experiments are from the legendary statisticians Karl Pearson and R.A. Fisher. Pearson's solution was to average and combine correlation coefficients (Pearson, 1904) while Fisher's one was a new statistic that combined p-values from different studies; thus, statistically independent tests of the same hypothesis (Fisher, 1932). The review of these statistics goes beyond the aim of this thesis. However, it is important to emphasize the simple principle behind them: as Fisher argued, if hypothesis tests are combined, their cumulative power is greater. Therefore, combining several findings from empirical experiments results in reaching the much sought research goal: statistical significance.

Conventionally, modern era of meta-analysis is considered to begin with Gene Glass. He introduced a modern definition of meta-analysis, described as "the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings" and proposed an estimator based on this concept (Glass, 1976). The estimator, identified as "g", is built as a standardized measure of some relevant effect, based on the difference between the average effects within the *experimental* and *control* groups:

$$g = \frac{\overline{X}_e - \overline{X}_c}{S}$$

Where $\overline{X}_e$ is the average of some relevant measure of effect in different experimental groups, $\overline{X}_c$ represents the average in the control groups and S is the standard deviation of this measure as seen in the control groups. If $0.2 < g < 0.5$, there is a small effect, whereas if $0.5 < g < 0.8$, there is a medium effect and, if $g > 0.8$, there is a big effect. Of course, this estimator is more suitable to experimental sciences, medical sciences and, in general in any experiment-based field of research (this being very rare in economics). But, again, it can be useful to understand the intuitive principles underlying meta-analysis: integrating results from different experiments means synthesizing them, increasing the statistical power and significance over single experiments and obtaining a more reliable measure or value for the considered effect.

Three studies are generally given credit for introducing meta-analysis in economics: Smith and Kaoru (1990), Walsh et al. (1989) and Stanley and Jarrell (1989). In their paper, Stanley and Jarrell propose using the same statistical tool traditionally used to produce econometric estimates, thus regression analysis, to summarize, integrate estimates and explain variations among them, and provide the basis of what they define "regression analysis of regression analyses", namely *meta-regression analysis*. Smith and Kaoru apply this methodology to summarize hundred empirical estimates of consumer surplus from travel cost method to valuate recreation benefits and they recognize that synthesis of different studies should account for some types of heterogeneity (factual or methodological) and some statistical issues, like heteroskedasticity (very common) and non-independence of observations. Walsh et al., similarly, analyze estimates of recreation value from travel cost and contingent valuation (CV) methods; their work is relevant as it is the first to recognize the potential application of meta-analysis to benefit transfer (Johnston et al., 2015). These studies are considered key contributions to the widespread of meta-analysis in economics as they have, respectively: laid the groundwork for the use of meta-regression analysis; detected the main issues related to the employment of this tool and recognized the opportunity to use meta-analysis not only to synthesize data but also to explain their variation; and identified the possible application of meta-regression in benefit transfer studies. After this pioneering studies, the adoption of meta-analysis in economics has grown exponentially over the last years, and it is still growing (Stanley and Doucouliagos, 2012).

It should be noted that two of these first applications are based on primary studies applying typical methods of non-market valuation, which nowadays are widely used within the field of environmental and resource economics. This branch of economics has also known a huge increase in use of meta-analysis, that today is commonly employed to harmonize data deriving from hundred choice experiments, contingent valuations, travel cost and hedonic price applications and so on. According to Nelson and Kennedy (2009), since 1990, over 300 meta-analyses have summarized empirical results in environmental and resource economics, many of them incorporating benefit transfer applications.

2.1.2 Meta-analysis: objectives, key issues and challenges

To sum up and further clarify the basic concepts underlying the previous sections, a simple question should be asked: "what is meta-analysis?"; or, more pragmatically: "how can meta-analysis be shortly defined?". Starting from the Glass' definition, meta-analysis is a statistical analysis which, with different possible methods, models and criteria, combines the results of multiple scientific studies on the same topic, with various purposes; typically:

- synthesizing values or measures, increasing statistical significance and accuracy of results;

- explaining differences among estimates (this meaning understanding and controlling for heterogeneity and biasing factors).

Single studies, although well conducted, are subject to be highly influenced by a variety of contextual and methodological factors, as well as by the size and characteristics of their samples: this means that even the most accurate experiment ever made can potentially result in a flawed estimate or value, which can be more or less distant from the "true" value in population. Meta-analysis gives the opportunity to identify and control such influential factors, or just to increase the statistical power and significance of the estimate by aggregating more data. Several meta-analysis techniques exist: conventional statistical methods like weighted averages, correlation coefficients, factorial analysis of variance (ANOVA), indicators based on difference between means (e.g., Glass' g, Cohen's d, Hedges' g), vote counting, etc.; various forms of meta-multicriteria analysis; epistemological or expert analysis; rough set analysis (Bergh et al., 1997). However, most of the work conducted in economics has been concerned with moderator variables to control for heterogeneity and differences across primary studies; this amounts to saying that the most common tool employed within this field is meta-regression analysis (Bergh et al., 1997; Stanley and Jarrell, 1989).

Among the others, Johnston et al. (2015) and Nelson and Kennedy (2009) report the main reasons a meta-analysis may be conducted for, many of them already mentioned: synthesizing values in a systematic manner to improve measures and draw new conclusions, identifying and managing outliers, understanding and explaining data heterogeneity and its impact on variation of estimates, controlling for methodological differences, providing "combined" estimates (e.g., calculating weighted means), providing a within-sample predicted values for the dependent variable under a particular set of conditions or an out-of sample prediction (benefit transfer), summarizing results of a single empirical study that has produced multiple estimates, identifying and correcting publication bias.

Of course, performing a meta-analysis, whatever the method used, is very complex. The meta-analyst has to address some key challenges, typical of that kind of analysis. A non-exhaustive list, obtained by comparing different sources (e.g., Borenstein et al., 2009; Johnston et al., 2015), include

- *identifying suitable source studies*: primary studies analyzed within the meta-analysis must be comparable in terms of dependent variable, commodity definition or contextual factors, and differences among them must be able to be realistically controlled with moderator variables or other statistical approaches; in general, a trade-off should be considered among quality, commensurability and number of data;

- *identifying a consistent definition for the dependent variable*: sticking with the concept of "commensurability", it is essential to give a consistent definition to the dependent variable used within the analysis, in terms of measure and functional unit, uses and values component taken into account and depicted commodity, paying particular attention to pooling different kind of estimates, in order to make it comparable across studies and to obtain theoretically consistent results; this is especially true in environmental economics and non-market valuation (this aspect will be further explored later);
- *identifying contextual and methodological variables* that can explain variations within the dependent variable;
- *addressing potential statistical complications* related to sample selection, heteroskedasticity, heterogeneity and correlation of primary data (both within- and between-studies).

It is necessary to account for and address these issues with the appropriate statistical tools by carefully designing the analysis and collecting data, in order to conduct a significant meta-analysis.

2.1.3 Effect size

Basic concepts

Within the field of meta-analysis, with "effect size" it is generally intended the measure expressing the findings of primary studies about some relevant effect or empirical phenomenon. Therefore, when conducting a meta-regression analysis, it corresponds to the dependent variable. Kelley and Preacher (2012) define the effect size as "a quantitative reflection of the magnitude of some phenomenon that is used for the purpose of addressing a question of interest". Typically, the effect size is a statistic or parameter which describes the extent (size) of such phenomenon. Transposing the concept onto economics, the effect size will be a measure of the *economic effect* of some variables, thus the estimate of an economic association (Stanley and Doucouliagos, 2012). As Stanley and Doucouliagos argue, there is a difference between *statistical* effects and *economic* effects: statistical effects are "unitless measures of an association between two variables", whereas economic effects measures the main effect of *economic interest*. They report a list of the most common effect sizes in economics, like zero-order or partial correlations, elasticities and t-statistics; environmental economics typically employs marginal values expressed in monetary terms (e.g., WTPs). While it is true that, in economics, effect sizes are generally computed from regression coefficients, use of unprocessed coefficients for meta-analysis purposes is ruled out, except when the scale and measure is identical, since effect sizes must be comparable across studies (again, Stanley and Doucouliagos, 2012). In fact, differences in the functional form of regressions, which define the interpretation of regression coefficients, may make them not commensurable.

Basic models

Before introducing meta-regression analysis, a brief but essential overview on statistical foundations of effect size in meta-analysis is needed.[1] Effect size can be modeled as either *fixed-effect size* (FES) or *random-effect size* (RES). Suppose design and estimation features of single primary studies, as well as contextual factors, do not affect the expected value of estimates. The fixed-effect size model postulates that estimates share a common, or *fixed*, effect size: in other terms, it hypothesizes an homogeneous effect, not varying across studies; this may seem restrictive, however it is demonstrated that FES models have better small sample properties (Rhodes, 2012). Given such assumptions, a basic FES model aggregating estimates from a population of N studies can be specified as follows:

$$Y_i = \beta + e_i, \quad i = 1, 2, \ldots, N \tag{2.1}$$

where β is the population mean of effect size, such that $\beta_1 = \beta_2 = \cdots = \beta_N = \beta$ (thus, the effect is fixed), and e is the measurement or sampling error. Therefore, only the latter explains deviations of single estimates from the common effect value.

Alternatively, there may be reason to believe that there is heterogeneity in the effects (thus, they vary across population of studies, so that each effect size is modeled as a random draw from a distribution of effects) and to suppose that such heterogeneity is not measurable using regressors. Formally, heterogeneity is modeled as $Y_i = \beta_0 + u_i$, where β_0 is the mean of a super-distribution of variable effects and u_i is a random term capturing the unmeasured heterogeneity (Borenstein et al., 2009). Adding these terms to equation (2.1) yelds:

$$Y_i = \beta_0 + u_i + e_i, \quad i = 1, 2, \ldots, N \tag{2.2}$$

which is the RES model. Technically, the RES model is more appropriate when there is excess heterogeneity, thus more than expected by sampling

[1] Where not otherwise indicated, the following information is based on the contribution of Johnston et al. (2015).

error alone (Stanley and Doucouliagos, 2012).

Based on these simple models, FES and RES weighted means of the effect size β, which are employed within meta-analysis techniques as summary statistics, can be calculated. They are given by (Borenstein et al., 2009)

$$\overline{\beta} = \frac{\sum w_i Y_i}{\sum w_i}, \quad i = 1, 2, \ldots, N \tag{2.3}$$

where w_i are the weights assigned to primary estimates according to their precision (inverse of variance) and differently computed in FES and RES models. Of course, such statistic is just a rough estimator of common effect size, and do not account for heterogeneity.

2.2 Meta-regression analysis

2.2.1 Controlling for heterogeneity

As anticipated in section 2.1.2, heterogeneity in meta-data may be originated by different factors. More technically, Christensen (2003) identifies two basic causes of heterogeneity, factual and methodological. Factual heterogeneity occurs when there are real differences in effects among primary studies (due to contextual factors, commodity characteristics, etc.); this kind of heterogeneity tends to be exacerbated when more "comprehensive" samples are selected. Methodological heterogeneity, instead, depends on the differences in design, methods and characteristics of primary studies. To give some examples of factors underlying data heterogeneity:

- natural variation of economic, social and ecological phenomena;
- country/region of study and site-specific features;
- sample characteristics, specifically socio-demographic aspects;
- time period;
- product characteristics, in case of market analysis;
- valuation method employed;
- model specification (econometric model and techniques);
- functional form;
- dependent variable measure;

- presence/absence of independent variables;
- other study-specific features;

and so on. As these factors may variate across studies, this results in the mentioned heterogeneity and variation of the estimate and related findings. Therefore, especially in economics, it is often not sufficient to synthesize effect sizes by applying conventional statistical methods and calculating basic summary statistics, coefficients and indicators; it is necessary to explain heterogeneity and control for it, in order to better understand the factors influencing the analyzed phenomenon and their magnitude, and to compute more accurate estimates for the related effect size. Moreover, if unaccounted, heterogeneity can bias estimates of meta-analysis, like omitted-variables do in any simple regression analysis, causing the regression coefficients to be biased. Which is why the most common statistical methodology employed in economic meta-analysis is meta-regression analysis, or MRA (Bergh et al., 1997; Stanley and Jarrell, 1989).

2.2.2 Models for meta-regression analysis

Indeed, the most obvious solution to account for heterogeneity is modeling a regression analysis having as dependent variable the estimated effect size reported by each primary study and employing the necessary independent variables (in this context also called *moderator variables*). Florax et al. (2002) provides the general statistical form of a meta-analytical problem, which can also be applied to outline the structure of a meta-regression model:

$$Y = f(P, X, R, T, L) + e \tag{2.4}$$

where Y is the reported estimate (e.g., hedonic prices for noise levels[2]), P encompasses the causes of the outcome (e.g., aircraft noise levels), X are the characteristics of the set of objects affected by P in order to determine the outcome (e.g., age of houses), R includes the methodological factors specific to each primary study, T indicates the time period, L is the location of study and e is an error term.

By modeling equations (2.1) and (2.2) in the form of (2.4) to account for heterogeneity, the basic fixed-effects (FE) and random-effects (RE) meta-regression models are obtained. Starting from the FE model, it is specified as follows:

$$Y_i = \beta + \alpha_1 X_i + e_i \tag{2.5}$$

[2]The examples in brackets are provided by Nelson and Kennedy (2009).

where X_i is the vector of measured and observable variables which are believed to influence the effect size in primary studies and α_1 is the vector of the related regression coefficients, whose estimates are usually computed using generalized least-squares estimators (GLS) instead of the more common OLS, in order to account for heteroskedasticity.[3] This model assumes only observed or measured heterogeneity is present.[4] Because most study descriptors (moderator variables) are specified as binary dummies, usually the intercept has a convenient interpretation as the expected effect-size for the null case.

Suppose, instead, there is both measured and unmeasured heterogeneity in the true effect sizes. The measured heterogeneity is again captured by the X_i variables. Denote the unmeasured sources of variation by a study-level vector Z_i. Unless the omitted Z variables are orthogonal to the included regressors, the parameter estimates are biased (Rhodes, 2012). Given the above, the RE model[5] can be specified as:

$$Y_i = \beta_0 + \alpha_1 X_i + u_i + e_i \tag{2.6}$$

where u_i is a random term capturing the part of $Z_i \alpha_2$ that is orthogonal to X_i (again, Rhodes, 2012). The sum of u_i and e_i, which are assumed to be independent, is the composite error term ν_i, whose variance is $\nu_i^2 = \sigma_i^2 + \tau^2$. An estimate of τ^2 is obtained by exploiting the known σ_i^2 values, using either an iterative maximum likelihood procedure or a non-iterative moments-estimator. The RE model can also be estimated by GLS, using the inverse of ν_i^2 as weights (Nelson and Kennedy, 2009). The provided FE and RE specifications and related information are based on the contribution of Johnston et al. (2015).

The RE meta-regression model is more appropriate when also unobserved or unmeasured heterogeneity is present, which is the case with most empirical phenomenon. However, as it is well-known in econometrics, if the random term and independent variables are correlated (this often happening in economic MRA, especially when standard errors are included among

[3]This aspect will be further addressed in section 2.2.4.

[4]Here the term "fixed-effects" should be regarded carefully: it doesn't mean that the primary studies share a common effect (as theoretically intended in the general FES model) but, instead, that differences in effect sizes can be totally explained, leading back multiple effects to a single one.

[5]More properly, this is a *mixed-effects* model, since both measured and unmeasured or unobservable heterogeneity is present: thus, heterogeneity is both explained with regressors (like in fixed-effects models) and captured by the random term u_i (like in random-effects models). However, here a decision was made to follow the usual practice in literature, where in the most of cases it is still indicated as *random-effects* model. Therefore, with "RE model" will be henceforth intended the mixed-effects model.

regressors), the estimates might be biased. Nevertheless, many researchers still prefer RE model, because it usually better fits meta-data and better explains heterogeneity, despite the depicted issue. This aspect will be deepened in section 2.2.4.

Publication bias

The simplest application of meta-regression analysis consists in its employment to identify and correct the so-called *publication bias*. It is originated by publication selection, which is that process "of choosing research papers, or their results, for statistical significance" (Stanley and Doucouliagos, 2012) or for their consistency with conventional views. Card and Krueger (1995) individuate three causes of publication bias in economics:

- reviewers and editors may be predisposed to accept papers consistent with the conventional view;
- researchers may use the presence of an expected (because in accordance with previous studies) result as a model selection test;
- everyone may possess a predisposition to treat "statistically significant" results more favorably.

This causes the scientific findings of published sources, as well as any synthesis or meta-analysis of them, to be biased.[6] The issue can be handled both graphically and with meta-regression models (Stanley and Doucouliagos, 2012).

Indeed, the first and most common step to deal with publication bias is building a funnel graph: this is a scatterplot of the estimates of effect size (x-axis) versus their precision (y-axis). Precision can be proxied both with the inverse of the standard error and the square root of the sample size. The visual evaluation of the plot's symmetry respect to the mean value, obtained averaging the top 10% of the most accurate estimates, gives a first rough information about publication bias: in its absence, the observations would be randomly and uniformly distributed around such mean value (Stanley, 2005).

Another approach is to perform a regression between the effects reported by primary studies and their standard errors (or variances); if, based on such regression, it turns out that the magnitude of the estimates is somehow dependent on their precision, publication bias is present. In fact, researchers

[6]Good examples of publication bias come from Doucouliagos and Stanley (2009) and Doucouliagos and Stanley (2012).

having small samples and, therefore, low precision estimates may employ econometric techniques and different model specifications to get larger estimates that would be statistically significant; on the other hand, researchers having large samples and accurate estimates will be satisfied with smaller values (again, Stanley, 2005). The regression model for publication bias is provided by Stanley and Doucouliagos (2012):

$$Y_i = \beta_0 + \beta_1 SE_i + e_i \tag{2.7}$$

where SE_i is the standard error of estimate Y_i. As it has obvious heteroskedasticity, the equation should be weighted with the inverse of standard errors themselves, becoming:

$$t_i = \beta_1 + \beta_0(1/SE_i) + \varepsilon_i \tag{2.8}$$

where t_i is the t-statistic of estimate Y_i and $\varepsilon_i = e_i/SE_i$. This specification can be estimated with OLS. Alternatively, equation (2.7) can be directly estimated with weighted least squares (WLS), again using the standard errors as analytical weight. Based on equation (2.8) or on the WLS version of equation (2.7), two hypothesis tests can be conducted:

- *Funnel asymmetry test*, or FAT ($H_0 : \beta_1 = 0$), to test whether or not there is publication selection; logically, recalling what discussed above, if the coefficient representing the impact of the standard error on the effect magnitude is significant, such bias is present.

- *Precision-effect test*, or PET ($H_0 : \beta_0 = 0$), to test whether there is a genuine effect or not, since β_0 can be interpreted as the effect magnitude, once controlling for the bias arising from publication selection.

If, as result of the PET, the null hypothesis is rejected, the estimate of $\beta_0(\hat{\beta}_0)$ from equation (2.8) is one of the best corrections for publication bias (Moreno et al., 2009). However, in this case (thus, if the PET is passed), it is good practice to estimate β_0 replacing the standard error with variance (i.e., its square) in equation (2.7); such estimate is called *precision-effect estimate with standard error*, or PEESE, and provides an even better estimate of the underlying “true” effect (Stanley and Doucouliagos, 2012).

Since standard errors and variances of effect size estimates are not always available or can be derived from information provided from primary studies, the square root of sample size can serve as a rough proxy for precision, as it does in funnel plots, and be used within the FAT-PET model in the place of them.[7] Although the sample size do not provide complete information on estimates variation, statistical theory has long proven that its square root is strongly and inversely related to publication bias (Berlin, Begg, and Louis, 1989), as well as highly correlated with the standard error (Stanley, 2005). Likewise, the sample size (i.e., the square of its square root) can be used within the PEESE model (Stanley and Doucouliagos, 2012).

2.2.3 Data collection

Identifying studies

While selecting primary studies to include within a MRA, the meta-analyst should be very careful not to introduce himself potential bias in the meta-analysis (Stanley, 2005). In this context, this means selecting an amount of studies which approximate the *population* of studies[8] or, at least, can be considered a *representative sample* of such population; the latter should be enough, especially when the population of studies is sufficiently large (Abreu, De Groot, and Florax, 2005). Indeed, it is not always conceivable to collect the totality of studies related to a certain topic or effect size, given the wideness of some research fields or the level of heterogeneity characterizing them: sometimes it is more practical to define sort of "system boundaries", thus following specific selection strategy (e.g., select only studies published after a certain year, focusing on commodities with defined characteristics, on a chosen group of countries or on a given model or method). Of course, also in this case, it is important to conduct an accurate search, which is as comprehensive as possible (i.e., which includes most of the available studies meeting the chosen criteria) and to state clearly the followed strategy.

However, it must be considered that, introducing more studies and observations for statistical and robustness purposes also means introducing more heterogeneity within the sample, this leading to include more regressors in the MRA. This issue is defined the "N versus K problem" by Moeltner et al. (2007). In a few words, a kind of balance between these aspects should be found.

The studies included in any type of meta-analysis must comply with some essential requirements[9] (Stanley and Doucouliagos, 2012): for instance, they

[7]Several studies adopt this strategy. Useful examples come from Yeh and Hirsch (2023) and Printezis, Grebitus and Hirsch (2019)

[8]Nelson and Kennedy (2009) point out that in most of environmental economics applications of MRA, the studies included in the analysis do not approximate their population.

[9]Stroup et al. (2000) also provide an useful checklist of these requirements.

must be so similar that their differences can be coded; they must be empirical studies reporting a statistical estimate (not necessarily regression-based); studies employing the same data and reporting the same estimates, although from different authors, should be avoided; it is not possible to combine estimates from binary regressions (e.g., probit or logit models), with estimates from continuous variable studies. A debate on whether or not include unpublished research, as well as low-quality studies, is still ongoing.

Data selection

Generally, the data collected from primary studies are intended to code the research dimensions of each of them, for the already discussed purposes, and belong to three different categories (Stanley and Doucouliagos, 2012):

- *Essential variables*: mainly, reported estimates' standard errors or sample sizes, used to "weight" their precision.[10]
- *Typical variables*: variables that code the already discussed methodological and factual differences of primary studies.
- *Value-added variables*: these are variables which are not collected from and do not vary across the data of a given study (thus, they are study-invariant), but across studies; by including them, meta-analysis can add new and relevant information, unavailable to the original studies. Examples of these variables can be the average income level, unemployment rates, year the study was published, average year of data, information about authors and so on.[11]

Data correlation: key aspects

As collected meta-data are very extensive and complex, the danger of data correlation is high. This amounts to say that primary estimates of effect sizes may not be independent of one another, violating the fundamental assumption underlying regression analysis and, therefore, its validity. Such correlation may arise both *within-study*, if it involves estimates from the same study, and *between-study*, if it involves estimates from different studies

[10] As explained in section 2.2.2 when discussing about publication bias, sample sizes can be employed as proxies for precision

[11] Coming to the fish market, which is the research topic of the analysis of this thesis, an example may be the following: suppose primary studies focus either only on farmed or only wild fish; single studies cannot account for the variation in consumers' WTP due to differences in the harvest method whereas, by coding this aspect, a MRA can provide this new information.

(Johnston et al., 2015; Nelson and Kennedy, 2009). There are several possible reasons for this to happen and, also, different solutions.[12]

Firstly, some primary studies may use the same data sources (e.g., the same dataset or a public survey) or they may share an observable characteristic, such as an identical functional form, the omission of a key explanatory variable or the same study location. This is a case of between-study correlation which can be dealt with by including the observable as a regressor (moderator variable). Also, most primary studies in economics produce more than one estimate of effect-size. A solution to this case of within-study correlation can be either using only one estimate per study or averaging the estimates from the same studies. However, this reduces the degree of freedom of the MRA and, then, its statistical power. Therefore, more appropriate solutions and techniques should be implemented while designing the MRA, thus in the phase of regression specification and choice of the econometric model. The most common modeling strategies used to deal with data correlation are: cluster-robust modeling, panel modeling and hierarchical (or multilevel) modeling. These strategies will be discussed more in detail in the following section.

2.2.4 Specification and modeling: further topics

In order to put together and clarify what has been said until now, the meta-regression model introduced with equations (2.5) and (2.6) can be rewritten in an extended form:

$$Y_i = \beta_0 + \beta_1 precision_i + \alpha_1 X_i + \delta_1 K_i + \gamma_1 R_i + e_i \tag{2.9}$$

where β_0 is the intercept, $precision_i$ is a general term representing the variable employed to proxy estimates' precision, X_i is the vector of moderator variables depicting the features of the commodity or amenity, K_i and R_i capture, respectively, other contextual factors influencing the effect magnitude (including location of study, time period and sample characteristics) and the methodological aspects related to study design, and e_i is the consuete error term. It can be noticed that this is a FE specification; a model for RE meta-regression is obtained by adding a random error term, u_i, to the equation. As already mentioned in section 2.2.2, often, but not always, the moderator variables are specified as dummies. In any case, regardless of the type of variables employed, there are different approaches to select regressors for the

[12] A more exhaustive and detailed explanation of data dependance is provided by Nelson and Kennedy (2009).

MRA model (Stanley and Doucouliagos, 2012); and the main two strategies are:

- *General-to-specific* (G-to-S): at first, all possible explanatory variables are included. Then, the least statistically significant variable is removed, one at time, until only statistically significant variables remain.

- *All-inclusive*: all possible explanatory variables are included.

Approaching with G-to-S strategy also helps to avoid multicollinearity (thus, independent variables partially containing the same information) and to ease model interpretation. However, some relevant variables may be excluded, reducing its capacity to explain the analyzed phenomenon. A sort of balance of this trade-off should be found.

Starting from the general model of equation (2.9), the following step of every meta-regression analysis involves identifying the best-fitting specification to the case study and the arising common econometric problems. In order to facilitate reading and outline the essential specification choices, these will be arranged in sub-sections based on the related issue.

Heterogeneity: FE vs. RE

It may seem obvious to state that the presence of heterogeneity within meta-data is the main reason for conducting a meta-regression analysis, given the possibility to explain such heterogeneity. Although the presence of excess heterogeneity in economic research may be taken for granted, the conventional meta-approach is to test explicitly for it (Stanley and Doucouliagos, 2012). This can be done both with the Q-test, either applying its formula or calculating the sum of squared errors of a MRA run with t-values on precision, unless it is widely known to have low power (Sidik and Jonkman, 2007), or with the I-squared statistic.[13]

In any case, the issue of heterogeneity is strictly related with the choice between fixed- and random-effects models. Indeed, in an ideal approach, such a choice should be made based on presence and type of heterogeneity: if no heterogeneity, simple FE model; if only explainable heterogeneity, FE model with moderator variables; if only unexplainable heterogeneity, simple RE model; if partially explainable heterogeneity, RE model with moderator variables (also known as ME, or mixed-effects, model). As already said (section 2.2.2), the latter suits and explains better the most cases of economic

[13]The information for the application of these statistical tests are provided, respectively, by Cooper and Hedges (1994) and Higgins and Thompson (2002).

phenomena; however, the well-known issue of correlation between independent variables and random term, which especially arises when a proxy for precision is included within the model (Stanley and Doucouliagos, 2012), makes this choice not so obvious. Only as long as the random errors are not correlated with the explanatory variables (thus, *i.i.d.*), the regression estimates will be asymptotically consistent (unbiased) and normally distributed (Davidson, 2000). Endogeneity can be verified with the Hausman specification test,[14] which compares the consistency of FE and RE estimators (Johnston et al., 2015; Stanley and Doucouliagos, 2012). However, Nelson and Kennedy (2009) state that: "the advantages of random-effects estimation are so strong that this estimation procedure should be employed unless a very strong case can be made for its inappropriateness", thus if the biases which it generates are modest, not important for the purposes of the specific MRA or are overtaken by a better fitness to data.

Heteroskedasticity

The term "heteroskedasticity" refers to heterogeneity of variances within a vector of random variables. Generally, this is the case of most meta-data: indeed, differences in sample size and estimation procedures of effect size at a primary study level (Nelson and Kennedy, 2009), as well as heterogeneity and variation in impacts of factors underlying single effects, may be the cause of heteroskedasticity. In regression analysis, if heteroskedasticity is present, the common OLS (*ordinary least-squares*) estimators are not efficient, although still unbiased, while GLS (*generalized least squares*) are also efficient (Goldberger, 1964).

There are different ways to detect heteroskedasticity: for example, with either the Breusch-Pagan test or the White test,[15] or by plotting fitted values versus regression residuals (if residuals become more spread out as fitted values increase, assuming a typical "cone" shape, this is a sign of heteroskedasticity). Also, there are different ways to account for it. A typical correction involves weighting effect size estimates in the meta-regression according to their estimated reliability, so that the most reliable ones have greater weight; this can be done either explicitly weighting the regression model or by implementing a WLS routine, using the chosen precision proxy as analytical weight. WLS (*weighted least-squares*) are a specialization of GLS[16] in which knowledge of the variance of observations is incorporated into the regression through a weight matrix.

[14] See Hausman (1978).
[15] See, respectively, Breusch and Pagan (1979) and White (1980).
[16] And a generalization of OLS.

Of course, the most appropriate weights would be standard errors or variance estimates, but they are not always provided by primary studies; in this case, some alternative solutions are possible (Nelson and Kennedy, 2009):

- approximating them with the *delta method*;
- constructing weights using estimates' t-statistics;
- proxying variances with sample sizes.

Another solution is employing heteroskedasticity-consistent standard errors (HCSE)[17] which, unless still biased, improve upon OLS estimates, or clustered standard errors (see the sub-section below), which are robust to both data correlation and heteroskedasticity.

Correlated estimates

Possible sources of correlation among primary estimates have already been discussed in section 2.2.3. As anticipated, there are three modeling techniques commonly used to deal with this issue (Johnston et al., 2015; Nelson and Kennedy, 2009; Stanley and Doucouliagos, 2012):

- *Cluster-robust modeling*: this means clustering multiple estimates by any dimension within which they're thought to be correlated (e.g., they're extracted by the same dataset or published by the same study or author). Once decided on the clustering variable, statistical packages will calculate cluster-robust standard errors. Usually, this does not significantly affect the results but standard errors and, consequently, t-values; therefore, some variables may not be statistically significant anymore.
- *Unbalanced panel modeling*: thus, building a model structured to account, within single regression, for the ith estimate of the sth study (or of the sth group: panels can be based on different criteria, not necessarily on individual studies; this amounts to say that estimates can be clustered with several possible grouping schemes, depending on the dimension within which they're correlated). Both random- and fixed-effects models can be used: if a FE regression is performed, it is possible to model it with a fixed intercept for each panel (but variating across panels) or with a fixed intercept for the whole regression and s dummy

[17]Also known as White or Huber-White standard errors. See White (1980).

variables for each group/study; whereas, the RE model provides a fixed intercept for the whole regression and s random error terms for each group/study, representing the unobserved study effect.

- *Hierarchical (or multilevel) regression modeling*: thus, building a model structured on different regression levels, corresponding to data which are also stratified at more than one level. In the discussed case, the coefficients of level 1 (estimate level) regression represent the overall impact of the related regressors for the ith estimate of the sth study/group and are obtained from level 2 (study/group level) regressions, of which they are the dependent variables. Within level 2 regressions, the impact of the regressors at the estimate level (i) and study/group level (s) on the dependent variable at level 1 are accounted for by separate specific coefficients. Multilevel models, differently from panel models, are in practice always based on random-effects principle, with all the well-known technical issues arising.

There is no good reason to prefer any one of the hierarchical/multilevel, panel or cluster-robust modeling; the most convenient, based on the research specificity and available data, should be used (Nelson and Kennedy, 2009).

2.2.5 Applications

Meta-regression analysis can provide several useful applications. Clearly, the most obvious, which has already been discussed, is quantifying impacts of contextual factors and research dimensions on the dependent variable, thus explaining heterogeneity, using regression coefficients' estimates. However, MRA's possibilities go far beyond this.

For instance, it is also possible to calculate a value for the dependent variable which represents the effect size corrected for heterogeneity and publication bias. This can be done by solving the model equation setting the precision proxy equal to zero and choosing a benchmark value for the other independent variables, based on specificities of the analyzed case study (Stanley and Doucouliagos, 2012).[18]

In economic research, these two applications lay the foundation for interpretation of socio-economic phenomena and policy making, gives us information about preferences and market trends, serve to test competing theories (through comparisons with the broader literature and primary studies themselves) and so on. But MRA can also be employed to guide and model new

[18]Not average values: in that case, the model would just give back the average effect size reported by primary studies, thwarting the research efforts.

research processes, by identifying gaps in empirical strategies and methodologies and suggesting variables to be included within studies which may have a significant and explanatory effect.

Very interesting is the application of MRA for *within-* and *out-of-sample forecasts*. Indeed, meta-regression models can be used to produce either within-sample predicted values for the dependent variable, by setting specific values for regressors, as done when calculating the corrected effects, or out of sample forecasts of effect size (Nelson and Kennedy, 2009; Stanley and Doucouliagos, 2012); in the second case, MRA coefficients are used to predict the likely direction of the relationship under investigation forward in time (e.g., for the next 5-10 years) or to do some inference in terms of space, using data from a certain region to infer values for other regions (V. Smith and Pattanayak, 2002; Stanley and Doucouliagos, 2012); this is also known as *benefit transfer* and it is very common in non-market evaluation to infer environmental values: the MRA model is adapted as a transfer function to calculate values for the dependent variable in contexts with different characteristics, which can be controlled with moderator variables.[19] Within- and out-of-sample prediction properties of meta-regression analysis can also be exploited to verify internal and external validity of single models.[20] Internal validity is tested based on within-sample forecasts and employing prediction statistics such as root mean squared error (RMSE), mean absolute percentage error (MAPE) and median absolute percentage error (MDAPE), which compare model-predicted values with observed values. Whereas, external validity, thus the capacity of the models to explain phenomena independently on their collocation in space and time, is trivially tested by comparing forecasted values for a certain year or region with studies conducted in that year or region, if available (Johnston et al., 2015; Stanley and Doucouliagos, 2012).

2.3 MRA in environmental economics

2.3.1 Overview

The use of meta-analysis in environmental economics has become increasingly common, following the general trend of this kind of analysis in economics. In this specific context, it mainly consists in a meta-regression

[19] An useful example of benefit transfer application of meta-regression analysis is provided by van Houtven et al. (2007).

[20] In general, the overall explanatory power of a meta-regression model is computed with the conventional R^2 or adjusted R^2.

analysis aimed to pool estimates from similar non-market evaluation studies, in order to, among others: (1) explain and quantify impacts of different independent parameters, which vary across studies, samples and sites, or relatively to commodities characteristics, on environmental values, (2) generate a predictive function of such values for a change in an amenity of interest, (3) employ this function for applications in benefit transfer studies (Johnston et al., 2015).[21] More technically, recalling regression models depicted in the previous sections with equations (2.4), (2.5), (2.6) and (2.9), on the left-hand side (dependent variable) we have an estimate from a wide variety of possible evaluation methods (such as WTPs from stated preferences, hedonic prices or wages, travel costs, elasticities and so on), while on the right-hand side we find the typical moderator variables explaining its variation (Johnston et al., 2015; Nelson and Kennedy, 2009).

As for the general case, the meta-analysis in environmental economics has several policy implications and it can be essential for decision making processes. For instance, it provides: more reliable and robust estimates of environmental commodities; further and more detailed information to decision makers relatively to trends and consumers' preferences, which can be exploited to address markets towards sustainable pathways; models allowing to make assessments in contexts were primary empirical analysis are for some reason not possible; and so on. Since the beginning of application of meta-analysis in economics (thus, since 1989), numerous studies have been carried out with this methodology. As of 2009,[22] out of the several hundred analysis already conducted, at least one-third belonged to the field of environmental and resource economics (Nelson and Kennedy, 2009).

2.3.2 Typical issues and challenges

The key issues of meta-analysis in environmental economics trace those of meta-analysis in general (refer to section 2.1.2). However, meta-analysis

[21]Meta-analysis provides more robust and accurate transfer functions than methodologies typically used in benefit transfer, as it summarizes and synthesizes estimates from different studies which actually differ for site or other contextual factors, including and controlling such aspects with specific variables. Consequently, results are also more statistically robust and less sensitive to characteristics and attributes of individual studies than methodologies that, for instance, build benefit transfer methods based on single studies (Johnston et al., 2015).

[22]These are the last available statistics about the amount of meta-analysis in environmental economics, and they come from Nelson and Kennedy (2009). However, in 2020, the updated version of the MAER-Net reporting guidelines (Havránek et al., 2020) found that between 2013 and 2020 more than two-thousand new studies had emerged within the broader economics.

(and meta-regression analysis) in environmental economics is particularly sensitive to specific aspects, one above all the identification of a consistent definition for the dependent variable. This issue occur across two key dimensions (Johnston et al., 2015).

The first one is linked to *commodity consistency* or, to be more precise, to environmental amenity definition. In environmental analysis, primary studies might refer to commodities apparently similar, but actually defined in substantially different ways. For example, when evaluating river health, this can be outlined and assessed in terms of healthy waterways for recreation, healthy vegetation, healthy birdlife, healthy fish stocks, and so on.[23] As in that case very different concepts are being measured, the computed values, whatever the method employed, are not commensurable and, therefore, cannot be pooled within the same meta-analysis. However, some solution are possible: the most common is including regressors explaining such differences.

The second issue is linked to *welfare consistency*. The problem arises when values based on welfare measures with different theoretical properties are pooled together. Consequently, the variations in the dependent variable reflects variations in such properties, more than in the value itself or in the characteristics of the environmental amenity. Often, this inconsistency comes from different valuation techniques or methods employed by primary studies (Nelson and Kennedy, 2009; V. Smith and Pattanayak, 2002). For example, including in the meta-analysis estimates from both revealed and stated preferences also means pooling Hicksian and Marshallian welfare measures. The appropriateness of pooling these measures is still the subject of debate (again, Nelson and Kennedy, 2009; V. Smith and Pattanayak, 2002). Other sources of welfare inconsistency are pooling willingness to pay (WTP) and willingness to accept (WTA) estimates and differences in payment vehicles or length of the payment period. In general, great attention should be paid when pooling measures with different underlying value components. Commodity and welfare inconsistency threat the commensurability principle which, as already said, is one of the foundations of a good meta-analysis.

Other challenges are related to typical characteristics of environmental commodities originating data heterogeneity (such as resource specific features, policy context, involved environmental issues and sites of interest, etc.) and data correlation. Thus, an in-depth knowledge of these characteristics is required, in order to select adequate moderator variables and corrections.

[23]This example refers to the work of Rolfe and Brouwer (2013), who conduct a meta-regression analysis on environmental valuation (based on choice experiments) of river health in Australia; they solve the depicted issue of commodity consistency as suggested here, thus by including within their model some dummy variables representing difference in definitions across studies.

2.3.3 Common features and criticalities

Nelson and Kennedy (Nelson and Kennedy, 2009), in a well-known and widely cited article, report a summary of 140 meta-regression analysis in environmental economics,[24] trying to categorize them, highlight the common features across the studies and evaluate their completeness. Based on their work, the purpose of this section is to define the outlines and synthesize the characteristics and criticalities of previous meta-regression analysis in environmental economics. The summarized features are represented in barplots.[25]

Primary estimates

Most of meta-analysis are based on welfare measures from stated preferences (33), hedonic prices or wages (22), or elasticities (20); also, many of them combine data from stated and revealed preferences (29): the appropriateness of this has already been discussed in section 2.3.2. **Figure 2.1**.

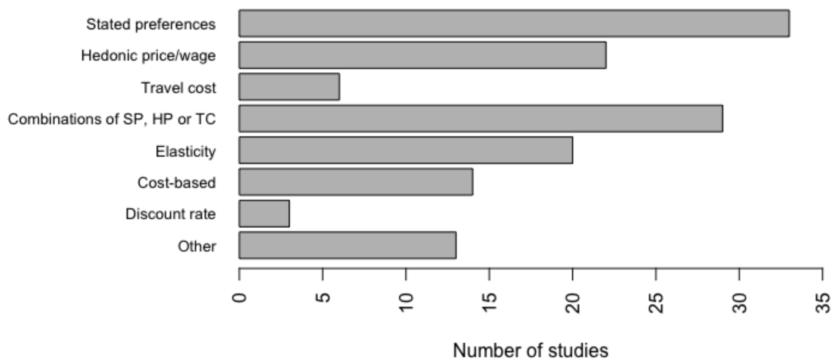


Figure 2.1: meta-regression analyses in environmental economics by primary estimate.

Primary data types

Survey-based data (50) are by far the most used, followed by various combinations of data types (42), which threat their consistency, or of pub-

[24]The authors' search results in over 300 meta-analysis studies in environmental economics published, or still unpublished, as of 2009. After their examination (for details please refer to the original article), 140 MRA were selected and included within the summary.

[25]*N.B.* in some categories, the count of the bar values sum to more than 140, since there are studies exhibiting more features simultaneously (multiple entries).

lic surveys and aggregate data (16), also raising some important issues of heterogeneity and non-independence (V. Smith and Pattanayak, 2002). In general, pooling data of different type or from different sources, casts doubts about commensurability of variables, measures, commodity definitions and context characteristics. **Figure 2.2**.

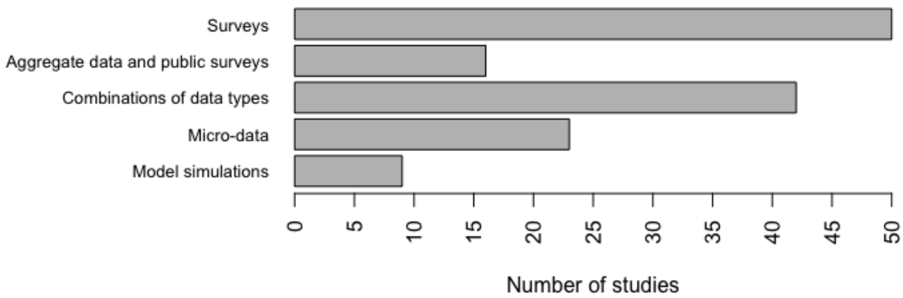


Figure 2.2: meta-regression analyses in environmental economics by primary data type.

Regression estimation

Regarding the estimation methods employed (**Figure 2.3**), 39 studies estimate an OLS model with White or Huber-White standard errors for heteroskedasticity; the OLS estimator is the most common, followed by the WLS (36). When weighted regressions are estimated, employed weights (**Figure 2.4**) vary across studies (5 of them use standard errors as weights, 11 use the variances, 13 use weights based on primary studies sample size; the remnants use other and less common weighting parameters). Some studies still estimate simple OLS. Also other estimation methods are used (e.g., maximum likelihood estimation in case of mixed-effects models), always depending on the choices made when specifying the model.

Heteroskedasticity and non-independence treatments

Heteroskedasticity is mainly dealt with White or Huber-White standard errors (50) and weighted regressions (44); some studies (7) use the Newey-West standard error, which is designed for stationary time-series data, thus not appropriate for meta-regression analysis and many others (46) do not report any treatment, which is also not recommended (**Figure 2.5**). Concerning data correlation (**Figure 2.6**), the most common treatments are: using single observation per primary study (30); and panel models (38), with

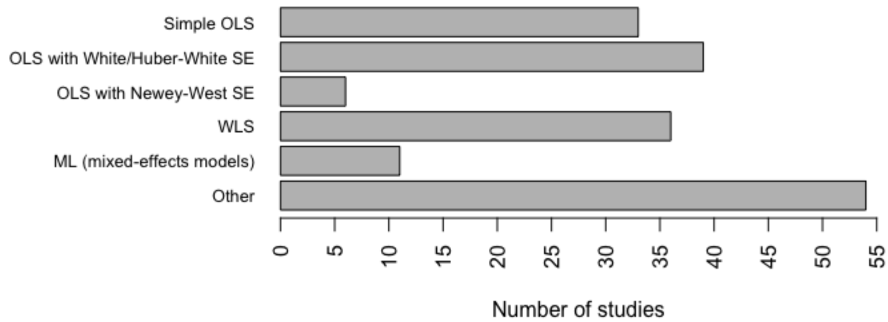


Figure 2.3: meta-regression analyses in environmental economics by estimation method.

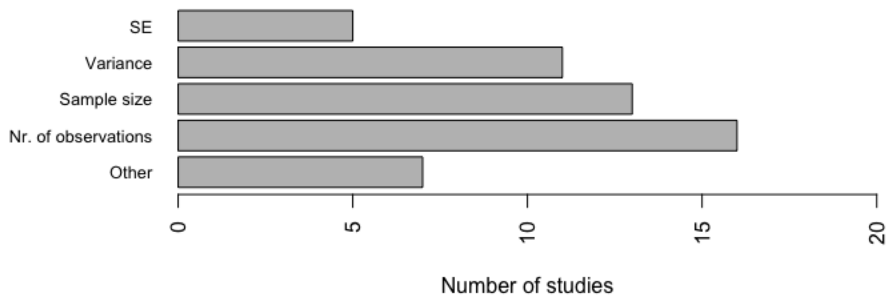


Figure 2.4: meta-regression analyses in environmental economics employing weighted regressions by weighting parameter.

a slight preference for RE (22) over FE models (16). Twenty-eight studies estimate multilevel models ore use other type of controls, whereas forty-seven of them do not report any treatment.

The preference for RE models confirms, as anticipated in sections 2.2.2 and 2.2.4, that RE model, unless potentially biased due to the widely discussed issue of correlation between random term and regressors, is still preferred, perhaps for its capacity to better fits data and explain heterogeneity.

Applications

Coming to the applications of estimated models (**Figure 2.7**), 21 studies exploit them as a transfer function for benefit transfer, which is the most typical application of MRA in environmental economics. However, most of studies (46) only provides within-sample predictions. Fifteen studies analyze and correct publication bias, and 17 compare its own results with other meta-

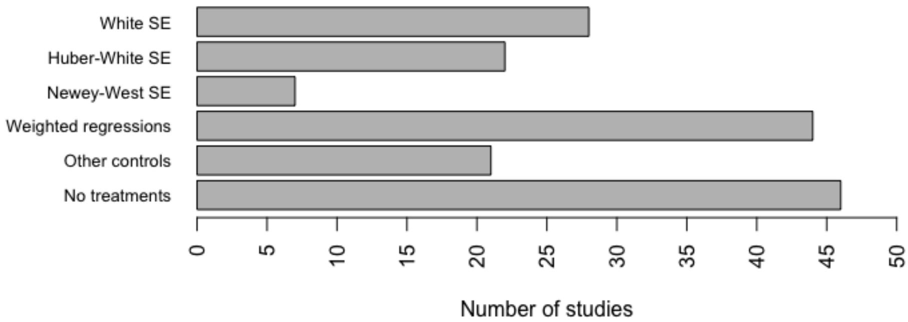


Figure 2.5: meta-regression analyses in environmental economics by heteroskedasticity treatment.

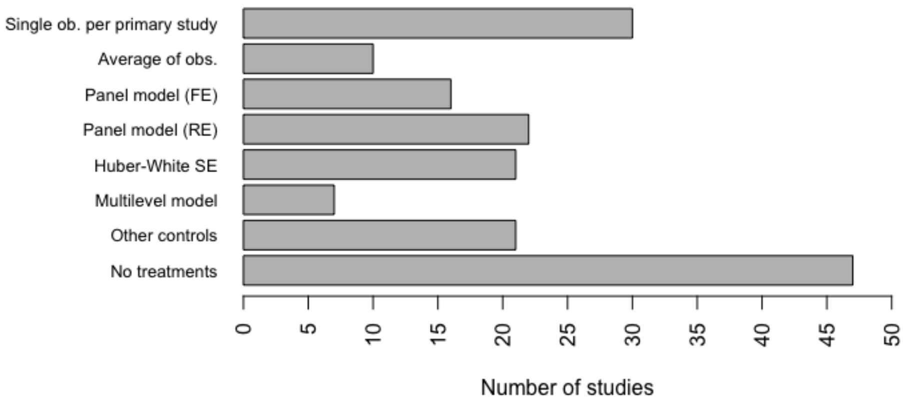


Figure 2.6: meta-regression analyses in environmental economics by non-independence treatment.

analyses. An high number of studies do not report any application (51).

Completeness score and final remarks

As a rough quantitative measure of overall quality or completeness of the analyses, Nelson and Kennedy code the attributes of the studies taken into account (thus, the characteristics mentioned above, as well as the employed tests, reporting procedures, etc.) and calculate a score for each one of them, ranging from 2 to 23. The resulting average score has a mean value of 9.6 ($SD = 4.0$, $median = 9$).

As suggested by the completeness score, the analysis of the common features of these 140 meta-regressions and the quantitative summary, there

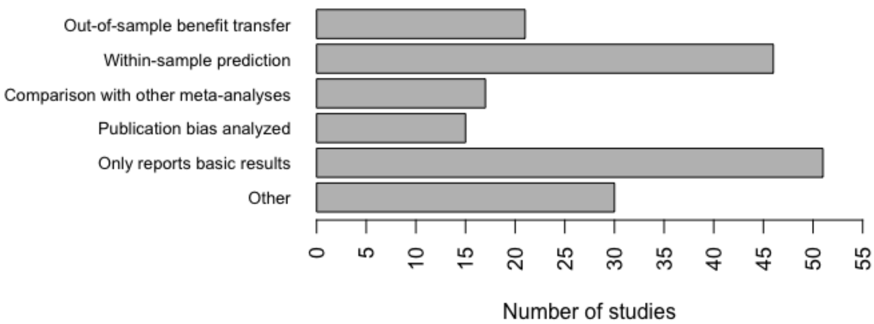


Figure 2.7: meta-regression analyses in environmental economics by application.

is great room for improvement for the meta-analysis in environmental economics. In part, this is due to "the relative newness of meta-regression analysis as part of the economist's tool-kit of econometric methods" (Nelson and Kennedy, 2009), both in this specific field and in economics in general. Methods employed by economists do not fully account for a variety of data-related issues and their analysis are often not complete or detailed enough.

To sum up and supplement what has been said in this section about MRA in environmental economics,[26] the major shortcomings to be solved are related to: consistency issues originated by pooling different welfare measures; commensurability of data and variables, heterogeneity and non-independence issues arising from using combination of data types or aggregate data and public surveys; incompleteness of the analysis (many studies lack of providing at least some primary data, selection criteria, statement about the model employed in primary studies, preliminary analysis such as data plots or weighted means, or additional tests such as outliers examination); insufficient use of formal test for model specification, heterogeneity, heteroskedasticity and so on; the neglect by some studies of possible heteroskedasticity or correlation issues and the consequential lack of corrections, given that, within the sample of studies analyzed by Nelson and Kennedy (2009), 46 out of 140 ignore possible heteroskedasticity, 47 ignore possible correlated observations and 21 ignore both problems; use of not adequate corrections for heteroskedasticity (like the Newey-West standard error); a not sufficiently high average degree of fit of the meta-regressions (calculated with adjusted R^2), which corresponds to a not sufficient explanation of data heterogeneity.

[26]The additional information provided in this paragraph also come from the summary of Nelson and Kennedy (2009).

2.4 MRA in this thesis

2.4.1 Research context

The meta-regression analysis conducted in this thesis builds on a systematic literature review carried out by Javier Cantillo, Juan Carlos Martín and Concepción Román[27] and presented in their article "Discrete choice experiments in the analysis of consumers' preferences for finfish products: A systematic literature review" (Cantillo, Martín, and Román, 2020). The authors *qualitatively* analyze and summarize 39 studies assessing consumers' preferences for finfish products and employing discrete choice experiments as evaluation method, with a particular focus on issues related to sustainability. The review aims to understand the main factors impacting the consumers' behavior towards these products, thus the drivers of their preferences, by screening the selected papers in order to individuate the most important product attributes, compare and synthesize the willingness to pay for each of them, identify peculiarities linked to geographical areas and species considered and exploit the summary of these findings to "obtain important insights for the industry and academics", thus both for policy making and future research. This thesis follows in the footsteps of this review, but performing a *quantitative* analysis with the so far described methodology of meta-regression analysis, in order to, not only synthesize WTPs estimates based on statistical foundations, but also explain differences across them, identify the most influential factors, control for research dimensions, location and characteristics of studies and products. The purpose is building a model as complete and comprehensive as possible, which can be widely applied to support decision making in fish policies and industries. Only the meta-data employed in this kind of analysis provides the variability and information needed to reach the set target.

Before continuing with the details about the analysis structure, the model and the specification choices, in the next chapter, a brief overview on choice experiments and their application within the context of the fish market, on the specific features of MRA of choice experiments and seafood products and on previous studies, is necessary.

2.4.2 Discrete choice experiments

Discrete choice experiments (DCE) are a stated preference method allowing to describe decision makers' choices among a set of alternatives (the

[27]At that time researchers at the Institute of Tourism and Sustainable Economic Development at the University of Las Palmas de Gran Canaria (Spain).

choice set) over hypothetical scenarios, and to collect data allowing to quantify them in monetary terms (Train, 2009); usually, this is a *willingness to pay* (WTP) or *willingness to accept* (WTA). DCEs are a popular tool in environmental economics and they are often employed for evaluation of non-market goods or attributes and value components of consumer goods, as in the case of their application to the fish market. Indeed, within the context of this thesis, decision makers are represented by consumers, expressing their individual preferences with consumption choices over an hypothetical market.

The economic theories underlying DCEs are the *random utility theory* (RUT) and *Lancaster's theory*. The random utility theory (Marschak, 1960; McFadden, 1974; Thurstone, 1927) suggests that individuals use to associate their choices to an utility level and act following the behavioral rule of utility maximization; thus, they choice the alternative which provides them the highest utility. Following the same principle, consumers, within a basket of goods, select and buy the product(s) whose characteristics offer them the highest utility. On the other hand, the Lancaster's theory (Lancaster, 1966) complement the RUT as theoretical foundation of DCEs stating that any product or service can be described using a set of attributes.

Choice experiments are based on surveys specially designed, in which respondents face a definite number of choice sets composed by a definite number of alternatives. In each of these alternatives, the product or non-market good is described in different ways. Indeed, a set of attributes (e.g., origin, environmental label, water quality, wildlife population, distance to site, depending on the type of product), whose levels vary across alternatives and choice sets,[28] is employed to describe products or environmental amenities assuming different combinations; the only attribute that must necessarily be present is the price. In each choice set, respondents can choose one option, score or rank them, or indicate the best and the worst, depending on the study approach; they are requested to repeat such action in the following choice sets. The number of choices, alternatives and attributes is determined based on the specific design and research purpose of single experiments.

Since the first of the mentioned approaches (choice of an option), which is what can be properly defined as choice experiment, is "unequivocally consistent with the underlying theory of welfare economics" (Pearce and Özdemiroğlu, 2002) and known to be the most common in environmental economics, as well as the only one employed by primary studies of the meta-

[28]For example, the attribute "origin" may vary from "domestic" to "imported" or "made in E.U."; similarly, the attribute "water quality" may vary on a scale from 1 to 5 and the attribute "wildlife population" may assume the value of 5, 10 or 15 deers.

analysis conducted in this thesis, the discussion will henceforth focus exclusively on it.

Therefore, be i the respondent facing J alternatives ($j = 1, 2, \ldots, J$) described by a set of K attributes ($k = 1, 2, \ldots, K$). The utility he (or she) obtains from alternative j is U_{ij}. Recalling the RUT, respondent i will eventually choose alternative j if, and only if, $U_{ij} > U_{iq}\ \forall q \neq j$. Recalling the Lancaster's theory, instead, it can be affirmed that the utility U_{ij} derive from the specific combination of attributes of alternative j. Thus, such utility can be expressed with the formula:

$$U_{ij} = \beta_0 + \beta_1 x_j + \beta_2 (y_i - C_j) + \varepsilon_{ij} \tag{2.10}$$

where x_j is a vector of K attributes ($k = 1, 2, \ldots, K$) of alternative j, y_i represents the income of individual i, C_j is the cost of alternative j, β_0, β_1 and β_2 are the coefficients and ε_{ij} is an error term capturing factors which affect utility, but that cannot be described by the included variables. Since utility, for its definition, underlies individual choices but cannot be directly observed (Louviere, Flynn, and Carson, 2010), the error term ε_{ij}, which describes this unobservable part of utility, is random. Therefore, the utility can also be expressed in terms of a deterministic component, V_{ij}, and a stochastic one, ε_{ij}. Thus, equation (2.10) can be simplified, becoming:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{2.11}$$

This model is also called *random utility model* (RUM). On the basis of what has been said above, the probability that individual i chooses alternative j is:

$$\begin{aligned} P_{ij} &= Prob(U_{ij} > U_{iq}\ \forall q \neq j) \\ &= Prob(V_{ij} + \varepsilon_{ij} > V_{iq} + \varepsilon_{iq}\ \forall q \neq j) \\ &= Prob(\varepsilon_{iq} - \varepsilon_{ij} < V_{ij} - V_{iq}\ \forall q \neq j) \end{aligned} \tag{2.12}$$

thus, the probability that each random term, $\varepsilon_{iq} - \varepsilon_{ij}$, is below the observed quantity $V_{ij} - V_{iq}$. Since, as mentioned, ε_{ij}, which can be also simply defined as the difference between true utility, U_{ij}, and the part of utility that the researcher can observe and capture in V_{ij}, is random, its characteristics, as well as its distribution, critically depend on the researcher's representation of the choice situation (Train, 2009). Technically, the specification of the joint density of the vector of random terms of each choice, $f(\varepsilon_i)$, while calculating the choice probability (2.12), depends on the discrete choice model employed (e.g., conditional logit, mixed logit, probit, etc.); thus, on the researcher's assumptions on the unobserved part of utility and, more broadly, on his interpretation of choice dynamics and behavioral patterns of individuals within the population I.

For example,[29] the *conditional logit* (CNL)[30] assumes the random terms to be *i.i.d* extreme values for all i, thus exhibits the property of independence from irrelevant alternatives. Although this assumption provides a very convenient form for the choice probability, it is not consistent when specific substitution patterns exist.[31] In these cases, it is more appropriate to opt for the *mixed logit* (ML), which allows the unobserved factors to follow any distribution, therefore being more flexible and accounting for correlation effects (Cantillo, Martín, and Román, 2020; Train, 2009).

Once the data from the choice experiment are collected and the model is estimated, it is possible to compute the marginal WTP (mWTP), thus the price premium respondents are, on average, willing to pay for the presence of a certain attribute or for an increase in his level, as the negative ratio of the attribute's coefficient to the price coefficient (Train, 2009). The mWTP can be defined as "the marginal rate of substitution between the attribute and the price attribute in the indirect utility function" (Mariel et al., 2021) and, therefore, it can be specified, for each alternative j, as:

$$mWTP_{x_{jk}} = -\frac{\partial V_j / \partial x_{jk}}{\partial V_j / \partial C_j} \tag{2.13}$$

where x_{jk} refers to the relative attribute. Based on a careful examination of equation (2.13), it can be stated that the mWTP is nothing but a monetary quantification of the utility that consumers gain or loss for the presence of attribute k within the chosen product, given the price they have to pay. If the attributes enter utility linearly, the mWTP, as anticipated, can be calculated as (Mariel et al., 2021):

$$mWTP_k = -\left(\frac{\beta_k}{\beta_C}\right) \tag{2.14}$$

[29]For a more complete and in-depth discussion on discrete choice models, reference is made to Train (2009).

[30]Often referred to as *multinomial logit* (MNL). Indeed, in literature the two terms are used interchangeably, although, from an econometric point of view, a distinction should be made. In the MNL, the utility depends on the characteristics of individuals (e.g., gender and age), while the CNL allows choices to be explained by the characteristics of alternatives, i.e., the product attributes (Mariel et al., 2021). Therefore, "conditional logit" is the correct term to label the model employed within DCEs since, for the intrinsic features of this method, it is the only one able to explain respondents' choices.

[31]Another criterion followed by researchers is mathematical conveniency and consistency with data.

where β_C is the price coefficient.[32] This measure is extremely useful in environmental and market evaluation and in policy making. Indeed, although being sensitive to the choice model employed, experimental design, data collection and sample characteristics, DCE is firmly grounded on a well-known, commonly accepted and tested behavioral theory and on econometric models very familiar to analysts. This theoretical robustness ensure its capacity to represent reality in a sufficiently accurate way and to provide the means for its interpretation. Therefore, it can be successfully used to capture consumers' preferences, market trends and, in general, individuals' attitude towards specific topics, such as sustainability. These information, expressed in monetary terms, can be exploited to drive markets and behaviors in the desired or appropriate direction.

2.4.3 MRA on discrete choice experiments

In section 2.3 it is stated that meta-regression analysis in environmental economics has its proper characteristics, issues and challenges, and such aspects are briefly analyzed. Similarly, it can be stated that a meta-regression analysis focusing exclusively on effect sizes from discrete choice experiments (to be specific, on WTP) is an even more special case, having as well its own features and key aspects.

First of all, recalling the discussion at the beginning of section 2.3.1 and equation (2.9), a MRA on DCEs can be described as a regression analysis having on the left-hand side (dependent variable) a WTP (or WTA) and on the right-hand side (independent or moderator variables) some kind of regressors capturing the characteristics of experiments and commodities underlying single estimates. These are: variables coding the presence (or the description) of attributes linked to the related WTP (X_i), usually dummy variables; regressors depicting the specific characteristics of the product or environmental amenity (P_i), which vary depending on the study and/or estimate; variables capturing other contextual factors (such as the study year and location, K_i), study design and methodological features (R_i), and sample characteristics (S_i), such as average income, age or education level. This way, equation (2.9) can be slightly modified, becoming:

[32] The mWTP can be calculated using either the mean value of the price coefficient or its median value; in the first case the *mean mWTP* is obtained, whereas in the second case the result is the *median mWTP*. Median mWTP is usually employed when the price coefficient has a log-normal distribution, which is very frequent in random parameter logit and, in general, mixed logit, since it represents a more reliable central tendency measure for this kind of distribution (Mariel et al., 2021).

$$WTP_i = \beta_0 + \beta_1 precision_i + \alpha_1 X_i + \pi_1 P_i + \delta_1 K_i + \gamma_1 R_i + \omega_1 S_i + e_i \quad (2.15)$$

where I $(i = 1, 2, \ldots, I)$ is the population of studies[33] and e_i is the error term.

Of course, variables coding study design and methodological factors are strictly related to the evaluation method employed by primary studies, therefore here resides one of the main difference distinguishing a MRA on choice experiments from any other performed in environmental economics. Indeed, within the vector of regressors R_i, can be included, among the others, the moderator variables coding

- the *choice model* employed: in addition to the already mentioned conditional logit and mixed logit, many others are commonly used in DCEs, such as the specifications of mixed logit, like the random parameters logit, generalized mixed logit and latent class model, the random effect logit or the probit model;
- if respondents have received an *information* regarding insights of the commodity or the attributes taken into account, which may have influenced their choices;
- whether the surveys employ *hypothetical* or *non-hypothetical* methods (such as the so-called "real choice experiments", where participants actually buy the chosen product(s));
- if the surveys have been *implemented online*, since such delivery method can affect WTP estimates, although findings on this point and on the wider topic of the appropriateness of online questionnaires are inconsistent and conflicting across literature;
- the number of *choice sets*, *alternatives* and *commodity's attributes*.

The importance of taking into account the above enlisted study characteristics mainly resides in the existence of some bias inherent to choice experiments and well-known in literature. Surely, the most discussed is the "hypothetical bias": use of hypothetical methods may cause declared WTP to be different to actual WTP, as a consequence of an inconsistent behavior of respondents due to the hypothetical nature of the experiment (Hensher, 2010;

[33]Here is still considered the simplest case where each study produces only one WTP estimate.

Lusk and Schroeder, 2004; Pearce and Özdemiroğlu, 2002); previous meta-analyses found out that WTPs from stated preference methods (thus, including DCEs) can be up to two or three times higher than WTPs from revealed preference methods (e.g., Murphy et al., 2005). To mitigate this problem, different solutions have been implemented across studies. Indeed, in addition to the mentioned real choice experiments, several possible options, both *ex-ante* (cheap talk scripts, honesty oaths, training, use of real-life settings like supermarkets, in case of DCEs on food products) and *ex-post* (statistical techniques of data screening, to individuate implausible responses) can be applied. It is also well established that providing information about the valued good, its features or production process may affect respondents' attitudes ("information bias"), as well as their WTP (Ajzen, Brown, and Rosenthal, 1996; Mariel et al., 2021; Yeh, Hartmann, and Hirsch, 2018). Some DCEs provide this kind of information during the experiment (or split the sample in "informed" and "uninformed") purposely to investigate the impact on consumers' choices of a better knowledge of the product or the related environmental issues (e.g., Alfnes et al., 2006; Bronnmann and Asche, 2017; Bronnmann and Hoffmann, 2018). Other frequently mentioned bias are: the "country-of-origin bias" (Yeh, Hartmann, and Hirsch, 2018), highly correlated to ethnocentrism; and the ones triggered by the design dimensions of the experiment (number of choice sets, alternatives, attributes and levels), which may imply fatigue or learning effects (Lusk and Schroeder, 2004), if long and complex questionnaires are delivered, and influence the frequency of status quo choices (Oehlmann et al., 2017), especially when an opt-out alternative is present.

Coding the presence or description of the attributes to which single WTP estimates refer is also a distinctive feature of MRA on DCEs. This can be done in different ways, depending on the type of analysis and WTP employed as dependent variable. For example, if a meta-regression on marginal WTPs for a common attribute of primary studies is conducted, this kind of moderator variables (above indicated as X_i) code the difference in the description of that attribute or its levels. Alternatively, a meta-regression analysis can be conducted, rather than on marginal WTPs for an attribute, on total WTPs for a product, whose value encompasses the market price of the product plus the price premium due to the presence of *all* or *some* of the exclusive attributes. In this case, the moderator variables indicate, precisely, the presence or absence of the attributes taken into account within the product and, therefore, within the estimated value, with the aim to quantify their impact on the WTP.

One further challenge when performing a MRA on choice experiments concerns the dependent variable itself, thus the WTP. Like any monetary measure, it is influenced by the purchase power and the inflation rates specific to single countries and years, as well as by the currency. Therefore, WTP estimates from different countries, years and currencies must be led back to a comparable measure, which one can assume expresses the same value, once eliminated such influences and fluctuations. Although different methods exist to adjust WTP values, a viable alternative is using the *purchase power parities* (PPPs).[34] These are defined by the OECD as "the rates of currency conversion that try to equalize the purchasing power of different currencies, by eliminating the differences in price levels between countries" (OECD, 2023) and can be employed to convert monetary values in a common currency (US dollar), by simultaneously adjusting them for the purchase power. The obtained values, expressed in PPP$, must, at a rate, be adjusted to the inflation rates for a base year. Then, values can be re-converted in a different currency, if needed, using the conventional conversion rates.

A challenge which also concerns the dependent variable is about proxying the accuracy of the estimates, thus finding a parameter to be included as the variable generally indicated as $precision_i$ in equation (2.15). When it is not the primary study itself to provide the estimates' standard error or confidence interval, computing it without full access to the employed dataset or, at least, to the regression coefficients' covariance is not possible. In these cases, the accuracy can be proxied with the sample size (thus, the number of respondents) or the number of observations (usually calculated as $sample\ size \times number\ of\ choice\ sets$).

Last issue to be mentioned is linked to the functional unit. Different choice experiments may employ different functional unit to express the WTP for the attributes of the product or environmental amenity which is the object of their analysis. By way of example, DCEs evaluating water quality of rivers may measure it in terms of kilometer of healthy river, percentage of healthy river or percentage of river length in good health; consequently, the WTP estimate(s) will be likewise expressed in, e.g., dollars per kilometer of healthy river, percentage of healthy river or percentage of river length in good health, depending on the study.[35] Coming to the fish market, instead, WTP for, say, the presence of an eco-label may be expressed in terms of price premium per kg, package of 250g, can, number of slices, number of sushi rolls, restaurant dish and so on. This is a problem when conducting a MRA on these studies, since, as mentioned several times, estimates must be comparable.

Some meta-regression studies on food products (e.g., Dolgopolova and

[34] For example, Bastounis et al. (2021) employ this method for their meta-analysis.

[35] This is the case of the already cited study (see note 23 in section 2.3.2) from Rolfe and Brouwer (2013); the authors converts all estimates in dollars per kilometer of healthy river.

Teuber, 2018; Printezis, Grebitus, and Hirsch, 2019; Yeh and Hirsch, 2023) solve the problem by transforming marginal WTP in *percentage WTP* (sometimes, *%WTP* or *WTPP*), with the following formula:

$$\%WTP = \left(\frac{WTP_{product\ with\ attribute\ k} - product\ base\ price}{product\ base\ price} \right) \times 100 \quad (2.16)$$

where $WTP_{product\ with\ attribute\ k}$ denotes the price for the product with the considered attribute (thus, the base price plus the mWTP for attribute k). Percentage WTP indicates the price premium (relatively to the base price of the product), *in percentage*, that consumers are willing to pay for a certain attribute; as it is expressed in relative terms, the obtained value is comparable across studies, this allowing to avoid, at the same time, issues linked to currency, measure and functional unit. Another option may be converting product packages in their corresponding weight, so as to be able to calculate the WTP per one kilo, one pound, etc., with a simple ratio. In any case, meta-analyst must find a scientifically robust method to lead back the different functional units to a common one.

2.4.4 Attributes and specific features of seafood products

Of course, when setting up and performing a meta-regression analysis, the specific research context and characteristics of the object of study, in addition to those of the study itself and the employed methods, also count and influence the regression model and structure.

As discussed in the previous section (2.4.3), part of the moderator variables of a MRA, when primary studies are DCEs or, generally, based on stated or revealed preferences methods, code for the presence and/or description of the product's attributes included within the experiment, survey or analysis. After reviewing several studies (mainly choice experiments) on consumers' preferences on seafood products, a list of the most common and relevant attributes, some of whom also reported by Cantillo, Martín and Román (2020), was made, with the focus on sustainability. The list is given below.

- *Origin*: within a globalized market, were commodities are commonly traded, imported and exported, the origin of the product can affect consumers' choices, also when taking into account seafood, with an expected preference for local products. Moreover, it is correlated with sustainability issues like those deriving from transportation. The location where finfish and other species are farmed or captured is very frequently included in choice sets as the attribute "domestic" (if it accounts for the difference among products harvested in the country of study and imported ones) or "local" (if it refers to products harvested within the specific area or region of study). Sometimes the attribute about origin is included in terms of the presence of a label certifying the provenance, the so-called "*country-of-origin* label", or COOL.[36]

- *Harvest method*: this attribute, when referred to seafood, mainly accounts for whether a wild captured or a farmed product is preferred; in this case, the expected outcome is a preference for wild fish. During the last decades, aquaculture and fish farming in general has exponentially increased its production volume, and this trend is still ongoing. Thus, nowadays it flanks traditional fishery within the fish market, with both of them arising sustainability and health issues but also presenting several opportunities.

- *Production method* and *organic*: exactly because of the just mentioned sustainability issues, production methods alternative to conventional aquaculture are spreading. Examples are the IMTA (*integrated multi-trophic aquaculture*), the CCA (*closed-containement aquaculture*) and the well-known, and already employed in agriculture, organic production.

- *Sustainable certification label* or *eco-label*: still talking about sustainability, the environmental issues related to fishery and aquaculture led to the emergence of several certification schemes and linked labels. The best known and most common, worldwide, are the MSC (Marine Stewardship Council) certification, which only applies to fishery products, the ASC (Aquaculture Stewardship Council) certification which, instead, only applies to aquaculture products, and the Naturland certification. In some studies, this attribute is expressed in terms of a generic eco-label (with two levels, yes/no or certified/not certified) or sustainability claim; this difference in format can potentially affect responses. In general, a preference for certified products is expected.

- *Nutrition*, *health* or *safety label*: environmental issues and mass production also arise uncertainties concerning nutritional benefits of seafood products, as well as their healthiness and safety. Consequently, some

[36]An example of study focusing on the COO label is Yeh and Hirsch (2023).

products are marked with labels or claims regarding their nutritional properties (usually, related to the contents of omega-3 fatty acids), health benefits (e.g., for brain or heart function) or food safety (e.g., absence of contamination from chemicals).

- *Other trustworthy labels*: the evolution of food market and related issues have resulted in an increase in number and types of certifications. Their presence on products may improve trust of consumers towards them, this also affecting their willingness to pay. The most common label for food market, excluding eco-labels, is the fair-trade label, which certifies the respect of rights of workers and communities involved in supply chains, according to the standards of the homonymous association. Other trustworthy labels are choice labels, quality labels, local or authenticity labels; there are also labels that are related to animal welfare, geographical indication and genetically modified organisms (GMO).

- *Product form*: seafood, in particular finfish, can be sold in various forms. The fillet and loin are the noblest and most prized cuts, for which the highest willingness to pay is expected, but fishes can also be sold in form of steak, tail cut, round cut, slices, sushi or as a whole fish.

- *Product presentation*: In addition to fresh seafood, in supermarkets consumers can usually find frozen, smoked, canned or ready-to-cook products. Packagings can also differ. Both product form and presentation have an impact on sustainability, since they lengthen supply chains and increase industrial processes.

- *Type of feed*: fish feed is also involved in the mentioned environmental issues related to aquaculture. Also, production of fish feed is becoming increasingly difficult and expensive, due to the geo-political events affecting supply chains and the pressure on resources. Therefore, some producers are experimenting alternative types of feed, like plant-based feed, insect-based feed and feed based on algae or other micro-organisms.

Other attributes which are present within the fish market are also common in other food sectors. These are, for instance, *product size*, *place of purchase* (whether in a specialized store or a supermarket) and *brand* (whether the product is marked with a specific brand or supermarkets' own brand).

The listed attributes can be included in choice experiments from primary studies, depending on the focus of the analysis, which allows to calculate the related WTP. However, when modeling a MRA on such WTPs, some of them are also coded as product characteristics (thus, as the moderator variables above indicated as P_i); in fact, not all primary studies consider, for instance, both fresh, frozen and smoked or both wild and frozen fish product, nor they take into account all the possible cuts and presentations. A single DCE could not provide information about the impact of these features on consumers' choices if they are not included as attributes, as these are study-invariant. A meta-regression analysis, instead, by coding them as product characteristics, exploits the variation across studies to obtain such informations. This is an example of the *value-added variables* mentioned in section 2.2.3. Another essential product characteristic to *necessarely* be coded when performing a MRA concerning any effect size within the context of fish market, including WTPs from choice experiments, is the fish species; this is rarely a proper attribute, but can both vary within (if the estimates for attributes are provided for different species) and across studies, as the variable "country of study" does. It is possible to limit the analysis to finfish species[37] or to extend it to the whole seafood market, then also including mollusks, shellfishes, crustaceans, echinoderms and other marine animals (e.g., jellyfishes).

2.4.5 Previous applications of MRAs on (sea)food products

In this section, some previous MRA studies carried on within the broader food sector and focusing on WTP as effect size are explored and exploited as examples of the different possible approaches.

Analysis set-up and dependent variable

When discussing the specificities of a MRA conducted on choice experiments (section 2.4.3), it was briefly explained that different kind of set-up exist, depending on the type of WTP employed as dependent variable.

First, the analysis can be conducted on a common attribute across primary studies. For example, Printezis, Grebitus and Hirsch (2019) apply meta-regression analysis to a set of 35 research papers providing, among the others, estimates on consumers' WTP for local food. A wide variety of product types is taken into account, including meat from different animals and with different cuts, fruit, vegetables, dairy products, beverages (both alcoholic and non-alcoholic), bakery products, sweets and ice cream; only a

[37] As done by Cantillo, Martín and Román (2020) and in the meta-analysis of this thesis.

seafood product is included (tilapia). Also, different evaluation methods are considered: mainly DCEs (86% of the studies), but also other methods such as auctions and CVs. Marginal WTP (or WTPs, in case of studies reporting multiple estimates for more than one sample or product) for an attribute referable to the concept of "local", although described in different ways, is (are) drawn from each primary study and selected as dependent variable. The authors made the original choice to employ a "double" dependent variable, by defining the mWTP measure both in terms of standardized \$/lb and percentage WTP, computed with equation (2.16). Coming to independent variables, in addition to those coding for study design, evaluation method, contextual factors and sample characteristics, several dummies are included to code for the food category to which the related estimate refers (animal product, produce or processed product) and the definition of "local" (if it is defined as grown or produced within the state or a specific region/province, as locally grown or produced, or using a state/region logo/label). The computed regression coefficients quantify the impact of socio-economic factors (age, income, gender), country and year of study, econometric model, experiment design and attribute definition on consumers' mWTP for local food; these are valuable results for policy making. Moreover, by setting the precision proxy (in this case, the square root of the number of participants) equal to zero (thus, correcting for the publication bias), choosing benchmark values for the other independent variables and adding the constant, the authors are able to employ their model to calculate a synthesized and correct value for such marginal WTP; starting from this, they also make some within-sample predictions, experimenting different combination of study dimensions.[38] Yeh and Hirsch (2023) also investigate consumers' preferences on the origin of food products, but in terms of *country-of-origin label* (COOL), thus the label indicating their provenance and addressing issues of traceability and safety. The authors perform a meta-regression analysis on 59 primary studies which focus on identifying the mWTP(s) for the presence of the attribute COOL on analyzed products by following similar principles as in the study mentioned above. Indeed, they also employ as dependent variable the percentage WTP for COOL, include moderator variables coding for the attribute definition (if the estimates refer to COOL in general, domestic country-of-origin or foreign country-of-origin, if the foreign country concerned is a developing or developed country, and so on) and exploit the estimated model to calculate a correct value for the effect size and to make within-sample predictions. However, the set of independent variables employed is more detailed in accounting

[38]This kind of applications are better and generally explained in section 2.2.5.

for study location (dummies are included for almost each continent),[39] experiment design (number of attributes, levels of both price and COOL attribute, choice sets, alternatives and blocking versions are coded) and survey method (web survey or personal interview). Moreover, differently from the previously summarized MRA, primary studies are only selected among DCEs. Given such an in-depth analysis, the authors' choice not to code for sample characteristics, in order to quantify the impact of, say, income or education of consumers on their preferences, seems rather peculiar. Also in this case a wide variety of food products is considered. Another example of MRA conducted on mWTPs for a common attribute is from Bastounis et al. (2021), who focus on the influence of environmental sustainability labels on willingness to pay for food. As in the last case, included studies are only DCEs and several types of food products are taken into account, although seafood products are more represented (11). However, this study differs from the previous ones because the mWTP, as dependent variable, is expressed in terms of PPP$/kg rather than in percentage terms: the authors adjust values to price levels, inflation and currencies exchange rates (see section 2.4.3) and standardize them to a single measurement unit (per kilogram). The set of independent variables is very essential, and mainly code for attribute description (content of sustainability labels, thus if they refer to organic production, sustainable production or both; if they are proper certifications or just generic claims; and format, thus if they are in form of a text, a logo or both) and sample characteristics. This because the authors only focus on the impact of these two categories of factors on consumers' preferences, in confirmation of the importance of the research question and aim when modeling the structure of a meta-regression analysis.

To sum up, when focusing MRA on a specific attribute, the dependent variable will be the mWTP estimate for that attribute and the regression coefficients of moderator variables will quantify the variation of such price premium depending on primary studies' features, including, typically, attribute description. The estimated model can be solved for the dependent variable by choosing appropriate values for regressors, so to calculate a synthesized and correct value for the effect size or to test different scenarios.

Secondly, recalling again section 2.4.3, a meta-regression analysis can also be conducted on total WTPs for a product, rather than on marginal WTPs for a common attribute. An useful example comes from Smetana, Melstrom

[39]This also depends on the geographical areas taken into account and the focus of the analysis; for instance, the analysis of Printezis, Grebitus and Hirsch (2019), to which the comparison is made here, only include a dummy variable indicating whether the study is conducted in the US or not, but this is justified by the fact that the authors aims to point out the difference between U.S. consumers and consumers from the rest of the world.

and Malone (2022), whose analysis is closely related to the one of this thesis, since it examine variation in willingness to pay for seafood, although only focusing on farm-raised (aquaculture) products and also including non-DCE studies (to be more specific, CV and HV[40] studies) and species different from finfish. As said, their MRA differ from the three aforementioned studies, on a methodological level, because it employs as dependent variable the total WTP for products in place of the marginal WTP for single attributes. The authors calculate it from the data provided by primary DCE studies using the following formula:[41]

$$tWTP_{prod} = \frac{\sum \beta_k x_k}{\beta_C} \tag{2.17}$$

where x_k includes all the relevant k attribute measures and the alternative-specific constant[42] for the product line, and β_k and β_C are the coefficients from the RUM.[43] It can be noticed that this is a generalization of equation (2.14), in which multiple attributes, as well as the value of the original product, can be taken into account to compute a *total* WTP. For primary studies which do not provide the necessary data (coefficients) and directly report values for the mWTPs, the authors calculate the total WTP by summing the “base price” of the product, obtained either from the average price level of the experiment or through a web search based on the year and country of study, and the reported mWTPs, thus using the formula:

$$tWTP_{prod} = b + \frac{\beta_k}{\beta_C} \tag{2.18}$$

where $\frac{\beta_k}{\beta_C}$ is the mWTP for a particular attribute. If this kind of set-up is chosen for the analysis, the product attributes captured by the total WTP must be coded as moderator variables. Equations (2.17) and (2.18) allow to collect up to $K + 1$ product WTP from each primary study, depending on the focus of the analysis: this means that it is the researcher itself to select

[40]Hedonic valuation, thus the valuation method based on hedonic prices.

[41]Data transformation and coding is described in detail by the authors themselves in section 2.2 (pp. 482-484) of their publication.

[42]The alternative-specific constants (ASCs), are constant parameters included for each alternative which assume value 1 if the related alternative is chosen and 0 otherwise (Klaiber and Haefen, 2019). Within an utility function, their estimated coefficients represent the utility and, then, the value, that individuals assign to that choice and is not captured from the attributes; in this sense, it has the role of regression intercept (e.g., Khan et al., 2019). Therefore, the ASC for the product line here mentioned is interpreted as the utility of the base product and employed to calculate the base price component of the total WTP.

[43]See equations (2.10) and (2.11).

the attributes to be encompassed by the WTP estimate. Therefore, given that such measure is "artificially built" for research purposes and does not represent a real product, the dependent variable, unlike the previous studies, cannot be interpreted with economic significance. Instead, the mentioned moderator variables, coding for the presence of certain attributes within the product WTP, can be interpreted as the correct and synthesized value of their marginal WTP, whereas the other independent variables quantify, as in any MRA, the impact of methodological or contextual factors on the willingness-to-pay. Going back to the case of study, the authors include and code 8 attributes from 45 studies, thus: fresh, local, domestic, live (some non-finfish species are sold alive), processed (which is a catch-all attribute indicating if the product is sold prepared or pre-cooked in any way), environmental certification (which encompasses different eco-labels), IMTA certification and home consumption (relative to away-from-home consumption, i.e., a restaurant). As the analysis takes into account both finfish and non-finfish species, two dummy variables are included for bivalve and crustaceans. The variables describing study characteristics are the conventional ones.

Thirdly, and finally, a MRA can be conducted on mWTPs for different attributes, but imputable to the same concept and, therefore, pooled together. For example, the analysis from Li and Kallas (2021) aims to quantify the price premium that consumers are willing to pay for sustainable food (meat, fruit, vegetables, dairy products, seafood and drinks are taken into account), generally intended. From each primary study of their database (80 studies) they extract a (or more) mWTP estimate(s) for an attribute describing the concept of "sustainability"; these are: local, environmental friendly (EF), fair-trade, organic, animal welfare. Such estimates, transformed in percentage WTPs, are jointly employed as dependent variable of the meta-regression. Of course, moderator variables coding for the sustainable attribute corresponding to each estimate are included, in addition to those accounting for study features, sample characteristics and type of food.

Not too differently from the first three studies, the computed coefficients of the moderator variables quantify the impact of the way the attribute is specified (although in this case they are different aspects of sustainability, rather than different definitions of the same attribute), as well as the research, contextual and socio-economic dimensions on the mWTP. Moreover, likewise in the first and the second case of study, the authors use the estimated model to calculate an "overall" WTP premium for sustainable food, achieving the research purpose stated at the beginning of their paper.

The analyzed possible set-ups are all equally valid and viable when planning the meta-regression analysis. The choice among them should only be driven by the final aim of the study, in confirmation of the importance of the research question when modeling and structuring the analysis itself.

Data heterogeneity

As mentioned in the previous sections, different degrees of data heterogeneity can be introduced within an MRA. Within the summarized cases of study, this depends on the variety of food types, valuation methods and countries taken into account.

Starting from the food types, all the analyzed studies which consider more than one food type (Bastounis et al., 2021; Li and Kallas, 2021; Printezis, Grebitus, and Hirsch, 2019; Yeh and Hirsch, 2023) have a very heterogeneous sample, jointly encompassing meat, fish, dairy products, beverages, sweets and so on. Consequently, all of them also include within their MRA a set of two or more moderator variables to account for such diversity. Coming to Smetana, Melstrom and Malone (2022), which focus the analysis only on farmed seafood products, as already pointed out the authors consider both primary studies on finfish species and on non-finfish species, such as molluscs, crustaceans and others. This aspect is very important for studies focusing on such sector of food industry. Although including very different species allows to better describe and understand purchasing behaviors, it is also true that, as highlighted by Nguyen et al. (2015), consumers' preferences between species vary significantly: both in the sense that they assign different intrinsic values to different species and in the sense that "a specific seafood attribute may be a quality cue for a given seafood alternative but not for all seafood species". This may be even more exacerbated when comparing finfish and other seafood species, as they are products which radically differ for taste, cooking methods and diffusion across regional markets. As a result, it is possible that including both these types of products within the analysis, excessively increase heterogeneity level, compromising its capacity to explain data or making it necessary to extend the model adding more regressors, which not always improves it. The R^2 from Smetana, Melstrom and Malone (2022) is, indeed, relatively low for three out of the four models employed, and the regression coefficients itself indicate that consumers valuate very differently bivalves, crustaceans and finfish species.

Similar considerations can be done in regard to valuations methods and countries taken into account. Two of the summarized MRAs (Li and Kallas, 2021; Printezis, Grebitus, and Hirsch, 2019) include both DCEs and studies conducted with other stated preference methods (such as CVs), whereas one of them (Smetana, Melstrom, and Malone, 2022) also include estimates from revealed preference methods (HVs). Concerning the countries, the analysis from Yeh and Hirsch (2023) is worldwide (as it includes studies conducted in North America, Europe, Asia, Oceania and Africa, although not equally distributed) and same can be said about the analyses from Li and Kallas (2021) and Bastounis et al. (2021) (in which Africa is substituted by Central America). Printezis, Grebitus and Hirsch (2019) include studies from North and Central America, Oceania and Europe, whereas Smetana, Melstrom and Malone (2022) only include North American and European studies. While pooling estimates from different valuation methods arises the welfare consistency issues discussed in section 2.3.2, regarding the inclusion of more or less countries, applies the same as for the inclusion of more or less seafood species. Taking into account all continents allows to carry out an in-depth analysis of differences in consumers' preferences at a global level, which is consistent with the existence of a globalized market, but narrowing it down to few countries or to a continent simplify data explanation and, sometimes, gives more reliable and useful results for policy making.

In conclusion, as for every MRA, a trade-off exist between data heterogeneity, number of observations and explanatory power. These aspects have already been addressed in section 2.2.3.

Models, data selection and econometric aspects

The structure of meta-regression analyses conducted within the food sector, as in the general case, is roughly the same. The studies usually begin with an introduction, providing the framework for the analysis. Then, both the method employed by primary studies (e.g., DCE) and for the MRA are explored; within this section, all the information related to regression form, precision proxy, heteroskedasticity and correlation treatments, etc., are provided. Next, the search strategy for data collection is explained,[44] the sources are cited, a summary of descriptive statistics of meta-data is given and, in most of cases, a list of primary studies accompained by basic information and moderator variables with their definition is included. If performed, results from preliminary analyses and tests (including publication bias corrections) are shown and, finally, also results from the MRA are presented and discussed. Among the summarized studies, all of them exhibit such features, except for the one from Bastounis et al. (2021), who do not clearly specify the regression model employed.

As discussed in sections 2.2.2, 2.2.4 and 2.3.3, the possible model specifications for a MRA are many. A single meta-regression study may employ more than one, in order to test different modeling strategies, identify which of them best fits and explains data and highlight the relative advantages

[44] Usually, this is done by following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and templates (Moher et al., 2009).

and disadvantages. For instance, Smetana, Melstrom and Malone (2022) use four different model specifications, i.e., they run four MRAs. The first one is specified as a conventional OLS regression which, as known, is rarely the best choice for this kind of analysis but, however, can be presented as the benchmark case. The second model, instead, employs a WLS estimator, using the square root of sample size as the analytical weight, to account for the difference in estimates' precision. The authors also create an additional specification by modifying the latter by further weighting on the inverse of the number of estimates from each study, so that each of them has the same weight and proportionately influences the parameters. Finally, a hierarchical version of the model, using random-effects regression, is estimated. With the OLS and WLS specifications cluster-robust standard errors are reported for the parameters. As can be noted, the issue of data correlation is accounted with both cluster-robust and hierarchical modeling, while heteroskedasticity is also accounted with cluster-robust standard errors and with the employment of WLS estimators. The model that best fits data is the modified WLS (thus, the third one), with an R^2 of 0.493, followed by, in order, the OLS model ($R^2 = 0.297$), the random-effects hierarchical model ($R^2 = 0.240$) and the simple WLS specification ($R^2 = 0.182$), which, instead, show a low level of fit to the data.

Some meta-regression studies, on the other hand, focus on a single and more simple model specification and, in parallel, use other meta-analysis methods and econometric tests to extract information from collected data, synthesize them and explain their variability. This is the case of Li and Kallas (2021), who choose a basic WLS specification for the only employed model[45] which, however, is part of a wider and complete analysis. Such analysis includes: an assessment of primary studies quality and risk of bias, according to the criteria provided by the Cochrane Collaboration (Higgins, 2011); a check for publication bias (whose presence is confirmed) with a funnel plot and the Egger's test;[46] an outlier analysis; a subgroup analysis of seven subgroups created based on the publication year, food type, sustainable attributes, valuation methods, continent and socio-demographic characteristics, which identifies the mean percentage WTP and explores the level of heterogeneity of each group.[47] Despite of what said above, the MRA shows a good degree of fit to data, with an adjusted R^2 of 0.82.

[45] They upgrade such a basic model by conducting a Monte Carlo permutation test to reduce the type I error and improve the accuracy of p-values.

[46] The Egger's regression test (Egger et al., 1997) follows the same principles of the FAT-PET-PEESE discussed in section 2.2.2.

[47] The heterogeneity level is tested with the I-squared statistic, mentioned in section 2.2.4.

With regard to the other cited studies, Printezis, Grebitus and Hirsch (2019), likewise Smetana, Melstrom and Malone (2022), employ in parallel several specifications for their MRA. To be more specific, they specify three models, using for all of them the WLS as estimator, but different types of standard errors; respectively, heteroskedasticity robust standard errors, cluster-robust standard errors and bootstrapped standard errors. In addition, for each regression, they estimate two versions of the model, one using the sample size and one using its square root as the analytical weight. Thus, in this study data correlation and heteroskedasticity are accounted with cluster-robust modeling; indeed, the authors focus on the second model for their discussion, as cluster-robust standard errors address both issues. However, all the employed specifications show a relatively high goodness-of-fit ($0.386 < R^2 < 0.413$). Yeh and Hirsh (2023) also make the same choice to rely on cluster-robust modeling, but they only estimate two WLS models employing two different clustering criteria for the standard errors, with the first based on primary studies and the second on study authors. The selected analytical weight for the MRAs is the square of the inverse of the square root of the sample size. Also in this case the R^2 is high (0.52). Finally, as mentioned above, Bastounis et al. (2021) do not clearly specify the model specification employed. In any case, they conduct two separate meta-regressions to account for, respectively, attribute (eco-labelling) features ($R^2 = 0.0745$) and socio-demographic characteristics of the samples ($R^2 = 0.5641$). Coming to the preliminary analyses and tests, Printezis, Grebitus and Hirsch (2019) and Yeh and Hirsch (2023) check for publication bias, which is confirmed in the first case and rejected in the second one, both with the graphical method (funnel plot) and the FAT-PET-PEESE regression. Yeh and Hirsch (2023) also report average relative WTPs from several subgroups based on the moderator variables of the MRA, likewise Bastounis et al. (2021) who, instead, create subgroups for three pre-specified food cathegories and also carry out a quality assessment, according to the checklist of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) (Bridges et al., 2011; F. Johnson et al., 2013).

It can be noted that even just a sample of five MRA studies shows a great variety of possible model specifications and employable meta-analysis techniques. As said when discussing about analysis set-up and the choice of the dependent variable, there is not a reason to prefer a research strategy over another, since everything depends on the framework of the analysis and the available data. The aim of the modeling phase should be avoiding known econometric issues and increase as much as possible the degree of explanatory power of the meta-regression. Of course, the integration of the main analysis with tests and applications of other methods from the wealth of tools of meta-analysis, if well managed, helps the analysts in their specification choices and improve the results and clearness of the study.

2.4.6 Findings from previous MRAs on (sea)food products

If in section 2.4.5 some previous MRAs conducted within the food sector served as examples of the different possible applications from a "technical" point of view, in this section the main findings from the same MRAs will be summarized, in order to better define the research context of the present analysis and to be able, in a second step, to make a comparison of results. Special attention will be paid to sustainability topics and the finfish products.

Sustainable attributes

Starting right from the consumers' perceptions and attitudes towards sustainability, defined in general terms, in the food market, Li and Kallas (2021) show that the overall price premium they are willing to pay for products with sustainable attributes is about 29.5%. Recall here that the attributes the authors include in the framing of "sustainability" and, therefore, within their model are: local, environmental friendly (EF), fair-trade, organic and animal welfare. The results from this analysis also show that, among such attributes, fair-trade, which they use as the benchmark variable, is the most valued by consumers. Following, in order, organic, animal welfare, local and EF. However, it has to be specified that the coefficients of organic and animal welfare are not statistically significant. This first result seems to indicate that consumers assign more "weight" to the social component of sustainability, than to the environmental one. In addition, it should be noted that the WTP for seafood, in general, is one of the lowest among food types, together with drinks.

The study from Smetana, Melstrom and Malone (2022) is the only other cited one that takes into account a wider definition of sustainability, thus includes more sustainable attributes, although mostly different from those of the analysis of Li and Kallas (2021). Moreover, the authors, as known, only focus on seafood products. Coming to the results, they show that consumers are willing to pay 4.28$ more per pound (or, equivalently, 8.97€ more per kilogram)[48] for fresh fish, 3.20$/lb (6.70€/kg) more for domestically harvested seafood, 4.25$/lb (8.90€/kg) more for eco-labelled products,

[48]Within this section, monetary values are converted in €/kg using simple annual average exchange rates for the study year. Source: European Central Bank (ECB); rates available at `https://sdw.ecb.europa.eu/browse.do?node=9691296`.

4.66\$/lb (9.75€/kg) more if an IMTA certification is present, and 7.42\$/lb (15.54€/kg) less for live seafood (this is a common product presentation for some crustacean or bivalve species). Processed and local attributes are not statistically significant. Another interesting finding from this study regards seafood species: bivalves and crustaceans are, on average, more valued than finfish species.

The studies from Yeh and Hirsh (2023), Printezis, Grebitus and Hirsch (2019) and Bastounis et al. (2021), instead, focus on a single sustainable attribute, as mentioned in the previous section. Respectively, they are: country-of-origin labelling (COOL), local and eco-labelling. By setting benchmark values for the variables of the estimated model, Yeh and Hirsh (2023) calculate that the price premium average consumer is willing to pay for the presence of a domestic COO label on plant-based products is about 71%. Going into detail, consumers from Oceania and Africa are willing to pay 39% more; the value raises by 28% for animal-based products and by further 87% if US consumers are considered. As expectable, the presence of a foreign, rather than domestic, COOL has a negative impact on WTP. In general, these values are much higher than the one resulting from the simple mean of the included WTP estimates (24%), this demonstrating the importance of controlling for heterogeneity and study features with regression methods. Similarly, the analysis from Printezis, Grebitus and Hirsch (2019) results in a marginal WTP for local products varying within a range of 1.70\$/lb to 2.08\$/lb (3.35€/kg to 4.10€/kg, or 41.4% to 52.2%). For studies conducted before 2011 on non-US consumers, using non-hypothetical methods, the mWTP for local produce assumes the value of 1.89\$/lb (3.72€/kg); focusing on the US increases this value by 0.79 (1.56), while considering studies post 2011 decreases it by 0.81 (1.59). Consumers, on average, prefer processed products (with a marginal impact on WTP of 1.55\$/lb, or 3.05€/kg), whereas the way the attribute local is defined and described on labels and packages has not statistical significance. Finally, Bastounis et al. (2021) do not provide any within-sample prediction or synthesized value for the WTP, but they show that it is higher for eco-labels carrying the message "organic" than for eco-labels about environmental sustainability in general, and that their format (if text or logo) and type (if certifications or not) has no influence on purchasing behaviors.

To sum up, it can be stated that results from previous meta-regression studies on food market show a positive and relatively high WTP for sustainable products, indicating that consumers, generally, are concerned by sustainability issues and they are willing to pay more for products whose production process addresses such issues. Moreover, the results seems consistent across studies in terms of the sign of the marginal WTP, although its magnitude varies, mainly because of different meta-data employed or study boundaries. The discussed results are summarized in **Table 2.1**.

MRA	Product	Attribute(s)	mWTP (€/kg)	%WTP
Yeh and Hirsh (2023)	Meat, seafood, cereals, dairy products, fruit, vegetables, eggs, wine, beer, other	COOL (domestic)	-	71%
Printezis, Grebitus and Hirsh (2019)	Meat, fish, bakery products, dairy products, fruit, vegetables, eggs, beverages, sweets, other	Local	3.72	42%
Li and Kallas (2021)	Meat, seafood, dairy products, fruit, vegetables, wine, beer, coffee, other	Local, EF, fair-trade (overall)	-	29.5%
Smetana, Melstrom and Malone (2022)	Seafood from aquaculture	Fresh, domestic, eco-label, live, IMTA cert.	8.97, 6.70, 8.90, 15.54, 9.75	-

Table 2.1: average mWTP or %WTP for sustainable attributes as reported by previous meta-regression analyses on food products, using benchmark values. Only statistically significant values.

Country of study

It has already been said that focusing on the US market increases the WTP for local food of 1.56€/kg, according to Printezis, Grebitus and Hirsch (2019), and that consumers from North America are willing to pay more than consumers from Oceania and Africa for the presence of a domestic COO label on the products they purchase (Yeh and Hirsch, 2023). Results from the latter also show that the marginal WTP for the analyzed attribute increases, relatively to the African or Oceanian market, by 67% if European studies are considered and by 107% if Asian studied are considered. On the other hand, findings from Smetana, Melstrom and Malone (2022) are in contrast with the ones from Printezis, Grebitus and Hirsch (2019), as they indicate that US consumers' WTP is by 2.22\$/lb (4.65€/kg) lower than non-US consumers. Li and Kallas (2021), instead, confirm the primacy of North America, Europe and Asia over Oceania, with the first three continents having very similar marginal values.

As said for the findings on sustainable attributes, the difference in results among the discussed MRAs, which in this case also involves the sign, in addition to the magnitude, of marginal WTPs, is due to different research boundaries and primary studies collected.

Year of study

The year in which the primary studies were conducted is taken into account by three analysis out of the five summarized in this literature review. However, only one of them estimates a statistically significant coefficient. Indeed, at the beginning of the present section it was mentioned how the results from Printezis, Grebitus and Hirsch (2019) show an increase of the marginal WTP for local food of 1.59 €/kg in studies carried out after 2011. In Li and Kallas (2021), instead, the estimated value is very small and not significant; the same can be said for Smetana, Melstrom and Malone (2022) although, in this case, two model specifications employed but not considered for the discussion (as their goodness-to-fit was not sufficiently high) report a significant and negative impact of the year of study on the dependent variable.

Socio-demographic characteristics of the sample

As to the socio-demographic characteristics of the sample (**Table 2.2**), some of the considered MRAs point out a significant impact of gender on the WTP. Female consumers, generally, have an higher WTP than male consumers in Smetana, Melstrom and Malone (2022) (0.17\$/lb more, or 0.36€/kg more, for sustainable seafood), Li and Kallas (2021) (46.7% more for sustainable food) and Bastounis et al. (2021) (28.25\$/kg or 23.89€/kg more, for eco-labelled food). On the other hand, Printezis, Grebitus and Hirsch (2019) report a negative value, although very small (0.04\$/lb less, or 0.08€/kg less, for local food). The findings about education level are, instead, scarce. Only Bastounis et al. (2021) estimate a statistically significant coefficient, which imply a marginal WTP of -28.81\$/kg, or -24.36€/kg, for consumers with an undergraduate or higher degree. Interestingly, the average age of the sample is not significant in any of the summarized MRA studies. Also the income, which is expected to have a major impact on consumers' WTP is only significant in Smetana, Melstrom and Malone (2022), with a positive estimated marginal value of 0.07\$/lb, or 0.15€/kg. However, it should also be said that collecting good quality data on respondents' income is very difficult because of a general reluctance to share this kind of information and that the relatively low amounts involved for the purchase of this kind of products may justify a reduced impact of income. Moreover, sometimes the income variable is correlated with other regressors causing multicollinearity problems, this leading to exclude it from the models.

Variable	Yeh and. Hirsh (2023)	Printezis, Grebitus and Hirsh (2019)	Li and Kallas (2021)	Smetana, Melstrom and Malone (2022)	Bastounis et al. (2021)
Gender	Not reported	− 0.08 €/kg	+ 46.7%	+ 0.36 €/kg	+ 23.89 €/kg
Education	Not reported	Not reported	Not significant (−)	Not significant (+)	− 24.36 €/kg
Age	Not reported	Not significant (−)	Not reported	Not reported	Not significant (+)
Income	Not reported	Not reported	Not significant (+)	+ 0.15 €/kg	Not significant (+)

Table 2.2: impact on mWTP or %WTP for sustainable attributes of socio-demographic characteristics as resulting from previous meta-regression analyses on food products. When not statistically significant, the sign of the estimated coefficient is reported in parenthesis. Gender: female over male.

Study features

The selection of study characteristics to include in a MRA widely varies, depending on the specificities of primary studies and research question. Here, an attempt was done to categorize some of them (**Table 2.3**).

Starting from the delivery method of the questionnaire, which is likely to influence the WTP estimates, as stated in section 2.4.3, a differentiation can be made between online (or web-based) surveys and in-person surveys. Smetana, Melstrom and Malone (2022) report that, if the questionnaire is delivered online, the marginal WTP for sustainable seafood increases by 2.14$/lb, or 4.48€/kg; whereas, Bastounis et al. (2021) report a conflicting result, with the value increasing by 3.48$/kg, or 2.94€/kg for in-person surveys. Yeh and Hirsh (2023) consider the variable within their MRA, but the estimated coefficient is not statistically significant.

Concerning the different impact of hypothetical and non-hypothetical methods, which is also expected to be relevant (again, see section 2.4.3), actually it is not significant in any of the meta-regression studies taking it into account, thus the ones from Smetana, Melstrom and Malone (2022), Li and Kallas (2021) and Printezis, Grebitus and Hirsch (2019).

Another study feature which is expected to impact on WTP estimates is the valuation method employed. The only cited MRA which consider such variable is the one from Printezis, Grebitus and Hirsch (2019). The authors show that studies conducted with the DCE method report, on average, a marginal WTP of 2.00$/lb higher, or 3.94€/kg higher, than studies carried out with other methods. Such result is consistent with previous studies (e.g., Garcia, De Magistris, and Nayga Jr, 2012; Grebitus, Lusk, and Nayga Jr, 2013).

Yeh and Hirsh (2023) conduct an in-depth analysis on the influence of the way a choice experiment is structured on the estimated WTP. They find out that it increases as the number of attributes, levels and alternative increases. Moreover, the WTP estimates are higher if an opt-out option is included within the choice sets. This is in line with Lusk and Schroeder (2004), who state that these kind of survey features may affect respondents' behavior. Also Printezis, Grebitus and Hirsch (2019) include the number of attributes as moderator variable within their models, but the estimated coefficient is not statistically significant.

2.4.7 Characteristics and added value of this thesis

In the framework depicted in the last sections falls the meta-regression analysis performed in this thesis. However, it differs from the previous sum-

Variable	Yeh and. Hirsh (2023)	Printezis, Grebitus and Hirsh (2019)	Li and Kallas (2021)	Smetana, Melstrom and Malone (2022)	Bastounis et al. (2021)
Online	Not significant (+)	Not reported	Not reported	+ 4.48 €/kg	− 2.94 €/kg
Hypo-thetical	Not reported	Not significant (−)	Not significant (−)	Not significant (+/−)	Not reported
Valuation method	Not reported	3.94 €/kg	Not reported	Not reported	Not reported
Year of study	Not reported	− 1.59 €/kg	Not significant (−)	Not significant (−)	Not reported
Nr. of attributes	+ 25%	Not significant (−)	Not reported	Not reported	Not reported

Table 2.3: impact on mWTP or %WTP for sustainable attributes of study features as resulting from previous meta-regression analyses on food products. When not statistically significant, the sign of the estimated coefficient is reported in parenthesis. In Smetana, Melstrom and Malone (2022) results from different specifications are sometimes conflicting. Valuation method: DCE over other methods. Year of study: in Printezis, Grebitus and Hirsh*(2019) and Li and Kallas (2021) it is coded as a dummy variable indicating if the study was conducted, respectively, after 2011 or before 2008; in Smetana, Melstrom and Malone (2022) it is coded as a discrete variable corresponding to the exact year. Number of attributes: in Yeh and Hirsh (2023), it is coded as a dummy variable = 1 if the attributes are more than 3 and = 0 otherwise; in Printezis, Grebitus and Hirsh*(2019), it is coded as a discrete variable corresponding to the exact number of attributes.

marized studies for two main reasons. First of all, as of now and to actual knowledge, MRAs focusing on the sustainable attributes in the specific market of finfish species, both considering wild and farmed products, are not available in literature.[49] Indeed, informations about this particular sector of the seafood industry are, at the moment, only synthesized and reviewed by qualitative studies or can be inferred from broader and more general analyses, where the specificities of finfish products are not coded in the regression models and, therefore, the resulting knowledge of consumers' preferences towards them is not accurate.

Secondly, the present MRA considers a very comprehensive variety of sustainable attributes; thus, those directly related to sustainability issues and those more generally connected to production methods, but also having an impact on the latters. These attributes are separately analyzed and, then, compared.

To sum up, our strategy consists, on one hand, in narrowing down the analysis, by building an homogeneous sample not only in terms of the mentioned products characteristics, but also in terms of country of study (only European countries are taken into account) and evaluation methods (primary studies all employ DCE), this aiming to obtain more reliable results, which can also be more informative and applicable in policy making; on the other hand, in expanding aspects of sustainable production accounted for, in order to have a wide overview over it, although in a limited market.

These concepts will be deepened and further discussed in the following chapters.

[49]For example, Smetana, Melstrom and Malone (2022), who also focus on seafood and whose analysis share several primary studies with the analysis of this thesis, actually only take into account products from aquaculture (i.e., farmed products) and also include species different from finfish.

Chapter 3

Data and methods

3.1 Analysis overview and structure

The analysis of this thesis, as widely explained in the previous chapters, focuses on exploring consumers' preferences and attitudes towards sustainability of finfish products. In accordance with the same principles of the reference qualitative study for the present work, thus the literature review of Cantillo, Martín and Román (2020), the data chosen as drivers of these aspects, to be synthesized and integrated within the meta-analysis, are willingness to pay (WTPs) from discrete choice experiments (DCEs). Therefore, only marginal values of WTPs for sustainable attributes of finfish products, obtained with such specific valuation method, part of the broader family of the stated preference methods, were included in the final dataset. However, differently from the cited literature review, this analysis only considers European studies. The choice to focus research on a specific product of seafood market and to narrow it down to a single continent has already been explained and is motivated by the purpose of capturing the specificities of consumers' attitude towards a well defined commodity[1], in a geographical area characterized by more homogeneous, though still varied preferences. This aiming to produce more concrete and valuable results for decision making.

The meta-analysis was structured as follows: relevant papers were collected, both from the list of studies employed by Cantillo, Martín and Román (2020) and a new search on the SCOPUS database; after a selection process based on the guidelines of the updated version of the PRISMA[2] statement

[1]Including the totality of seafood products would have excessively expanded the study boundaries and, moreover, preferences vary significantly between them; this involving the intrinsic value assigned to each species, as well as the type and the impact of the relative product attributes (Nguyen et al., 2015).

[2]"Preferred Reporting Items for Systematic Reviews and Meta-Analyses".

(Page et al., 2021), data on mWTP estimates were extracted through a full text review of papers and the moderator variables coded; the data were treated to avoid (or, at least, mitigate) multicollinearity issues;[3] summary statistics were calculated; an outlier analysis, based on the Rosner's test, was conducted; the Breush-Pagan and the White tests for heteroskedasticity, as well as the Hausman specification test, were also conducted; two intercept-only regressions were estimated using the weighted RE and WLS cluster-robust specifications and the grand-mean for the sample was provided; a complete check for publication bias was carried out with both the graphical method (funnel plot) and the FAT-PET-PEESE regression; finally, the meta-regression analysis was performed; and the results reported and discussed.

Since two different meta-regressions were estimated, the summarized analysis was conducted twice identically.

3.2 Collection of primary studies

As mentioned just above, the relevant papers reporting the mWTP measures of interest, to be included in the MRAs, were collected from two sources. First, the list of studies employed by Cantillo, Martín and Román (2020), whose search[4] also exploited the database of SCOPUS, focusing on the time period ranging from 2000 to 2019 (thus, at the time of their work, the last 20 years), was screened, and only those suitable to the present analysis for contents, quality, reported data and country of study, were considered. This primary search resulted in 18 papers included within the final dataset, thus the majority of the total number. Then, a new search was conducted,[5] again on SCOPUS, to cover the years between 2020 and 2023. The search output consisted in 99 records; the abstract of all of them was screened, and 79 studies were excluded for several reasons. Of the 20 remaining eligible articles, only one was finally added to the dataset. Therefore, the definitive list amounts to 19 primary studies. Since the data were further skimmed when coding the variables of the MRAs, for totally random patterns one of these studies was not employed for the analysis. The criteria driving data

[3]This was done employing correlation matrices and calculating the variance inflation factors (VIFs) to detect concerned variables.

[4]For more information on the search of Cantillo, Martín and Román, reference is made to their article.

[5]The boolean terms employed for this new search are very similar to those used by Cantillo, Martín and Román; they are: (fish OR aquaculture OR seafood OR (farmed AND wild) OR salmon OR tilapia OR trout OR turbot OR seabass OR seabream) AND ((choice AND experiment) OR (discrete AND choice) OR (stated AND choice) OR (conjoint AND analysis)) AND (consumer OR public).

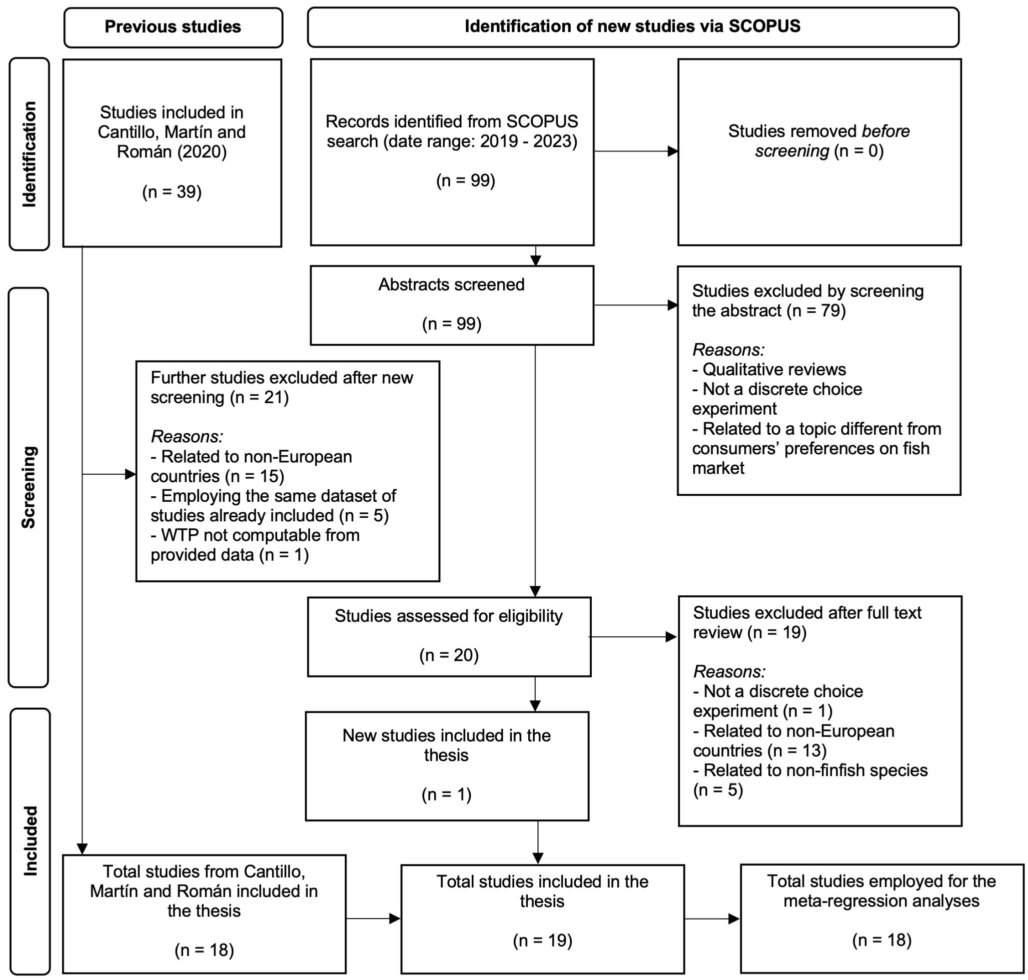


Figure 3.1: the PRISMA flow chart showing the search and selection process of this analysis.

selection were the same for both the screening of the papers from Cantillo, Martín and Román and the new search: studies not using DCE, focusing on non-European countries or on seafood species different from finfish, were excluded. The described search process is reported using the flow chart in **Figure 3.1**, which is based on the PRISMA template. Each step of each phase is individually represented, and the related number of studies involved, as well as the reasons of exclusion, are also indicated.

The actual list of studies, complete with information on study features, species and product characteristics considered (harvest method, form and presentation), country of study, sample size and MRA in which they are included can be found in **Table A.1** of **Appendix A**.

3.3 Statistical tests and preliminary analyses

The MRA was preceded by a series of statistical tests and additional analyses, aimed at supporting specification choices, cleaning data and providing some basic information about them and on potential bias.

For starters, as recommended in any statistical analysis, a test for outliers was conducted in order to detect and exclude from the dataset values not belonging to the distribution which would have skewed models estimation and calculations in general. The tool chosen for this task was the Rosner's test (Rosner, 1975, 1983): given k suspected extreme values within a sample of n observations, the Rosner's test compute several statistics R_k, representing the highest deviates from the simple mean of the sample, by progressively reducing its size from n to $n - k + 1$; then, it compares the obtained R_k statistics with the critical values, based on the Student's t-distribution, to conclude how many and which of the suspected extreme values are actually outliers.

The other statistical tests performed have already been mentioned in the methodology review, section 2.2.4. These are: the Breush-Pagan test and the White test (Breusch and Pagan, 1979; White, 1980), to check for heteroskedasticity; and the Hausman specification test, to determine the appropriateness of the employment of RE models (again, see section 2.2.4). The Breush-Pagan test and the White test build on the same principle, that is, fitting an auxiliary regression in which the response variable is the vector of squared residuals of the main regression and the explanatory variables are those from the original model. If the residuals are somehow dependent on the regressors, this meaning their variances are heterogeneous, the null hypothesis is rejected and heteroskedasticity is assumed. The test statistics are asymptotically distributed as a chi-square with degrees of freedom equal to the number of parameters of the auxiliary regression. Since the White test also detects non-linear forms of heteroskedasticity, both the tests are conducted on the OLS version of the estimated models. They were also flanked by a visual analysis of residual plots. Concerning the Hausman test, instead, it is essentially an hypothesis test evaluating the consistency of an estimator relatively to an alternative one. Its test statistic is based on the difference between the variances of coefficients of the alternative specifications, in this case the FE and RE models, and is distributed as a chi-square with degrees of freedom equal to the number of coefficients. If a difference is found out, this is because RE estimates are inconsistent due to the bias produced by the endogeneity of one or more independent variables.

In addition to the enlisted statistical tests, two preliminary regressions were performed. One was specified as an "intercept-only" model, thus without explanatory variables, and estimated with weighted RE[6] and WLS cluster-robust methods, using for both of them the square root of the sample size as the analytical weight. These specifications are equivalent, respectively, to the RES and FES models outlined in section 2.1.3 (Hox, Moerbeek, and Van de Schoot, 2017); therefore, the computed intercept coefficients, representing the grand-mean of the sample, should approximate the values of the RES and FES weighted means, giving first synthesized estimate of consumers' mWTP for sustainable attributes of finfish products. These estimated values were further corrected for publication bias by performing the second of the preliminary regressions, thus the one from the FAT-PET-PEESE procedure, which was also complemented by funnel plots. Since the procedure was discussed in the previous chapter (section 2.2.2), here it is only mentioned and the information about model specification is provided: this is a WLS cluster-robust, because of the obvious heteroskedasticity characterizing the regression. The chosen proxy for precision is the square root of the sample size, which also serves as the analytical weight, likewise in the case of the intercept-only models.

3.4 Meta-regression analysis models and methods

3.4.1 "Production" MRA and "Sustainability" MRA

As anticipated, two separate meta-regressions were estimated for the analysis of this thesis. This choice addresses the need for making more homogeneous subsamples, for commodity consistency reasons,[7] firstly, and to answer to two different research questions, exploring different spheres of consumers' preferences, secondly. The meta-regressions were called "Production" MRA and "Sustainability" MRA, respectively. They only differ for the considered sustainable attributes, whose mWTP estimates are employed as dependent variable, and are almost identical under all the other aspects (except for one independent variable and the number of observations).

The "Production" MRA investigates those characteristics of finfish products which influence consumers' choices but are related to production process[8] rather than the sustainable aspect specifically. Nevertheless, such char-

[6] With a random intercept at a study-level. Both the conventional maximum likelihood method (FEML) and the restricted maximum likelihood method (REML) were employed.

[7] Since, as it will be explained shortly, mWTP estimates for different attributes are pooled together and employed as dependent variable.

[8] Cantillo, Martín and Román (2020), for instance, when summarizing findings about

acteristics also have relevant implications for sustainability of the supply chain, although selected for various reasons.[9] The research goal pursued is to understand how consumers' preferences towards production features can be exploited or targeted in market policies aiming to increase sustainability of the seafood sector.

On the other hand, the "Sustainability" MRA focus on attributes directly linked, and easily attributable by consumers when making buying decisions, to sustainability issues. Be warned that here an extensive definition of "sustainability", encompassing all the aspects of this concept, including the social dimension and the human health, is employed. In this case, the research aims to identify the most valued attributes when it comes to sustainability of the shopping basket.

Keeping these two focus in parallel and comparing the obtained results allows to describe and model in a more realistic way the purchase behavior of consumers, because taking into account more than one influential factor on their choices and, likewise, to understand how consumers' preferences differently meet such factors or which are the product attributes and individual characteristics modifying the impact of the latters. Of course, this is done by employing the tools of meta-analysis, with the purpose of controlling for differences in primary studies and improve their findings.

The meta-regressions were estimated using the statistical software R (R Core Team, 2022).

3.4.2 Dependent variable(s)

The MRAs of this thesis employ as dependent variable the estimated marginal values of consumers' WTP for sustainable attributes of finfish products. As anticipated in the previous section, such mWTPs refer to different attributes which can be imputable to the same concept,[10] the latter varying according to the MRA, as the related attributes. This kind of analysis

---

the harvest method, state that "consumers usually prefer wild fish over farmed fish. The preferences for wild products occur for different reasons: consumers often describe farmed fish as being less healthy and with lower quality when compared to wild fish, while other key elements that have conditioned the image and acceptance of aquaculture fish are the comparatively lower costs, perception of an artificial-like product, and lack of information on sustainable farming practices". Other characteristics of finfish production which, according to the literature, influence purchase decisions are product presentation (e.g., fresh vs. frozen or canned) and domestic production.

[9]Among which, according to the case and the consumer profile, can be included a sensitivity to environmental issues.

[10]This assuring commodity consistency.

set-up[11] follows the example of Li and Kallas (2021).

Before including them within the datasets used for the regressions, values were standardized. This was necessary, since the original estimates came from different countries and years, and were expressed in different currencies (mainly, sterling, euro and norwegian krone) and functional units (from 100g to 1kg). The currency standardization process involved the cited purchase power parities (PPPs). WTP estimates were first converted in a common currency (US dollar), by simultaneously eliminating differences in the purchase power of countries, with the PPPs conversion rates.[12] Then, the obtained values were adjusted for the inflation level,[13] using 2021 as the base year.[14] Finally, PPPs rates were applied again to convert values to euros for all estimates. Concerning functional units, instead, they were led back to euro per kg with a simple ratio.

The final number of WTP measures was 85 for the "Production" MRA and 133 for the "Sustainability" MRA. Since many papers report more than one estimate, these numbers are larger than the number of primary studies; respectively, 15 and 12. In a few cases, were not directly reported or computed in a form not consistent with the analysis and the other estimates, marginal values for WTPs were (re-)calculated based on the regression coefficients provided.

More details about the attributes considered in each of the MRAs and their definition are given in the following sub-sections.

"Production" MRA

For the "Production" MRA, the chosen attributes were:

- *domestic*, defined from the included studies in terms both of a simple indication of the domestic origin on products' packagings and a proper country-of-origin label (COOL);[15]

- *harvest method*, consistently defined across studies as wild catch, having farmed products as reference;

[11] The different possible set-ups are discussed in the literature review and methodology chapter, section 2.4.5.

[12] Retrieved from the database of the International Monetary Found (IMF, 2023).

[13] Inflation rates calculated by the U.S. Bureau of Economic Anlaysis and retrieved from the Federal Reserve Economic Data (FRED) database (BEA, 2023).

[14] At the present time, 2021 is the last year for which the conversion rates and the national accounts indexes available are complete.

[15] Sometimes a specific foreign country is used as benchmark within choice sets.

- *fresh*, also consistently defined as fresh and not processed fish by all studies;[16]

- *production method* (benchmark), defined as a decrease (expressed in percentage) in environmental pressure due to a change towards an IMTA production system.[17]

Of course, dummies coded as = 1 if the estimate refers to the related attribute and = 0 otherwise were included within the MRA among the independent variables, to explain differences in the underlying feature captured by the WTP measures and further assure commodity consistency. The dummy for the production method was excluded because employed as benchmark, in order to avoid collinearity and for interpretative reasons.

"Sustainability" MRA

Coming to the "Sustainability" MRA, instead, the attributes encompassed by the mWTP estimates serving as dependent variable are:

- *specific eco-label*, defined as the presence of an eco-label from a specific labelling scheme, namely MSC, ASC, Naturland or AB,[18] depending on the study;

- *generic eco-label*, indicating a sustainability claim uncorrelated with any labelling scheme;[19]

- *organic*, simply defined as organic production in all studies;

- *fair trade*, corresponding to a generic fair trade claim;[20]

- *nutrition claim*, defined as a statement about the nutritional properties of the product, included as part of its label and involving the omega-3 fatty acids or the protein content in the totality of studies;

[16]The benchmark can be frozen or smoked fish.

[17]In this case the definition is consistent across studies because this attribute is considered in only two MRAs; it was nevertheless included since the estimated mWTPs come from four samples corresponding to four countries, this assuring the independence and variability of data.

[18]The latter refers to an eco-label designed for the French market, indicating that the product (only farmed fish) is organic at 95% minimum.

[19]Although format and contents may vary across studies, most of them are based on such market standards.

[20]Only one study reports it.

- *health claim* (benchmark), defined as a statement concerning the health benefits given by regularly consuming the product and involving the heart or brain functions;[21]

As in the case of the "Production" MRA, dummy variables were included to code the related attribute and, therefore, the difference in the depicted feature. Here, the benchmark attribute is *health claim*.

3.4.3 Independent variables

Two sets of independent (moderator) variables were defined for the MRAs. Indeed, in order to test robustness of results across different specifications and increase their resiliency to multicollinearity, each MRA was estimated in form both of a "complete" and a "restricted" model. Following the categorization of the methodology review chapter, section 2.4.3, the moderator variables are outlined for both the complete and the restricted model. Where not otherwise indicated, the variables are coded as dummies, = 1 if the feature is present in the related estimate (or in the related study) and = 0 if not. The proxy for precision, in all specifications, is the square root of the sample size (indicated as sqrt(n)); this choice was driven because the majority of primary studies, as often happens, do not report the standard errors of the WTP estimates (only four studies calculate them).

There are moderator variables which correspond to *harvest method* (*wild*) and *fresh* (the latter having the same name in both cases). These variables, when reported as moderator variables, have to be considered to code such features at a study-level, whereas, when reported as attribute variables, code them at an estimate-level.[22]

It has to be pointed out that, taking the other MRAs on food products summarized in section 2.4.6 as a sample, this is the first case in which the impact of two specific study features, that is, the focus on the implementation of innovative products, rather than on the actual market, and the informations provided to respondents, is investigated using meta-data.

The employed moderator variables are also reported in **Table 3.1** and more details are given on their coding. The expected sign is based on the

[21]Reported by only one study but based on five samples from five countries.

[22]For example, the dummy *harvest method* is equal to 1 only if the related mWTP is for wild fish, and 0 otherwise. Whereas, in a study reporting eight estimates for totally different attributes, but only taking into account wild fish, the dummy *wild* is always equal to 1. This is intended to capture variation among studies and exploit it to obtain new information about the features behind this kind of variables. Elsewhere in the text they are also called *value-added* variables.

findings from the review of Cantillo, Martín and Román (2020), the results of the previous MRAs or, when not available, on the information which can be found in literature and the assumptions made for the present research.

Complete model

In addition to the mentioned dummies coding for the attributes (X variables), the other independent variables of the complete model are enlisted below.[23]

- P variables (product features): finfish species, harvest method, product form (fillet, steak, ...), product presentation (fresh, frozen, canned, ...). Among the species taken into account, seabream and seabass were merged into a single variable since, having similar taste and similar price, at least within the European market, they are considered to meet also similar preferences.

- K variables (contextual factors): country and year of study. The country of study should be defined more properly "region of study", since the involved countries were split in sub-groups according to the United Nations geoscheme for Europe,[24] and a dummy variable was created for each of them. To be more specific:

  - Norway, UK and Ireland are part of *Northern Europe*;
  - France and Germany are part of *Western Europe*;
  - Italy and Spain are part of *Southern Europe.*

  The year of study, instead, because of collinearity issues given by both the continuous and dummy coding, was transformed into a categorical variable:

  - if the year of study is between 2001 and 2010 (first decade), = 1 ;
  - if the year of study is between 2010 and the mean of the sample, = 2 ;[25]
  - otherwise, = 3.

[23]The income values were also standardized with the same procedure employed for the WTP estimates and described in the previous section.

[24]Part of the publication "Standard country or area codes for statistical use (M49)", by the United Nations Statistical Division (UNSD, 2023).

[25]Thus, 2015 for the "Production" MRA and 2016 for the "Sustainability" MRA.

- R variables (study design): informed/uninformed estimate, survey delivery method, study classification,[26] number of attributes and number of choice sets (using as threshold values the median of the dataset, respectively, 4 and 8).

- S variables (sample characteristics): mean age (using 40 years old as threshold value) and income (continuous variable).

The only difference between the two MRAs to be highlighted here is that within the complete model of "Sustainability" MRA the dummy coding the number of choices was excluded because of collinearity.

Restricted model

As can be guessed, the restricted model is based on the complete specification, but it differs from the latter because some variables are merged and others are excluded.

The dummies coding for survey delivery method, number of attributes, number of choices and mean age of the sample are not included in the restricted version. Whereas, finfish species, product forms and product presentations are aggregated. Therefore, the new regressors are:

- a dummy indicating if the estimate refers to salmon, cod, seabream or seabass, improperly named *big 5*;[27]

- a dummy indicating that the product is sold in form of a fillet or steak (thus, not as a whole fish), named *cut*;

- a dummy indicating that the product was processed in any way (frozen, smoked, canned or prepared to be reasy-to-cook), named, precisely, *processed*.

The other variables are identical to the ones of the complete model.

[26]Cantillo, Martín and Román (2020) categorize the reviewed studies in "market competition" and "market innovation"; here is reported the definition that the authors themselves provide in their paper for such categories: " [. . .] the first category looks to identify the most important factors affecting the buying decision of consumers in real markets, while the second addresses the level of motivation that consumers have for the implementation of new products that are not still available in real markets or they are not still well-known by the consumers".

[27]Improperly because the so-called "big 5", thus the species dominating the seafood market, are: cod, haddock, salmon, tuna and prawns (Tetley, 2016). Here this label was employed for convenience.

Variable	Coding	Expected sign
Dependent variable		
mWTP	Marginal WTP (continuous variable) in €/kg for sustainable attributes of finfish products	None
Precision		
sqrt(n)	Square root of the sample size	None
X variables (attributes)		
"Production"		
Domestic	If the estimate refers to the attribute = 1, = 0 otherwise	+
Harvest method	If the estimate refers to the attribute = 1, = 0 otherwise	+
Fresh	If the estimate refers to the attribute = 1, = 0 otherwise	+
Production method	If the estimate refers to the attribute = 1, = 0 otherwise	Benchmark
"Sustainability"		
Specific eco-label	If the estimate refers to the attribute = 1, = 0 otherwise	+
Generic eco-label	If the estimate refers to the attribute = 1, = 0 otherwise	+
Organic	If the estimate refers to the attribute = 1, = 0 otherwise	+

(Continues)

(Continued)

Variable	Coding	Expected sign
Fair-trade	If the estimate refers to the attribute = 1, = 0 otherwise	−
Nutrition claim	If the estimate refers to the attribute = 1, = 0 otherwise	+
Health claim	If the estimate refers to the attribute = 1, = 0 otherwise	Benchmark
P variables(product features)		
Salmon, cod, sea-bream, seabass*	If the estimate refers to the species = 1, = 0 otherwise	+
Other species	If the estimate refers to the species = 1, = 0 otherwise	Benchmark
Wild	If the product is from wild catch = 1, = 0 otherwise	+
Farmed	If the product is farmed = 1, = 0 otherwise	Benchmark
Fillet	If the product is sold in fillet form = 1, = 0 otherwise	+
Steak	If the product is sold in steak form = 1, = 0 otherwise	−
Whole	If the product is sold as whole fish = 1, = 0 otherwise	Benchmark
Frozen	If the product is sold frozen = 1, = 0 otherwise	−
Smoked	If the product is sold smoked = 1, = 0	+

(Continues)

(Continued)

Variable	Coding	Expected sign
	otherwise	
Canned	If the product is sold canned = 1, = 0 otherwise	−
Ready-to-cook	If the product is sold ready-to-cook = 1, = 0 otherwise	+
Fresh	If the product is sold fresh = 1, = 0 otherwise	Benchmark
K variables (contextual factors)		
Northern Europe	If the country of study is Norway, UK or Ireland = 1, = 0 otherwise	+
Southern Europe	If the country of study is Italy or Spain = 1, = 0 otherwise	−
Western Europe	If the country of study is France or Germany = 1, = 0 otherwise	Benchmark
Year of study	$2001 \leq$ year ≤ 2010 = 1, $2010 \leq$ year < sample mean** = 2, = 3 otherwise	−
R variables (study design)		
Informed	If participants received prior information = 1, = 0 otherwise	+
Uninformed	If participants have not received prior information = 1, = 0 otherwise	Benchmark
Online	If the survey was deliv-	+

(Continues)

(Continued)

Variable	Coding	Expected sign
	ered online = 1, = 0 otherwise	
In-person	If the survey was delivered in-person = 1, = 0 otherwise	Benchmark
Market innovation	If the study belongs to market innovation category = 1, = 0 otherwise	+
Market competition	If the study belongs to market competition category = 1, = 0 otherwise	Benchmark
Nr. of attributes	If the number of attributes > 4 = 1, otherwise	+
Nr. of choice sets	If the number of choices > 8 = 1, = 0 otherwise	−
S variables (sample characteristics)		
Mean age	If the mean age of the sample ≤ 40 = 1, = 0 otherwise	+
Income***	Sample mean income (continuous variable) = 0 otherwise	+

Table 3.1: list of variables of the meta-regression analyses, with coding and expected sign. *one dummy per species; seabream and seabass are merged into a single variable (since meeting similar preferences). **2015 in the “Production” MRA, 2016 in the “Sustainability” MRA. ***Where not reported by the primary studies or not computable from the information provided, data on income were imputed based on the national statistics of the country of study; sources: Eurostat (for EU countries/UK), Statistik sentralbyra (for Norway).

Models shortcomings

Looking at sections 2.4.3 to 2.4.6 from the previous chapter and making a comparison with the outlined models, it can be noticed that not all product or study features enlisted there are taken into account in the MRAs. This decision, made in the modeling phase (indeed, data for some of such features were actually collected), comes from the need of simplifying specifications and facilitate their interpretation, the relevance of the variables in view of the research questions and an evaluation of the quality and quantity of the available data.

For instance, the identification of the impact of using different choice models was considered too specific and not of interest for the present research; whereas, including a dummy to account for hypothetical and non-hypothetical methods would have been possible but, supposedly, not very significant, since only 3.3% of the originally collected estimates resulted from non-hypothetical experiments.

The small number of observations is also behind the exclusion of some product attributes or features from the analysis; this is the case of the presence of safety or other trustworthy labels, the product size, the brand and place of purchase, the type of feed and certain product forms and presentations (e.g., tail cut, slices, sushi).

Coming to the socio-demographic characteristics, a selection was made among the four possible variables to be included (gender, age, education and income), again with the aim to simplify the models; age and income were chosen because more relevant for the research and more likely to be significant in the context of study.

3.4.4 Specifications

Various specifications were tested to individuate the best MRAs. As a result, it was decided to report two specifications for each model of each MRA: an OLS and a RE regression[28] with a random intercept at a study-level. The OLS specification was chosen, although the well justified doubts arising for its employment in meta-analysis, as the base case. The RE model, instead, was chosen for its capacity to explain both observed and unobserved heterogeneity; to go into more detail, to account for the unmeasurable study effect on the estimates, this also representing a control for data correlation. Indeed, as anticipated, the RE specification was modeled with a specific

[28] Recall here that, as specified in the methodology review chapter, this thesis follows the usual practice in literature, using the term "random-effects" for what is more properly a *mixed-effects* model, as in this case.

random intercept for each study, consisting in a fixed part, common to all of them, and a random error term, varying across them. The results from the Hausman test[29] allowed the employment of random-effects. The described specification can be expressed with the following formula:

$$mWTP_{is} = \theta_s + \beta_1 sqrt(n)_i + \alpha_1 X_i + e_i \tag{3.1}$$

where $\theta_s = \beta_0 + u_s$ is the random intercept, β_0 is the fixed part and u_s is the random term specific to the study s. The vector of regressors (and related coefficients) $\alpha_1 X_i$ can be expanded to include all types of moderator variables shown in equation 2.15 and summarized in the previous sections. This RE model was estimated twice with both the maximum likelihood (FEML) and the restricted maximum likelihood method (REML).

Panel or hierarchical specifications were not taken into account since data resulted not adequate to be treated as panel or, in general, for these kind of modeling. Cluster-robust models were estimated using a WLS routine, but they suffered of severe multicollinearity problems and other relevant bias; moreover, the almost total absence (except for one case) of heteroskedasticity within data, made useless the employment of such estimation method.

[29] Such results, as well as those from the heteroskedasticity tests, will be presented in the result chapter.

Chapter 4

Results

4.1 Summary

The results from the meta-regression analyses performed for this thesis work are presented in the following sections. Descriptive statistics for both the "Production" and "Sustainability" models are reported in section 4.2; information like the number of studies and observations for each variable and the simple mean of the mWTP of single attributes considered are included, together with a forest plot showing the statistics at a study-level. Section 4.3 provides the outcomes of the outlier and the other statistical tests (for heteroskedasticity and model specification). In section 4.4 the results of the intercept-only models and those of the publication bias analysis (thus, the FAT-PET-PEESE procedure) are reported. Finally, section 4.5 summarizes the findings from the main MRAs.

4.2 Descriptive statistics

4.2.1 Attributes

"Production" MRA

As shown in **Table 4.1**, the attributes having more "weight" within the "Production" MRA and, therefore, within the measure employed as dependent variable, are *domestic* and *harvest method*, whose WTP estimates are reported by 9 and 8 studies (out of 15) and consist in 41% and 38% of the observations, respectively. The attributes having the higher average mWTP, instead, are *fresh* (7.26 €/kg) and, again, *domestic* (7.60€/kg). *Harvest method* and *production method* settles on lower and very similar values. The simple mean for the complete sample (85 observations) is 5.83 €/kg; this is

Attribute	Nr. of studies	Nr. of obs.	Share	mWTP
Domestic	9	35	0.412	7.26 (2.99)
Harvest method	8	32	0.376	4.78 (3.19)
Fresh	4	6	0.071	7.60 (3.99)
Prod. method	2	12	0.141	3.57 (2.35)
Overall	15	85	1	5.83 (3.36)

Table 4.1: Descriptive statistics of sustainable attributes, "Production" MRA. The number of studies actually sums to more than 15 because of multiple entries. Values in €/kg, standard deviations in parenthesis.

Attribute	Nr. of studies	Nr. of obs.	Share	mWTP
Specific eco-label	7	26	0.195	2.06 (1.68)
Generic eco-label	2	34	0.256	1.54 (1.47)
Organic	5	5	0.038	2.63 (1.38)
Fair trade	1	2	0.015	2.49 (1.61)
Nutrition claim	3	51	0.383	1.61 (1.34)
Health claim	1	15	0.113	0.76 (0.87)
Overall	12	133	1	1.64 (1.44)

Table 4.2: Descriptive statistics of sustainable attributes, "Sustainability" MRA. The number of studies actually sums to more than 12 because of multiple entries. Values in €/kg, standard deviations in parenthesis.

the additional price that consumers are willing to pay, on average, for such finfish production characteristics. These values are also reported in **Figure 4.1**, where the dots represent the median mWTP and the error bars indicate the standard deviation.

"Sustainability" MRA

Coming to the "Sustainability" MRA, the most represented attributes are *generic eco-label*, *specific eco-label* and *nutrition claim*. The latter has the highest share of observations (38%) which, however, come from only 3 primary studies (out of 12); the same can be said for *specific eco-label* (2 studies). *Generic eco-label*, instead, is considerably represented also in terms of the number of studies reporting estimates for it (7). The mean values of mWTPs are closely spaced and significantly lower than those from the sample of the "Production" MRA. In any case, the largest means are

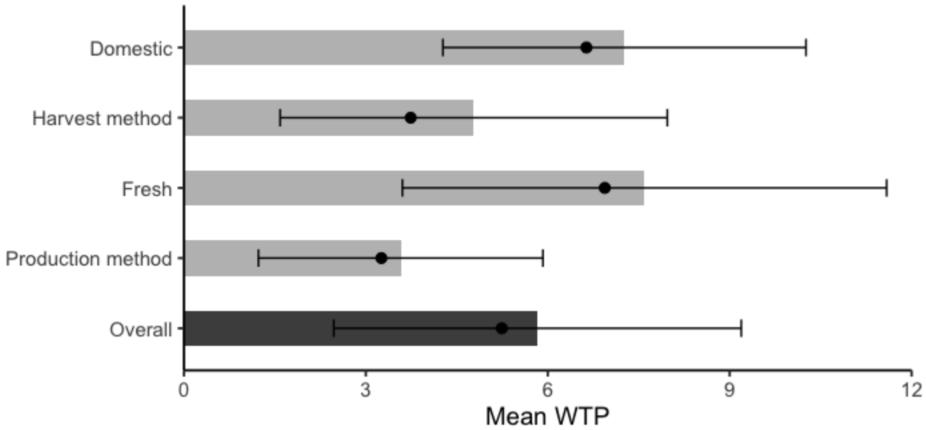


Figure 4.1: Barplot of mean mWTP (in €/kg) of sustainable attributes, "Production" MRA. The dots represent the median value and the error bars the standard deviation.

observed for *organic* (2.63 €/kg), *fair trade* (2.49 €/kg) and *specific eco-label* (2.06 €/kg), whereas the value obtained averaging all estimates (133) is 1.64 €/kg. As for the descriptive statistics of the "Production" MRA, the described information are also reported in **Table 4.2** and **Figure 4.2**. Two additional graphs (**Figure 4.3**) are included for both MRAs, in order to better visualize the distribution of the mWTP estimates; density lines are plotted, and density values are indicated on the y-axis.

4.2.2 Independent variables

"Production" MRA

Starting, again, with the "Production" MRA, the data concerning the moderator variables are very heterogeneous. The finfish species most present in primary studies are salmon (33%), among the "big 5", and other species (34%); the most common product form and presentation are, respectively, fillet (91%) and fresh (47%), while the variables coding the harvest method are equally distributed. Regarding the study features, most estimates comes from online surveys (82%), uninformed respondents (69%) and papers belonging to the market competition category (59%); the number of attributes is > 4 in only 11% of the estimates and also the number of choice sets is > 8 in a few cases (25%). Almost half of the observations come from Western Europe (46%, Northern and Southern Europe follow settling both around

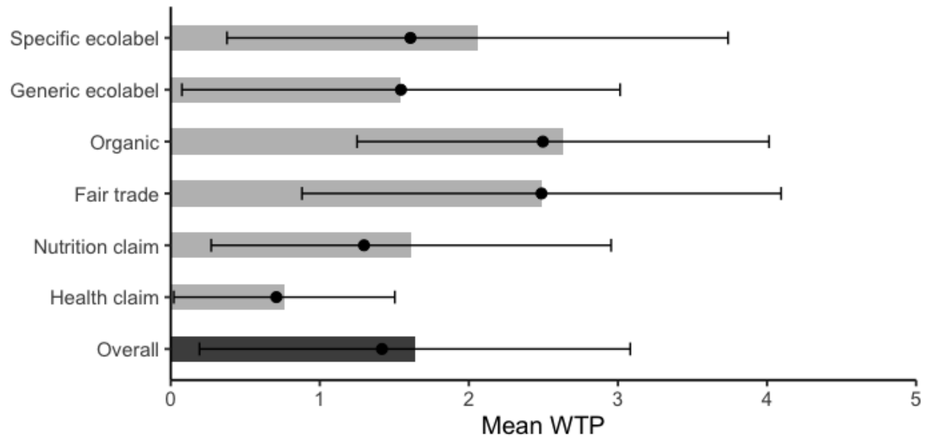
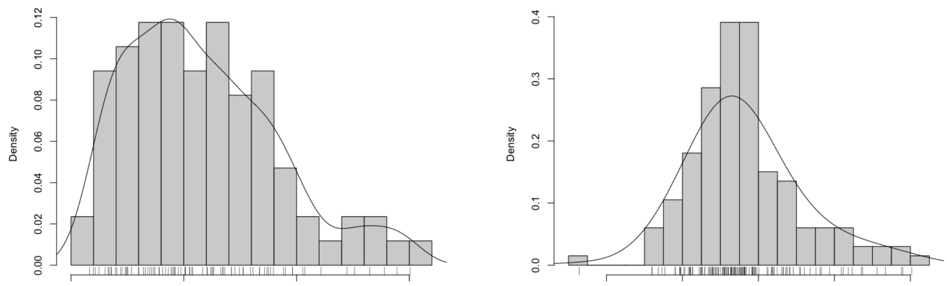


Figure 4.2: Barplot of mean mWTP (in €/kg) of sustainable attributes, "Sustainability" MRA. The dots represent the median value and the error bars the standard deviation.



(a) *"Production" MRA* **(b)** *"Sustainability" MRA*

Figure 4.3: mWTP distributions with density lines.

25%). The means for year of study and precision (sqrt(n))[1] are 2015 and 21. Finally, over 78% of the sample is older than 40 years; the mean income is 33901 €/year. Values are reported in **Table 4.3**.

Variable	**Nr. of studies**	**Nr. of obs.**	**Share/mean**
sqrt(n)	-	-	21.005
Salmon	6	28	0.329
Cod	4	10	0.118
Seabream/ seabass	6	20	0.235
Other species	6	29	0.341
Wild	8	45	0.529
Farmed	7	40	0.471
Fillet	13	77	0.906
Steak	4	16	0.188
Whole	2	3	0.035
Frozen	6	15	0.176
Smoked	4	10	0.118
Canned	2	7	0.082
Ready-to-cook	1	20	0.235
Fresh	8	40	0.471
Northern Europe	5	24	0.282
Southern Europe	6	22	0.259
Western Europe	9	39	0.459
Year of study	-	-	2015 (3.54)
Informed	6	26	0.306
Uninformed	11	59	0.694
Online	9	70	0.824
In-person	6	15	0.176
Market innovation	7	35	0.412
Market competition	8	50	0.588

(Continues)

[1] I.e., the square root of the sample size.

(Continued)

Variable	Nr. of studies	Nr. of obs.	Share/mean
Nr. of attributes (> 4)	5	9	0.106
Nr. of choice sets (> 8)	5	21	0.247
Mean age (≤ 40)	3	18	0.212
Income	-	-	33901 (10264)
Tot. MRA	15	85	1

Table 4.3: Descriptive statistics of independent variables, "Production" MRA. In some cases the number of studies and/or estimates sums to more than 15 and 85 because of multiple entries. When the mean value is reported, the standard deviation is in parenthesis.

"Sustainability" MRA

The descriptive statistics of the "Sustainability" MRA also paint a very heterogeneous picture, although different from the previous case. In this case, indeed, most of estimates refers to other species than those of the "big 5", which are poorly represented (salmon 9%, cod 9%, seabream/seabass 16%), farmed products (60%), in form of a fillet (87%). There are much more observations about ready-to-cook fish (51%); the following most common presentations are smoked and fresh, both with 16%. On the other hand, the study features do not differ substantially from those of the "Production" MRA, although more estimates coming from uninformed respondents (97%) and from DCEs having a number of choice set > 8 (45%) are present; indeed, the majority of the observations are obtained with online surveys (92%) and are reported by papers belonging to the market innovation category (59%). The distribution of the countries of study is also very similar (but Southern Europe is more represented, with 38%), as well as the mean year of study (2016); however, the estimates are less precise (sqrt(n) = 18, on average). The mean respondent is more than 40 years old (86%) and has an annual income of 32389 €. Values are reported in **Table 4.4**.

Variable	**Nr. of studies**	**Nr. of obs.**	**Share/mean**
sqrt(n)	-	-	17.632
Salmon	3	12	0.090
Cod	2	12	0.090
Seabream/ seabass	4	22	0.165
Other species	7	88	0.662
Wild	4	53	0.398
Farmed	9	80	0.602
Fillet	9	116	0.872
Steak	2	17	0.128
Whole	2	2	0.015
Frozen	4	11	0.083
Smoked	4	21	0.158
Canned	1	15	0.113
Ready-to-cook	1	68	0.511
Fresh	6	22	0.165
Northern Europe	3	24	0.180
Southern Europe	5	50	0.376
Western Europe	8	59	0.444
Year of study	-	-	2016 (2.01)
Informed	3	4	0.030
Uninformed	9	129	0.970
Online	5	123	0.925
In-person	7	10	0.075
Market innovation	4	54	0.406
Market competition	8	79	0.594
Nr. of attributes (> 4)	4	12	0.090
Nr. of choice sets (> 8)	7	60	0.451
Mean age (≤ 40)	2	19	0.143

(Continues)

(Continued)

Variable	Nr. of studies	Nr. of obs.	Share/mean
Income	-	-	32389 (4866)
Tot. MRA	12	133	1

Table 4.4: Descriptive statistics, "Sustainability" MRA. In some cases the number of studies and/or estimates sums to more than 12 and 133 because of multiple entries. When the mean value is reported, the standard deviation is in parenthesis.

4.2.3 Primary studies

The two forest plots in **Figure 4.4** and **4.5** show the main statistics at a study-level, for both MRAs. They are structured as follows: on the left, the list of primary studies (identified with authors and year); in the middle, the actual forest plot, where the dots represent the mean mWTP for the attributes considered in the related MRA, error bars represent the variation of the estimates within single studies (minimum and maximum) and the dashed line indicates the sample average; on the right, the mean mWTP of each study, the minimum and maximum in values (in parenthesis) and the sample size.

Concerning the magnitude of the WTP estimates, of course, again applies the consideration done based on the descriptive statistics: it is, on average, significantly higher in the "Production" sample. However, in both MRAs the values reported by single studies seem to be nested around such average, although extreme and conflicting findings are present, especially upwards. The variation interval is in few cases very large and this is probably due to the different number of observations provided and by pooling together distinct attributes in the dependent variable; in most cases it is acceptable, this being another element reassuring about commodity consistency. Finally, the sample size varies considerably, but it results generally smaller in the "Sustainability" MRA.

4.3 Statistical tests

The outcomes of the statistical tests performed before carrying on with the preliminary analysis and the proper MRAs (specifically, the Rosner's test

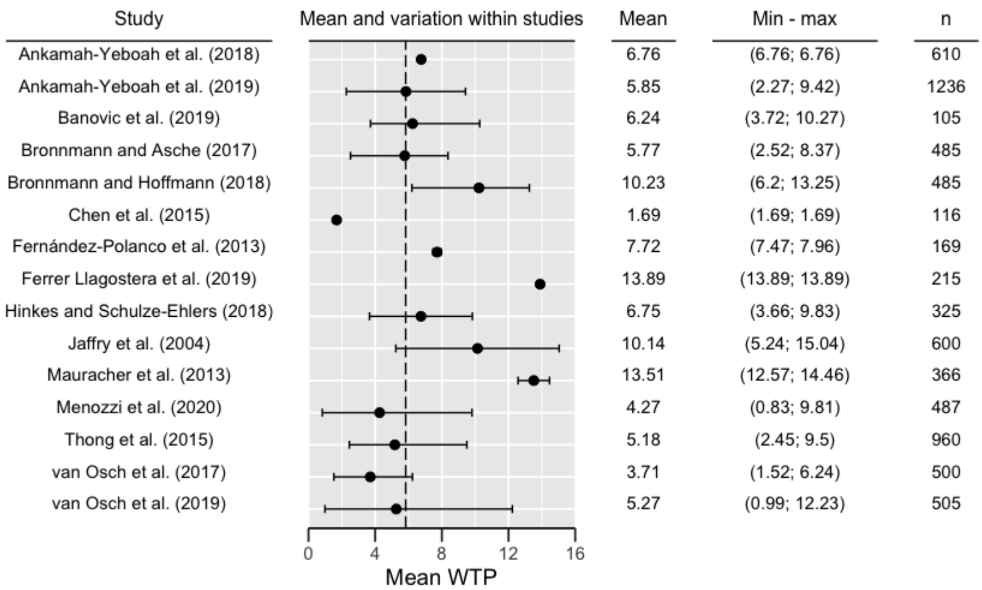


Figure 4.4: Forest plot, "Production" MRA. mWTP values in €/kg.

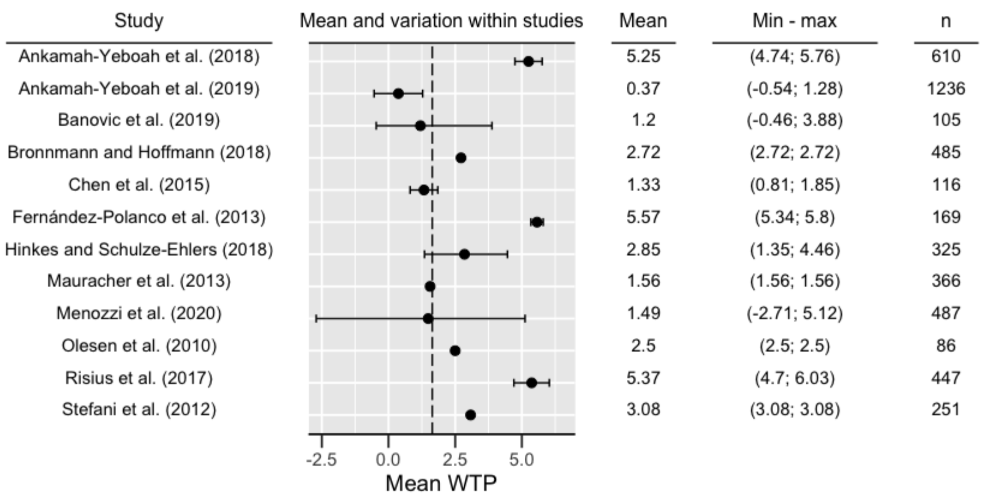


Figure 4.5: Forest plot, "Sustainability" MRA. mWTP values in €/kg.

for outliers, the Breush-Pagan and the White tests for heteroskedasticity and the Hausman specification test), are reported here.

As a result of the Rosner's test, 3 out of 3 suspected extreme values (visually identified with a boxplot) were actually confirmed to be outliers in the "Production" MRA sample; whereas, in the "Sustainability" MRA sample, out of the 16 analyzed values, 12 were excluded because outliers. This led to the definitive number of observations; as already said, 85 and 133, respectively.

The outputs from the Breush-Pagan and the White tests are, instead, included in **Appendix B.1**. No one of the calculated test statistics is significant (thus, the null hypothesis cannot be rejected), except that in the restricted version of the "Sustainability" MRA; this means that only in this case there is enough evidence of heteroskedasticity. A confirmation comes from the residual plots, again available in Appendix (**Figure B.1** of **Appendix B.2**): residuals of the mentioned model become more spread out as fitted values increase, assuming the typical "cone" shape which is a clear sign of variance heterogeneity. The Hausman test also results in not significant test statistics; therefore, since this indicates that no difference is found out between the FE and RE specification, it *should* allow to employment of the latter without introducing any bias. Such outcomes contributed to the specification choices reported in section 3.4.4.

4.4 Preliminary analysis

4.4.1 Intercept-only models

In data and methods, section 3.3, it was anticipated that, to obtain first rough synthesized measure of the consumers' WTP for sustainable finfish products, an "intercept-only" model was estimated for each MRA; the coefficient of its intercept, indeed, calculated employing weighted RE (with a random intercept at a study-level, as in the main MRAs) and WLS cluster-robust estimators, should provide an approximation of the RES and FES weighted means, thus a more reliable value than the simple mean reported within descriptive statistics (section 4.2.1). The results are reported in **Table 4.5** and **4.6**.

As it can be seen, all estimated coefficients are statistically significant, across all specifications and for both MRAs; values are lower in the WLS regressions. Variances of random-effects vary but are always nonzero, this being first indication of the presence of a study effect, although the interclass correlation coefficients (ICCs) are very small (ranging from 0.02 to 0.09). Of

Variables	REML		FEML		WLS cluster-robust	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	6.775***	0.739	6.734***	0.703	5.708***	0.534
Random effects	*Var.*	*SD*	*Var.*	*SD*		
Intercept	5.18	2.276	4.484	2.118		
Residual	176.62	13.29	177.293	13.315		
Model info						
R^2	-		-		0	
Adj. R^2	-		-		0	
Pseudo-R^2 (FE)	0		0		-	
Pseudo-R^2 (tot.)	0.03		0.02		-	
AIC	449.2		450.39		-	
BIC	456.53		457.72		-	
ICC	0.03		0.02		-	
Groups	15		15		15	
Obs.	85		85		85	

Table 4.5: Regression output of the intercept-only model, "Production" MRA. Significance levels: $*p < 0.1$; $**p < 0.05$; $***p < 0.01$. sqrt(n) is employed as weight.

course, the model fit indicators are also small (because of the absence of regressors).

The regression output shows that consumers are willing to pay between 5.71 and 6.77 €/kg more in the case of the "Production" MRA, thus for finfish from production processes meeting their preferences, and between 1.68 and 2.73 €/kg more for products having sustainable attributes. Actually, this values are not too different from those obtained calculating the simple means; the possible reasons are that there are not influential factors or, in general, heterogeneity sources (this being highly implausible, also considering the calculated R^2) or that further corrections are needed.

4.4.2 Publication bias

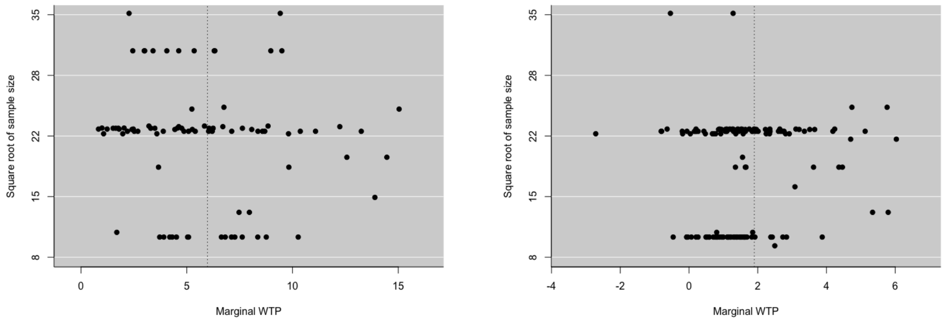
One first correction to estimated values may be controlling for publication bias. The graphical method (funnel plots) and the FAT-PET-PEESE procedure are employed to detect it.

Variables	**REML**		**FEML**		**WLS cluster-robust**	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	2.728***	0.539	2.722***	0.517	1.682***	0.21
Random effects	*Var.*	*SD*	*Var.*	*SD*		
Intercept	2.703	1.644	2.432	1.559		
Residual	26.353	5.134	26.338	5.132		
Model info						
R^2	-		-		0	
Adj. R^2	-		-		0	
Pseudo-R^2 (FE)	0		0		-	
Pseudo-R^2 (tot.)	0.09		0.08		-	
AIC	467.33		467.89		-	
BIC	476		476.56		-	
ICC	0.03		0.02		-	
Groups	12		12		12	
Obs.	133		133		133	

Table 4.6: Regression output of the intercept-only model, "Sustainability" MRA. Significance levels: $*p < 0.1$; $**p < 0.05$; $***p < 0.01$. sqrt(n) is employed as weight.

Starting from the visual analysis of funnel plots (**Figure 4.6**), it is immediately evident that one of them **(a)** is more asymmetric than the other one **(b)**, since the dots are not equally distributed on the sides of the dashed line, representing the mean value of the top 20% most precise estimates, and are more spread out. This is a sign of possible publication bias in the "Production" MRA, which may make sense, since the attributes considered there are those providing the strongest and most widely accepted findings in literature. Publication selection may be due to a tendency to a achieve consistency with the conventional view and is confirmed by the FAT-PET-PEESE regression.

The output is reported in **Table 4.7** and **4.8**. The regression coefficient of the precision proxy (sqrt(n)) from the FAT-PET regression of the "Production" MRA is significant, although at a 0.1 significance level and with a very small value. The null hypothesis should still be rejected, but this means that the effect of publication bias is limited. Anyway, and since the PET is passed (thus, there is a genuine effect also after controlling for the bias), the PEESE estimate is calculated: it represents the consumers' WTP corrected



(a) *"Production" MRA* **(b)** *"Sustainability" MRA*

Figure 4.6: Funnel plots. The dashed line indicates the mean value obtained averaging the top 20% most precise estimates (5.98 for the "Production" MRA and 1.90 for the "Sustainability" MRA). mWTP estimates in €/kg.

for publication bias. Because of the small magnitude of the latter, the estimated value (6.40 €/kg) lies in the same interval of the ones obtained with the intercept-only models. In the "Sustainability" MRA, instead, as guessed from the funnel plots, the coefficient of sqrt(n) is not statistically significant.

To sum up, the presented results show that publication selection should not have had a concerning impact on the data and, consequently, on the consistency and reliability of the models based on them.

4.5 Findings from the meta-regression analyses

Finally, in this section the findings from the main meta-regression analyses are summarized. Based on the specifications outlined in data and methods, section 3.4.4, the OLS and RE regressions were estimated and their outputs are reported here for each MRA ("Production" and "Sustainability") and each model (complete and restricted); the OLS specification serves as the simplest and reference case, while the RE estimation is assumed to be more suitable to the analysis and its results to be more robust. In any case, those findings which are consistent across specifications[2] can be considered the most reliable. Indeed, as just said, when data and econometric issues are so complex, like in meta-regression analysis, a good expedient to evaluate robustness of results is comparing their variation and significance when

[2]Not only between OLS and RE regressions, but also between REML and FEML estimations.

Variables	**FAT-PET**		**PEESE**	
	Coeff.	*SE*	*Coeff.*	*SE*
Intercept	7.096***	1.14	6.396***	0.953
sqrt(n)	- 0.06*	0.033	-	-
$(\text{sqrt(n)})^2$	-	-	- 0.001	0.001
Model info				
R^2	0.0116		0.0098	
Adj. R^2	- 0.0003		- 0.0021	
F	3.343*		2.018	
Prob. > F	0.089		0.177	
Groups	15		15	
Obs.	85		85	

Table 4.7: Regression output of the FAT-PET-PEESE model, "Production" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. WLS cluster-robust; sqrt(n) is employed as weight.

Variables	**FAT-PET**		**PEESE**	
	Coeff.	*SE*	*Coeff.*	*SE*
Intercept	1.61**	0.82	1.769***	0.535
sqrt(n)	0.004	0.041	-	-
$(\text{sqrt(n)})^2$	-	-	- 0.0002	0.001
Model info				
R^2	0.0002		0.0009	
Adj. R^2	- 0.0074		- 0.0067	
F	0.008		0.036	
Prob. > F	0.929		0.853	
Groups	12		12	
Obs.	133		133	

Table 4.8: Regression output of the FAT-PET-PEESE model, "Sustainability" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. WLS cluster-robust; sqrt(n) is employed as weight.

estimated with different methods.

4.5.1 "Production" MRA

Complete model

The output from the estimation of the complete model of the "Production" MRA is reported in **Table 4.9**. All specifications seem to good fit the data ($0.5889 < R^2 < 0.69$), but the REML has the best relative quality (based on AIC and BIC); indeed, although with the FEML the estimated variance for random-effects is close to zero, the positive values from the REML itself and the ICC (= 0.50) suggest the presence of an unobserved study effect, confirming what seen in the intercept-only models. Nevertheless, and most interesting, the estimated coefficients are very similar between both REML and FEML, and OLS and RE in general. Values for *domestic* and *smoked* are significant in all three specifications, and those for *fillet*, *steak* and *age* are always significant in two of them.[3]

To quantify, based on these results, consumers are willing to pay from 3.22 to 3.48 €/kg more for domestically harvested finfish, from 3.38 to 3.51 €/kg more for smoked products, 4.40 €/kg less if it is sold in form of a fillet and 2.23 €/kg less if it is sold in form of a steak. Younger consumers show an higher mWTP, from 2.68 to 2.99 €/kg. The number of choice sets produce a significant and negative coefficient (− 4.80 €/kg) only in the OLS regression.

Restricted model

Similar considerations can be done, about the goodness-to-fit, relative quality and estimation parameters, for the restricted model (**Table 4.10**), although the ICC and, as expectable, the R^2 are generally lower. However, in this case, 6 out of 6 significant coefficients, some of which were not significant in the complete model, are so in all specifications. They are: *domestic*, *wild*, *cut*, *processed*, *informed* and *classification*.

According to the restricted model, consumers are willing to pay between 3.04 and 3.25 €/kg more for domestic products and between 6.48 and 6.77 €/kg less if it is not sold as a whole fish, similarly to the complete model. But, differently from it, they are also willing to pay from 4.24 and 4.34 €/kg more for wild fish and from 1.16 to 1.87 €/kg more for processed products

[3] *N.B.* except that in the restricted model of the "Sustainability" MRA, coefficients from OLS and FEML specifications are the same, since the latter fails to estimate random-effects. However, it is anyway reported because standard errors are different and this can have an impact on statistical significance.

Variables	**OLS**		**REML**		**FEML**	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	7.713	7.829	5.087	14.851	7.713	6.578
Sqrt(n)	- 0.055	0.177	- 0.025	0.352	- 0.055	0.149
Domestic	3.219***	1.239	3.478***	1.197	3.219***	1.041
Har. method	- 0.753	1.782	- 0.801	1.729	- 0.753	1.497
Fresh	0.307	2.078	0.461	2.085	0.306	1.746
Salmon	- 1.35	1.143	- 0.405	1.422	- 1.35	0.96
Cod	- 1.339	1.32	- 1.17	1.479	- 1.339	1.109
Seab -	- 0.949	1.329	- 0.317	1.457	- 0.949	1.116
Wild	2.14	2.815	2.423	3.838	2.14	2.365
Fillet	- 4.403***	1.698	- 3.531	2.482	- 4.403***	1.426
Steak	- 2.228*	1.228	- 1.337	1.39	- 2.228**	1.032
Smoked	3.382**	1.344	3.514*	2.123	3.382***	1.129
Frozen	0.231	3.742	- 0.975	6.496	0.231	3.144
Canned	1.719	1.435	1.882	2.17	1.719	1.205
R.-to-cook	1.564	2.157	1.38	3.718	1.564	1.812
North. Eu.	- 1.339	1.076	- 1.63	1.05	- 1.339	0.904
South. Eu.	- 0.979	1.227	- 1.308	1.188	- 0.979	1.031
Year	2.073	3.053	3.315	5.28	2.073	2.565
Informed	2.467	1.795	2.454	1.714	2.467	1.508
Online	- 4.344	6.282	- 6.485	10.72	- 4.344	5.278
Classification	1.452	1.39	1.999	3.493	1.452	1.168
Nr. attrib.	2.804	2.253	2.095	3.367	2.804	1.893
Nr. c. sets	- 4.805*	2.759	- 3.373	4.067	- 4.805**	2.318
Age	2.683	1.67	2.994*	1.611	2.683*	1.403
Income	- 0.257	0.591	- 0.543	0.583	- 0.257	0.496
Random effects			*Var.*	*SD*	*Var.*	*SD*
Intercept			5.839	2.416	0	0
Residual			5.902	2.429	4.585	2.141
Model info						
R^2/Adj. R^2		0.59/0.42	-/-		-/-	
Pseudo-R^2 (FE/tot.)		-/-	0.39/0.69		0.59/0.59	
AIC/BIC		423/486	385/451		425/491	
ICC		-	0.50		0	
Obs. (groups)		85 (15)	85 (15)		85 (15)	

Table 4.9: Regression output of the complete model, "Production" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

(thus, not fresh). Moreover, it is detected a positive impact on mWTP if respondents receive prior information about insights of the product and its production process (from + 3.09 to + 3.37 €/kg) and if it is a market innovation (from + 2.30 to + 2.55 €/kg).

Variables	**OLS**		**REML**		**FEML**	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	5.805	3.979	5.311	4.739	5.805	3.611
Sqrt(n)	0.049	0.06	0.056	0.083	0.049	0.055
Domestic	3.039**	1.195	3.253***	1.173	3.039***	1.085
Har. method	- 0.677	1.63	- 0.711	1.613	- 0.677	1.479
Fresh	1.493	1.824	1.305	1.844	1.493	1.655
Big 5	- 1.388	0.988	- 0.732	1.119	- 1.388	0.897
Wild	4.242***	1.57	4.344**	1.808	4.242***	1.424
Cut	- 6.769***	1.843	- 6.48***	2.151	- 6.769***	1.672
Processed	1.869**	0.815	1.612*	0.954	1.869**	0.739
North. Eu.	- 0.415	0.935	- 0.784	0.945	- 0.415	0.849
South. Eu.	0.487	0.876	0.237	0.892	0.487	0.795
Year	0.008	0.697	0.299	0.847	0.008	0.632
Informed	3.368***	1.037	3.088***	1.196	3.368***	0.941
Classification	2.303**	1.071	2.549*	1.464	2.303**	0.972
Income	0.035	0.537	- 0.185	0.537	0.035	0.487
Random effects			*Var.*	*SD*	*Var.*	*SD*
Intercept			1.502	1.226	0	0
Residual			5.95	2.439	5.223	2.285
Model info						
R^2/Adj. R^2		0.53/0.44	-/-		-/-	
Pseudo-R^2 (FE/tot.)		-/-	0.44/0.56		0.53/0.53	
AIC/BIC		414/453	402/444		416/457	
ICC		-	0.20		0	
Obs. (groups)		85 (15)	85 (15)		85 (15)	

Table 4.10: Regression output of the restricted model, "Production" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

4.5.2 "Sustainability" MRA

Complete model

Coming to the "Sustainability" MRA, the first thing to be noticed is that the complete model fits definitely better the data. This is due, supposedly, to the different patterns and complexity in the underlying dataset, which can only be captured by a more detailed model. Indeed, in this case, the R^2 is high in all specifications (ranging from 0.39 to 0.52). As for the complete model of the "Production" MRA, the FEML fails to estimate the random-effects, but the REML highlights the presence of a random study effect (positive, although small, variance and ICC = 0.31).

The results (**Table 4.11**) indicate that consumers' WTP increases by 0.75 €/kg if a nutrition claim is present and by 1.01 to 1.04 €/kg for salmon; the preference for smoked products is confirmed also in the case of this MRA (from + 2.90 to + 3.09 €/kg), as well as the positive impact of the innovative products (from + 3.38 to + 3.85 €/kg). The mWTP also increases if the southern Europe is considered (+ 1.24 to + 1.26 €/kg), as the study year increases (from + 3.85 to + 3.95 €/kg for each group of the categorical variable) and if the number of attributes is > 4 (from + 6.73 to + 6.80 €/kg); on the other hand, it considerably decreases if the survey is delivered online (from − 10.41 to − 10.77 €/kg). Very interesting is that the coefficient of *income* is significant (from + 1.13 to + 1.21 €/kg): as explained in section 2.4.6, an impact of income is expected when it comes to purchase behavior but, especially in cases where involved amounts are low, this is rarely observed. Finally, the negative estimated coefficient for *informed* (from − 4.15 to − 4.62 €/kg) is conflicting with the findings from the "Production" MRA.

Restricted model

As anticipated, the restricted model shows a low degree of fit to the data, in the case of the "Sustainability" MRA. Indeed, the R^2 for fixed-effects is very low, both in the RE (0.03 and 0.05) and in the OLS specification ($R^2 = 0.2146$, Adj. $R^2 = 0.1062$) and, based on the AIC and BIC values, there is not one of them to be preferred. The random-effects, instead, result to have a greater impact than in the previous outputs (the FEML estimator succeed to estimate them and there is also an high ICC); this, as a confirm of what said, suggests that the data are not sufficiently explained by the model.

Despite this concerns, it should be highlighted that the OLS version of this model is the only one estimating a significant coefficient for *specific ecolabel* (+ 0.90 €/kg). The other significant coefficients are those of *nutrition claim*

Variables	OLS		REML		FEML	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	- 3.944	4.426	- 4.765	5.158	- 3.944	3.97
Sqrt(n)	- 0.17	0.179	- 0.137	0.218	- 0.17	0.161
Spec. eco-lab.	0.528	0.406	0.538	0.406	0.528	0.364
Gen. eco-lab.	0.649	0.487	0.644	0.487	0.649	0.437
Organic	0.702	1.15	0.774	1.157	0.702	1.032
Fair trade	- 0.014	1.049	- 0.005	1.047	- 0.014	0.941
Nutr. claim	0.755*	0.407	0.751*	0.407	0.755**	0.365
Salmon	1.014**	0.467	1.036**	0.469	1.014**	0.419
Cod	0.283	0.49	0.299	0.49	0.283	0.439
Seab -	0.299	0.434	0.306	0.434	0.299	0.389
Wild	- 0.143	0.421	- 0.159	0.421	- 0.143	0.377
Fillet	- 1.291	2.584	- 0.895	3.014	- 1.291	2.318
Steak	1.472	1.669	1.664	2.123	1.472	1.497
Smoked	3.088*	1.748	2.897	2.24	3.088**	1.568
Frozen	- 0.179	2.667	- 0.555	3.315	- 0.179	2.392
Canned	2.438	1.762	2.238	2.252	2.438	1.58
R.-to-cook	3.887	3.272	3.978	4.482	3.887	2.934
North. Eu.	0.206	0.304	0.229	0.307	0.206	0.272
South. Eu.	1.238***	0.386	1.262***	0.39	1.238***	0.346
Year	3.95**	1.723	3.847*	2.33	3.95**	1.545
Informed	- 4.151**	1.941	- 4.617**	2.316	- 4.151**	1.741
Online	- 10.411***	1.512	- 10.766***	2.063	- 10.411***	1.356
Classification	3.377*	1.959	3.814	2.35	3.377*	1.757
Nr. attrib.	6.726***	1.627	6.805***	2.156	6.726***	1.459
Age	- 0.143	0.431	- 0.11	0.436	- 0.143	0.387
Income	1.132***	0.41	1.212***	0.437	1.132***	0.367
Random effects			*Var.*	*SD*	*Var.*	*SD*
Intercept			0.566	0.752	0	0
Residual			1.244	1.115	1.003	1.002
Model info						
R^2/Adj. R^2		0.52/0.40	-/-		-/-	
Pseudo-R^2 (FE/tot.)		-/-	0.39/0.58		0.52/0.52	
AIC/BIC		432/510	434/515		434/515	
ICC		-	0.31		0	
Obs. (groups)		133 (12)	133 (12)		133 (12)	

Table 4.11: Regression output of the complete model, "Sustainability" MRA. Significance levels: $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

(from + 0.76 to + 0.81 €/kg), *southern Europe* (from + 0.82 to + 1.03 €/kg), *year* (− 1.02 €/kg, only in the OLS specification) and *income* (+ 0.99 €/kg), and they are consistent with the results from the complete model (except that for *year*). The regression output is reported in **Table 4.12**.

4.5.3 Predicted values

As a conclusion of the analysis, the estimated models were employed to calculate a value for the dependent variable representing the mWTP for the considered attributes corrected for heterogeneity and publication bias. This was done by setting the precision proxy to zero and the moderator variables to a benchmark value, before solving the equation; practically, this is a special case of within-sample prediction. Some of the previous MRAs summarized in section 2.4.6 also provide this information.

With regard to the "Production" MRA, values were chosen to represent a common product within the European market[4] (salmon) produced with the commonly preferred process (domestic fish from wild catch) and sold without any treatment (fresh whole fish). The reference context was identified in a western country and recent years ($\geq$ 2015), and the consumer profile in an (uninformed) individual having more than 40 years and an annual income equal to the mean annual income of EU citizens in 2022 (35220 €/year). The remaining study design variables were set to zero, except for *online*. As a result, it was estimated that this average consumer in the depicted market and context (selected because as general as possible) is willing to pay between 11.79 and 13.02 €/kg more[5] for finfish with the depicted production characteristics.

On the hand of the "Sustainability" MRA, the benchmark values were almost the same. Of course, the product outline slightly changes: it is taken into account the presence of a specific ecolabel, nutrition and health claim, and both a fair trade and organic certification; this is the most sustainable finfish product the actual market can offer. For the latter, our average consumer is willing to pay from 3.21 to 4.32 €/kg more than for a "base" product.[6]

It is interesting to notice that in both cases the obtained values are higher

[4]The *real* market, thus the *classification* variable was set = 0.

[5]The value changes according to the model employed for the prediction: complete OLS = 12.245 €/kg; complete REML = 11.792 €/kg; restricted OLS = 12.663 €/kg; restricted REML = 13.017 €/kg.

[6]Complete OLS = 4.322 €/kg; Complete REML = 3.211 €/kg. In this case, the predictions from the restricted model were not considered; because of its low degree of fit to data, results are biased.

Variables	**OLS**		**REML**		**FEML**	
Fixed effects	*Coeff.*	*SE*	*Coeff.*	*SE*	*Coeff.*	*SE*
Intercept	0.543	1.908	- 2.573	4.531	- 1.015	2.941
Sqrt(n)	0.079	0.055	0.083	0.157	0.065	0.097
Spec. eco-lab.	0.898*	0.479	0.564	0.412	0.573	0.396
Gen. eco-lab.	0.717	0.584	0.661	0.494	0.675	0.475
Organic	0.624	1.286	0.466	1.163	0.202	1.091
Fair trade	1.013	1.093	0.033	1.06	0.058	1.014
Nutr. claim	0.811*	0.491	0.756*	0.413	0.772*	0.397
Big 5	0.44	0.464	0.488	0.399	0.446	0.382
Wild	- 0.206	0.494	- 0.154	0.434	- 0.126	0.414
Cut	1.85	1.42	1.597	3.063	1.522	1.968
Processed	0.327	0.425	0.063	0.358	0.093	0.344
North. Eu.	- 0.157	0.354	0.192	0.307	0.096	0.291
South. Eu.	0.24	0.366	1.032***	0.36	0.818**	0.327
Year	- 1.021*	0.613	- 1.087	2.31	- 1.038	1.371
Informed	1.553	1.078	- 1.381	2.77	- 0.632	1.747
Classification	0.129	0.573	0.773	1.992	0.63	1.199
Income	- 0.252	0.4	0.99**	0.469	0.614	0.404
Random effects			*Var.*	*SD*	*Var.*	*SD*
Intercept			8.836	2.972	2.818	1.679
Residual			1.283	1.133	1.192	1.092
Model info						
R^2/Adj. R^2		0.21/0.11	-/-		-/-	
Pseudo-R^2 (FE/tot.)		-/-	0.03/0.88		0.05/0.72	
AIC/BIC		478/530	455/510		465/520	
ICC		-	0.87		0.70	
Obs. (groups)		133 (12)	133 (12)		133 (12)	

Table 4.12: Regression output of the restricted model, "Sustainability" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

than those from the simple average, intercept-only models and publication bias correction, this confirming the great importance meta-regression analysis has, not only to extract new informations about the influential factors on the effect size, but also (and mainly) to include and control for heterogeneity: this way, it is possible to exploit data by reducing bias and allowing them, and the scientific research behind them, to actually explain and represent the economic phenomena correctly.

The enlisted and summarized results will be discussed and explored in the following chapter.

Chapter 5

Discussion

5.1 General considerations

This study pools together very heterogeneous variables and encompasses a wide and complex variety of aspects, despite the strict criteria followed in data collection and in the set-up phase. However, the research efforts were well rewarded since results are very interesting in view of the interpretation of the dynamics of the specific sector of finfish market, the explanation of consumers' purchase behavior and attitudes towards sustainability of the related supply chain and the support to policy making. The painted picture is a quite comprehensive overview of factors influencing buyers' decisions when it comes to the environmental impacts of finfish products, while remaining focused on well defined context, in terms of preferences, cultural and geographic characteristics (the European market) and in terms of the analyzed commodity.

The most important finding from the present research is that consumers are actually willing to pay a positive price premium for more sustainable products, and for some single correlated attributes. At the end of the previous chapter was reported that the mWTP for finfish from preferred production processes, which also have (mainly positive) implications in reducing their impacts, ranges from 11.79 to 13.02 €/kg, whereas the estimated value for products ensuring the environmental and social sustainability of the supply chain, as well as their healthiness, is lower, but still positive and significant (between 3.21 and 4.32 €/kg). This is very important because it means that a sensitivity to the related issues, and a preference for products addressing them, exists[1] and is measurable in monetary units; therefore it can be

[1]As mentioned before, in the assumptions made for this research the attributes from the "Production" MRA (thus, the production characteristics) are chosen by consumers for

exploited and targeted to even more sustainable pathways, in the broader framework of the ecological transition of global markets.

The analysis of influential factors on consumers' preferences is also very interesting and seems to leave room for sustainable policies and developments. To sum up, tendencies to locally harvested finfish, avoiding the highly polluting transportation processes, and to products sold without pre-cuts, in this case avoiding not necessary industrial activities, are detected. A positive attitude towards market innovation also arises from the results, this allowing researchers and industry experts to find new solutions to the mentioned issues without renouncing to profitability. Another relevant aspect is that the informations provided to respondents about the insights of the offered products and the environmental problems of the global fishery actually have an impact on their behavior, although of different sign according to the considered attributes; in any case, this is also significant because it means that awareness campaigns and diffusion of this kind of information would be able to generate a more conscious consumption. Finally, an increasing trend in terms of willingness to pay for sustainable products across years and generations, also detected in some of the models, indicates that, as anticipated above, an higher level of consciousness already exists.

In conclusion, the research questions proposed in section 3.4.1, thus how to target preferences towards production features to increase sustainability of the seafood sector and which are the most valued sustainable attributes of finfish, can be answered. These considerations will be further deepened in the following sections.

5.2 Discussion of results

5.2.1 Sustainable attributes and production process

Therefore, concerning the sustainable attributes and the production process, only the variables for *domestic*, within the "Production" MRA, and *nutritional claim*, within the "Sustainability" MRA, are consistently significant across specifications. This last result can be interpreted as a not relevant impact of the way sustainability is defined in consumers' behavior. In other words, there would be a generalized preference for environmental friendly or healthy products, no matter the specific aspect taken into account.

Focusing on the two mentioned variables, instead, if the strong relevance of *domestic*, statistically significant at 1% in all regressions[2] with a quite high

several and very different reasons, among which can be included sustainability aspects.

[2]Except than in the restricted OLS model.

marginal value (between 3.04 and 3.48 €/kg), was expectable, that of *nutrition claim* over other more markedly defined and usually impacting sustainable attributes, although with a lower mWTP (from 0.75 to 0.81 €/kg) and significance level, is rather surprising. Indeed, as stated by Cantillo, Martín and Román (2020), the origin is found to be the most important attribute in buying decisions involving finfish in many studies, with a preference for local products. This might be due to well-known behavioral patterns[3] leading to have more trust on local commodities and on domestic supply chains and production processes, or simple ethnocentrism. Nutrition claims are less commonly considered influencing factors if compared with other attributes, but the findings from the present research can be explained. The increasing and severe problems, at a global level, with dietary habits and food quality, have shifted public focus on such issues and on the importance of an healthy and balanced diet; this is particularly true with regard to finfish, which is one of the main sources of omega-3 fatty acids, essentials for the human organism. Moreover, other cases in the food sector literature can be found in which this kind of claims are preferred over the environmental performance of the product.[4]

In this review of the results regarding sustainable attributes, it should also be pointed out that one of the specifications of the "Sustainability" MRA (the OLS version of the restricted model) estimated a significant coefficient for *specific eco-label*, with a value of 0.90 €/kg. This is also notorious in literature and subject of many studies; such eco-labels (like those from ASC, MSC, Naturland, etc.) are well-known by consumers and trustworthy, since the associations behind them have been active for years in the field of environmental certifications. Therefore a preference for products involved in these labelling schemes can be justified.

5.2.2 Product characteristics

With regard to the relevance of the other product characteristics,[5] it should be noted, first, that the latters are more important when consumers

[3]There are several reasons behind this behavior, such as: cultural distance; other socio-cultural factors (e.g., values, norms, media, reference groups); political, economic or ethical aspects; affective meanings; and so on (Luomala, 2007).

[4]Rudd, Pelletier and Tyedmers (2011), although their analysis focuses on health claims.

[5]The "P variables" are taken into account in this case. Recall that dummies corresponding to *harvest method* and *fresh* are included among these, but coded at a study-level. They are reported here because, since present in both MRAs, help to highlight the difference in the relevance of these aspects when evaluating the sustainability of a product or its more tangible features.

are evaluating the product itself, thus the features deriving from its production process, rather than its sustainability. Indeed, there are more significant results among the related coefficients of the first MRA. A possible interpretation is that, if sustainability is taken into account, the focus is shifted on different and correlated features. This may seem obvious, but the little relevance of product characteristics when evaluating its sustainability is a remarkable finding because, given the role such aspects actually have in the environmental impacts of the supply chain, a stronger link between these two should also be present in buying choices, to assure more sustainable purchase behaviors.

Therefore, starting from the results of the "Production" MRA, consumers are willing to pay a price premium of more than 4 €/kg for wild fish respect to farmed fish, whereas they have a negative mWTP for pre-cut finfish respect to the product sold as whole. The preference for wild fish is also identified from previous research, which also investigated the motivations behind it. Claret et al. (2014) study consumers' beliefs towards harvest methods and find out that wild fish is generally considered to be safer, healthier, fresher, of better quality, more nutritious, tastier and providing more guarantees than farmed fish. However, this outlines an aversion for seafood from aquaculture. The negative marginal values for finfish fillet (between 3.53 and 4,40 €/kg less) and steak (2.23 €/kg less), and for pre-cuts in general (between 6.48 and 6.77 €/kg less) were, instead, unexpected and partially conflicting with the expected signs indicated in section 3.4.3, **Table 3.1**. Indeed, fillet should be the most valued cut of finfish as well as a time-saving solution for cooking it. These findings could also be red in accordance with buying choices driven from the naturalness of the product. Of course, this is only a possible interpretation. In any case, this tendency is very interesting and exploitable for policy making, in view of streamlining industrial processes.

Coming to the product presentation, a preference is detected for smoked fish, common to the "Production" MRA and the "Sustainability" MRA with not significant differences in mWTPs.[6] This can be due to the perception of this treatment as of "excellence", to the great diffusion of such presentation in the cuisine of northern Europe or to the link with an highly consumed and appreciated species, the salmon.

The latter represents the last noteworthy variable to be mentioned here about product features. Salmon is the only finfish species whose coefficient resulted significant over the other species, although with a low mWTP (about 1 €/kg). It has also to be specified that such finding could be biased by the extra attention of public opinion on the environmental issues of salmon

[6] However, there are not provable preferences of the other presentations over fresh fish.

farming and, therefore, to the higher value assigned by consumers to its sustainable attributes. As a confirmation, the coefficient is only significant in the "Sustainability" MRA. Nevertheless, this poses no obstacle to state that environmental policies should exploit this trend and focus on salmon market.

5.2.3 Other aspects

As anticipated at the beginning of this chapter, many other interesting findings arises from the estimated models. Two of the most valuable are those regarding the information provided to participants to the experiments and the study classification, both involving both MRAs. The first one is somehow a conflicting result, since the computed sign for the mWTP is different according to the considered regression. When called to choose for the production features of finfish, receiving information about the characteristics of the product itself, its certifications and sustainability aspects raises the WTP of consumers by 3.09 to 3.37 €/kg; on the other hand, the same kind of information decreases their WTP by 4.15 to 4.62 €/kg if it is the environmental impact to be evaluated. It seems that communication about the mentioned insights encourages the consumption and increases the intrinsic value assigned to the finfish in one case, while acts as a deterrent in the other case, by also reducing the willingness to pay for the offered product. There are different possible explanations for the observed behavior, but it is plausible that contents of such disclaimers generate major concerns about the depicted environmental issues when attention is focused on them, this modifying the meaning of the provided information and the cognitive and affective response of consumers to the point that even the sustainable attributes, labels and features associated to the product and part of the communication are not trusted anymore. In a few words, the effect of negative information about the impacts of fishery and aquaculture has more "weight" than any solution addressing them in consumers' choices. In addition, it is also possible that consumers consider eco-labels and certifications as an indication of the quality of the product, while do not give credit to the sustainable practices they express.

From the considerations just made there are two conclusions to be drawn. First of all, that the awareness campaigns and consumer education are potentially effective, since it is unquestionable, based on the results of the present research, that having more information on the commodities object of buying decisions actually influences their outcome;[7] secondly, that when sustainabil-

[7]This confirms the previous literature about the "information bias", already cited in the

ity is the topic, communication strategies should be revised.

Moving on to the study classification, results from the regressions show that the estimated mWTP is higher when the study belongs to the market innovation category, whatever the sustainable attributes and purchase decisions are considered,[8] with a value ranging from 2.30 to 3.80 €/kg. Since, as mentioned in data and methods, at the actual knowledge this is the first analysis investigating the difference in willingness to pay and, thus, consumers' attitudes, towards sustainable attributes of products innovative and still not available and products already present in the real market, this finding provides new information to literature and very concrete and practical suggestion for research and development: market innovations find no obstacles in consumers behavioral patterns.

Regarding the influence of the socio-demographic characteristics of consumers, it is interesting to notice the relevance of age, even if its coefficient is only significant in the "Production" MRA. Indeed, consumers less than 40 years old are willing to pay, on average, up to 3 €/kg more for finfish products having specific production features. Also, an increasing trend over the years of mWTP for sustainability of finfish products is observed in the "Sustainability" MRA. This bodes well for the effectiveness of policies aimed at reducing impacts of seafood consumption in the future and suggests to plan new ones. Finally, the positive and significative estimated value for the income, again in the "Sustainability" MRA, is also remarkable, since the influence of such factor is usually negligible when involved monetary amounts are so low. This indicates that households' budget constraints are relevant when choosing product features which go beyond its basic attributes and increase its price. More simply, the latter is still one of the most important elements of evaluation.

Another important contextual factor giving significant results is the country of study. Consumers from southern Europe are willing to pay slightly more than those from the rest of the continent for sustainable attributes and eco-labels of finfish products. Some previous studies suggest that there are differences in consumer ethics of southern and northern countries (Polonsky et al., 2001), that southern consumers have better knowledge about how food is actually made, because the reference population is closer to farming and the reality of food production (Krystallis and Ness, 2005), and an higher level of awareness about traceability of the supply chains, which represent for them a criterion to follow in buying decisions (Giraud and Halawany,

literature review and methodology chapter (Ajzen, Brown, and Rosenthal, 1996; Mariel et al., 2021; Yeh, Hartmann, and Hirsch, 2018).

[8]Thus, in this case, the estimates are homogeneous and agree between MRAs.

2006). This is a possible interpretation. Alternatively, people from southern countries may be willing to pay more because this actually has an impact on the quality and/or sustainability level of products purchased, whereas this does not apply in northern countries, where products are more regulated and almost all of them is eco-labelled or certified. In any case, the resulting consideration is that even within a market supposed to be more heterogeneous in terms of cultural background and generalized preferences, some differences exist, and should be taken into account.

Of course, attention must be payed to the study design as well. The negative impact of the number of choice sets, in the "Production" MRA, is consistent with the broader literature on choice experiments which explains this behavior as a consequence of fatigue or learning effects (Lusk and Schroeder, 2004). The increase in estimated values in case of an high number of attributes, instead, can be due to the perception of a product described with several features as a good product, and is confirmed by previous MRAs. In conclusion, the significant coefficient of the delivery method, indicating that the declared mWTP is by more than 10 €/kg lower for online questionnaires, is also consistent with the literature, although, as mentioned, findings about the sign of such correlation are conflicting; it should be specified that some of the DCEs from primary studies where conducted in real contexts (e.g., supermarkets) and respondents were allowed to visually inspect finfish products. This could have generated, as a side effect, a propensity to purchase or an higher willingness to pay. Such interpretation suggests that the impact of the delivery method also depends on the context and object of study. The described results about *number of attributes* and *online* are from the "Sustainability" MRA.

In any case these findings are helpful for future research in order to set-up more accurate experiments, avoiding to introduce bias in the designing phase.

5.2.4 Production features vs. sustainability

As a complement of the discussion above, further deepening on the parallel between the results from the two MRAs is necessary. It is evident that the factors influencing consumers' choices are different according to the criterion which drives purchase and the aspects which are taken into account. This confirming the assumption underlying the research questions of this thesis, thus that behavioral patterns and cognitive mechanisms follow different pathways in the case of the selection of production features or sustainable attributes. Since in actual buying decisions these mechanisms activate simultaneously, it is important to exploit experimental processes and the tools typical of economic research to explore them separately and better understand the dynamics behind them.

For example, to sum up what already said, product characteristics, such as the cut and the presentation, and the harvest method are less relevant when the choice criterion concerns sustainability of finfish; other factors like the species, the geographical area and the budget, instead, come into play in this case. Simple considerations like these, or like those related to the attributes themselves, are very valuable for the outlined research sector.

Also the more general difference in the overall value obtained when solving the estimated models, which suggest that consumers give greater weight to the more tangible characteristics of the product itself than to sustainability of the supply chain, are important. These points will be discussed again in a more practical way when analyzing policy implications.

5.2.5 Comparison with previous literature

A comparison and a dialectics with previous literature is also necessary, in order to integrate research and contribute to its development. Here, the MRAs summarized in section 2.4.6 are taken into account. Of course, the comparison was possible only for the common variables, given also the specificity and newness of the MRAs performed in this thesis.

First of all, the marginal WTP for *domestic* is in line with the value estimated by Printezis, Grebitus and Hirsch (2019) (3.72 €/kg),[9] this meaning that there is no difference between finfish and the broader category of food products in general with regard to such attribute. Whereas, the values from Smetana, Melstrom and Malone (2022) for attributes like *fresh*, *domestic* and *eco-label* are definitely higher (that for *domestic* is more than double). The comparison with Yeh and Hirsch (2023) and Li and Kallas (2021), instead, is more difficult and can only be approximated, given that they report estimates in %WTP. Anyway, a rough conversion of coefficients from the MRAs of this thesis was made, using the formula (2.16) and the mean price for salmon within the European market in 2021.[10] While the estimate of Yeh and Hirsch for COOL is also higher than the value reported here for *domestic* (71% against 25%), the overall premium price calculated by Li and Kallas (29.5%), although considering simultaneously attributes here split in the two MRAs, lies in the same interval of the provided predicted values (21.89%;

[9] Although their research focuses, more specifically, on the attribute *local*, which can be considered a special case of *domestic*.

[10] From the report of the European Market Observatory for Fisheries and Aquaculture Products (EUMOFA, 2022). The employed value is the average mean price of salmon in the top three consuming countries of fish products.

93.74%) and is more similar to those of the "Production" MRA (ranging from 21.89% to 25.05%) . Of course, it should be specified that, except for Smetana, Melstrom and Malone (2022), all other studies take into account a wide variety of food products and this creates great theoretical difference with this analysis. In any case, similarities with their results are detected.

Coming to the other variables, the positive estimated coefficient for the year of study is in agreement with Printezis, Grebitus and Hirsch (2019), who also find a positive correlation between the mWTP for local food and the time period, although they employ a dummy instead of a categorical and obtain a lower value (1.50 €/kg against more than 3 €/kg). Same can be said for the impact of income, investigated by Smetana, Melstrom and Malone (2022): it has a positive impact and this matches the findings of this analysis, but the mWTP is lower (0.15 €/kg against 1 €/kg). Results for *age*, instead, cannot be compared since none of the summarized MRA estimate significant coefficients.

Finally, concerning the influence of study design, values for *online*, as expectable, are very different or conflicting. Bastounis et al. (2021) also report a negative mWTP when this delivery method is employed, but of lower magnitude (− 2.94 €/kg against − 10 and more €/kg); on the other hand, Smetana, Melstrom and Malone (2022) estimate a positive mWTP. This is in accordance with the theory supported here, thus that the impact of using online surveys depends on the contextual factors and considered commodities. The results about the number of attributes are confirmed by Yeh and Hirsch (2023), whose findings show an increase in marginal values when more attributes are included in the experiment.

5.3 Policy implications

As mentioned several times, the main objective of analyses like this is supporting decision making and suggest policies applicable in the reference sector. The information about the consumers' preferences and attitudes towards more sustainable finfish and its production features, and the quantification of their magnitude in monetary terms are core elements for intervening in the market, by exploiting observed trend or creating new ones, in order to reduce environmental impacts of production, promote sustainable consumption and respect international commitments and policy papers like the Sustainable Development Goals (SDGs).

In this perspective, some possible policies based on the results of the meta-regression analyses performed in this thesis are proposed here. Policies recall the research questions outlined in data and methods, section 3.4.1, which, indeed, were built as specifically oriented to such objective. It was decided to present them in form of an "agenda", to ease reading and make an example of how to manage findings from scientific research and transform data in intervention plans. Of course, this is not an exhaustive list, and single policies may be different or, in any case, should be deepened.

1. *Encourage domestic consumption:* a tendency to prefer locally harvested finfish already exist. It should be possible to "cluster" consumption of countries and communities based on the domestic species, with the support of advertising and communication to consumers, which exploit their affection and trust in local products. A re-organization of import and export driven by surplus of production, if present, could be a complementary option, through institutional means. These changes would be balanced by the higher WTP for products but, on this point, further analyses are needed. The aim is reducing transportation, which nowadays is still one of the first sectors for greenhouse gases emissions in Europe,[11] and food processing. In addition, this would enhance biodiversity and multiculturalism, because fostering the consumption local finfish, limiting the introduction of alien species and favoring the preservation of regional cuisine.

2. *Relaunch and support local market of fresh fish, streamlining industrial activities and supply chains:* the detected consumers' preference for fresh and not processed products (here meaning not pre-cut, frozen, canned, etc.) would give plenty of room for market strategies that shorten the distance between producer (or fisheries), retailer and consumer, improving the environmental performance of seafood industry.

3. *Exploit the effectiveness of awareness campaigns and consumer education:* since results from the analysis confirmed the actual influence of informed consumption on buying choices, attitudes and willingness to pay, broader policies for consumer education should be scheduled. This translates into public awareness campaigns, "talking labels" and so on. Such intervention would support the market of sustainable, eco-labeled or healthy products, assuring the growth of the sector and profitability for compliant companies. It is important to point out that communication strategies should be revised: actual information provided to con-

[11]The emissions from *transport and storage* recorded in EU in the last quarter of 2022 were 102051 thousand tonnes of CO_2 equivalent, 10.87% of total. Source: Eurostat, data available at `https://ec.europa.eu/eurostat/databrowser/view/ENV_AC_AIGG_Q__custom_2691128/bookmark/table?lang=en&bookmarkId=4bb9ab20-296b-4119-88e9-580ea7741c0a` (visited on 06/14/2023).

sumers with regard to environmental impacts and principles of labelling schemes act more as a deterrent than as a nudge towards sustainable products.

4. *Link nutritional benefits of healthy products to their sustainability:* in the framework of these informative campaigns, or of corporate strategies, linking the sustainability of a product with its healthiness and nutritional properties could be a winning move for both promoting consumption of sustainable finfish and supporting the market with new revenues, also having positive effects on health and related public expenses. Moreover, it is often true that sustainable products have better nutritional properties: for example, new experimental types of fish feed, which reduce the pressure on resources, provide more omega-3 fatty acids to farmed fish. Therefore, also in this case, this is a communication and low-cost strategy.

5. *Improve social acceptance of aquaculture:* both aquaculture and traditional fishery arise concerns for their environmental impacts. However, developments in aquaculture practices are very promising and, as anticipated in the introduction, fish farming can potentially better address sustainability issues. The problem, here, is that planning such interventions and modifications of fish industry would also collide with consumers' attitudes, given that they express a preference for wild products, for the already enlisted reasons. Therefore, strategies to enhance the image of aquaculture in public opinion must be implemented.

6. *Focus on salmon market:* although in section 5.2.2 was stated that the results about the preference for salmon, among the considered species, may be biased due to the increasing attention payed to issues generated by its farming and the subsequent great value assigned to attributes addressing them, exploiting such popularity to offer more sustainable products could be rewarding both in terms of environmental protection that of profitability. Sustainable attributes of salmon, indeed, are repaid by consumers with an higher WTP and, in general, salmon results to be one of the most common and appreciated seafood species in Europe.[12]

7. *Invest in research and development:* the predisposition of consumers towards new products and market innovation in general should promote

[12] From the report of the European Market Observatory for Fisheries and Aquaculture Products (EUMOFA, 2022).

investments in research and development of solutions to the environmental issues of finfish industry and provide the necessary fundings with higher prices. Examples of innovations can be alternative fish feed (e.g., insect-based), to reduce resource consumption, harvesting of new species, new production methods (e.g., IMTA), but also new regulations, certification standards or the association of organic production to species only traditionally farmed.

8. *Ride the trend of sustainable consumption and educate younger generations:* the positive trend towards sustainability detected in the last years, also relatively to finfish products, should be encouraged both with informative campaigns directed to individuals as consumers and through the national education system.

9. *Diversify policies at a regional level:* the differences in preferences and purchase behavior observed in the European sub-regions suggest that these policies should be tailored accordingly to the geographical and cultural context, both at an international level (e.g., EU) and a national and regional level.

10. *Support purchase of sustainable products:* given the impact of households' budget constraints on buying choices, especially when it comes to selection of sustainable attributes, purchase of products (including finfish) having such attributes should be not only promoted, but also supported, with subsides to consumers or to retailers; analysis like that of this thesis help to quantify the magnitude of subsides. Another option is a greater diversification of sustainable products to meet different income levels.

11. *Associate production features and sustainability in consumers' choices*: it was demonstrated that, when called to valuate sustainability of finfish, consumers' focus is shifted and product characteristics like harvest method, cut and presentation are less relevant. Therefore, it should be better emphasized the importance of such attributes also for the environmental impacts of production, in order to make their choices more sustainable.

12. *Associate production features and sustainability in actual products*: to better support the association of these two aspects in consumers' choices and increase sales of sustainable finfish, products actually offered on the market with sustainable attributes (e.g., eco-labels, organic certification, etc.) should also have the most sustainable production features

(e.g., unprocessed, non-pre-cut and fresh). Regulations imposing the compliance with such requirements to producers demanding for certifications is an example of the means which can be employed to reach this goal.

The European market is a favourable context for the implementation of the outlined policies. The institutional framework of the European Union, which already has ambitious targets and programs directed to sustainability, provides relevant regulatory and political tools to be employed. Top-down policies, regulations and directives, as well as the dialectic and competition among Member States, can be some of the means to speed up the process.

However, the enlisted actions can be put in place by both institutional entities, as part of public policies aimed at granting environmental protection and ecological transition, and single companies, as part of corporate strategies aimed at adapting to new markets and increasing profits, in a long-term view.

Any losses in production volumes or due to the changes in supply chains, indeed, could be compensated by the higher prices or by subsidy mechanisms, at the beginning, and by a new subsequent growth, then; providing information for such evaluations is one of the goals of analysis or meta-analysis detecting consumers' preferences and quantifying WTPs. Of course, further studies are needed on this point to support these policies.

5.4 Limitations of the study and future research

This research was planned to be as complete and informative as possible, but some limitations are present and should be highlighted.

From a technical point of view, the decision to perform two separate MRAs had strong reasons behind; in particular, those of commodity consistency and model performance. Also theoretical reasons drove this choice. However, estimating a full model based on the complete dataset and including all attributes, would give additional and interesting results; for example, the impact of the introduction of such a significant variable like *domestic* among the other sustainable attributes would suggest different interpretations. Still on modeling issues, the inclusion of interaction variables also would add new information and capture more complex dynamics; here, the example refers to *information* and its interaction with sustainable attributes and other relevant regressors.

In addition, although even in this case there were specific motivations and research questions underlying the choice, the analysis could be extended to non-European studies, to account for the preferences across the global market, this having policy implications as well. The resultant increase in the number of observations and data heterogeneity would also allow to avoid multicollinearity issues and take into account variables which have been excluded from the present analysis exactly because collinear or not significant since not having enough variation (e.g., other attributes like fish feed or brand, specific dummies for ASC and MSC eco-labels, other species, use of hypothetical methods, number of alternatives, choice model, and so on).

These and other developments are commissioned to future research. Further analyses on the policy side, as already mentioned, would also be required to concretize suggested actions, better understand how to employ estimated values in compensations and subsides, and assess economic feasibility.

Indeed, the present study aims to promote new research on the sustainability of finfish market and lay the groundwork to improve models, experimental design and analysis tools, to have even more quality standards and produce even more robust and reliable results.

Chapter 6

Conclusion

The results of the present thesis are relevant for two main reasons. First of all, they confirm the important role of meta-analysis and, specifically, meta-regression analysis. The incredible amount of findings, estimates, statements, interpretations and information produced by scientific research, which we referred to with the term “jungle of findings”, can actually invalidate or slow down its progress. This also applies to economic research, where results and values from studies conducted on the same topic or issue are sometimes conflicting. Thus, the concerns of James Heckman can be agreed. It has been a basic and imperative principle of scientific method, since its devising at the end of 16th century, that of accurately collecting data, repeating experiments (by changing their conditions and features, if necessary) and testing new ideas and theories, also arisen based on the observed outcomes, until a clear and unquestionable result is obtained. In the discussed case and, in general, in non-experimental sciences, this translates into somehow putting order among the several findings available in literature about the object of study and giving a codified and rational explanation to their differences. This is where meta-analysis or meta-regression analysis comes to the rescue, the latter allowing to create a comprehensive and realistic model, built on those preceding it, describing the phenomenon under consideration, while also extracting new information and correcting old ones. Coming to the analysis of this thesis, it provided values for WTPs more reliable than those obtained by simply comparing and averaging the estimates from similar studies, clarified the direction (sign) or magnitude of preferences for certain attributes, where still uncertain (e.g., towards the product form), and harmonized and synthesized findings originally produced with inconsistent contextual factors and product features, employing an econometric model able to explain the influence of these factors. Having a summarized and easily interpretable portrayal of the economic phenomenon of interest, whose correlations are also quantified in monetary terms, is essential for an appropriate and effective policy making, from the point of view of both planning and results.

This represents the second main reason of the relevance of findings from this research. It highlighted and estimated values, in terms of WTP, for the trends of finfish market, detected the most important product features or sustainable attributes for consumers, modeled the differences derived from the geographical area and/or age, identified those aspects where interventions are needed (e.g., communication) and possible obstacles (e.g., aversion to aquaculture) or favourable elements (e.g., propensity to innovation) to future developments. These are highly valuable (and indispensable) information for decision making. By recalling the research questions, such findings can be resumed as follows.

How consumers' preferences towards production features can be exploited or targeted in market policies in order to increase sustainability of the seafood sector? The positive attitude towards domestic, unprocessed and fresh finfish can lay the groundwork for economic and industrial policies aimed at streamlining production processes and shortening supply chains by relaunching local markets, with the final purpose of reducing the share of emissions, pollutants, waste and resource demand coming from transportation and manufacturing activities. Moreover, consumers' response to awareness and informative communication confirms their effectiveness and should encourage their employment to further promote sustainable consumption, associate preferred attributes with the related impacts, improve acceptance of aquaculture, whose new practices are promising in view of the mitigation of environmental damages and pressure on stocks, and modify purchase behaviors, where needed. The generalized propensity to innovative products supports the investments in research and development, and that to salmon products suggests to focus on such sector. Finally, the heterogeneity in preferences across countries and generations indicates that policies should be diversified according to these factors.

Which are the most valued attributes when it comes to sustainability of the shopping basket? Surprisingly, the nutrition benefits are more important for consumers than other attributes strictly related to sustainable production, e.g., eco-labels. This suggests to employ market strategies to connect these two aspects in purchase decisions and better promote the latters, in order to increase resulting revenues and incentivize companies to adopt sustainable practices and follow standards. Also the link of features like cut, product presentation and harvest method to sustainability should be emphasized since they appear to be less relevant in buying choices involving the sustainable attributes of finfish, although actually having implications in terms of environmental impacts. Moreover, since income results to be influential when looking at the sustainability of purchased products, financial support to households and producers or the diversification of the offer of sustainable products is needed.

The findings of the performed analysis are also relevant to guide new research, especially relatively to the design of discrete choice experiments conducted within the field of fish market, given the information provided about the impact on respondents' WTP of study features like the number of attributes or the delivery method. But, in general, this study can pave the way for further exploration of sustainability issues of this industry and the implementation of the related policies. In particular, it was already pointed out that its value added is the focus, although the considered aspects are several and widely analyzed, to finfish and the European market. This makes the obtained findings more accurate and applicable, because tailored on the specific context and object of study. The past experience of research, economic policies, technological development, international agreements and, more broadly, attempts to deploy solutions to the main issues which, during the last decades, followed the common thread of environmental protection, climate change and resource depletion, has taught us that such solutions must be sector- and context-specific and must take into account a great variety of influential factors. The scope of this thesis is consistent with these considerations.

To sum up, even if the outlined picture of sustainability of food and, therefore, seafood production is far from rosy, the prospects to trigger change are good. The *inertia* of the contemporary economic and social system, as well as of the set-up of its interaction and material exchange with the environment is hard to counter or deviate. However, it is possible and can only be done by acting both with a top-down and bottom-up approach. This means, on one hand, planning policies and regulations at an international and, then, national or regional level and, on the other hand, targeting preferences towards more sustainable consumption in order to influence markets and exert pressure on companies and producers from below. It was already mentioned how the geo-political context of the EU is extremely suitable to such interventions, which also led to the choice of focusing on the European market, and how finfish products, especially those from aquaculture, are very promising for improving the impacts of food industry and ensuring nutrition security. Of course, to reach these goals four things are necessary: political will, favourable starting conditions of consumers' attitudes, accurate information about them (and the related environmental issues) and prospect of profitability from the alternative market system. This research detected the presence of the second one, provided the required information and identified signals of the fourth. It demonstrated that effective policies exist. Now it all depends on the political will.

Bibliography

Abreu, M., H.L.F. De Groot, and R.J.G.M. Florax (2005). 'A meta-analysis of beta-convergence: The legendary two-percent'. In: *Journal of Economic Surveys* 19(3), pp. 389–420.

Ajzen, I., T.C. Brown, and L.H. Rosenthal (1996). 'Information bias in contingent valuation: effects of personal relevance, quality of information, and motivational orientation'. In: *Journal of Environmental Economics and Management* 30(1), pp. 43–57.

Alfnes, F. et al. (2006). 'Consumers' willingness to pay for the color of salmon: a choice experiment with real economic incentives'. In: *American Journal of Agricultural Economics* 88(4), pp. 1050–1061.

Ankamah-Yeboah, I., J.B. Jacobsen, and S.B. Olsen (2018). 'Innovating out of the fishmeal trap: The role of insect-based fish feed in consumers' preferences for fish attributes'. In: *British Food Journal*.

Ankamah-Yeboah, I., J.B. Jacobsen, S.B. Olsen, et al. (2019). 'The impact of animal welfare and environmental information on the choice of organic fish: An empirical investigation of German trout consumers'. In: *Marine Resource Economics* 34(3), pp. 247–266.

Banovic, M. et al. (2019). 'A cross-cultural perspective on impact of health and nutrition claims, country-of-origin and eco-label on consumer choice of new aquaculture products'. In: *Food Research International* 123, pp. 36–47.

Bastounis, A. et al. (2021). 'The impact of environmental sustainability labels on willingness-to-pay for foods: A systematic review and meta-analysis of discrete choice experiments'. In: *Nutrients* 13(8).2677.

BEA (2023). *Gross Domestic Product: Implicit Price Deflator [GDPDEF], retrieved from FRED, Federal Reserve Bank of St. Louis.* URL: `https://fred.stlouisfed.org/series/GDPDEF` (visited on 04/16/2023).

Bergh, J.C.J.M. van den et al. (1997). *Meta-analysis in environmental economics.* Vol. 14. Economy and Environment. Kluwer Academic Publishers.

Berlin, J.A., C.B. Begg, and T.A. Louis (1989). 'An assessment of publication bias using a sample of published clinical trials'. In: *Journal of the American Statistical Association* 84(406), pp. 381–392.

Borenstein, M. et al. (2009). *Introduction to meta-analysis*. Wiley.

Breusch, T.S. and A.R. Pagan (1979). 'A simple test for heteroscedasticity and random coefficient variation'. In: *Econometrica* 47(5), pp. 1287–1294.

Bridges, J.F. et al. (2011). 'Conjoint analysis applications in health—A checklist: A report of the ISPOR Good Research Practices for Conjoint Analysis Task Force'. In: *Value Health* 14, pp. 403–413.

Bronnmann, J. and F. Asche (2017). 'Sustainable seafood from aquaculture and wild fisheries: Insights from a discrete choice experiment in Germany'. In: *Ecological Economics* 142, pp. 113–119.

Bronnmann, J. and J. Hoffmann (2018). 'Consumer preferences for farmed and ecolabeled turbot: A North German perspective'. In: *Aquaculture Economics and Management* 22(3), pp. 342–361.

Canfora, I. (2016). 'Is the short food supply chain an efficient solution for sustainability in food market?' In: *Agriculture and agricultural science procedia* 8, pp. 402–407.

Cantillo, J., J.C. Martín, and C. Román (2020). 'Discrete choice experiments in the analysis of consumers' preferences for finfish products: A systematic literature review'. In: *Food Quality and Preference* 84.

Card, D. and A.B. Krueger (1995). 'Time-series minimum-wage studies: a meta-analysis'. In: *The American Economic Review* 85(2), pp. 238–243.

Chen, X., F. Alfnes, and K. Rickertsen (2015). 'Consumer preferences, ecolabels, and effects of negative environmental information'. In: *AgBioForum* 18(3), pp. 327–336.

Christensen, P. (2003). *Topics in meta-analysis*. Report 692. Institute of Transport Economics, Oslo.

Claret, A. et al. (2014). 'Consumer beliefs regarding farmed versus wild fish'. In: *Appetite* 79, pp. 25–31.

Cooper, H. and L.V. Hedges, eds. (1994). *The handbook of research synthesis*. Russell Sage Foundation.

Crippa, M. et al. (2021). 'Food systems are responsible for a third of global anthropogenic GHG emissions'. In: *Nature Food* 2, pp. 198–209.

Davidson, J. (2000). *Econometric theory*. Wiley.

DNV (2023). *Marine Aquaculture Forecast to 2050*. URL: `https://www.dnv.com/focus-areas/offshore-aquaculture/marine-aquaculture-forecast/index.html` (visited on 06/15/2023).

Dolgopolova, I. and R. Teuber (2018). 'Consumers' willingness to pay for health benefits in food products: A meta-analysis'. In: *Applied Economic Perspectives and Policy* 40(2), pp. 333–352.

Doucouliagos, H. and T.D. Stanley (2009). 'Publication selection bias in minimum-wage research? A meta-regression analysis'. In: *British Journal of Industrial Relations* 47(2), pp. 406–428.
— (2012). 'Are estimates of the value of a statistical life exaggerated?' In: *Journal of Health Economics* 31(1), pp. 197–206.
Egger, M. et al. (1997). 'Bias in meta-analysis detected by a simple, graphical test'. In: *BMJ* 315(7109).
El Bilali, H. et al. (2019). 'Food and nutrition security and sustainability transitions in food systems'. In: *Food and energy security* 8(2).e00154.
EUMOFA (2022). *The EU Fish Market - 2022 Edition.* Tech. rep. European Commission.
FAO (2022a). *The State of World Fisheries and Aquaculture 2022. Towards Blue Transformation.* Tech. rep. Rome: Food and Agriculture Organization.
— (2022b). *World Food and Agriculture – Statistical Yearbook 2022.* Tech. rep. Rome: Food and Agriculture Organization.
Fernández-Polanco, J., S.M. Loose, and L. Luna (2013). 'Are retailers' preferences for seafood attributes predictive for consumer wants? Results from a choice experiment for seabream (Sparus aurata)'. In: *Aquaculture Economics and Management* 17(2), pp. 103–122.
Fisher, R.A. (1932). *Statistical methods for research workers.* 4th. Oliver and Boyd.
Florax, R.J.G.M., P. Nijkamp, and Willis K.G., eds. (2002). *Comparative environmental economic assessment.* Edward Elgar Publishing.
Galgani, L. et al. (2019). 'Impacts of marine litter'. In: *Frontiers in Marine Science* 6.208.
Garcia, A., T. De Magistris, and R.M. Nayga Jr (2012). 'Importance of social influence in consumers' willingness to pay for local food: are there gender differences?' In: *Agribusiness* 28(3), pp. 361–371.
Giraud, G. and R. Halawany (2006). 'Consumers' perception of food traceability in Europe'. In: *98th EAAE Seminar 'Marketing Dynamics within the Global Trading System: New Perspectives', Chania, Crete, Greece.* No. 736-2016-50823. European Association of Agricultural Economists.
Glass, G.V. (1976). 'Primary, secondary, and meta-analysis of research'. In: *Educational Researcher* 5, pp. 3–8.
Goldberger, A.S. (1964). *Econometric theory.* Wiley.
Grebitus, C., J.L. Lusk, and R.M. Nayga Jr (2013). 'Explaining differences in real and hypothetical experimental auctions and choice experiments with personality'. In: *Journal of Economic Psychology* 36, pp. 11–26.
Hamilton, H.A., E. Brod, et al. (2016). 'Investigating cross-sectoral synergies through integrated aquaculture, fisheries, and agriculture phosphorus as-

sessments: A case study of Norway'. In: *Journal of Industrial Ecology* 20(4), pp. 867–881.

Hamilton, H.A., R. Newton, et al. (2020). 'Systems approach to quantify the global omega-3 fatty acid cycle'. In: *Nature Food* 1(1), pp. 59–62.

Hausman, J.A. (1978). 'Specification tests in econometrics'. In: *Econometrica* 46, pp. 1251–1271.

Havránek, T. et al. (2020). 'Reporting guidelines for meta-analysis in economics'. In: *Journal of Economic Surveys* 34(3), pp. 469–475.

Hensher, D.A. (2010). 'Hypothetical bias, choice experiments and willingness to pay'. In: *Transportation Research Part B: Methodological* 44(6), pp. 735–752.

Higgins, J.P.T. (2011). *Cochrane handbook for systematic reviews of interventions. Version 5.1. 0 [updated March 2011]. The Cochrane Collaboration.* Aviable at www.cochrane-handbook.org.

Higgins, J.P.T. and S.G. Thompson (2002). 'Quantifying heterogeneity in a meta-analysis'. In: *Statistics in medicine* 21(11), pp. 1539–1558.

Hinkes, C. and B. Schulze-Ehlers (2018). 'Consumer attitudes and preferences towards pangasius and tilapia: The role of sustainability certification and the country of origin'. In: *Appetite* 127, pp. 171–181.

Hough, C. (2022). *Regional review on status and trends in aquaculture development in Europe – 2020.* Tech. rep. FAO Fisheries and Aquaculture Circular No. 1232/1. Rome: Food and Agriculture Organization.

Hox, J.J., M. Moerbeek, and R. Van de Schoot (2017). *Multilevel analysis: Techniques and applications.* Third edition. Quantitative methodology series. routledge.

IMF (2023). *Implied PPP conversion rate.* URL: `https://www.imf.org/external/datamapper/PPPEX@WEO/OEMDC/ADVEC/WEOWORLD` (visited on 04/16/2023).

ISSF (2023). *ISFF Glossary, under "finfish".* URL: `https://www.iss-foundation.org/glossary/finfish/` (visited on 04/14/2023).

Jaffry, S. et al. (2004). 'Consumer choices for quality and sustainability labelled seafood products in the UK'. In: *Food Policy* 29(3), pp. 215–228.

Jarzębowski, S., M. Bourlakis, and A. Bezat-Jarzębowska (2020). 'Short food supply chains (SFSC) as local and sustainable systems'. In: *Sustainability* 12(11).4715.

Johnson, F.R. et al. (2013). 'Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design good research practices task force'. In: *Value Health* 16, pp. 3–13.

Johnston, R. et al., eds. (2015). *Benefit transfer of environmental and resource values: A guide for researchers and practitioners.* The economics of non-market goods and resources 14. Springer.

Kelley, K. and Preacher K.J. (2012). 'On effect size'. In: *Psychological Methods* 17(2), pp. 137–152.

Khan, I. et al. (2019). 'Public attitudes, preferences and willingness to pay for river ecosystem services'. In: *International Journal of Environmental Research and Public Health* 16(19).3707.

Klaiber, H.A. and R.H. von Haefen (2019). 'Do random coefficients and alternative specific constants improve policy analysis? An empirical investigation of model fit and prediction'. In: *Environmental and Resource Economics* 73, pp. 75–91.

Krystallis, A. and M. Ness (2005). 'Consumer preferences for quality foods from a South European perspective: A conjoint analysis implementation on Greek olive oil'. In: *International Food and Agribusiness Management Review* 8(1030-2016-82535), pp. 62–91.

Lancaster, K.J. (1966). 'A new approach to consumer theory'. In: *Journal of Political Economy* 74(2), pp. 132–157.

Li, S. and Z. Kallas (2021). 'Meta-analysis of consumers' willingness to pay for sustainable food products'. In: *Appetite* 163.105239.

Lizin, S. et al. (2022). 'The state of the art of discrete choice experiments in food research'. In: *Food Quality and Preference* 104678.

Llagostera, P.F. et al. (2019). 'The use of insect meal as a sustainable feeding alternative in aquaculture: Current situation, Spanish consumers' perceptions and willingness to pay'. In: *Journal of Cleaner Production* 229, pp. 10–21.

Louviere, J.J., T.N. Flynn, and R.T. Carson (2010). 'Discrete choice experiments are not conjoint analysis'. In: *Journal of Choice Modelling* 3, pp. 57–72.

Luomala, H.T. (2007). 'Exploring the role of food origin as a source of meanings for consumers and as a determinant of consumers' actual food choices'. In: *Journal of Business Research* 60(2), pp. 122–129.

Lusk, J.L. and T.C. Schroeder (2004). 'Are choice experiments incentive compatible? A test with quality differentiated beef steaks'. In: *American Journal of Agricultural Economics* 86(2), pp. 467–482.

Mariel, P. et al. (2021). *Environmental valuation with discrete choice experiments: Guidance on design, implementation and data analysis*. SpringerBriefs in Economics. Springer.

Marschak, J. (1960). 'Stanford symposium on mathematical methods in the social sciences'. In: ed. by K. Arrow. Stanford University Press. Chap. Binary choice constraints on random utility indications, pp. 312–329.

Mauracher, C., T. Tempesta, and D. Vecchiato (2013). 'Consumer preferences regarding the introduction of new organic products. The case of

the Mediterranean sea bass (Dicentrarchus labrax) in Italy'. In: *Appetite* 63, pp. 84–91.
McCausland, W.D. et al. (2006). 'A simulation model of sustainability of coastal communities: aquaculture, fishing, environment and labour markets'. In: *Ecological Modelling* 193(3-4), pp. 271–294.
McFadden, D. (1974). 'Frontiers in econometrics'. In: ed. by P. Zarembka. Academic Press New York. Chap. Conditional logit analysis of qualitative choice behavior, pp. 105–142.
Menozzi, D. et al. (2020). 'Consumers' preferences and willingness to pay for fish products with health and environmental labels: Evidence from five European countries'. In: *Nutrients* 12(9).2650.
Moeltner, K., K.J. Boyle, and R.W. Paterson (2007). 'Meta-analysis and benefit transfer for resource valuation-addressing classical challenges with Bayesian modeling'. In: *Journal of Environmental Economics and Management* 53(2), pp. 250–269.
Moher, D. et al. (2009). 'Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement'. In: *Annals of internal medicine* 151(4), pp. 264–269.
Moreno, S.G. et al. (2009). 'Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study'. In: *BMC medical research methodology* 9, pp. 1–17.
Murphy, J.J. et al. (2005). 'A meta-analysis of hypothetical bias in stated preference valuation'. In: *Environmental and Resource Economics* 30, pp. 313–325.
Nelson, J.P. and P.E. Kennedy (2009). 'The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment'. In: *Environmental and Resource Economics* 42, pp. 345–377.
Nguyen, T.T. et al. (2015). 'Consumer willingness to pay for quality attributes of fresh seafood: A labeled latent class model'. In: *Food Quality and Preference* 41, pp. 225–236.
OECD (2023). *Purchasing power parities (PPPs)*. URL: `https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm` (visited on 04/03/2023).
Oehlmann, M. et al. (2017). 'Uncovering context-induced status quo effects in choice experiments'. In: *Journal of Environmental Economics and Management* 81, pp. 59–73.
Olesen, I. et al. (2010). 'Eliciting consumers' willingness to pay for organic and welfare-labelled salmon in a non-hypothetical choice experiment'. In: *Livestock Science* 127(2-3), pp. 218–226.

Osch, S. van, S. Hynes, S. Freeman, et al. (2019). 'Estimating the public's preferences for sustainable aquaculture: A country comparison'. In: *Sustainability* 11(3).569.

Osch, S. van, S. Hynes, T. O'Higgins, et al. (2017). 'Estimating the Irish public's willingness to pay for more sustainable salmon produced by integrated multi-trophic aquaculture'. In: *Marine Policy* 84, pp. 220–227.

Page, M.J. et al. (2021). 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews'. In: *BMJ* 372(71).

Parker, R.W. et al. (2018). 'Fuel use and greenhouse gas emissions of world fisheries'. In: *Nature Climate Change* 8(4), pp. 333–337.

Pearce, D. and E. Özdemiroğlu (2002). *Economic valuation with stated preference techniques: Summary guide*. Tech. rep. Department for Transport, Local Government and the Regions (United Kingdom).

Pearson, K. (1904). 'Report on certain enteric fever inoculation statistics'. In: *British Medical Journal* 3, pp. 1243–1246.

Polonsky, M.J. et al. (2001). 'Consumer ethics in the European Union: A comparison of northern and southern views'. In: *Journal of Business Ethics* 31, pp. 117–130.

Printezis, I., C. Grebitus, and S. Hirsch (2019). 'The price is right!? A meta-regression analysis on willingness to pay for local food'. In: *PLoS one* 14(5).

R Core Team (2022). *R: a language and environment for statiatical computing*. R Foundation for Statistical Computing. Vienna, Austria.

Reisch, L., U. Eberle, and S. Lorek (2013). 'Sustainable food consumption: an overview of contemporary issues and policies'. In: *Sustainability: Science, Practice and Policy* 9(2), pp. 7–25.

Rhodes, W. (2012). 'Meta-analysis: An introduction using regression models'. In: *Evaluation Review* 36, pp. 24–71.

Risius, A., M. Janssen, and U. Hamm (2017). 'Consumer preferences for sustainable aquaculture products: Evidence from in-depth interviews, think aloud protocols and choice experiments'. In: *Appetite* 113, pp. 246–254.

Rolfe, J. and R. Brouwer (2013). 'Design effects in a meta-analysis of river health choice experiments in Australia'. In: *Journal of Choice Modelling* 5(2), pp. 81–97.

Rosner, B. (1975). 'On the detection of many outliers'. In: *Technometrics* 17, pp. 221–227.

— (1983). 'Percentage points for a generalized ESD many-outlier procedure'. In: *Technometrics* 25, pp. 165–172.

Rudd, M.A., N. Pelletier, and P. Tyedmers (2011). 'Preferences for health and environmental attributes of farmed salmon amongst southern Ontario

salmon consumers'. In: *Aquaculture Economics and Management* 15(1), pp. 18–45.
Saarinen, M. et al. (2012). 'Life cycle assessment approach to the impact of home-made, ready-to-eat and school lunches on climate and eutrophication'. In: *Journal of Cleaner Production* 28, pp. 177–186.
Sidik, K. and J.N. Jonkman (2007). 'A comparison of heterogeneity variance estimators in combining results of studies'. In: *Statistics in medicine* 26, pp. 1964–1981.
Skirtun, M. et al. (2022). 'Plastic pollution pathways from marine aquaculture practices and potential solutions for the North-East Atlantic region'. In: *Marine Pollution Bulletin* 174.113178.
Smetana, K., R. Melstrom, and T. Malone (2022). 'A meta-regression analysis of consumer willingness to pay for aquaculture products'. In: *Journal of Agricultural and Applied Economics* 54(3), pp. 480–495.
Smith, B.G. (2008). 'Developing sustainable food supply chains'. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1492), pp. 849–861.
Smith, V.K. and Y. Kaoru (1990). 'Signals or noise? Explaining the variation in recreation benefit estimates'. In: *American Journal of Agricultural Economics* 72, pp. 419–433.
Smith, V.K. and S.K. Pattanayak (2002). 'Is meta-analysis a Noah's ark for non-market valuation?' In: *Environmental and Resource Economics* 22, pp. 271–296.
Stanley, T.D. (2001). 'Wheat from chaff: Meta-analysis as quantitative literature review'. In: *Journal of Economic Perspectives* 15, pp. 131–150.
— (2005). 'Beyond publication bias'. In: *Journal of Economic Surveys* 19(3), pp. 309–345.
Stanley, T.D. and H. Doucouliagos (2012). *Meta-regression analysis in economics and business*. routledge.
Stanley, T.D. and S.B. Jarrell (1989). 'Meta-regression analysis: A quantitative method of literature surveys'. In: *Journal of Economic Surveys* 3, pp. 161–170.
Stefani, G., R. Scarpa, and A. Cavicchi (2012). 'Exploring consumer's preferences for farmed sea bream'. In: *Aquaculture International* 20, pp. 673–691.
Stroup, D.F. et al. (2000). 'Meta-analysis of observational studies in epidemiology: A proposal for reporting'. In: *Journal of American Medical Association* 283, pp. 2008–2012.
Subasinghe, R., D. Soto, and J. Jia (2009). 'Global aquaculture and its role in sustainable development'. In: *Reviews in aquaculture* 1(1), pp. 2–9.

Tetley, S. (2016). 'Why the Big 5? Understanding UK seafood consumer behaviour'. Doctor of Philosophy (PhD) thesis. University of Kent.

Thurstone, L. (1927). 'A law of comparative judgement'. In: *Psychological Review* 34, pp. 273–286.

Train, K.E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

UNSD (2023). *Standard country or area codes for statistical use (M49)*. URL: https://unstats.un.org/unsd/methodology/m49/ (visited on 06/12/2023).

Van Houtven, G., J. Powers, and S.K. Pattanayak (2007). 'Valuing water quality improvements in the United States using meta-analysis: Is the glass half-full or half-empty for national policy analysis?' In: *Resource and Energy Economics* 29(3), pp. 206–228.

Walsh, R.G., D.M. Johnson, and J.R. McKean (1989). 'Issues in nonmarket valuation and policy applications: A retrospective glance'. In: *Western Journal of Agricultural Economics* 14, pp. 178–188.

Werner, S. et al. (2016). *Harm caused by marine litter – MSFD GES TG marine litter: thematic report*. Tech. rep. Joint Research Centre (European Commission).

White, H. (1980). 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity'. In: *Econometrica* 48(4), pp. 817–838.

Winther, U. et al. (2020). 'Greenhouse gas emissions of Norwegian seafood products in 2017'. In: *SINTEF Ocean: Trondheim, Norway.*

Yeh, C.H., M. Hartmann, and S. Hirsch (2018). 'Does information on equivalence of standards direct choice? Evidence for organic labels from different countries-of-origin'. In: *Food Quality and Preference* 65, pp. 28–39.

Yeh, C.H. and S. Hirsch (2023). 'A meta-regression analysis on the willingness-to-pay for country-of-origin labelling'. In: *Journal of Agricultural Economics.*

Ziegler, F. et al. (2022). 'Greenhouse gas emissions of Norwegian seafoods: From comprehensive to simplified assessment'. In: *Journal of Industrial Ecology* 26(6), pp. 1908–1919.

Appendix A

List of studies

Study	Species	Har. met.	Form	Presentat.	Country	Info	Online	Classification	S. size	MRA
Ankamah-Yeboah et al. (2018)	Other (trout)	Farmed	Fillet, whole	Fresh, smoked, frozen	Germany	No	Yes	Innovation	610	B
Ankamah-Yeboah et al. (2019)	Other (trout)	Farmed	Fillet, whole	Fresh, smoked, frozen	Germany	No	Yes	Competition	1236	B
Banovic et al. (2019)	Other (amberjack)	Farmed	Fillet, steak	Fresh, canned, smoked	France, Germany, Italy, Spain, UK	No	Yes	Innovation	≈ 105	B
Bronnmann and Asche (2017)	Salmon	Wild, farmed	Fillet	Fresh, frozen	Germany	Both	No	Competition	485	P
Bronnmann and Hoffmann (2018)	Other (turbot)	Wild, farmed	Fillet	Fresh, frozen	Germany	Both	No	Competition	485	B
Chen et al. (2015)	Cod	Wild, farmed	Fillet	Fresh	France	No	No	Competition	116	B
Fernández-Polanco et	Seabream	Wild, farmed	Fillet steak,	Fresh	Spain	Yes	No	Competition	169	B

(Continues)

(Continued)

Study	Species	Har. met.	Form	Presentat.	Country	Info	Online	Classification	S. size	MRA
al. (2013)			whole							
Llagostera et al. (2019)	Seabream	Wild, farmed	Whole	Fresh	Spain	No	Yes	Innovation	215	P
Hinkes and Schulze-Ehlers (2018)	Other (tilapia, pangasius)	Farmed	Fillet	Frozen	Germany	No	Yes	Innovation	325	B
Jaffry et al. (2004)	Salmon, cod, other (tuna, haddock)	Wild, farmed	Fillet, steak	Fresh, smoked, frozen, canned	UK	No	No	Innovation	600	P
Mauracher et al. (2013)	Seabass	Farmed	Whole	Fresh	Italy	Yes	No	Innovation	366	B
Menozzi et al. (2020)	Salmon, cod, seabream, seabass	Wild, farmed	Fillet	Fresh, ready-to-cook	France, Germany, Italy, Spain, UK	No	Yes	Competition	487	B
Olesen et al. (2010)	Salmon	Farmed	Fillet	Fresh	Norway	Yes	No	Competition	86	S

(Continues)

(Continued)

Study	Species	Har. met.	Form	Presentat.	Country	Info	Online	Classification	S. size	MRA
Risius et al. (2017)	Other (trout)	Farmed	Fillet	Smoked	Germany	No	No	Competition	447	S
Stefani et al. (2012)	Seabream	Farmed	Whole	Fresh	Italy	No	No	Competition	251	S
Nguyen et al. (2015)	Salmon, cod, seabream, other (tuna, sole)	Wild, farmed	Fillet, steak, whole	Fresh	France	No	Yes	Competition	960	P
van Osch et al. (2017)	Salmon	Farmed	Fillet	Fresh	Ireland	Yes	Yes	Competition	500	P
van Osch et al. (2019)	Salmon, seabream	Farmed	Fillet	Fresh	Italy, UK, Ireland, Norway	Yes	Yes	Innovation	505	P

Table A.1: List of primary studies included in the meta-regression analyses. Some of them originally take into account also other species, here only those considered for the analysis are reported. When more samples for different countries are employed, the reported sample size is a mean. S = "Sustainability" MRA, P = "Production" MRA, B = both.

Appendix B

Results of statistical tests

B.1 Outputs

Breush-Pagan test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
19.4	0.729	24
Restricted model		
14.4	0.419	14
White test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
22.5	0.999	48
Restricted model		
19.9	0.867	28

Table B.1: Outputs of the Breush-Pagan and White tests, "Production" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Hausman test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
4.318	1	24
Restricted model		
11.448	0.6505	14

Table B.2: Output of the Hausman test, "Production" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

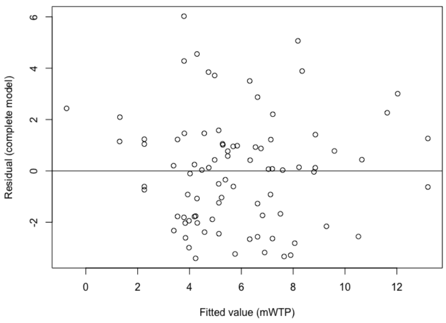
Breush-Pagan test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
18.1	0.837	25
Restricted model		
45.2***	0.0001	16
White test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
18.8	1	50
Restricted model		
46.8**	0.0442	32

Table B.3: Outputs of the Breush-Pagan and White tests, "Sustainability" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

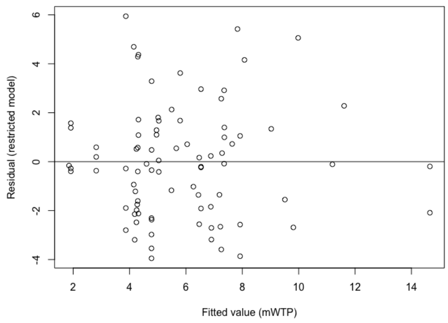
Hausman test		
Test statistic	*p-value (Prob. > ChiSq)*	d.f.
Complete model		
4.929	1	25
Restricted model		
1.115	1	16

Table B.4: Output of the Hausman test, "Sustainability" MRA. Significance levels: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

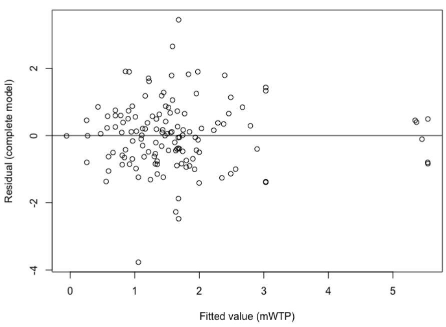
B.2 Residual plots



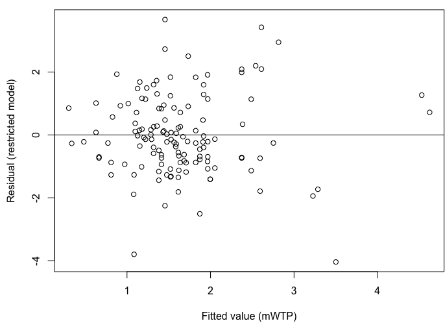
(a) *Complete model, "Production" MRA*



(b) *Restricted model, "Production" MRA*



(c) *Complete model, "Sustainability" MRA*



(d) *Restricted model, "Sustainability" MRA*

Figure B.1: Residual plots of the performed MRAs.