Qualitative and quantitative methods in sustainable development

Iwona Olejnik
Editor

PUEB PRESS
POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS

NAWA
POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE
cenet sie
AKADEMICKIE PARTNERSTWA MIĘDZYNARODOWE

Qualitative and quantitative methods in sustainable development

Iwona Olejnik
Editor

PUEB PRESS

Poznań University of Economics and Business

Poznań 2021

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

CENETSIE

AKADEMICKIE PARTNERSTWA MIĘDZYNARODOWE

REVIEWER
Martina Hanová

COVER DESIGN
Piotr Gołębniak

MANAGING EDITOR
Marta Dobrecka

PROOFREADER
Bernadeta John-Jankowska

DTP: eMPi²
Reginaldo Cammarano

Publication financed by Polish National Agency for Academic Exchange
Project *Central European Network for Sustainable and Innovative Economy*,
no. PPI/APM/2019/1/00047/U/00001

eISBN 978-83-8211-072-2
https://doi.org/10.18559/978-83-8211-072-2

POZNAŃ UNIVERSITY OF ECONOMICS AND BUSINESS PRESS
ul. Powstańców Wielkopolskich 16, 61–895 Poznań, Poland
phone: +48 61 854 31 54, 61 854 31 55
www.wydawnictwo.ue.poznan.pl, e-mail: wydawnictwo@ue.poznan.pl
postal address: al. Niepodległości 10, 61–875 Poznań, Poland

CONTENTS

PREFACE

In order to conduct research and analysis on sustainable development, it is worth introducing the concept of this idea first. Sustainable development is the idea that "human societies must live and meet their needs without compromising the ability of future generations to meet their own needs. It contains within it two key concepts:

1) the concept of 'needs', in particular the essential needs of the world's poor, to which overriding priority should be given;
2) the idea of limitations imposed by the state of technology and social organization on the environment's ability to meet present and future needs" (*Report of the World Commission*).

Thus, sustainable development is a way of organizing society so that it can exist in the long term. This means that the protection of the environment and natural resources as well as social and economic equity should become the most important standards guiding the behavior of all individuals.

The Agenda for Sustainable Development 2030 defines 17 goals that are to be achieved by the world by 2030. They were adopted by all UN Member States in 2015, as part of the 2030 Agenda for Sustainable Development which set out a 15-year plan to achieve the goals.

The goals address the global challenges we face, including poverty, inequality, climate change, environmental degradation, peace and justice. As a result, they concern achievements in 5 areas: people, planet, prosperity, peace, partnership. They are presented in Table 1.

Systematic research and comprehensive analyses allow to monitor the implementation of the sustainable development goals. Systematic analysis of the changes taking place is the basis for coordinating programs aimed at sustainable development.

Obviously, when you are interested in the selected issue of sustainable development, it is worth using data from the secondary sources in the first place. The main secondary resources on sustainable development include international reports, government reports and statistics, reports from institutions or organizations, as well as articles in scientific journals, newspapers and magazines.

Table 1. Sustainable development—goals

Goal	Content
1	**End poverty** in all its forms everywhere: *economic growth must be inclusive to provide sustainable jobs and promote equality*
2	**End hunger**, achieve food security and improved nutrition and promote sustainable agriculture: *the food and agriculture sector offers key solutions for development, and is central for hunger and poverty eradication*
3	**Ensure healthy lives** and promote **well-being** for all at all ages
4	Ensure inclusive and equitable **quality education** and promote lifelong learning opportunities for all
5	Achieve **gender equality** and empower all women and girls: *gender equality is not only a fundamental human right, but a necessary foundation for a peaceful, prosperous and sustainable world*
6	**Avoid wasting water**. Ensure availability and sustainable management of water and sanitation for all: *clean, accessible water for all is an essential part of the world we want to live in*
7	Ensure access to **affordable, reliable, sustainable and modern energy** for all
8	Promote sustained, inclusive and sustainable economic growth, full and productive **employment and decent work for all**
9	**Build resilient infrastructure**, promote inclusive and sustainable industrialization and foster innovation
10	**Reduce inequality** within and among countries: *to reduce inequalities, policies should be universal in principle, paying attention to the needs of disadvantaged and marginalized populations*
11	**Sustainable cities and communities.** Make cities and human settlements inclusive, safe, resilient and sustainable: *there needs to be a future in which cities provide opportunities for all, with access to basic services, energy, housing, transportation and more*
12	Ensure **sustainable, responsible consumption and production** patterns
13	**Climate action.** Take urgent action to combat climate change and its impacts
14	**Life below water.** Conserve and sustainably use the oceans, seas and marine resources for sustainable development
15	**Life on land.** Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss: *plant a tree and help protect the environment*
16	**Peace, justice and strong institutions.** Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
17	Strengthen the means of implementation and revitalize the Global **Partnership for Sustainable Development**

Source: (Global indicator framework; https://www.un.org/sustainabledevelopment/sustainable-development--goals/).

Quite often, however, the problem you are interested in, may concern a very specific issue or area in which no research has been conducted so far, e.g., research on equality in a specific enterprise or group of enterprises, or on wasting water in households in a certain region of your interest. Then primary research methods need to be used.

*

How to gather data and how to analyse them? This e-book presents a few selected methods that will allow you to answer these questions at least partially. Among the data collection methods presented, in this book, we have chosen both: qualitative, in particular focused group interviews, and quantitative—based on a questionnaire. In terms of data analysis methods, we present three methods: factor analysis, structural equation modelling and data envelopment analysis. The examples presented in this book relate to sustainable development, e.g.: sustainable consumption, ecological culture, better nutrition, agricultural development and many more.

Acknowledgements
The book was created thanks to the involvement of many people from several universities: University of Zagreb, Croatia, The D. A. Tsenov Academy of Economics, Bulgaria, as well as the Poznań University of Economics and Business, Poland.

I would like to sincerely thank all the authors for their commitment. In the first part, concerning primary research, Atanaska Reshetkova and Iwona Olejnik prepared texts on designing and conducting qualitative research. In this part of the e-book, Sylwester Białowąs, Blaženka Knežević, Magdalena Stefańska and Iwona Olejnik also prepared texts on quantitative research, introducing the reader to the levels of measurement, the principles of constructing the questionnaire, sample selection and methods of presenting research results.

The second part of the e-book presents three selected methods of analysing primary or secondary data. Blaženka Knežević, Todor Krastevich, Atanaska Reshetkova, Agnieszka Sapa, Katarzyna Smędzik-Ambroży, Magdalena Stefańska and Iwona Olejnik took up the challenge of explaining the essence of three methods of analysis: factor analysis, structural equation modelling and data envelopment in sustainable development research. Thank you all.

This publication would not have been written if it had not been for the international project, coordinated by Barbara Borusiak. It is thanks to her efforts and commitment that cooperation between universities from Croatia, the Czech Republic, Bulgaria, the Ukraine, Slovakia, Hungary and Poland has been established and strengthened. The project entitled Central European Network for Sustainable and Innovative Economy (CENETSIE) is financed by NAWA funds intended for the development of international cooperation. The project has been implemented in the years 2020–2022.

On behalf of the entire team of authors, I would like to thank the reviewer Ing. Martina Hanová, Ph.D., Slovak University of Agriculture in Nitra for the valuable comments.

We would also like to thank the employees of the Publishing House of the Poznań University of Economics and Business for their help in publishing the e-book, in particular the editor Marta Dobrecka.

We hope, as a team of authors, that thanks to international cooperation on the book, a series of joint research, organisational and teaching works will be continued.

We sincerely hope that our publication will be useful for both students and other interested people, especially those who would like to apply qualitative or quantitative research methods in the area of sustainable development problems on their own.

Iwona Olejnik

References

Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. Retrieved October, 2020 from https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202020%20review_Eng.pdf

https://www.un.org/sustainabledevelopment/sustainable-development-goals/ Retrieved January, 2021.

Report of the World Commission on Environment and Development: Our Common Future. Retrieved January, 2021 from http://www.un-documents.net/our-common-future.pdf

PART 1.

PRIMARY DATA COLLECTION

Qualitative and quantitative methods in sustainable development, pp. 13-40
https://doi.org/10.18559/978-83-8211-072-2/01

1.

QUALITATIVE METHODS

Iwona Olejnik Atanaska Reshetkova
Poznań University of Economics and Business

Atanaska Reshetkova
D. A. Tsenov Academy of Economics, Svishtov

Abstract: The main goal of the chapter is to present how to use qualitative methods in sustainability research. First, the theoretical basis of the methods is presented, i.e., the essence of qualitative methods, what differs them from quantitative methods, and their types.

The second part of this chapter covers designing and conducting a focus group interview (FGI): its essence and main stages, sampling, projection techniques and the script, as well as it contains the case study of ecological culture of Bulgarians.

The third part presents considerations necessary to conduct a qualitative research, i.e., the organizational aspects of FGI and the guidelines for the work of the moderator. Finally, the last part shows considerations concerning data analysis—using CAQDAS software.

Keywords: CAQDAS software, ecological culture, qualitative methods.

1.1. Methodology of qualitative research—the basics

1.1.1. Qualitative research—theoretical background

Market researchers are noticing a growing discrepancy between what clients declare officially and the actions they take. The implementation of qualitative research methodology, whose aim is to get to the "soul" of a consumer, recognize their emotions, hidden motives and needs, becomes essential to identify the actual market behaviours (Kaczmarek, Olejnik, & Springer, 2013). Qualitative methods make it possible to go beyond consumers' declarations and recognize their subconscious patterns of behaviour, as well as understand the reasons for their irrational decisions.

The main differences between qualitative and quantitative marketing research are presented in Table 1.1.

Table 1.1. Qualitative vs quantitative research—the main features

Criterion	Qualitative research	Quantitative research
main research questions	what? how? why?	how much? how many? how often?
tool of measurement	scripts: questions are the outline of an interview	questionnaire
sampling	non-random; purposive smaller samples (20–50 people)	random, quota bigger samples (200–1000 people and more)
influence of a researcher on the progress of the research	greater	smaller
possibility of quantitative generalization of the results for the population	no	yes
interpretation method	no statistics—more liberal and subjective	statistics—more objective

Source: (Creswell & Creswell, 2018).

The main features of qualitative research include:

– the guidelines for quantitative research are determined on it basis;
– it may deepen and extend the results of quantitative research;
– it makes it easier to get closer to natural, real world of the respondents;
– its course is audio-visually recorded (on cameras, or voice recorders);
– it is not representative; no statistical reasoning can be performed;
– it requires the use of non-standardized tools;
– the obtained information is presented in a descriptive way (thoughts, attitudes, feelings).

Qualitative research
The process of collecting and analysing non-numeric data (e.g., text that is a transcript of an interview, video, audio, photo) in order to recognize the essence of a specific phenomenon, understand concepts, learn about opinions or experiences related to a given topic.

In turn, considering the most important applications of qualitative research, we can use this method in the following situations:
- launching a product (modification and adjustment of the features: package, name, or image of a launched product);
- product positioning (discovering the most effective ways of sharing information on a product or service with potential customers);
- research on the habits connected with the use of a product (learning about the needs and motives of a given group of products or a brand);
- research on attitudes (learning about opinions, emotions, and associations connected with a tested product or concept);
- advertising message research;
- generation of ideas (creating new ideas—creative groups).

1.1.2. Qualitative research methods

There are a lot of qualitative methods in marketing research. The most important of them are presented in Figure 1.1.

Observation	Interview	Content analysis
– participating or not participating – open or hidden – structured, partially-structured or unstructured	– focus group interviews – individual in-depth interviews – dyads, triads, mini-groups – affinity group – narrative interviews – extended groups – panel of experts	– private documents (e.g. letters, e-mails, contracts, invoices, materials published in social media) – public documents (e.g. newspapers, magazines, official reports and reports, entries in social media)

Figure 1.1. Qualitative research methods

Source: Own elaboration.

Observation is a systematic collection and analysis of data, both verbal and non-verbal, about consumer behaviour. It is based on collecting data using sight and hearing. As a research method, it is a deliberate, systematic, selective and objective activity. This method allows you to understand how and why people behave in a certain way, which is often different from what they think. Therefore, the purpose of observation is to provide information about the real behaviour of people, about what their daily activities look like, i.e., the routine, often followed beyond consciousness. For most people, these activities include those that are done every day at home, on the way to the store (work), while shopping, while working, while spending free time, in the waiting room at the doctor's office, at the at the airport, in the car, reading the newspaper, using the Internet, etc. (Olejnik, 2011, p. 243). Observation can be a good method to be used in comparisons, including competitive ones. The arrangement of stores, merchandising, customer service rules are just examples of areas in which you can legally observe competitors using, e.g., hidden observation, i.e., mystery shoppers.

In the case of content analysis, the information contained in, e.g., newspapers, magazines, books, films, websites, blogs, social media, photos, art objects, song texts, e-mails are analysed.

The interviews, on the other hand, are based, in particular, on verbal data. A researcher-moderator conducts a conversation with one or more respondents on a specific topic, trying to find a solution to the research problem (more—see next part).

The most popular qualitative methods are focus group interview (FGI) and individual in-depth interview (IDI). The comparison between both of these methods is presented in Table 1.2.

Table 1.2. Focus group interview and individual in-depth interview—comparison of main features

FGI	IDI
1 interviewer: 6–8 respondents	1 interviewer: 1 respondent
more information in a shorter time, but less information from one respondent (is less in-depth)	less information in particular time, but more information from one respondent
during the interviews there are discussions among the respondents (interaction)	no discussions (contact) among the respondents
the recruitment is easy, because the respondents are easily available, e.g., housewives, students	there is difficult recruitment, e.g., nuclear engineer
there is the need to confront the opinions of particular respondents	there is also the need to obtain in-depth information from one person
we assume that the presence of others may stimulate the respondents to express their opinions	when we are worried that the presence of other people may block or influence the replies of the respondents

Source: (Creswell & Creswell, 2018).

In each method there are three main stages of research process: preparation, execution and data analysis with interpretation. In the first stage—preparation, we should:

– formulate the research problem and the aims of the research,

- create a research team,
- choose research methods and techniques,
- plan sample selection,
- construct research instruments and prepare additional materials.

In the second stage—performance execution, the researcher collects data and prepares material for analysis and interpretation. In the last phase, the material is analysed and interpreted. In addition, a presentation of the results is prepared for the client.

1.2. Designing a focus group interview

1.2.1. The essence and main stages of focus group interview

Focus group interview as a research method was used for the first time in sociology, by two scientists, Paul F. Lazarsfeld and Robert K. Merton (Merton, Fiske, & Kendall, 1962; Merton, 1987). They are considered to be the precursors of the application of this method. In the 1940s, they used focus group to test radio propaganda broadcasts (Barbour, 2007).

Focus group interview (FGI) is the most frequently used method of collecting qualitative data. The essence of FGI includes:

1) the data collected with this method are a result of the effort of all research respondents—so if somebody does not want to speak, the researcher should activate them;
2) the group is an informal gathering / collection of purposefully selected people, from whom we try to get information on a particular subject;
3) the group most often consists of 6 to 8 people and most often it is relatively homogenic;
4) the discussion is conducted by a trained moderator, whose task is to make the respondents provide their answers, using the prepared techniques and materials;
5) the FGI does not generate quantitative information, which might be extended on population;
6) the aim of an FGI is to record any behaviour that is in any way related to the researched subject (that is why the image and voice recording equipment is used at the interview).

What are the main stages of an FGI? If we know why we want to conduct an FGI[1], we have to define a research problem and research objectives. Then the next thing to do is to set the number of discussion groups and the rules for selecting the

[1] FGI should be used if it is required to: 1) study consumer attitude to products and services, brands, companies; 2) develop a strategy on the product or services positioning, 3) determine negative characteristics of products or services in consumers' perception; 4) find new ideas on the develop-

respondents. The third step is to determine the technical and location conditions of conducting the interviews. And then it is necessary to prepare a design (scenario) of the interview. Of course, similarly to quantitative research we should prepare a time schedule and a budget for the research. Conducting the interviews is the last, but not less important stage. Finally, we analyse and evaluate the results and prepare a report.

1.2.2. Sampling

It is more common for quantitative research to apply random sample selection, which allows us to obtain a representative sample of a given population, whereas for qualitative research the sample choice is most often purposive. It basically means selecting people who constitute a complete (diversified) collection of empirical cases, allowing us to thoroughly recognize a researched phenomenon and achieve our research objectives. For example, if we want to identify a decision-making process of people buying used clothes, we have to invite to our interview the respondents who have bought clothes in second hands shops.

In qualitative research **the sample choice is most often purposive.**

What are the **rules and criteria of participant selection**? One of the basic criteria in the sampling process are demographic criteria, e.g. gender, age, education, income. The criteria that result from the research objectives / purpose, e.g., we need respondents who segregate (or do not) garbage, are also very important for selecting respondents to the sample.

There are also a few additional criteria that are worth considering, e.g. assertiveness, creativity of participants. Preferred respondents are also people who have not recently participated in such surveys (this is due to the fact that such respondents remember the course of such research and may not react as "fresh" as the first-time participants—especially if the research concerns a similar product category or if the same projection techniques are used; on the other hand, if they have participated in such a study before, it is less stressful for them). What is more, people with education or experience in marketing, sociology, psychology or production / sale of the same products (and members from their immediate family) cannot participate either.

The number of cases (people, FGI groups) is often arbitrary determined. For example, when we research students' eating habits, we can assume that they vary depending on gender and accommodation (with family or on their own), so when

ment of products, services or brands; 5) find new ideas on the development of products or services; 6) test new products or new advertising materials, etc.

determining a sample, it is worth considering these criteria for the sample selection and size, e.g. research in 4 FGI groups, where the first group consists of female students living on their own, the second group—of female students living with family, the third group—male students living on their own, and the fourth one—male students living with their family.

There are many other ways of sampling, such as:

- **combining the so-called extreme cases**, e.g., having a discussion with 3 ardent critics of an X brand and 3 very loyal customers supporting the same brand;
- **random-purposive sampling**—when purposive sampling of individuals could create the sample too large for the needs, e.g., a list has been obtained of a few dozen elementary school maths teachers with up to 5 years teaching experience in a given area, and then, based on this list, a purposive sampling has been conducted for the FGI;
- **the "snowball" sampling**—the method is used with non-easily available respondents. It is basically reduced to finding a first respondent, who then indicates the researcher another person matching the research subject, e.g., it is highly probable that a fan of Olga Tokarczuk's books has friends who are also readers of the same writer, and a Labrador dog's owner presumably knows other owners of the same breed.

Beside the research objective it is also very important not to mix in one FGI „ordinary" customers (users) with people professionally related to a product and therefore having much more knowledge on the research subject. For example, in beer consumers' research, certain individuals should not take part in an FGI, namely: brewery industry employees, bartenders, alcoholic beverages salespeople, etc., which applies to their close relatives as well. Moreover, because of their profession, marketing specialists, marketing researchers, psychologists or sociologists are also not invited to participate.

Another essential thing in the process of qualitative research sampling is to determine the optimal sample size. How many groups and how many people should we interview? The size of a qualitative sample should be big enough to guarantee the recognition of most (or all if possible) observations concerning the researched subject. Therefore, the bigger the sample, the larger the chance for more complex recognition of, e.g., reasons for service quality satisfaction (or dissatisfaction), advertising concept evaluation, or consumer needs, as well as opinions and decision identification.

Summarizing, what should be the sample size in FGI? It depends on many factors—research problem, research objectives, budget, or time set for completing research and presenting its results. The size of a single focus group interview, where respondents can speak freely, is between 6 and 9 people, and FGIs are conducted in at least 2–3 groups.

1.2.3. Ecological culture of Bulgarians: case study (Part 1)

In 2012 Bulgarian Ministry of Environment and Water initiated research in an effort to understand the "ecological culture" of Bulgarians. The research consisted of two parts—qualitative research using focus group interviews, and quantitative research—using a standardized questionnaire. The main goal of the qualitative research was to study in detail the attitudes of Bulgarian citizens to ecology, ecological issues, and the measures undertaken by the government to solve these issues, as well as to better understand everyday behaviour regarding ecological issues and personal involvement in environmental topics. Four focus groups were organized in three major cities: Sofia, Plovdiv, and Veliko Tarnovo. For more details on demographics, see Table 1.3.

Table 1.3. Demographics of participants

Location	Sofia (1)	Sofia (2)	Plovdiv	Veliko Tarnovo
Age	20–35 years	35–55 years	35–55 years	20–35 years
Number	7	8	8	8
Gender	men – 3 women – 4	men – 4 women – 4	men – 4 women – 4	men – 4 women – 5
Education	higher education – 5 secondary education – 2	higher education – 5 secondary education – 3	higher education – 6 secondary education – 2	higher education – 3 secondary education – 5
Marital status	married with children – 1 divorced with children – 0 not married – 6	married with children – 5 divorced with children – 2 not married – 1	married with children – 5 divorced with children – 1 not married – 2	married with children – 1 divorced with children – 1 not married – 6

Source: Own elaboration.

In each focus group, the discussion went through three levels:
- everyday life (limit waste and save energy in households);
- nationwide issues (prevention of pollution of air, soil, water, and food);
- ideology (harmony between humans and nature).

The main results indicated that although the awareness of environmental issues is growing, the everyday "ecological culture" of Bulgarians is relatively low. People acknowledge the fact that Bulgarian society is not active enough, and they lack the motivation to contribute to public initiatives. Few people are engaging in waste sorting, and few are using energy-saving appliances. Most participants share the idea that there should be more activities to educate people on the main ecological issues, and this should start at school. People approve of the nationwide campaign to limit the use of plastic bags. Most participants are not aware of the government measures to protect the environment, and at the same time, they do not trust the institutions that should ensure environmental protection.

1.2.4. Projection techniques

In psychology projection means a transfer of our experiences, sensations, feelings, motives or ideas to another person or object. Projection allows us to express our behaviour, but without being always aware that we are describing ourselves. When we use projection techniques in marketing research, the respondents are asked to interpret the features and behaviour of the others. Thus, this is one of the ways to discover the truth about a consumer.

Table 1.4. Why do we use projection techniques?

- You can ask directly about a lot of things but sometimes the received answer is not true
- A respondent does not always want to reveal too much, e.g., all the reasons for buying a product, for fear of being judged by others or because he / she is simply not aware of them
- These techniques help protect respondents' self-esteem and allow us to avoid their emotional discomfort,
- We want to diversify the research
- We want to get independent opinions of each respondent without possible influence of a group on these opinions

Source: Own elaboration.

There are many kinds of projection techniques. In general, they can be divided into two groups:
- individual,
- collective.

Selected types of projection techniques with a brief explanation and examples are presented in Table 1.5.

Table 1.5. Projection techniques—example

Technique	Description	Example
Word association – free – controlled *(individual technique)*	words or images evoke some free associations, e.g., a logo, a name, a product, a service or a brand is associated with a colour, a tree, an object, an animal or a person	give a single word associated with "sustainability" *(free)* which of the following words (*here are these words*) do you most associate with "ecology" *(controlled)*?

Technique	Description	Example
Completion techniques – comic strips, – supplementary questions, – test of incomplete statements *(individual technique)*	they include finishing statements, opinions, stories, pictures the respondents are asked to complete the sentences with the first words that come to their mind.	first of all, I would tell the Minister for the Environment that .... What irritates me most about using ...., is .... I would convince my friends to sort the garbage by telling them that .... A person who does not sort garbage behaves .... and ....
Picture sorting *(individual and group technique)*	each person (or mini-group) gets a set of pictures of: – different people, among which they, e.g., choose a typical user of a product and describe him / her in detail – people, who are then to be sorted into two groups—supporters and opponents	divide these photos into two groups—people who always sort garbage and those who never sort it
Brand party *(group technique)*	this includes imagining all discussed brands at the same time. It helps determine the image of a brand in comparison with the competing ones	let us imagine all brands are people invited to a party – what do they look like? – how do they behave? – what are they chatting about? – how do they treat each other? – who is dominating? – who is staying aloof?
Collage construction *(group technique)*	usually, a set of pictures is prepared with various images of inquired associations it is also used with personification or defining a „user profile”	please choose from the photos and magazines on the table the images that will best show the world of a person who cares for the natural environment and lives ecologically

Source: Own elaboration.

1.2.5. The script

The script is a plan of the interview, stating the main points to be discussed at the interview, and the time set for discussing these points (Table 1.6). It is worth emphasizing that a script:

- is prepared by a moderator on the basis of the research objectives,
- is not a questionnaire,
- shows the direction in which the discussion should go,
- presents discussion points in a logical, top-down order,

- contains not too many questions; usually, there are 3–4 discussion points (main area of discussion) for an interview lasting between one and a half to two hours.

Table 1.6. The outline of a script

1. Introduction (5 min) – moderator's self-introduction – informing the participants on the research subject – informing about the rules of the discussion – informing the participants that the discussion is recorded – participants' self-introductions 2. Introductory questions—„a warm-up"—easy questions which everybody should be able to answer. For example, in research on the quality of food it can be: – *Please finish the following sentence: Good food means to me ....* – *What does it mean to you that some food is of good quality?* 3. Questions essential for achieving the research objective and supporting techniques (50–60 min) 4. Closing questions (15 min)

Source: Own elaboration.

We must remember that the questions essential for achieving the research objective should be open, e.g.: *If you were the manager, what would you change first?*, and non-threatening as e.g. a question with reproach: *Why didn't you give up smoking?*

1.2.6. Ecological culture of Bulgarians: case study (Part 2)

In the previous part of this case study, a brief description of a qualitative research initiated by the Bulgarian Ministry of Environment and Water was presented. The goal of the research was to study in detail the attitudes of Bulgarian citizens to ecology, ecological issues, and the measures undertaken by the government to solve these issues, as well as to better understand everyday behaviour regarding ecological issues and personal involvement in environmental topics. In this part, the planning stage of the focus group interview—the creation of questions and script are focused on.[2]

Before the start of each session, participants received a document explaining the privacy policy, and a form to express their consent for researchers to record, store, analyse, and present their opinions. When the discussion started, the moderator followed the script shown in Table 1.7.

[2] Please note that the questions and the script are not the ones used by the research agency that conducted the research. The given script is created specifically to fit the learning objectives of this chapter.

Table 1.7. Ecological culture of Bulgarians—script of the focus group

Script item	Estimated time
Introduction The moderator introduces himself and describes the purpose of the interview. Then he asks for permission to record the session with an audio and / or video recorder. He kindly invites participants to introduce themselves	5 min
Warm-up questions – What is your definition of ecology? – What do you hear people around you say about today's ecological problems? – What kind of people care about the environment, and why? How do you know that they care?	10 min
Everyday life – How can households limit the waste they produce? – Which of the above-mentioned measures do you apply at your home? – What sources of information about ecological issues do you typically use? – Do you participate in any nature preserving initiatives? – Does your household engage in a separate collection of waste? – What do you think about energy-saving appliances? Do you have any?	20 min
Nationwide measures – Who do you think is responsible for the prevention of pollution of air, soil, and water? *Probe:* Do you think that people / government should also be held responsible? – Are you aware of any initiatives of the government that are aimed at solving a major environmental problem? *Probe:* Do you think that there are none or they are not well promoted? – Do you think that the government does enough to protect the environment? *Probe:* What more should be done? – Can you name a recent activity—a government or citizens' initiative—aimed to protect the environment that you remember taking place in your hometown? – Do Bulgarians participate in initiatives intended to preserve nature for future generations? *Probe:* What are the most popular initiatives you have heard of? – Are people in Bulgaria educated on the environmental issues in the country and the whole world? *Probe:* What could be done to make them more educated?	20 min
Ideological questions – How do you understand the relationship between people and the environment? – Is there a harmony between humans and nature? *Probe:* What is causing this lack of harmony? How can this harmony be restored? – How important is the protection of the environment for the personal well-being of most Bulgarians? – What is the reason why some people ignore environmental problems?	20 min
Closing questions – Is there anything else you would like to share? – Does anyone have something to add that was not commented so far?	15 min

Source: Own elaboration.

1.3. Conducting a focus group interview

1.3.1. The organizational aspects of FGI

Inviting respondents to the interview we must remember about:
- „back-up participants"—as there may be people who will not come (e.g., while conducting and FGI with people who buy clothing in second-hand stores, out of 10 invited women only 6 joined, even though they had confirmed their attendance by phone the previous evening);
- incentive tools to make people come and participate in a discussion—taking part in an interview is often rewarded with a gift or some amount of money—larger or smaller, e.g. the respondent can receive a gift voucher for the value of 20–25 euros.
- the time and location of the interview—it should finish before 8.00 p.m. If we invite housewives with low income, we should not arrange for it to be in a luxury hotel;
- preparation of recording equipment (along with backup equipment).

1.3.2. FGI moderation

A person who conducts an interview is a moderator. His / her preparation, experience and predisposition are essential for a properly conducted interview (McDaniel & Gates, 2018).

Features of a good moderator
- shows respect for respondents
- expresses himself / herself clearly and loudly
- asks simple questions
- actively listens to answers
- demonstrates creativity and flexibility when asking questions
- uses various techniques involving all respondents
- adopts an open body posture
- conducts the discussion in a logical sequence, from general to specific questions

When asking questions, a moderator should:
- make further inquiry about an ambiguous thing said by a respondent. *Could you explain it to me again, in other words?*
- extend the statements *Could you tell me more about it?*, or *So you mean the biggest advantage of this product is its simplicity?*

How to invite a person to give us some opinion and extend conversations? The moderator could ask *do you agree*?, but it is quite closed question. So, you should open up that question to get a more interesting answer.

There are some examples of a polite way of inviting respondents into the discussion:

- *What do you think…?*
- *What is your opinion…?*
- *How do you feel…?*
- *I am interested to hear your opinion about this.*
- *I would like to hear your thoughts on…*
- *Would you like to add anything?*

On the other hand, what should researcher say when they want to interrupt—when someone else is talking and talking. He or she could say:

- *Excuse me,*
- *Do you mind if someone else adds to that,*
- *Sorry to interrupt but…, sorry for interrupting but…,* or *if I may interrupt for a minute…*

And when somebody has gone off the topic, and started talking about something else—not necessarily connected with the main topic, the moderator has to lead the conversation back to the topic, he / she can say something like this:

- *So anyway, getting back to my question…*
- *So anyway, where were we?*
- *As I was saying / asking…*

In summary, there are a few techniques of conducting group discussion, including: paraphrasing, confirmation and further extending questions.

The beginnings of paraphrasing statements:
I understand that ....
What you mean is ....
What you are saying is .... *(and repeat back what you heard),*
I understand that you are asking about ....
Do you mean ....? *(and then paraphrase what you heard),*
In other words, what you are saying is .... *(and summarize what you heard),*
I would like to confirm what you said ....
Can I just confirm that ....?

Paraphrasing is shortened restating of a person's words by a moderator. It should not contain more details than you have heard. If you use a paraphrase, you show that you are paying attention and understand what has been said and it highlights your interest in what a respondent has to say. Moreover, it allows you to check if you have correctly understood a respondent's intentions (if you have misunderstood, your respondent has a chance to correct your mistake and present the idea more clearly) and allows you to organize the content of a respondent's message.

On the other hand, we use clarification when we ask the respondent about some unclear answer. We then use sentences like: *Could you explain this to me again, in other words?* or *Can you explain once again what it means to you?*

The examples of further extending questions include: *Could you say something more about it? Could you give me more details?*

Finally, it is worth adding that when the topic is controversial, it does not mean, that moderator needs to avoid it. How to keep the conversation open and positive in this situation? The moderator could start like this: *I know not everyone agrees with it…* or *I know this is a contentious issue, but…*

When conducting FGIs we may come across various problems. These include the following: overeager debater, untalkative—'silent' participants, overactive or passive group and jokers. But of course, a well-prepared, qualified moderator can handle these issues.

Conducting face-to-face interviews is different from online projects. The basic differences in moderating interviews are presented in Table 1.8.

Table 1.8. Moderating online and face-to-face interviews

Online	Face-to-face
– the discussion may be conducted by more than 1 moderator; a second moderator is very useful – it allows you to generate more data in less time – length of the interview: maximum 90 minutes – much greater enthusiasm and involvement of the moderator is necessary – greater opportunities for communication between moderators, and moderators and the research client	– the interview is usually conducted by only 1 moderator – it allows you to generate less data in a longer time – length of the interview: about 1.5–2 hours – it is easier to control the involvement of all respondents – less communication opportunities between the moderator and his assistant, and between the moderator and the manager during the survey

Source: (Olejnik, Dębska & Zieliński, 2020).

1.4. Analysis of qualitative data using the CAQDAS programs

1.4.1. Fundamentals of qualitative data analysis

CAQDAS is an acronym for computer-assisted qualitative data analysis software. Qualitative data typically includes text, images, audio and video materials. The decision whether or not to use CAQDAS software in the process of analysing qualitative data depends on two considerations: first, whether the data processing is feasible by the researcher alone or it would take too much effort and time, and second, whether the researcher is willing to adapt the whole research process to the specific requirements of the software. The software package can facilitate the analysis but cannot do it alone—the researcher is the one with expertise who interprets the results in the light of research goals and hypothesis. It is essential to be aware of the fact that both the researcher and the software can have impact on the research process and its outcomes. The researcher should refrain from conveying his or her expectations and opinions on the interpretation of the results. At the same time, one should use CAQDAS packages and its various functionalities only when the outcomes have implications for the subject of matter. Some researchers believe that this type of software makes analysis outcomes more plausible because it is unbiased and brings some structure to the analytical process. Unfortunately, this is not true, especially when using too much automatic procedures leads to overlooking the complexity of the data.

Just like any other research project, qualitative research can use two main types of data:

- primary,
- secondary.

Primary data includes interview and focus group transcripts, created through the processing of the audio and video materials, and notes and other materials collected during the fieldwork. Secondary data includes all other sources of information that are relevant to the research—public documents, press releases, opinions in social networks, etc. Among these, some data sources can be analysed with CAQDAS packages (like interview transcripts and answers to open-ended questions), and others have fewer options for analysis (like audio and video materials and other types of visual materials). However, most types of data sources can be imported and managed in CAQDAS software packages, so the researcher can have all the information at one place.

Qualitative data is by definition unstructured or semi-structured and traditional statistical methods have a limited application in the process of analysing it. Instead, different methods to derive a meaning from the data are used. Some of the most common types are discussed here. Application of each method depends on the purpose of the study.

Discourse analysis

This analysis is focused on written or spoken language, as well as clues from the body language and the social context, in order to unveil the meaning of what is said. It helps understand the conversation in a rich cultural and personal life context and thus—it lets the researcher discover concealed meaning of used language.

Discourse analysis comprises different approaches and procedures that provide the researcher a perspective for interpretation rather than absolute answers. The interviewee can speak about the topic from many different perspectives and not be consistent in his/her answers. In the discourse analytical approach, it is important to understand what these points of view are and what is really meant, i.e., how culture and context affect the expressed opinions.

Narrative-based analysis

Narratives are stories that are used by researchers to understand the personal experiences of the "narrator"—the interviewee. It is feasible if the interview includes questions that presume a story-like answers, like explaining past experiences and biographical stories. This type of storytelling can occur between two people (in-depth interview) or in a group (focus group interview). Usually, the interviewer or the moderator helps participants construct the narrative by making suggestions and asking questions, so that the participant can proceed and finish the story. There are guidelines for analysing the narratives, and one of the following four approaches is used: narrative thematic analysis, structural analysis, dialogic / performance analysis, and visual narrative analysis (Butina, 2015).

Grounded theory

The approach followed in grounded theory is that all concepts and hypothesis in the research should be generated after the data is gathered and analysed, rather that prior to this, with the use of theoretical induction and deduction (Gibson & Brown, 2009, p. 27). It provides methods for both conducting the research and the analysis. It is more an exploratory and not conclusive research approach, although it can make use of both qualitative and quantitative data.

Qualitative content analysis

Just like previous types of analytical approaches, qualitative content analysis is also aimed at bringing to light the meaning of qualitative data. This analysis uses quantitative approach to analyse qualitative data—it classifies words or phrases and counts its frequencies in the analysed material. It can be used to analyse both primary and secondary data, verbal and visual. This analytical method includes:

1) creating coding frame with categories and subcategories of data carrying the same meaning,
2) dividing the analysed material into units of coding,
3) validating the coding frame,
4) classifying the units into created categories and subcategories (Schreier, 2012, p. 6).

1.4.2. Common steps in qualitative data analysis

Regardless of the approach to the analysis, there are some common steps in qualitative data analysis with CAQDAS packages that can be pointed out:

1. Transcribing the interviews.
2. Preparing and exploring the data.
3. Creating categories of data.
4. Reviewing the codes.
5. Interpreting and presenting the results.

The first step of qualitative data analysis is **the transcription of all audio and video materials**. This way, you can easily take notes on the printed transcript and use the digital file for further analysis (if one is needed). Different software packages require specific formatting of the text documents. Before starting the transcription, the researcher must check these requirements and choose an appropriate file format (.txt, .rtf or .doc). During the process itself some rules must also be followed in order to make the text 'readable' for the software, especially if one intends to make the coding process automatic. Creating a transcript of the recorded interview can be facilitated with the use of an appropriate software package that turns speech into text.

Preparing and exploring the data: next, you need to organize all available materials, including transcripts, field notes, and other documents that will be used during the analysis. Most CAQDAS software packages allow you to import various types of data files and to organize them in a research project. However, they differ in the extent to which visual and audio formats can be used in analytical procedures. It should be mentioned that sometimes the time needed to prepare the data is extensive. For example, if you have focus group interviews data, you have to revise a large number of pages in order to unify participants ID names, change font colour, bold or highlight text to make visual inspection of the text easier. In addition, if you use automated analytical procedures sometimes you need to change the structure of the text. You should consider whether you can skip some of the data preparation steps based on your plan for the analysis.

This step also includes a thorough reading of the materials and making any additional notes that will further help with the analysis. These notes can be stored as separate documents or embedded in the transcription file.

Creating categories of data: at this step, you create categories by using keywords or phrases that are linked to the research objectives as categories. There are different approaches to do this: you can use theory to list terms and words that are related to the research topic; you can also use past research to identify relevant keywords; finally, you can combine these approaches and add your own categories based on your experience during the initial stages of the research and the data preparation process. After you are ready with the category list, you have to read the transcript carefully and mark the text that falls in each category. This process is known as "coding".

Reviewing the codes: the main goal of interviews is to obtain a deep insight into the opinions shared on the topic of interest. Deriving meaning from the qualitative data is possible when we identify recurring themes and connect codes that seem to be interrelated.

Interpreting and presenting the results: at this last step, we relate the obtained meanings of the text to the initial research questions. The interpretation is always made in the light of the objectives of the study. A report with the findings is prepared and presented before the intended audience.

The main difference between analysis of data from an in-depth interview and focus group is that the latter includes an interaction between group members. This could be an important aspect of the study and the researcher could be interested in how participants collaborate and create a shared view on some topics.

An important element in the presentation of the research results is also the use of various **graphic forms**. They facilitate a better understanding of the research results—a good picture says more than 100 words. Examples of graphic forms that can be used in the presentation of the results of qualitative research are presented in Figures 1.2 and 1.3.

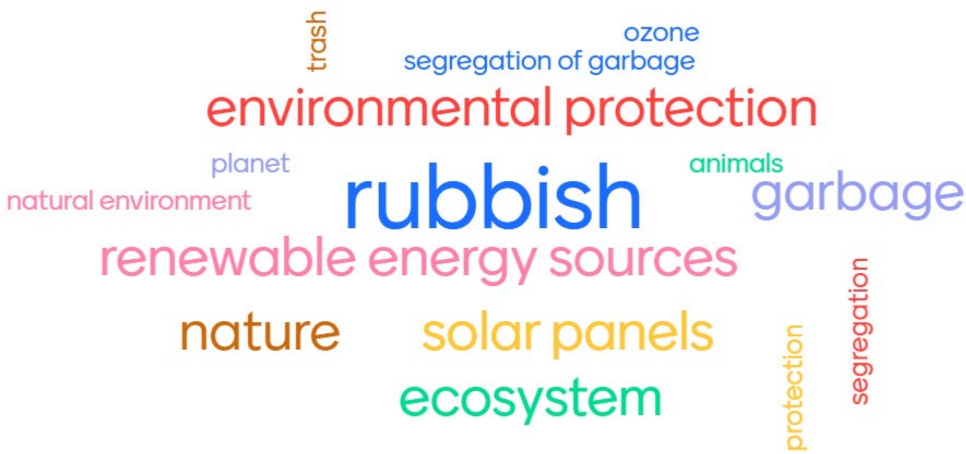


Figure 1.2. Word cloud related to the word "ecology"—results of the spontaneous associations test

Source: Own elaboration.

Figure 1.3. Advantages and disadvantages of renewable energy sources—example of a metaphorical visual display

Source: Own elaboration.

1.4.3. Analysing qualitative data with a CAQDAS software

There are two major groups of CAQDAS software packages that follow qualitative approach to the data analysis (Lewins & Silver, 2007):

– Code-based theory building software: this type of software enables thematic coding of pieces of data. These packages can either facilitate the researcher in reducing the data, allowing him to explore relationships between different themes, or help him to develop more detailed and sophisticated codes.
– Text retrievers and textbase managers: these packages provide the researcher with tools to analyse the text for specific words and to look for synonyms or words with close meaning, while also providing indexing for all words in the text, frequency tables, and key word in context retrieval.

The distinction between these types of software is not absolute: contemporary software packages allow both coding and text retrieving, while also offering complex text-based searching tools.

There are different software packages that can be used to analyse qualitative data, including ATLAS.ti, NVivo, Quirkos, MAXQDA, and many others.

Before we discuss an example on how you can manage and analyse qualitative data with CAQDAS software, consider the following tips on data preparation process:

1. Preparing the transcript
   – Coloured, bolded and underlined text can help highlight important moments of the interview but it cannot be used by the CAQDAS packages. This type

of text formatting can be applied to help the researcher quickly identify different respondents and topics.
- Some software packages can use full stops, question marks and exclamation marks to recognize the end of a sentence. This can make the automatic coding much easier. If the transcription is made with a separate tool that turns audio into text, it is necessary to thoroughly revise the punctuation.
- Sometimes the researcher may want to indicate different structures in the text by using new paragraphs and sections in the text document. CAQDAS packages differ in their ability to handle these text structures so before doing this, the researcher must check specifically for the chosen software. Again, even if this formatting is not useful for the package, it can be used in order to make the visual inspection of the text easier.

2. Identifying respondents, topics and questions.

When we analyse focus group data, we may want the CAQDAS software to identify each speaker in the discussion script. In order to automatically do so, the text needs to have a unique identification name for each participant that should be used every time this participant have said something. This identification can be the actual name of the participant or just a code (like RE12–20) that is used to replace the name for anonymity purposes. Using capital letters can make visual identification of each participant much easier and can improve the interpretation of answers. You should use identifiers for questions asked by the moderator/interviewer as this facilitates the detection of new topics in the text. These identifiers can be as simple as this: Q1-S (question one, interviewer Sonya) or Q1.F2 (question one, focus group two). It is important to make the identifiers somehow different from the short words in the text. For example, if your materials are in English you should avoid using 'or', 'on', 'to', etc. as identifiers for both respondents and questions, especially written with lower letters.

1.4.4. Example of CAQDAS software—Atlas.ti

Atlas.ti[3] is a popular software used for qualitative data analysis and mixed methods research in academic, market and user experience research. This program is a very powerful data analysis tool. It allows not only simple grouping and counting of data, but also the performance of advanced qualitative analyses leading to the creation of hypotheses and generation of theories (Niedbalski, 2014).

We choose this software to illustrate how you can manage your qualitative research project using CAQDAS package. Atlas.ti is easy to use and intuitive. Therefore,

[3] You can download a free trial here: https://atlasti.com/free-trial-version/ A trial license lets you explore these apps with no functional limitations for up to five days within a 90-days period (according to information of March 2021).

only a few selected elements will be shown below to help you get started with this program. We encourage people interested in expanding their competences to use the program independently, supporting themselves with user manual (Friese, 2019), numerous tutorials and other materials available on the Internet.

After launching Atlas.ti, you need to add the documents to be analysed (Figure 1.4). Atlas.ti allows you to analyse different types of data: large bodies of textual data (e.g., transcripts from interviews, but also in the case of content analysis: articles and all texts obtained, e.g., from the Internet), as well as graphical, audio or video data.

There are important options in the menu, including:

- documents,
- quotations,
- codes,
- memos.

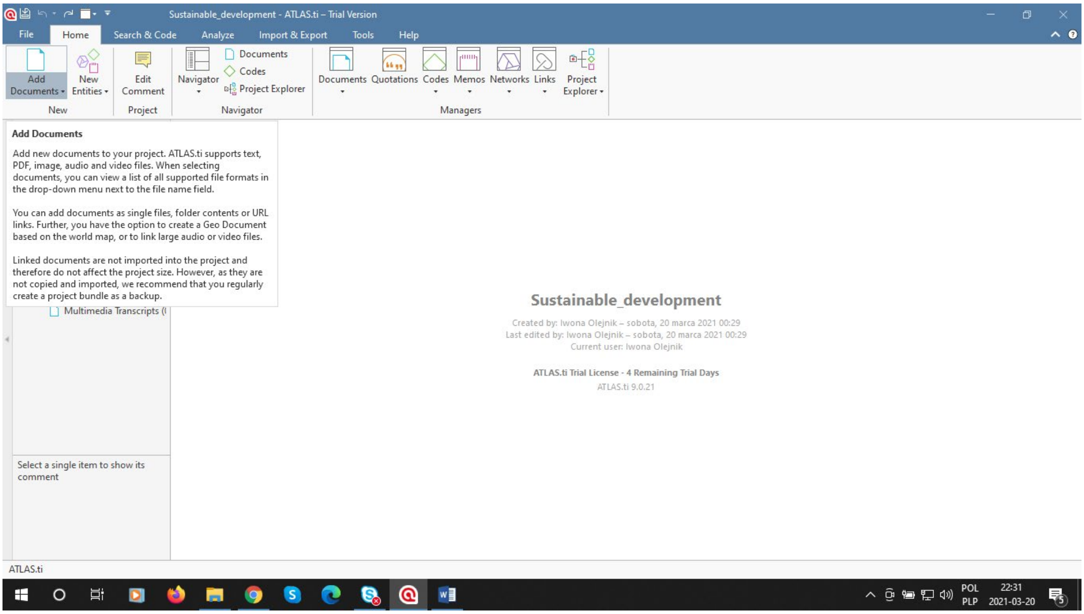


Figure 1.4. Adding documents in Atlas.ti 9—computer program window

Source: Own elaboration.

The analytical process begins with data coding, sorting and assigning them to specific topic category. The next step is to combine thematically coherent codes and search for relationships between them (Niedbalski, 2014). The basic unit of analysis is a piece of data that is separated according to the issue of interest of the researcher—**quotations**. It can be a single phrase (e.g., "experiences with services"), a line, a sentence or a paragraph. They are assigned certain, specific codes. Coding in Atlas.ti is (technically) a selection of a piece of text that is assigned specific labels—**codes** (Figure 1.5).

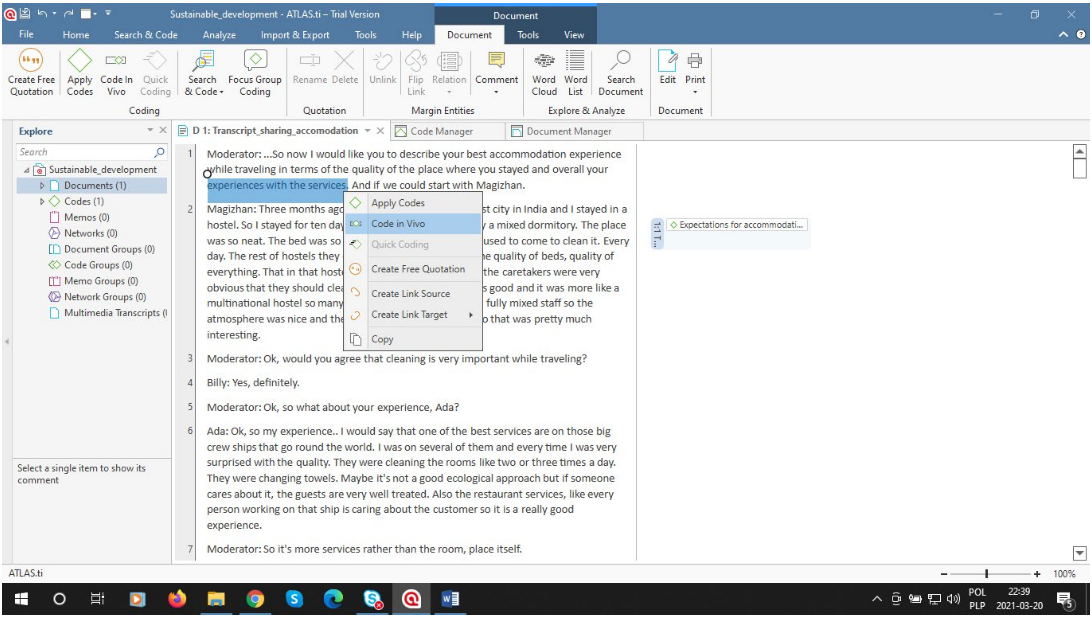


Figure 1.5. Data encoding—code in Vivo in Atlas.ti 9

Source: Own elaboration.

Atlas.ti allows you to encode data in several ways. These are mainly:

- open coding: creating a new one code based on data analysis or creation of one that has not yet been marked with any text (i.e., "new free codes") (*expectations for accommodations*, see Figure 1.5),
- code in-vivo: the code name will be a fragment of the original text (*experiences with the services*, see Figure 1.5),
- coding based on a list of codes already created: reusing the code that has already been created to mark a new piece of data with it (see Figure 1.6).

When creating codes, it is worth remembering that there can be an unlimited number of them, and a single code should be built rather of a small number of words (if it is too long, further analysis is facilitated by shortening its names and moving the content to the comment) (Figure 1.6). The Code Manager provides an overview of all codes and code groups. You can create new "free" codes, rename, delete, duplicate, merge or split codes.

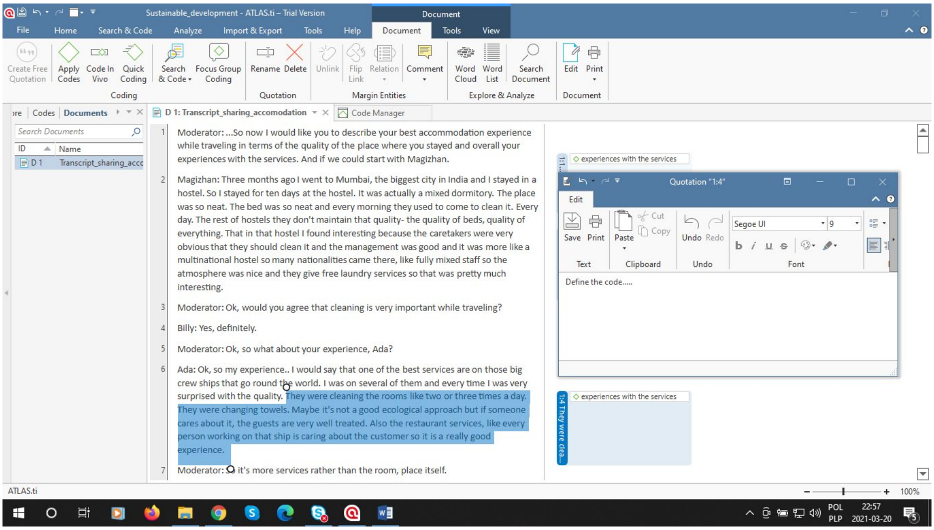


Figure 1.6. Data encoding—quotation and code definition

Source: Own elaboration.

Another basic and very useful function of the presented software is "memos" (Figure 1.7). It is some kind of a notebook in which you can write down all kinds of notes. These may be the researcher's thoughts related to the analysed text, being a "link" between the codes and the final research report (or final article, Ph.D. thesis etc.).

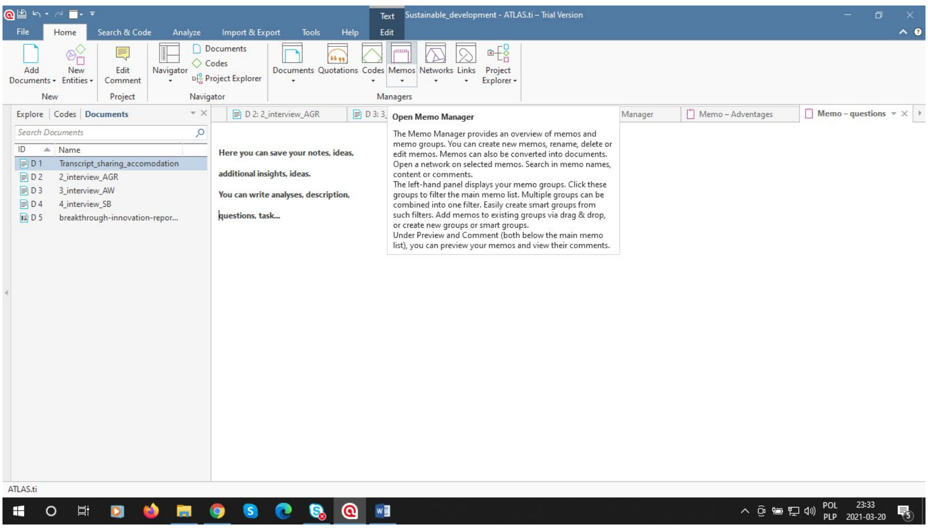


Figure 1.7. Memos in Atlas.ti 9

Source: Own elaboration.

We could also take other advantage of some of the tools. E.g., in the case of content analysis, if we have an article (or any other text) and we have not even read it yet, we just want to have a quick global overview and know what the main concept in this text is, we can create a word list or in a different layout—word cloud (Figure 1.8). This enables, among others, to analyse word frequencies, so we can see which ones are occurring more or less frequently.

For example, the analysis of the second chapter of this e-book (*Questionnaire design*) is presented below. We can notice that the dominant words in the text are *questionnaire* and *questions*, the former appearing 45 times in the text. It is worth adding that in this way we can analyse more than one document at the same time.

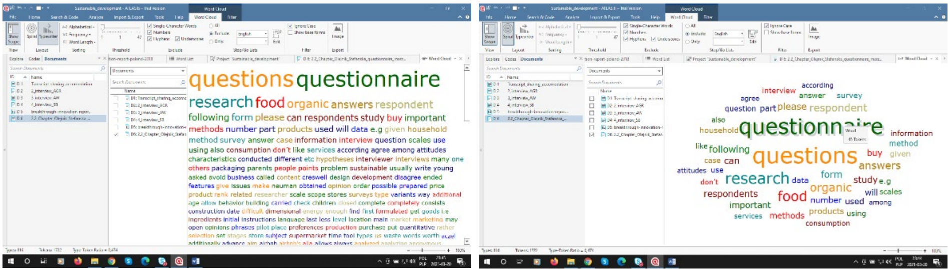


Figure 1.8. Word cloud in Atlas.ti 9—quick global data overview

Source: Own elaboration.

This tool also allows you to quickly find what context is a given word in, in the analysed text. The example below shows where the word "food" appears in the text.

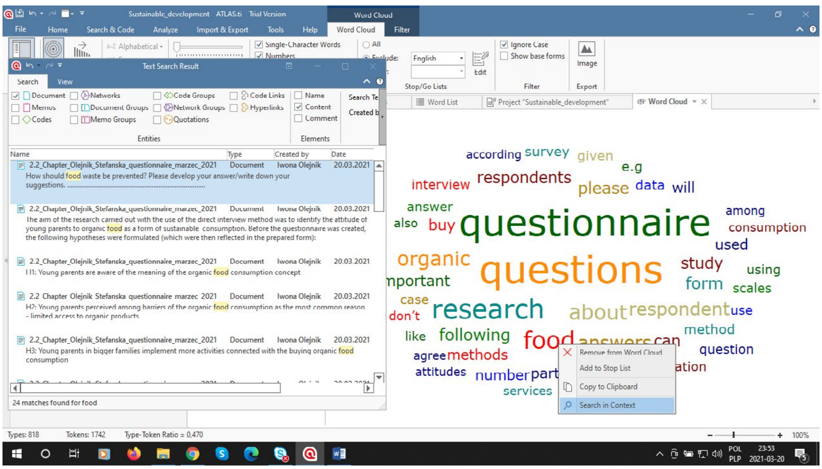


Figure 1.9. Search word in context

Source: Own elaboration.

In conclusion, only a few examples of the usefulness of the Atlas.ti software are presented above. It should be emphasized that this and other software have many additional functions useful in the analysis of qualitative data. However, it is worth emphasizing once again that a researcher conducting qualitative data analysis using CAQDAS software should remember that even the best software cannot "read between the lines". Therefore, such software without the participation of a researcher will not read the proper sense of the statement, e.g., a humorous or ironic tone. In this way, the information obtained can even be misinterpreted. Thus, the software used in qualitative research should be treated as supporting and not replacing data analysis (Kaczmarek et al., 2013).

Questions / tasks

Work individually to answer the following questions:

1. Do the questions used in the focus group interviews meet the recommendations given in this chapter?
2. Is the time provided for each item in the script enough to cover all the questions there?
3. Can you think of other questions that can be added to each category? Please, give an example of at least three other questions.
4. Please use 2–3 chosen projection techniques, which may be applied in this project.

Work in groups to answer the following questions:

5. Was the selected method of qualitative research appropriate to fulfil the goal of this research?
6. Considering the number of people interviewed, would it take more or less time to take individual interviews rather than focus group interviews?
7. Do you think that the main conclusions of these focus groups could be useful for questionnaire development for the qualitative part of this research?
8. What other demographics the researchers could have used when selecting the participants of the focus groups?
9. In the light of sustainability issues, how useful could focus group interviews be for private companies? Please, give an example of a managerial problem related to sustainability that could benefit from conducting a focus group.
10. Which of the following statements about in-depth interviews is correct?
    - The questions used in this type of interview have predefined answers.
    - There is a strict order in which the interviewer asks questions.
    - The interview involves one interviewer and a group of people (interviewers).
    - This method helps marketers reveal the real motives behind consumer attitudes and actions, which often are subconscious.

11. In which of the following situation is it appropriate to conduct an in-depth interview?
    - When the purpose is to gather representative data on the topic of interest.
    - When the researcher needs information that will be directly used to make an important managerial decision.
    - When one wants to generate a new product conception.
    - When there is a need to discuss topics of public interest.
12. All but one of the following statements refer to the focus group interview. Which one?
    - Focus groups can include either a small or large number of people (between 5 and 30).
    - A typical focus group consists of a relatively small number of people.
    - The discussion in a focus group interview is spontaneous and unstructured.
    - In a focus group interview, the discussion is led by a moderator.
13. Which one of the following do you consider the advantage of focus group interviews?
    - They can be used to gather representative data.
    - The results are easy to analyse and interpret.
    - They are useful for gaining insight into consumers' inner world.
    - Moderation of the discussion is easy.
14. What features should a moderator have?
15. What are the pros and cons of implementing qualitative online interviews as compared to face-to-face?
16. Prepare a recruitment questionnaire for research on a chosen topic:
    - How to reduce the "production of garbage" in households? The results are to be used in a promotional campaign—aimed at increasing consumer awareness.
    - Unethical consumer behaviour—identification, perception, possible countermeasures.
    - Perception of the corporate social responsibility.
    - Sharing economics—advantages and disadvantages for consumers (or companies / cities / society).
17. Prepare a fragment of a script for FGI research on a chosen subject (see—task 16)
18. Prepare a practical implementation of a chosen projection technique: animalization, personification, completion techniques, picture drawing, picture sorting, „brand party", shopping list analysis, „family game", collage construction; attention: the topic should be connected with sustainable development. Apply this technique by researching with your colleagues at the lecture / workshop / home. Present theoretically what this technique is about.

References

Barbour, R. (2007). *Doing focus groups*. London: Sage.

Butina, M. (2015). A narrative approach to qualitative inquiry. *Clinical Laboratory Science*, *28*(3), 190-196.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: Sage.

Friese, S. (2019). *Atlas.ti 8 Windows user manual*. Berlin: ATLAS.ti Scientific Software Development GmbH.

Gibson, W. J., & Brown, A. (2009). *Working with qualitative data*. London: Sage.

Kaczmarek, M., Olejnik, I., & Springer, A. (2013). *Badania jakościowe. Metody i zastosowania*. Warszawa: CeDeWu.

Lewins, A., & Silver, C. (2007). *Using software in qualitative research: A step-by-step guide*. London: Sage.

McDaniel, C., & Gates, R. (2018). *Marketing research* (11th ed.). Hoboken: Wiley.

Merton, R. K. (1987). The focused interview and focus groups: Continuities and discontinuities. *Public Opinion Quarterly*, *51*(4), 550-566. https://doi.org/10.1086/269057

Merton, R. E., Fiske, M., & Kendall, P. (1962). *The focused interview*. Illinois: The Free Press.

Niedbalski, J. (2014). Zastosowanie oprogramowania Atlas.ti i NVivo w realizacji badań opartych na metodologii teorii ugruntowanej. *Przegląd Socjologii Jakościowej*, *2*. Retrieved from https://www.ceeol.com/search/article-detail?id=4663

Olejnik, I. (2011). Metoda obserwacji – zastosowania w badaniach marketingowych. In K. Mazurek-Łopacińska & M. Sobocińska (Eds.), *Badania marketingowe – metody, nowe podejścia i konteksty badawcze* (pp. 242-249). Wrocław: Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, *236*.

Olejnik, I., Dębska, W., & Zieliński, K. (2020). Pandemia – szansą na rozwój zogniskowanych wywiadów grupowych on-line. In R. Romanowski (Ed.), *Marketing w czasach pandemii* (pp. 242-249). Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Schreier, M. (2012). *Qualitative content analysis in practice*. London: Sage.

Qualitative and quantitative methods in sustainable development, pp. 41-79
https://doi.org/10.18559/978-83-8211-072-2/02

2.

QUANTITATIVE METHODS

Sylwester Białowąs Blaženka Knežević, Iwona Olejnik, Magdalena Stefańska
Poznań University of Economics and Business

Blaženka Knežević
University of Zagreb

Iwona Olejnik
Poznań University of Economics and Business

Magdalena Stefańska
Poznań University of Economics and Business

Abstract: The main goal of the chapter is to present the basics of survey research that can be used in analyzes of sustainable development.

The first part presents the measurement levels. The basic characteristic of every variable is its level of measurement. It implies the following analysis and available techniques. This part introduces four levels of measurements: nominal, ordinal, interval and ratio, showing their characteristics and examples. Then the focus is on the implications of a given level of measurement on the possibilities of the statistical analysis.

The aim of the second chapter is to explain the foundations of preparing a questionnaire for the research on the issues related to sustainable development. An example of an organic food questionnaire is also provided.

The third part presents considerations necessary for the sampling process. The main goal is to present the basic methods of calculating the minimum sample size, as well as the methods of its selection. This section presents the arguments for conducting the study on a sample rather than on the entire population, and also several formulas enabling the calculation of the minimum sample size. A discussion of the most important methods of selecting respondents to the sample—both random and non-random, can also be found here.

The last two parts of this chapter, describe the ways of presenting the results of quantitative research. They describe first view of the variables including frequency distribution with charts, central tendency measures and cross-tabulation. Finally, the methods of presenting research results obtained on the basis of the Likert scale and other examples of data visualization schemes are presented.

Keywords: data presenting, levels of measurement, questionnaire, sampling.

2.1. Levels of measurement

2.1.1. Introduction

Thinking about measurement we usually connect the idea with physical items. We can easily imagine measuring width, length or weight. We have the proper tools to say that the table is 1,4 meters long and the mobile phone weights 120 grams. It may look different in economics, still we measure physical objects, but lots of the measurements are related to non-physical concepts, like happiness, financial literacy, attitudes toward clean energy or opinions about waste management.

The results of the measurements can be compared more or less precisely. Talking about the country of origin we can just tell if it is same or different. But when we measure salaries we can easily say that three thousand euro is not the same as two thousand euro; the difference is one thousand euro, and the three thousand is 1,5 times more than two thousand euro. So, the comparison is full of new details. How can we compare the measurement results and what statistic techniques can we use considering the level of measurement?

There are four levels of measurement:

1) nominal,
2) ordinal,
3) interval,
4) ratio.

2.1.2. Nominal, ordinal, interval and ratio levels of measurement

The nominal level of measurement is the lowest and simplest one. On this level we can assess the observation in terms of the category. Every observation may be assessed as belonging to one of the categories. It may not be attributed to more than one category.

According to such assessment we can confirm, if the elements belong to the same category, or if they belong to different categories.

Example of the nominal level of measurement:
- *Gender: male, female*
- *Believing in human driven climate change: yes, no*
- *Believing that one of the countries named below will be the first one to reach carbon neutrality (net carbon dioxide emission): Bulgaria, Croatia, Czech Republic, Hungary, Poland, Slovakia, Ukraine*
- *Sorting waste in your household: yes, no*

Usually, we use numbers for coding the categories, and in the case of nominal level for measurement, the numbers have no meaning. Using codes like:
1. *Bulgaria,*
2. *Croatia,*
3. *Czech Republic,*
4. *Hungary,*
5. *Poland,*
6. *Slovakia,*
7. *Ukraine,*
8. *Other,*

or
1. *Ukraine,*
2. *Slovakia,*
3. *Poland,*
4. *Hungary,*
5. *Czech Republic,*
6. *Croatia,*
7. *Bulgaria,*
8. *Other*

gives the same possibilities in interpretation.

What should remain constant is the last position of "other". "Other" is used very often if we cannot name all possible answers.

Data from a nominal scale should not be used for arithmetical calculations, it would be meaningless. It does not make sense to replace Czech Republic (represented by value 3) by combination of added Bulgaria (1) and Croatia (2). And there is no point in presenting the arithmetic mean from the countries from the above example.

If the variable has only two possible categories we call it a dichotomous variable. The distinction is important, as some statistical methods can be used for dichotomous variables and cannot be used for other nominal variables. For example, the question about believing in human driven climate change, with two possible answers (yes, no) is a dichotomous variable.

The ordinal level of measurement offers us more than nominal data. Using the ordinal scale, we not only compare if the elements are the same or different, but we can also order them according to the intensity or time. However, the data tell us nothing about the differences between the values.

Imagine the three households A, B and C. The households compare produced waste. Household B produced the most, whereas household C produced the least. However, we know nothing about the differences. Was household A more similar in waste production to B or C? Maybe it produced almost the same number of waste bins as B with a huge difference to C? This could also be completely different.

Example of ordinal level of measurement:
- *Performance in waste production (first, second, third)*
- *Liking this e-book (dislike very much, rather dislike, rather like, like very much)*
- *Supporting state investments in coal mining (fully against, rather against, rather accepting, fully accepting)*
- *Attitude toward sorting waste: very negative, rather negative, rather positive, very positive*

The numbers are important and have to be assigned to the verbal levels in ascending or descending order.

Interval level of measurement is even more useful. We not only compare if the elements are the same or different, or not only can we order them, but we are able to measure the distances between them. Data are interval, if the differences can be measured in units, and the units have equal intervals between them. The interval variables can have negative values.

Example of interval level of measurement:
- *Bank account balance (e.g., –270 EUR, 50 EUR, 70 000 EUR)*
- *Temperature of the sewage in main sewer collector in Celsius (e.g., 17, 21)*
- *The year you have started sorting waste in your household: (e.g., 2001, 2002, 2003…)*

We can calculate that Samantha's household (1995) started sorting waste six years earlier than Bob's household (2001).

Ratio level of measurement goes a step further. Data collected on this level allow for comparing elements, ordering them and measuring the distances between them. And additionally, the ratio level allows for calculating meaningful ratios of values along the scale. This is possible only if the scale has true and meaningful zero point. The zero value indicates in this case, that there is none of that variable. Of course, the ratio level variables cannot have negative values.

Example of ratio level of measurement:
- *Number of bumblebees visiting a given flower in one hour (e.g., 0, 5, 35)*
- *Weight of the waste produced in one month by the household (15 kg, 45 kg, 70 kg)*
- *Tons of unsorted waste reported by municipality of your city in 2020: (e.g., 3, 250, 781)*

2.1.3. Measurement levels and data analysis methods

The level of measurement has serious impact on how we can analyze them. You will know more about it in the next part of the book. Table 2.1 shows the central tendency measures and basic calculations available for given levels.

Table 2.1. Levels of measurement and basic calculations

	Nominal	Ordinal	Interval	Ratio
Mode, frequency distribution	+	+	+	+
Median, presenting order		+	+	+
Calculating mean, adding, subtracting			+	+
Calculating ratio				+

Source: Own elaboration.

Levels of measurement imply the permissibility of statistical techniques. When forming hypotheses and designing the questionnaire for the primary research we always have to plan the method of future analysis and check its feasibility.

Table 2.2. Permissible descriptive and inferential statistics for certain levels of measurements

	Descriptive	Inferential
Nominal	percentage, mode	chi-square, binominal test
Ordinal	percentile (e.g., median)	rank-order correlation, Friedman ANOVA
Interval	range, mean, standard deviation	product-moment correlations, *t*-test, ANOVA, regression, factor analysis
Ratio	geometric mean, harmonic mean	coefficient of variation

Source: Own elaboration based on: (Malhotra & Birks, 2003).

2.2. Questionnaire design

2.2.1. Stages in creating a questionnaire

The questionnaire is the basic research tool used in the primary quantitative research, both in the interview methods (in which the questions are asked, and answers are written by the interviewer) and in the questionnaire (in which the respondent reads the questions and writes down the answers himself). It is an indispensable research tool in many types of research: door-to-door interviews, executive interviews, mall-intercept interviews, telephone interviews, self-administered questionnaires, and mail surveys

(McDaniels & Gates, 2018). The questionnaire is a set of written questions formulated to elicit the desired answers related to the examined problem (Neuman, 2014).

A properly constructed questionnaire should meet many rules concerning both the content of the questions, the way they are formulated, the scaling of the answers, the order of the questions, and the graphic design (Churchill & Iacobucci, 2018). The questionnaire should encourage the respondent to participate in the survey and make the survey interesting for them. At the same time, it is designed to facilitate the work of the researcher (especially the interviewer) through instructions on how to ask questions and write down the answers.

The following stages of building the questionnaire can be distinguished:
1) specification of the purpose and subject of the questionnaire,
2) formulation of an initial list of questions,
3) initial scaling of the response,
4) initial check of the questionnaire,
5) construction of the sample questionnaire,
6) pilotage and possible modifications of the questionnaire,
7) preparation of the final version of the questionnaire.

When starting to build the questionnaire, one should remember about the assumptions of the research project, in particular, about defining who the respondents will be (their level of education, age, etc.) and what their knowledge is in relation to the subject covered by the research (general or specialist?)—it determines form of asked questions and the use of specific phrases.

The construction of the questionnaire, including the number of questions and their form, depends on the research method. The form will look slightly different in the interview method than in the survey method. These methods differ from each other, inter alia, in the form of contact with the respondent (direct, indirect) and the number of questions asked and answers received. An important determinant of the selection of the type and construction of the questionnaire is also the data that is to be collected—their scope and details, as well as the costs to be incurred, time and planned methods of analysing the obtained data.

Regardless of the intended research method and the type of questionnaire, certain elements are present in each of them. The questionnaire therefore consists of three basic parts:
1) introductory part—the header,
2) the main part—containing questions aimed at solving a research problem,
3) respondent's particulars—personal information (data about the respondent).

The introductory part—the header of the questionnaire—should include such elements as: the name and contact details of the organizer of the study, the title of the questionnaire, a short description of the purpose, nature and scope of the study, and possible indication of benefits for the respondent resulting from the participation in the study, ensuring that the survey is anonymous (optional), the date of return (in the case of forms returned by post), the name and surname of the interviewer and, possibly, information on how to complete the questionnaire.

2.2.2. Types of questions in the questionnaire

The main part of the questionnaire consists of the subject-related questions concerning the research problem. There should be enough questions to reflect the entire substantive scope of the planned research, and thus to enable the verification of the research hypotheses set out in the project, including the following categories of questions (Neuman, 2014, p. 317):

- behaviour (*How frequently do you ...., When did you last ....*),
- attitudes, beliefs, opinion (*What is the biggest problem facing ...., What do you think about ....*),
- expectations (*Do you plan to buy .... in the next 3 months*),
- self-classification (*Do you consider yourself to be ....*),
- knowledge (*About what percentage of .... It is legal to own ....*).

The questions in the questionnaire can be open or closed. Open-ended questions do not have any given options, i.e., the respondent formulates the answers in a completely arbitrary way. They are used when it is difficult to predict the answers of the respondents, when the number of these answers may be too high, or when they concern difficult issues, e.g.:

> How should food waste be prevented? Please develop your answer / write down your suggestions ....
> ....
> Please, finish the statement: „In second hands the most irritating thing is ....,
> because ...."

Closed-ended questions (with a scale) provide for the selection of answers from a prepared set of options predicted and determined in advance by researchers, i.e., from the so-called cafeteria. These questions are also called scaled questions. It should be remembered that the consequence of the selection of a specific scale are—after the research is conducted—the methods of processing and analysing the data obtained using them (Aczel, 2009). As a result, the so-called one-dimensional and multi-dimensional scales are applied (Neuman, 2014).

An example of a one-dimensional closed question is, e.g.:

> Do you think that biodegradable waste should be sorted separately? *Please underline 1 answer*: a) definitely yes, b) rather yes, c) rather not, d) definitely not.

On the other hand, when constructing complex questions, the following scales can be used:

- ranks: they give the opportunity to rank the most important categories for the respondent (things, attitudes, etc.) according to a given criterion,
- rank scale of summed up grades: it gives the possibility to rank the most important things, attitudes, etc. for the respondent by dividing a certain pool of points (usually 10 or 100),
- comparative scale of summarized features: the respondent assigns the listed features or objects, according to their preferences, numerical ratings so that they give the required sum (usually 10 or 100),
- semantic differential (semantic): it allows to obtain the respondents' opinions about a given object in the range between two opposite attitudes,
- positional: it consists of many individual words or phrases in any order, which are assessed by the respondents according to their preferences using the same ordinal scales,
- Stapel's: it was designed for situations in which some features of the tested objects do not have obvious opposites,
- Likert: it allows you to specify the degree of acceptance of a given statement or statements (from 1—strongly disagree, to 5 (or 7)—strongly agree).

As mentioned before, a properly constructed questionnaire should meet many rules regarding the structure of questions. First of all, "the language of the questionnaire should approximate the level of understanding of the respondents (Bougie & Sekaran, 2020, p. 147). So, they should:

- be short and simple (with some exceptions),
- avoid jargon, slang and abbreviations,
- be unambiguous in terms of the vocabulary used—understandable in the same way by each potential respondent,
- not suggest an answer; avoid emotional language and prestige bias,
- explain a priori specialist words and phrases,
- have comprehensive scales—containing all potential answer options,
- be formulated using homogeneous grammatical forms.

Questions should also be placed on the form in the correct order, logical from the respondent's point of view. In the case of research on consumer behaviour and attitudes in relation to various products and services:

- first, questions about the awareness of their existence are placed,
- then questions about behaviours, e.g., in terms of the volume of purchases of a given product and competing products, frequency of purchase, ownership, decision-making process,
- finally, questions about opinions and motives, as well as anticipating future purchases.

In addition, in the case of a questionnaire dealing with various issues, it is worth dividing the questions into sections—according to the thematic threads discussed, using additional spaces and subheadings. Different fonts for the content of the questions and for instructions for the interviewer or respondent are also worth using.

In the last part of the questionnaire—respondent's particulars (respondents' personal questions) are included. There are subjective questions characterizing the research unit. The lack of these part of questionnaire is a fatal error of the study, as it makes it impossible to assess its representativeness and to conduct segmentation. In the case of surveys conducted among individuals and households, this part includes questions about gender, age, occupation, number of people in the household, number of children, place of residence, income and others, while in the case of surveys conducted among enterprises these are questions about the industry, form of ownership, number of employees, location, period of the company's existence, scope of activity, position of the person giving the answer and others.

At the end of the questionnaire, you can also write acknowledgements to the respondent for participating in the survey, as well as any organizational information (by when and where to send the form). Additionally, if the research is performed using the interview method, the last element is the part for the interviewer, in which they write down: the date, time of the beginning and end of the interview, information about the interview conditions (respondent's reactions, presence of third parties, etc.), information about the questionnaire (e.g., which questions were difficult for the respondent, which were sensitive, etc.) and his / her signature.

When the questionnaire seems ready to start your research, it is worth re-evaluating it and answering the questions in Table 2.3.

Table 2.3. Questionnaire assessment

– What do we want to get with the question? – Will the respondent know the answer to a given question? What answers will we get? What is one solution to the research problem—what? – Are options of answers adequate to the question posed? – Is it possible to make the choice of the question easier (without losing information)? – Why is this question situated here? – What is the expected method of analysis? – What is the reliability of the data obtained? – Have we not asked about it before? – Are there any errors in the language notation? – Should not there be a filtering question in advance?

Own elaboration.

Before the actual examination, a so-called pilot study is conducted, i.e., a preliminary test that verifies the correctness of the prepared form. Usually, it is performed on a small pilot sample, e.g., 5–15 people, but depending on the complexity of the study, it may cover even a larger number of respondents. Its objectives are to determine the respondents' reactions to the objectives of the survey, to find out whether the questions included in the questionnaire pose difficulties for the respondents, and then to make the necessary changes in the questionnaire as well as write instructions for respondents and / or interviewers.

2.2.3. Organic food as a form of sustainable consumption: case study

The aim of the research conducted with the use of the direct interview method was to identify the attitude of young parents to organic food as a form of sustainable consumption. Before the questionnaire was created, the following hypotheses[1] had been formulated (they were then reflected in the prepared form):

H1: Young parents are aware of the meaning of the organic food consumption concept

H2: Young parents perceived limited access to organic products as the most common reason among barriers of the organic food consumption

H3: Young parents in bigger families implement more activities connected with the buying organic food.

The researcher started to design the questionnaire and prepared its draft (check below).

The data of institution that conducted the research

How and why is organic food important?

We would like to invite you—young parents—to share your opinions about the attitude to organic food consumption. Your answers are completely anonymous and will only be analysed in the form of aggregate results and not be used for marketing or any other purposes non-related to this survey. The survey takes around 15 minutes to complete.

1. Who buys organic food products in your household?
   a. I buy
   b. My partner
   c. Both
   d. Other person, *who* ……………

[1] Hypotheses are „predictions the researcher makes about the expected outcomes of relationships among variables. They are numeric estimates of population values based on data collected from samples" (Creswell & Creswell, 2018, p. 136).

2. What organic food products do you usually buy: ....
....
....

3. From what sources do you obtain information on organic food products available in the market? *Please sort by importance and enter the numbers in the boxes, from left to right*

a. My partner
b. Friends
c. Daily press
d. Internet
e. Family
f. Shop leaflets
g. Others

1… 2… 3… 4… 5… 6… 7…

4. What is the most important for you when you buy food *(please select the 3 most important factors from the following, and rank them by entering the response codes into the boxes, treating the first box from the left as the most important)*:

a. Price
b. Type of packaging (recycled or not)
c. Production place
d. Ingredients
e. Promotion
f. Size of packaging
g. Brand
h. Price reductions for a given product
i. Type of packaging (can be recycled or not)
j. Expiration date
k. Recommendation of the staff
l. Others.................

1........ 2........ 3........

5. How important are the following criteria for you when choosing food products? *Please divide the 100 points between these factors in such a way as to give the most points to the one you consider most important, and correspondingly less points to the less important ones. If you deem any factor to be invalid, please do not score it.*

Criteria	0–100
store location	
package size	
price	
manufacturer	
method of production	
country of origin	
ingredients used in production	
caloric content	
nutrient content	
others (please write your answer)	
Total:	100

6. Please divide 100 points between each of the following four pairs of different types of stores according to your preferences
supermarket (........) + convenience store (........) = 100
convenience store (........) + local market (........) = 100
local market (........) + discount stores (........) = 100
discount stores (........) + supermarket (........) = 100
7. Please mark your grades that best reflect your opinion on the mentioned characteristics of eco-chocolate. *Please insert an X*

tasty						distasteful
cheap						expensive
smells good						smells bad
nicely packaged						poorly packaged
very healthy						less healthy

8. How do you rate the features of our supermarket as listed below? *Please insert an X*

	I like it very much	**I rather like it**	**I don't really like it**	**I don't like it at all**
Parking places				
Location				
Appearance of the building				
Interior decor				
Wide of assortment				
Facilities for families with children				
Services provided by employees				
Atmosphere				
Level of prices				
Attractiveness of promotions				

9. A number of different views on different purchasing issues are presented below. For each of these opinions, please tell me to what extent you personally agree with them. *Using the scales from 1 to 5, where 1 means that you completely disagree with the opinion, 2: rather disagree, 3: neither agree nor disagree, 4: rather agree, and 5: that you fully agree, please respond to the following statements:*

	1–5
I always try hard to reduce misuse of goods and services (e.g., I switch off light and fan when I am not in the room)	
I sort food packaging to recycle it	
I avoid being extravagant in my purchase	
I avoid overusing / consumption of goods and services (e.g., take print only when needed)	
I choose to buy product(s) with biodegradable container or packaging	
I don't like to waste food or beverages	
I recycle my old stuffs in every possible way (e.g., distribute old clothes among needy people)	
I always search for products with symbols of organic food	
Being a parent made me more sensitive about food we consume	
I don't buy food from unknown sources	
I search for new organic products	
We buy organic food branded only by producers, not stores	
I believe in organic food	
I don't buy more than we need in our household	
I don't buy ready to eat only organic food	
I usually buy not processed organic food	
I put a lot of energy to find organic food I need	
I put a lot of energy to learn more about product ingredients	

10. The monthly percentage of expenditure on organic food in your household is around: ....%

And finally, a few words about your household—characteristics

1. Gender:
   a. female
   b. male.
2. Age ....
3. Job/occupation ....
4. Which of the following sentences describe the best the financial situation of your household. Our household income:
   a. allows us to purchase luxury goods,
   b. allows us to set a small financial surplus,
   c. allows us to satisfy everyday needs,
   d. is not enough, we spend savings,
   e. is not enough, we are indebted.
5. Are there children in your household? How many? ....
6. Place of living: ....

Thank you for your answers!

2.3. Population and sample

2.3.1. Difference between population and sample

When a researcher wants to explain some of the characteristics, trends or interdependencies in a certain group of people, animals, objects, events, countries or phenomena, the group is called a population. Thus, a **population** is the entire group about which conclusions from the research are drawn (Barrow, 2017).

When we collect data, we can decide to collect it from an entire population or from a sample. Usually, population size and geographical spread are problems that can significantly lower cost-efficiency of the research and that can increase complexity of the research management. If a population is large or difficult to reach geographical area, a researcher can use a sample to estimate or test hypotheses about population data. A **sample** is a smaller, specific group extracted from population that will be used to represent population in a research (McDaniel & Gates, 2018). Using samples allows researchers to conduct their studies more easily, and in a faster and more cost-efficient way.

Regarding collected data, we distinguish a **parameter** as a measure that describes the entire population and a **statistic** that describes the sample. Estimation or hypothesis testing is used to describe how a statistic differs from the population parameter. The difference between population parameter and a sample statistic is called a **sampling error**. The larger the sample, the lower the sampling error (Dean & Illowsky, 2013).

2.3.2. Determining sample size

Sampling is the process of obtaining information from a subset of a larger group. If we select an appropriate number of people for the research conducted on a sample, we can take results and project them to the larger group (Levy & Lemeshow, 2008). The sample should be a true miniature of the population, which interests us considering the research problem (Acharya, Prakash, Saxena, & Nigam, 2013).

The sample size is always smaller than the size of a population. There are some recommendations on how to determine how large a sample should be in order to ensure accuracy of research conclusions on targeted population (Kasiulevičius, Šapoka, & Filipavičiūtė, 2006).

Sample size depends on the following:

1) population size (N),
2) prevalence (p),
3) margin of error (e),
4) sampling confidence level or its z-score (z).

Population size (N) is the total number of items in targeted (researched) population.

Prevalence (p) is a proportion of a population that has a specific characteristic in a given time period. Prevalence can be used from previous survey results or observed by running a small pilot survey. If it is unknown, 0.5 can be used in calculation and it will give the largest possible sample size.

Margin of error (e) is a percentage that describes how close we can expect a survey result to be relative to the real population value.

Sampling confidence level shows reliability of the research; it is expressed as a percentage, which shows the level of certainty regarding how accurately a sample reflects the population within a chosen confidence interval.

In practice, the sample size may also depend on factors such as, e.g., the spatial scope and subject matter of the research, research method and technique, as well as the budget available for the research.

For the purpose of sample size calculation, sampling confidence level is converted into own z-score that is used in the formula by applying the following table of values:

Table 2.4. Confidence level and z-score

Confidence level (%)	z-score (±)
70.0	1.04
75.0	1.15
80.0	1.28
85.0	1.44
90.0	1.65
95.0	1.96
99.0	2.58
99.9	3.29

Source: Own elaboration.

One of recommended formulas (1) to calculate the size of the sample is:

Formula (1)

$$\text{Sample size, n} = \text{N} * \frac{\frac{Z^2 * p * (1-p)}{e^2}}{[N - 1 + \frac{Z^2 * p * (1-p)}{e^2}]}$$

The larger the population, the lower the targeted margin of error, and the higher the desired sampling confidence level, the larger the size of the sample. In addition, the accuracy of the research results will increase if we use a larger sample.

When the feature differentiating the respondents' answers (and then taken into account when assessing the representativeness of the sample) is a quantitative one (e.g., household income, expenses for a product), and the population on which we conduct the survey is infinite, the sampling will be independent (with return) and when the confidence interval is assumed based on the sample mean and the sample standard deviation (δ), the formula (2) can be used.

Formula (2)

$$n = \frac{Z_\alpha^2 \delta^2}{e^2}$$

But if the general population is formed by a specific number of N elements, and the distribution of the studied trait x in the population is close to normal, we can apply the following formula:

Formula (3)

$$n = \frac{\delta^2}{\frac{e^2}{Z_\alpha^2} + \frac{\delta^2}{N}}$$

Examples:

1. We want to research how many Internet users in Croatia search for cosmetic products online and we want to analyze what kind of information is the most important.
2. We want to research how retail companies in Croatia change business practices in time of pandemics.
3. We want to estimate the level of food waste and analyze sustainable consumer behavior patterns in rural parts of Poland.
4. We want to analyze information structure of advertisements published at our portal.

Table 2.5. Types and examples of probability sampling methods

Number of research (given examples)	Population description	Population size	Sample determinants (desired values)	Calculated sample size (according to Formula 1)
1	Internet users in Croatia	3 787 838 (Source: Eurostat data for 2019)	margin of error 5% confidence level 95%	385
2	number of active companies in retail industry in Croatia	49 120 (Source: Croatian Statistical Bureau)	margin of error 3% confidence level 95%	1 045
3	number of people living in rural areas in Poland	15 064 972 (Source: https://www.worldometers.info/demographics/poland-demographics/)	margin of error 3% confidence level 90%	757
4	number of advertisements in the last month	1000 (Source: own database of a company)	margin of error 5% confidence level 90%	177

Practical note: There are numerous sample size calculators available online for free (use)! Google and try them!

Source: Own elaboration.

2.3.3. Sampling method

There are two basic methods to select a sample from a population:
a) probability sampling,
b) non-probability sampling.

In **probability sampling** each member of population has a chance to be randomly selected as a member of the sample. This sampling method reduces the risk of sampling bias and gives a better ground for generalization of findings to the whole population. It is widely used in quantitative research. Probability sampling methods are simple random sampling, systematic sampling stratified sampling and cluster sampling. The types and examples of probability sampling are described in Table 2.6.

Table 2.6. Types and examples of probability sampling methods

Method	Characteristic	Examples
Simple random sampling	Each member of the population has an equal chance to become a member of the sample. Sampling takes into account the entire population.	*We want to research opinions of employees on working conditions in company A. The company has 500 employees, calculated sample size is 100. We use the database of employees and by random number generator we generate 100 random numbers to select employees with such ID from the database.*

Method	Characteristic	Examples
		We want to ask questions concerning opinions about local election preferences in some city, calculated sample size is 1,000. We use the register of inhabitants and randomly choose 1,000 people as targeted respondents of the survey.
Systematic sampling	Every member of the population is listed with a number, and we use some interval to select members of the sample.	*We want to research the opinions of employees on working conditions in company A. The company has 500 employees, calculated sample size is 100. We assign numbers from 1 to 500 to each employee and then we select every fifth employee as a sample member. The persons with numbers—5, 10, 15, 20 etc. are sample members.*
Stratified sampling	The population is divided into sub-population (strata) according to some characteristics and then based on proportion of the characteristics in population, we select a similar number of persons with particular characteristics for our sample.	*We want to observe consumer behavior in rural part of Poland. We check the proportion of females, the proportion of young / old, the proportions according to average income in the sample. Then we divide the population in sub-groups according to gender, age and income. Then we apply random sampling method or systematic sampling method to select a certain number of people from each sub-group we have created. The aim is to ensure same sub-group proportions in the sample. If there is 10% of young females with high incomes in population, then we want to have 10% of them in our sample.*
Cluster sampling	The population is divided into sub-groups and each subgroup has similar characteristic as the entire population. Then we randomly choose a subgroup or subgroups to represent the population.	*A retail company has 90 stores operating in various cities across country. In each store, there is a similar organizational structure and there is the same number of middle managers with similar structure regarding education and experience. So, if we want to perform a research on management attitudes towards business ethics in this company, we can randomly choose 5 stores and test managers for opinions.*

Source: Own elaboration.

Non-probability sampling means that members of a sample are selected by using some specific criteria. In this way, the sample is more convenient and accessibility to sample members is facilitated. Thus, non-probability sampling provides a cheaper way to conduct research. Anyhow, due to non-random selection of sample members, validity of generalization of results to an entire population is questionable. Non-probability sampling methods include convenience sampling, voluntary response sampling, purposive sampling and snowball sampling (Etikan, Musa, & Alkassim, 2016). Such sampling methods are often used for qualitative research when the purpose of research is to understand a small, specific or under-researched population. The types and examples of non-probability sampling are described in Table 2.7.

Table 2.7. Types and examples of non-sampling methods

Method	Characteristic	Examples
Convenience sampling	The sample includes persons who are the most accessible to researcher.	*Hotel management wants to examine loyalty of guests, so they ask their guests to fill in the questionnaire on check in or on check out.* *We want to examine consumer ethics of young adults. Therefore, the researcher asks students at the University to participate in the sample.*
Voluntary response sampling	It is based on easy access to respondents, but the researcher is not approaching respondents directly. The researcher places a call for the research and people volunteer to participate in the sample.	*We want to examine the attitudes of young people towards the quality of nightlife in the city. We place a call for research participation at local newspapers, on official web page of the city, and on social networks. Then young people voluntarily answer to the call. With such calls, we have to be aware that voluntary participation can mean that a certain person has a stronger opinion (positive or negative) towards researched topic, thus research can be highly biased.*
Purposive sampling	The members of the sample are chosen based on their characteristic or expertize as the most useful for achieving the research purpose and goal.	*We want to examine the expectation of digitalization in 10 largest retail companies in our country. We make contacts and organize interviews with 20 top managers in those companies.* *We want to research how mothers with small children are satisfied with HR policies in a firm. The sample will include only mothers that have kids under 5 years.*
Snowball sampling	The population is hard to access so we use participants to share a call for participation to others. We utilize a social network effect to reach the targeted sample size.	*We want to research the way of conducting business in social supermarkets as a new phenomenon in food supply chain. As there is no valid register of such organizations, we make the first contact with several managers of social supermarkets in our area and then ask them to share their contact information on other managers with us. We repeat this action with other social supermarkets' managers until we reach a certain number of members in our sample.*

Source: Own elaboration.

When designing research, we have to try to avoid sample biases. A **sample bias** is a situation in which some members of population have a greater chance to become sample members than others. In such a situation, the possibility of generalization of results to targeted population is limited. Findings from samples that are biased can be generalized only to the part of population with the same characteristics as the sample.

A sample bias can occur both in probability and in non-probability sampling methods. However, in probability sampling, bias, by definition, is usually lower than in non-probability sampling.

There are some suggestions on how to reduce sampling bias:
- clearly identify your survey goals,
- clearly define characteristics of your targeted population,
- strictly define the requirements for your target sample members, do it in accordance with the characteristics of the population,
- ensure potential respondents an equal chance of taking part in your survey,
- regularly check if the profile of respondents fits your survey's goals,
- include more respondents (enlarge sample size),
- make online questionnaire as short as possible and accessible form different devices,
- send several follow ups to potential sample members to motivate them to respond.

Stratified sampling: example

We intend to conduct research on food waste by young people in six selected countries in Central and Eastern Europe: Czech Republic, Hungary, Poland, Slovak Republic, Bulgaria and Croatia. It was hypothesized that the respondents' answers would depend on their age and gender. Using the Formula (1), having the data presented in the table 4 and assuming that p is 0.5, the margin of error is 2% and the confidence level is 95%, the minimum sample size was calculated at the level of 2,400 respondents. The question is, how many people from each country (by age and gender) should be included in the research sample?

To answer this question, data on the population of interest to us ($N = 14{,}685{,}442$ people) was collected and differentiated by age and gender (Table 2.8).

Table 2.8. Population of young people in selected countries in 2018

	20–24		25–29		30–34	
	women	**men**	**women**	**men**	**women**	**men**
Czech Republic	249,465	261,933	328,088	344,934	350,984	373,507
Hungary	276,356	293,067	302,709	323,479	296,272	312,200
Poland	1,058,343	1,101,924	1,296,586	1,344,572	1,518,295	1,561,848
Slovak Republic	151,195	158,815	188,303	196,327	206,880	216,384
Bulgaria	153,262	163,210	209,886	222,545	232,002	247,635
Croatia	117,329	123,054	117,514	122,789	130,048	133,702

Source: Demography and Population.

The percentage structure of the population was then calculated as presented in Table 2.9.

Table 2.9. Percentage structure of young people in selected countries in 2018 (in %)

	20–24		25–29		30–34	
	women	men	women	men	women	men
Czech Republic	1.70	1.78	2.23	2.35	2.39	2.54
Hungary	1.88	2.00	2.06	2.20	2.02	2.13
Poland	7.21	7.50	8.83	9.16	10.34	10.64
Slovak Republic	1.03	1.08	1.28	1.34	1.41	1.47
Bulgaria	1.04	1.11	1.43	1.52	1.58	1.69
Croatia	0.80	0.84	0.80	0.84	0.89	0.91

Source: Own calculation.

In the next step, considering the determined structure of the population and the minimum size of the sample, the number of women and men to be tested in each age group were determined. The results are included in Table 2.10.

Table 2.10. Number of respondents needed to be tested in each subgroup

	20–24		25–29		30–34	
	women	men	women	men	women	men
Czech Republic	41	43	54	56	57	61
Hungary	45	48	49	53	48	51
Poland	173	180	212	220	248	255
Slovak Republic	25	26	31	32	34	35
Bulgaria	25	27	34	36	38	40
Croatia	19	20	19	20	21	22

Source: Own calculation.

Then we apply random sampling method or systematic sampling method to select a certain number of people from each subgroup we created.

2.4. Variables—first view

2.4.1. Introduction

Researchers usually start their analysis with investigating single variables. Such an initial analysis may answer many basic questions, e.g., how many observations fall in the categories, what is the structure of values, what are the typical values or ranges, or what is the dispersity of data.

The first view of the variable includes checking its frequency table. This works well if the variable does not have a large number of values. From the frequency table we can easily read the minimum and maximum value, missing data, the pattern of frequencies and percentage.

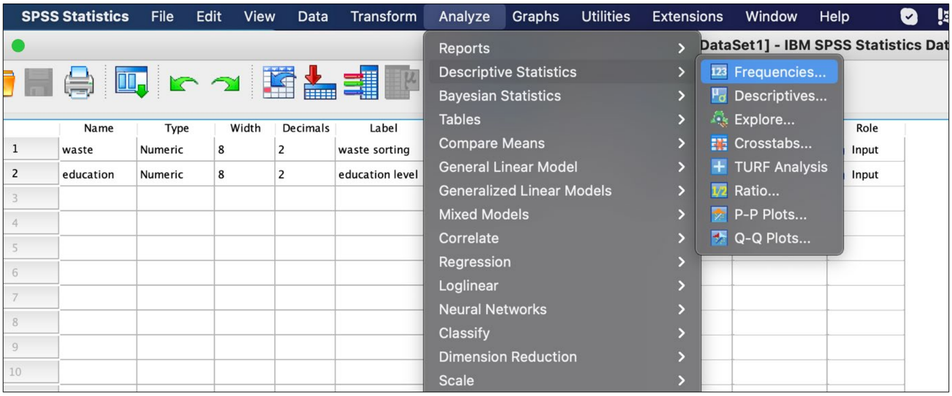


Figure 2.1. Screenshot of obtaining frequency table

Source: Own elaboration.

Frequencies

Statistics

waste sorting

N	Valid	160
	Missing	0

waste sorting

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00 no, everything goes to general waste	49	30,6	30,6	30,6
	2.00 we sort waste into not more than two categories	78	48,8	48,8	79,4
	3.00 we sort waste into more than two categories	33	20,6	20,6	100,0
	Total	160	100,0	100,0	

Figure 2.2. Screenshot of frequency table (waste sorting in households)

Source: Own elaboration.

Within the same frequencies command, we can obtain descriptive statistics of the variable. It just requires clicking the button Statistics and choosing the statistics appropriate for the measurement level of the investigated variable.

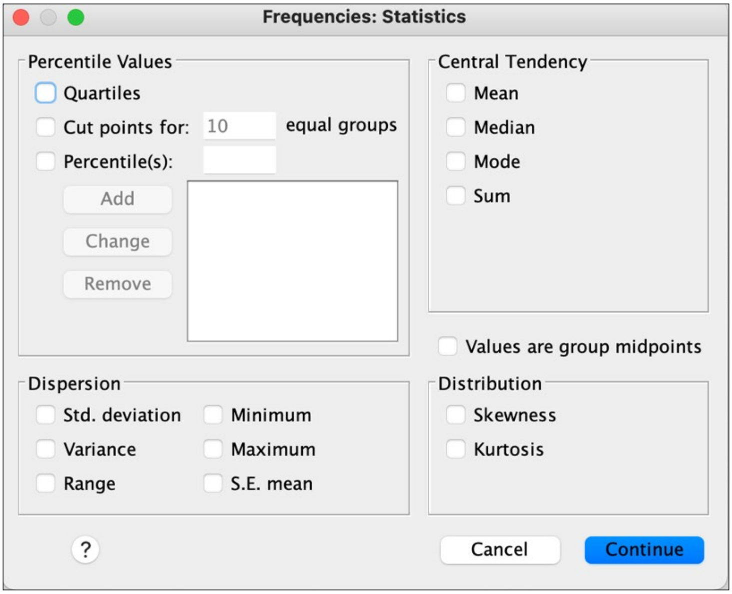


Figure 2.3. Screenshot of the dialog box of basic statistics within the Frequency command

Source: Own elaboration.

In Figure 2.4. and Figure 2.5. you can see the table including statistics for the ordinal and ratio level of measurement.

Statistics

Attitudes toward waste sorting

N	Valid	200
	Missing	0
Median		3,00
Mode		2
Percentiles	25	2,00
	50	3,00
	75	3,00

Figure 2.4. Descriptive statistics for ordinal variable (Frequency command)

Source: Own elaboration.

Statistics

sustainability

N	Valid	200
	Missing	0
Mean		6,33
Std. Error of Mean		,198
Median		6,00
Mode		6[a]
Std. Deviation		2,800
Variance		7,840
Skewness		,010
Std. Error of Skewness		,172
Kurtosis		-,162
Std. Error of Kurtosis		,342
Range		14
Minimum		0
Maximum		14
Sum		1266
Percentiles	25	4,25
	50	6,00
	75	8,00

a. Multiple modes exist. The smallest value is shown

Figure 2.5. Descriptive statistics for ratio variable (Frequency command)

Source: Own elaboration.

The frequency tables can be visualized with the chart. Same frequency procedure allows for producing charts (clicking the button charts). We can choose between:

– a pie chart,
– a bar chart,
– a histogram.

Pie charts are designed for nominal data. They can show the percentage of categories, but they are difficult to compare. They are not clear with the bigger number of categories. And even with the smaller number of categories we can be mistaken assessing the structure. So even for the nominal data the bar chart is really worth considering as the default one. To explain the issue, compare the structure in two pie charts presented below.

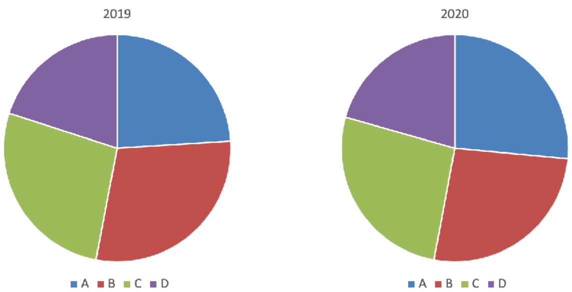


Figure 2.6. Pie charts

Source: Own elaboration.

Bar charts are designed for nominal and ordinal data. They show differences between frequencies of categories much better. The vertical axis can show counts or percentages. Compare the same data as presented above in the form of bar charts, the differences within the chart and between the charts are much clearer now.

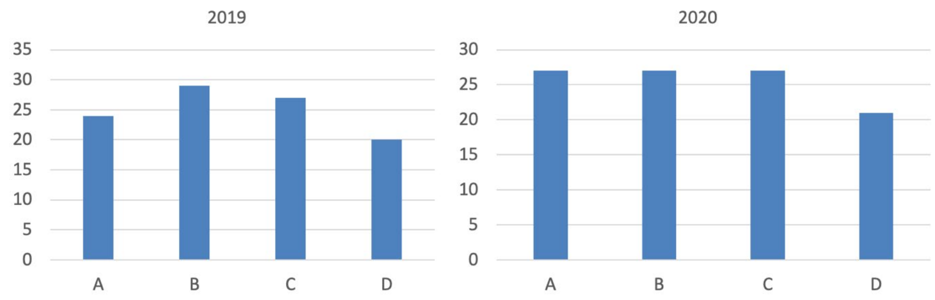


Figure 2.7. Bar charts

Source: Own elaboration.

The histograms are designed for interval and ratio levels of measurement. They are similar to bar charts, but bars are touching (continuous data), and every bar shows the frequency of observation within the interval.

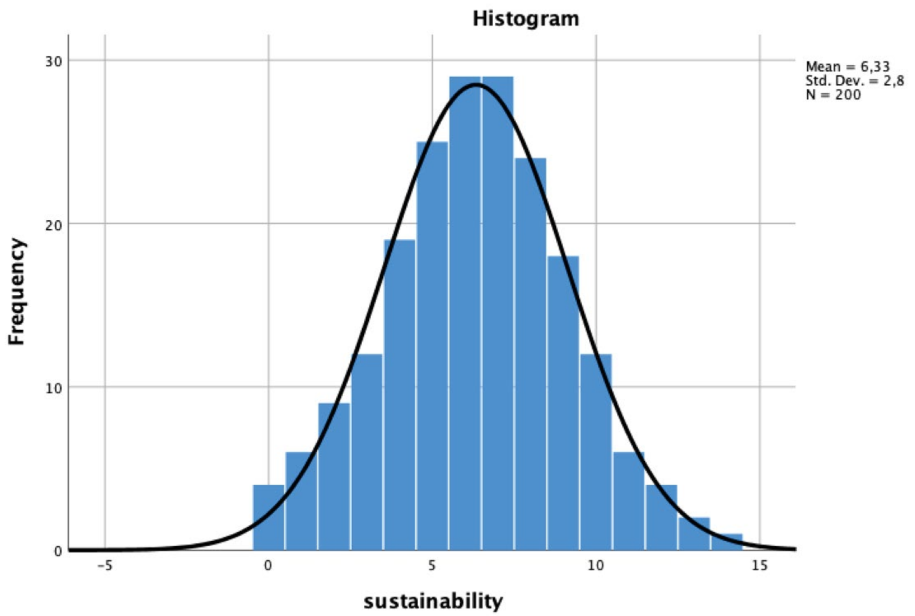


Figure 2.8. Histogram with the added normal curve

Source: Own elaboration.

2.4.2. Box and Whiskers chart

There is one more popular and very informative way to visualize variables. The chart is called box and whiskers chart, five numbers chart or box chart. What we need for constructing the chart is a minimum and a maximum, first, second and third quartile. The box is drawn having its bottom as the first quartile, the upper limit as the third quartile and the thick line within this range presenting the median.

The minimum and maximum are presented in the form of whiskers. It sounds very easy and can be hand-drawn, but SPSS definitely helps to produce it.

If the minimum (maximum) is considered an outlier, the whisker will present the smallest (highest) value not exceeding the 1.5 of interquartile range. And all the outliers will be shown as small circles (closer distance outliers) or asterisks (far distance outliers). You can produce the box and whiskers chart using the Chart menu in SPSS (both Charts / Legacy Dialog / Boxplot and Chart / Chart Builder / Boxplot lead to this), but a very handy way is available in Analyse / Descriptive Statistics / Explore. It produces a bunch of useful descriptive statistics, and two graphical summaries: Stem and Leaf (not described here), as well as Box and Whiskers chart. Just drop the investigated variable into the field of Dependent list. If you want to compare the results splitting your data (e.g., for females and males) you have to drop this variable into the factor list. There is one more field there, the Label Cases by. It tells SPSS how to name the outliers. If you leave this field empty, every outlier will get the number of rows from the data file, so you can spot it if needed. If you drop the variable there, the outliers will be named according to their value of the chosen variable.

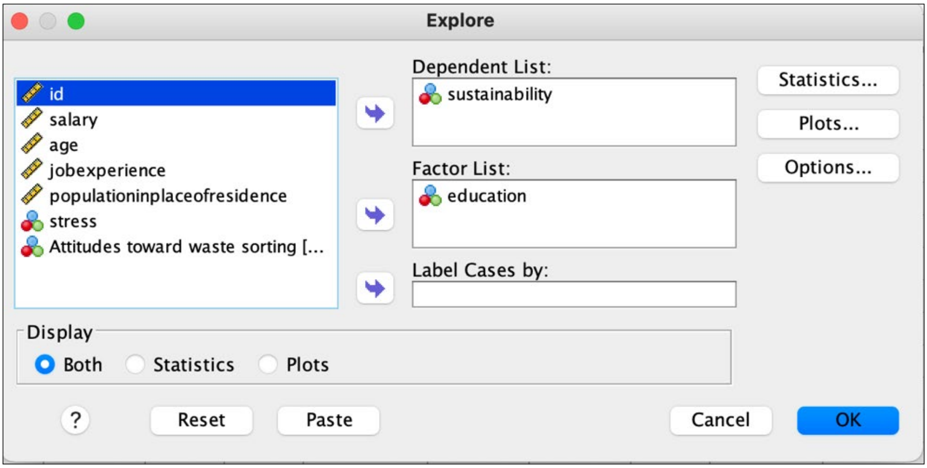


Figure 2.9. Dialog box for Explore (including Box and Whiskers chart)

Source: Own elaboration.

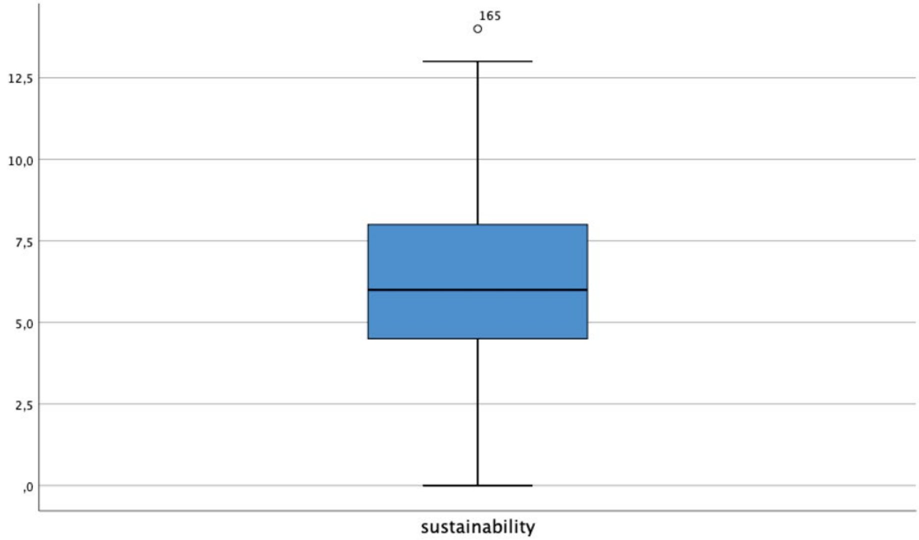


Figure 2.10. Box and Whiskers chart—single box

Source: Own elaboration.

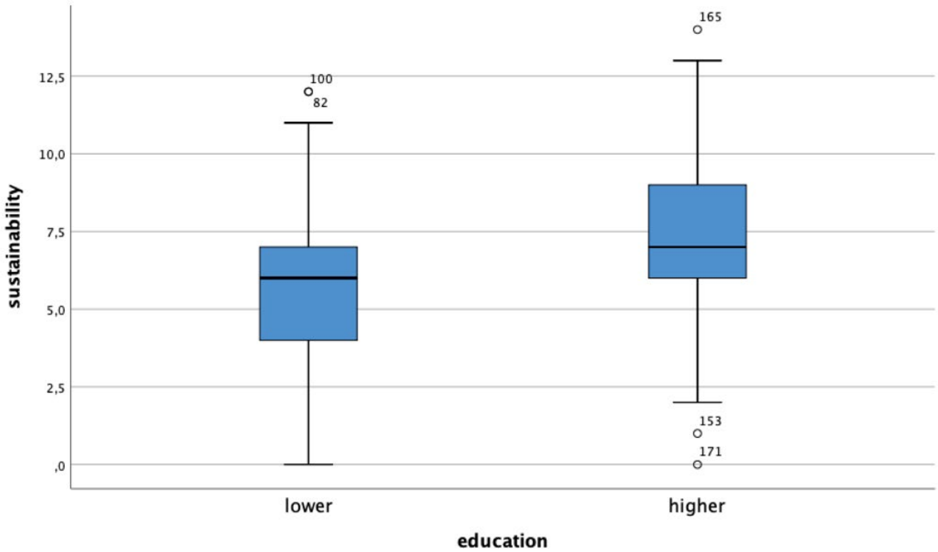


Figure 2.11. Box and Whiskers chart—comparison (more boxes)

Source: Own elaboration.

The comparison box shows that even if the difference between medians is small, the typical 50% of the people (inside the box) with higher education have better sustainability knowledge compared to the people with lower education levels. We can observe two outliers in the group with lower education (having much higher knowledge) and three outliers in the group with higher education (two much lower, one much higher).

2.4.3. Crosstabs: percentage

Tables are the most popular way to present data. But this popularity does not mean, that readers always understand them correctly. In this chapter we will focus on the tables presenting relationships between two categorical variables.

A simple table usually presents the dependent variable and all its values as rows. The independent variable constitutes columns. This allows to check the counts or percentage in the cell where row and column cross.

The basic table may be obtained in SPSS in Analyse / Descriptive Statistics / Crosstabs.

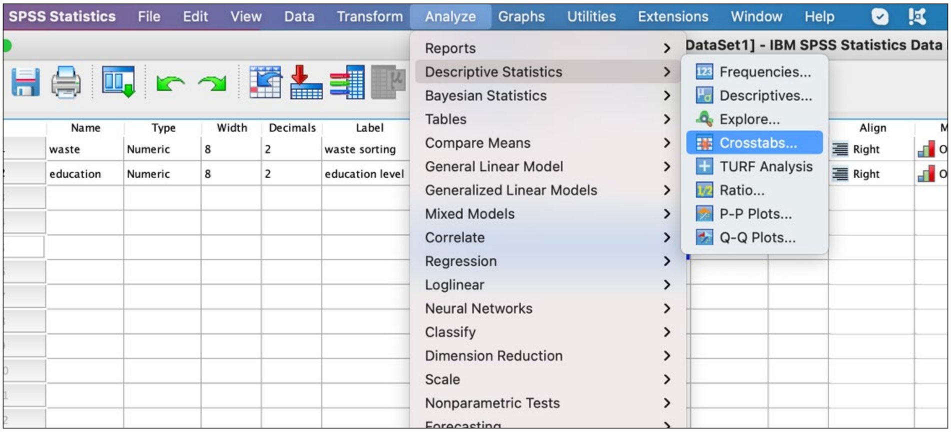


Figure 2.12. Obtaining the Crosstabs: path

Source: Own elaboration.

In the dialog box you have to drop the independent and dependent variable into fields called Row(s) and Column(s) as shown on the example in Figure 2.13.

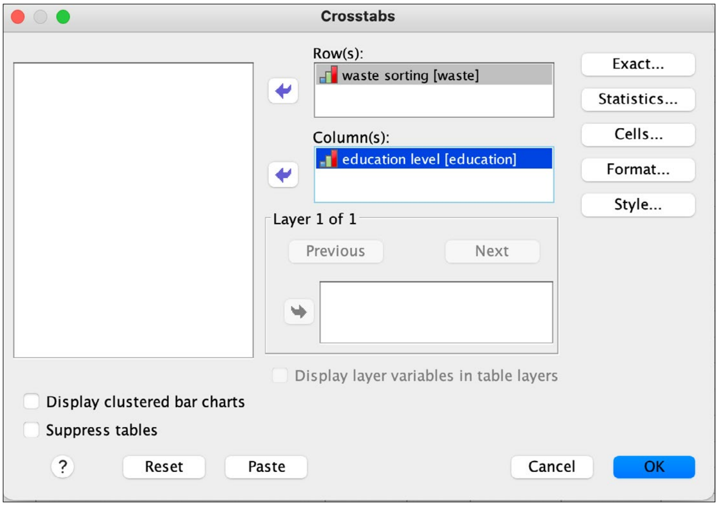


Figure 2.13. Dialog box of Crosstabs

Source: Own elaboration.

The default version produces a table with counts. Each cell of the table shows the number of observations meeting the condition of having a certain value of independent and certain value of dependent variable.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
waste sorting * education level	160	100,0%	0	0,0%	160	100,0%

waste sorting * education level Crosstabulation

Count

		education level			
		1.00 primary education	2.00 high school	3.00 university education	Total
waste sorting	1.00 no, everything goes to general waste	22	18	9	49
	2.00 we sort waste into not more than two categories	19	34	25	78
	3.00 we sort waste into more than two categories	10	12	11	33
Total		51	64	45	160

Figure 2.14. Crosstabulation with counts (default)

Source: Own elaboration.

We can read that 22 observations in the database meet both conditions: no sorting (all to general waste) and primary education. Below and on the right side of the table we can see the totals: the summarized observations from every row and from every column. We can read that we have 33 responses of sorting waste into more than two categories.

When presenting counts you can also display the expected values (the values if there are no dependencies between variables) or the differences between expected and observed values. Going further in this direction would lead us to chi-square, but it is not our intention.

Comparing counts may be difficult if the groups are not evenly divided. A much easier option is offered by presenting percentages. The table showing percentages are usually calculated as:

– row percentage,
– column percentage,
– table percentage.

The percentage direction in the table may be chosen in the dialog box within the procedure Crosstabs. This is under the button Cells. The default option is Observed, you can unclick it and choose the percentage.

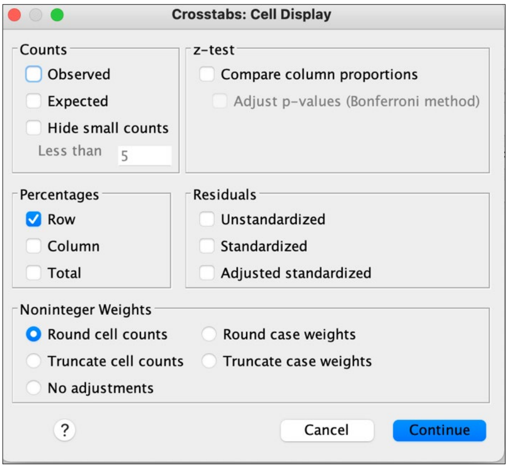


Figure 2.15. Dialog box of Cells options in the procedure of Crosstabs

Source: Own elaboration.

Row percentage tables take every row as the base for calculating percentages. It means that every row makes 100%. This shows the shares of all the values of the column variable in one single value of row variable. In Figure 2.16 it shows that 44.9% of all the people not sorting waste (everything goes to general waste) are people with primary education. As you can see every row sums up to 100%.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
waste sorting * education level	160	100,0%	0	0,0%	160	100,0%

waste sorting * education level Crosstabulation

% within waste sorting

		education level			
		1.00 primary education	2.00 high school	3.00 university education	Total
waste sorting	1.00 no, everything goes to general waste	44,9%	36,7%	18,4%	100,0%
	2.00 we sort waste into not more than two categories	24,4%	43,6%	32,1%	100,0%
	3.00 we sort waste into more than two categories	30,3%	36,4%	33,3%	100,0%
Total		31,9%	40,0%	28,1%	100,0%

Figure 2.16. Crosstabulation with row percent

Source: Own elaboration.

The tables displaying column percent are the most popular ones. They allow for comparing shares of row variable within the chosen value of column variable. This way we can easily read, in which group the waste management is more frequently accepted.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
waste sorting * education level	160	100,0%	0	0,0%	160	100,0%

waste sorting * education level Crosstabulation

% within education level

		education level			
		1.00 primary education	2.00 high school	3.00 university education	Total
waste sorting	1.00 no, everything goes to general waste	43,1%	28,1%	20,0%	30,6%
	2.00 we sort waste into not more than two categories	37,3%	53,1%	55,6%	48,8%
	3.00 we sort waste into more than two categories	19,6%	18,8%	24,4%	20,6%
Total		100,0%	100,0%	100,0%	100,0%

Figure 2.17. Crosstabulation with total percent

Source: Own elaboration.

The tables displaying table percent shows the percentage of observations meeting certain condition (having a given value of independent variable and given value of dependent variable) compared with the total number of observations in the table. Thus we have the 100% only if we sum up all the cells of the table.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
waste sorting * education level	160	100,0%	0	0,0%	160	100,0%

waste sorting * education level Crosstabulation

% of Total

		education level			
		1.00 primary education	2.00 high school	3.00 university education	Total
waste sorting	1.00 no, everything goes to general waste	13,8%	11,3%	5,6%	30,6%
	2.00 we sort waste into not more than two categories	11,9%	21,3%	15,6%	48,8%
	3.00 we sort waste into more than two categories	6,3%	7,5%	6,9%	20,6%
Total		31,9%	40,0%	28,1%	100,0%

Figure 18. Crosstabulation with table percent

Source: Own elaboration.

We can read from the table in Figure 2.18 that 13.8% of all observations meet both conditions: not sorting (all to general waste) and primary education. The totals below and on the right side of the table show the structure of column variable and row variable.

All the above tables (Figure 2.14, Figure 2.16, Figure 2.17, and Figure 2.18) were constructed from same database. Every table shows data according to certain expectations. Without knowing the direction of percenting the table one could be mistaken. Attention should be paid not only when constructing tables with percentages, but when reading someone's else results.

2.5. Visualization— Likert scale and some chosen charts

2.5.1. Visualization of the Likert scale

As mentioned in the chapter about questionnaire design, there are various scales helpful in collecting data. There are multiple ways to present data collected with these scales and showing them would probably lead to publishing a separate book. But at least the presentation of data from the Likert scale is worth mentioning as a highlighted example.

The Likert scale is a bunch of statements measured on the ordinal level. So, respecting the rules set in the previous chapter, the analysis should not go further than position measures. But the practice is different in this case. There is an assumption that the intervals within the scale are equal, and the mean and variance can be measured. This allows for various types of analyses, including the discriminant power of statements based on standard deviation or factor analysis based on variance.

This upgrade is often discussed and raises questions. I would suggest not upgrading the level of measurement if the scale uses less than five categories. The sample is small, and the scale uses verbal labels for all classes (not only for the extreme classes), or the distribution is heavily skewed.

There is one more often highlighted misunderstanding. The results from the Likert scale can be summed up and presented as mean or sum. This is the proper procedure and has nothing to do with the above-stated concerns.

Talking about the raw data collected using the Likert scale, the profile chart seems to be a proper tool for effective visualization. Furthermore, it can use both approaches: the ordinal and the interval assumption.

Such a chart displays the statements and the line with equally distributed points (units of the scale). For each statement, the mode or median (ordinal level of measurement approach) or mean (interval level of measurement approach) can be shown. Usually, all the points are connected by lines. This scheme allows for effective comparison of results in groups (e.g., results for females and males).

For better understanding, all the negative statements should be reversed, so every statement result on the right side of the chart means always agreeing more and having a stronger attitude.

The example below shows data about sustainable consumption behaviour collected using the questionnaire designed by Quoquab, Mohammad and Sukari (2019). There are two subgroups in the sample, the younger (red line) and the older (blue line) inhabitants of the camp. The lines are matching the means for every statement. It means, that the more to the right, the group agreed more with the statement. The gaps between lines show the differences between groups.

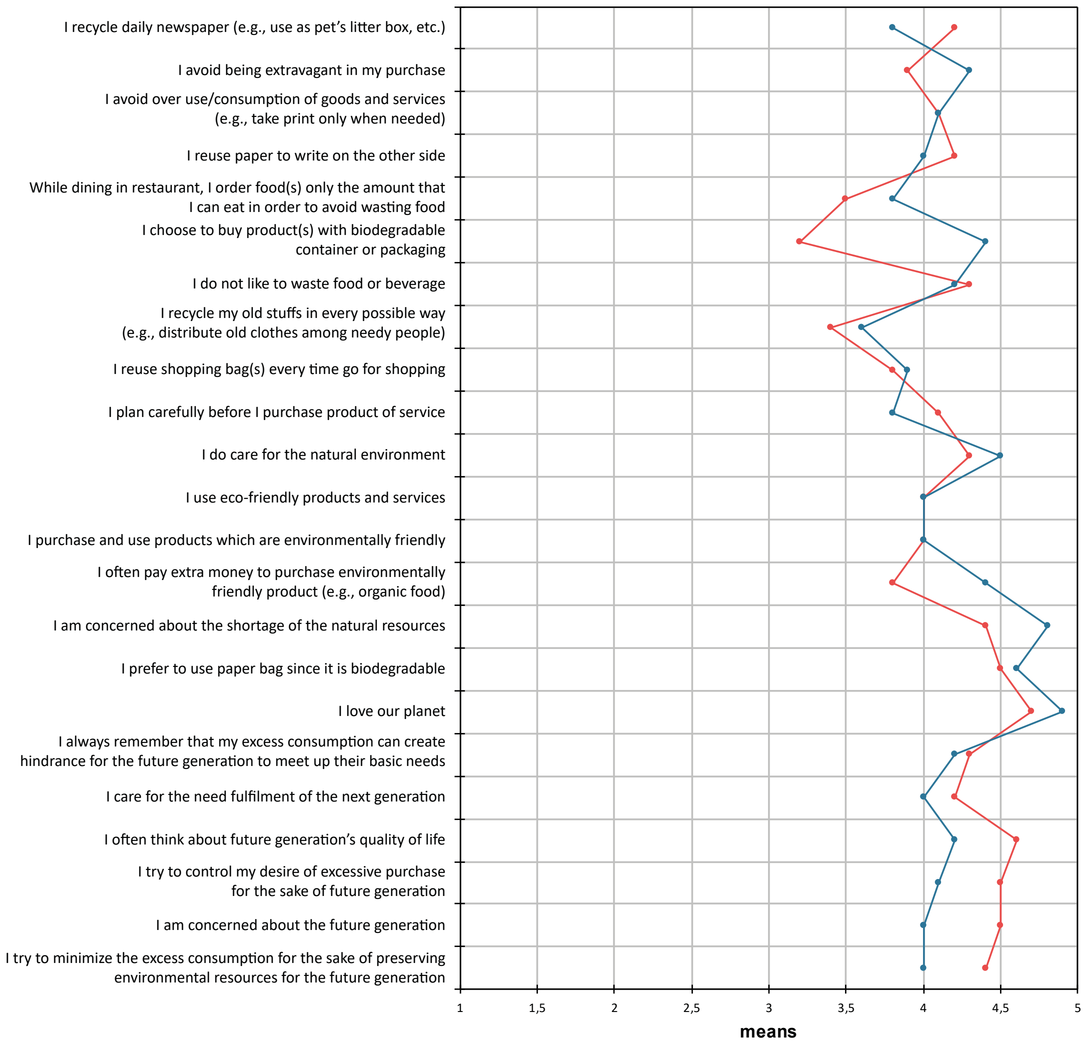


Figure 2.19. Profile chart from Likert scale

Source: Own elaboration.

2.5.2. Other examples of data visualization schemes

It is not our intention to decide which other charts are more practical, so three types of graphs are presented to show some options from the variety.

The scatterplot shows a relationship between two continuous variables. Every point on the scatterplot is one observation, and its X and Y measurements are the values of two variables.

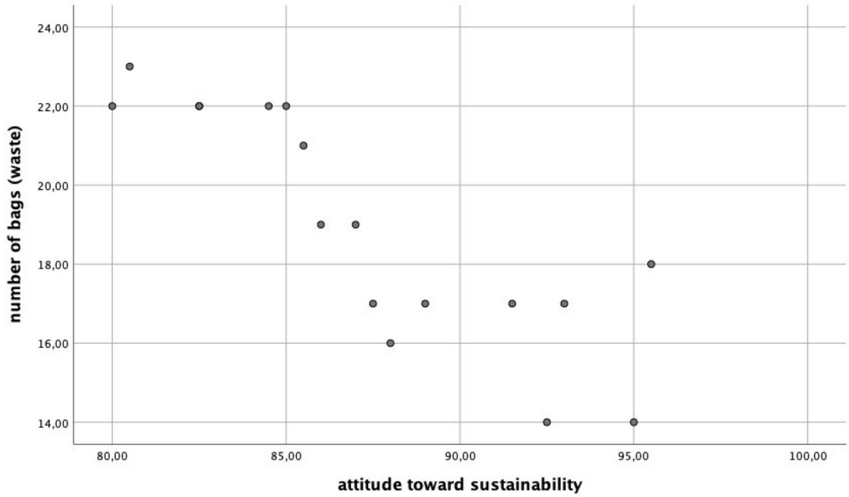


Figure 2.20. Scattterplot

Source: Own elaboration.

If the variables should suggest causation, the independent variable (cause) sets the horizontal axis, and the dependent variable (effect) sets the vertical axis. The point on the scatterplot can show additional information, through adding successive variables. Usually, additional information is brought by a difference of colour or shape of the points.

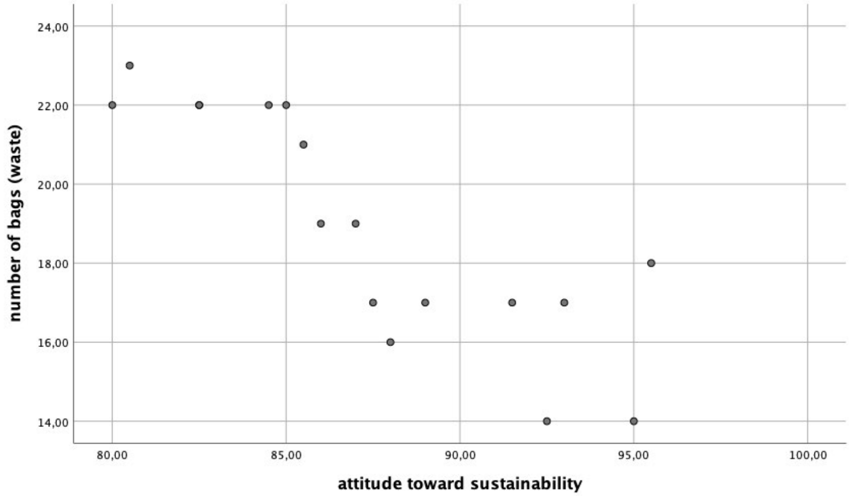


Figure 2.21. Scattterplot showing results for two groups

Source: Own elaboration.

Quite a popular extension of scatterplot is known as a bubble chart. In this case, an additional value (usually a continuous variable) is presented as the size of the point (bubble).

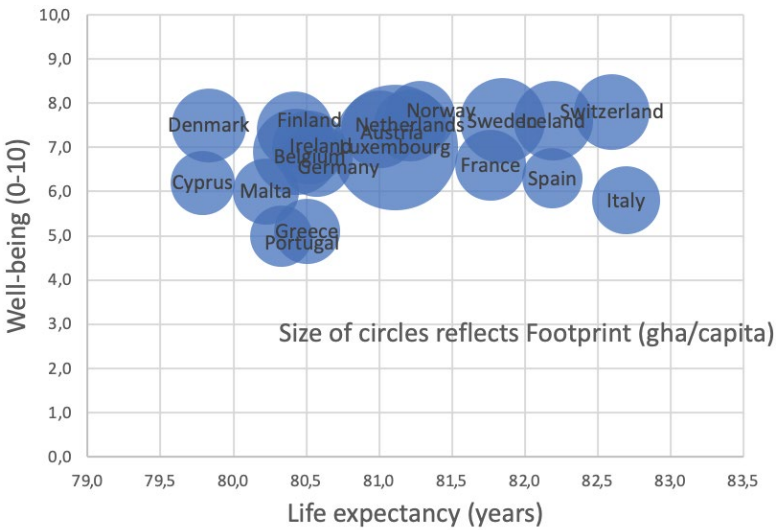


Figure 2.22. Scattterplot showing results for two groups

Source: Own calculations from Happy Planet Index data.

Questions / tasks

1. Which level of measurements gives more possibilities in analysis (more statistical techniques): ordinal or interval?
2. Which level of measurement are yearly CO_2 emissions in countries of the European Union (in tons) the examples of?
3. Can a median be applied for estimating the average emission?
4. Can mean and standard deviation be applied to describe the emissions in EU?
5. Give your own examples of variables measured on every level of measurement.
6. Considering the example presented in section 2.2.3, answer the following questions:
   - What scales were used in the questions?
   - Which questions and why may be irritating to consumers and will be skipped by them?
   - How can the questions be improved?
   - Create additional questions if necessary to verify the hypothesis.

- Create additional 3 questions that could be included in the part of characteristics.
- What methods can the data obtained in each of the questions be analysed with?

7. Face-to-face interview, telephone interviews, Web surveys or mail questionnaires—what types of research problems are they recommended for? Give some examples of methods.
8. Open-ended versus close-ended questions—what are the advantages and disadvantages of these questions? Give some examples of questions which can be used in the study on sustainable development.
9. How can the correctness of the questionnaire be further controlled before the pilot study?
10. How can you use the questions in the questionnaire to check if the respondent answers are reliable?
11. When designing the study on sharing economy, the researchers formulate, *inter alia*, the following research hypotheses:
    H1: The main reason to use Airbnb's services is to get to know the life of the inhabitants of a given country.
    H2: The consumer, having a choice of Airbnb and a hotel—assuming the same price—will choose Airbnb.
    H3: People who use Airbnb's services most often are people under 30, on low incomes, solo travellers.
    Prepare a fragment of the questionnaire that will verify the above hypotheses
12. Find an online questionnaire in scientific publications on selected aspects of sustainable development. Indicate what kind of questions and scales were used in that tool. Also make a substantive evaluation:
    - sharing economy,
    - ethical and unethical shopping behaviour of buyers,
    - sustainable development,
    - food waste,
    - shopping in second-hand shops,
    - renewable energy resources.
13. What topics are commonly threatening to respondents? How can a researcher ask about them?
14. In 2018, there were 640 third age universities in Poland, i.e., entities whose main goal is educational activity, integration and activation of older people in order to improve the quality of life and increase their participation in social life. A total of 113.2 thousand elderly people, including 95.4 thousand women, studied there. The greatest number of studying seniors are people aged 61–75 (71.9%), while people under 60 were 11.7%, and people aged 76 and over 16.4%. How

many retirees studying at such universities should be selected for the study on attitudes to the concept of lifelong learning assuming that the margin of error is 3% and confidence level 90%.

15. An average resident of a certain country in Europe saves an average of 6% of their salary. The average monthly salary at the beginning of 2021 was 1,450 euro. A certain financial institution intends to conduct a representative nationwide quantitative survey. It was hypothesized that the respondents' answers would depend on their average monthly savings.
    - Calculate how many respondents should be included in the sample of quantitative research, assuming a confidence level of 90% and the margin of error of 4% of the average savings. The statistical reports of banks also show that the average standard deviation of monthly savings in banks is 50 euro.
    - How will the sample size change if the confidence level is 95%?
    - How will the sample size change if the maximum acceptable error of estimate is reduced to 2%?
16. What sampling method should be used if the purpose of the quantitative study is to define forecasts for the sustainable development of air transport in selected countries of Central and Eastern Europe? The respondents in the survey are to be airport managers in the Czech Republic, Hungary, Poland, the Slovak Republic, Bulgaria, Croatia and Ukraine.
17. One international research agency intends to conduct quantitative research on the attitudes of people in several European countries towards the problems of sustainable development. It was assumed that the attitudes would depend on the level of education and age of the inhabitants. Select one country, find statistical data on a given society (in terms of education and age), then determine the minimum sample size and use stratified selection to determine the size of each sub-group.
18. What kinds of variables may be visualized using the bar chart?
19. Which chart would you suggest for the variable on the ratio level of measurement?
20. How can you interpret the boxes in the Box and Whiskers chart?
21. Explain the differences in three ways of presenting tables.
22. What characteristics of Likert scale should be considered before upgrading it from ordinal to the interval level of measurement?
23. In case of showing causation, how the independent and dependent variable should be presented in the scatterplot?
24. What is the difference between scatterplot and bubble chart?

References

Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it?. *Indian Journal Of Medical Specialities*, *4*(2), 330-333.

Aczel, A. D. (2009). *Complete business statistics* (7th ed.). New York: McGraw-Hill.

Barrow, M. (2017). *Statistics for economics, accounting and business studies*. Harlow: Pearson.

Bougie, R., & Sekaran, U. (2020). *Research methods for business. A skill building approach*. Hoboken: John Wiley & Sons.

Churchill, G. A., & Iacobucci, D. (2018). *Marketing research. Methodological foundations* (12th ed.). Nashville: Earlie Lite Book, Inc.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: Sage.

Dean, S., & Illowsky, B. (2013). *Introductory statistics*. Houston: Rice University, OpenStax College.

Demography and Population. Retrieved February, 2021 from https://stats.oecd.org/

Etikan, E., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4.

Happy Planet Index data. Retrieved May 19, 2021 from happyplanetindex.org

Kasiulevičius, V., Šapoka, R., & Filipavičiūtė, R. (2006). Sample size calculation in epidemiological studies. *Gerontologija*, *7*(4), 225-231.

Levy, P. S., & Lemeshow, S. (2008). *Sampling of populations: Methods and applications*. Hoboken: John Wiley & Sons.

McDaniel, C., & Gates, R. (2018). *Marketing research* (11th ed.). Hoboken: John Wiley & Sons.

Malhotra, N. K., & Birks, D. F. (2003). *Marketing research. An applied approach*. Harlow: Prentice Hall.

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches*. Harlow: Pearson.

Quoquab, F., Mohammad, J., & Sukari, N. N. (2019). A multiple-item scale for measuring "sustainable consumption behaviour" construct: Development and psychometric evaluation. *Asia Pacific Journal of Marketing and Logistics*, *31*(4), 791-816.

PART 2.

SELECTED METHODS OF DATA ANALYSIS

Qualitative and quantitative methods in sustainable development, pp. 83-116
https://doi.org/10.18559/978-83-8211-072-2/03

3.

FACTOR ANALYSIS IN SUSTAINABLE DEVELOPMENT RESEARCH

Iwona Olejnik Blaženka Knežević, Magdalena Stefańska
Poznań University of Economics and Business

Blaženka Knežević
University of Zagreb

Magdalena Stefańska
Poznań University of Economics and Business

Abstract: Too much data describing a given phenomenon requires synthesizing them. For this purpose, researchers can use various methods of analysis. Factor analysis is one of them.

In this section, first the basic theoretical aspects of factor analysis, as well as the stages of its use are described while presenting the essential minimum necessary to understand the essence of the method.[1] The second part presents an example of the use of this method in research on sustainable consumption. The last part of this chapter presents case study of the use of factor analysis in research on managers' ethics in retail industry.

Keywords: factor analysis, sustainable consumption, PRESOR scale.

[1] Readers interested in theoretical details related to this method are referred to the book: Aczel (2009, pp. 768–797).

3.1. Theoretical background

In social sciences, many concepts based on data that comprehensively describe a given problem have been created and developed thanks to the use of the factor analysis (FA) method. At the same time, in quantitative research, the number of variables is often so large that it may hinder the interpretation of the results. This especially concerns the situation when attitudes towards a certain phenomenon are studied with the use of many statements that the respondent evaluates according to their own perception. Many of these statements can be closely related to each other, so it may be reasonable to group them together. Factor analysis can be used for this purpose.

Factor analysis is one of multivariate data procedures. Its main purpose is data simplification "by reducing a large set of variables to a smaller set of factors or composite variables by identifying underlying dimensions of the data" (McDaniel & Gates, 2018, p. 454). The resulting unobservable (hidden) variables are called factors (or also constructs, structures and dimensions). They can also be used then in further analyses, e.g., in market segmentation or regression analysis (see case study no. 2 and 3).

In statistical packages, the "factor analysis" procedure, two methods that differ in terms of their assumptions can usually be found. They are factor analysis (FA) and principal-components analysis (PCA). The results obtained with their use are usually remarkably similar (Aczel, 2009).

FA is a relatively simple method used to analyse the structure of the studied phenomenon, i.e., the tested feature obtained, e.g., in quantitative research. It is worth adding that by definition the factor analysis should be conducted on the variables obtained from at least interval scale (this scale is the third level of measurement after the nominal and the ordinal scale). However, this method can also be used in the case of Likert scale popular in social research (preferably a minimum of 5 points). In addition, it needs to be remembered that in the study we should have ten times more observations in comparison with the variables that we want to include in the analysis (i.e., if we analyse 30 statements, the sample on which we conducted the study should not be smaller than 300 respondents) (Costello & Osborne, 2005; Field, 2009).

Thus, in FA it is assumed that each of the correlated variables is affected to a different extent by common factors that explain the observed correlation. Moreover, it is assumed that each of the explicit variables is a linear combination of hidden variables, and a specific factor, separate for each of the variables. The main problem in this method is the selection and determination of directly unobservable factors.

3.2. Factor analysis—research steps

The basic stages of factor analysis are presented in Figure 3.1. The first three stages in the application of factor analysis are associated with the process of selecting variables. In databases created on the basis of surveys, lack of data caused by various reasons is often encountered (McDaniel & Gates, 2018, pp. 116–120). Before starting factor analysis, the analyst should therefore examine the available database and prepare one that will be free from data gaps. This is because they could distort the obtained results. There are various ways to replace missing values, e.g.: excluding cases listwise, excluding cases pairwise, and replacing with mean.

The purpose of the second and third steps, i.e., matrix calculation and correlation assessment, is to check whether it makes sense to use factor analysis in the process of data analysis. At least some of the analysed variables should be correlated with each other—if it were not so, each variable would constitute one factor, which of course would undermine the sense of using this method. The assessment of the correlation matrix allows for the removal of variables that are not correlated with others, from further analysis.

Determination of how to deal with the lack of data

Calculation of the correlation matrix

Assessment of the correlation and removal of variables not correlated with others from the analysis

Identification of factors

Determination of the number of factors for further analysis

Factor rotation

Calculation of factor values and naming the factors

Figure 3.1. Stages of factor analysis

Source: Own elaboration.

For reference, Kaiser put the following values on the results:
0.00 to 0.49 unacceptable
0.50 to 0.59 miserable
0.60 to 0.69 mediocre
0.70 to 0.79 middling
0.80 to 0.89 meritorious
0.90 to 1.00 marvellous

Kaiser-Meyer-Olkin test (KMO) enables among others, the selection of variables and the assessment of their usefulness in factor analysis. This test is a measure of how suited our data is for Factor Analysis (*What is the Kaiser-Meyer-Olkin (KMO) Test?* The ideal situation is 0.8 for this test, but lower values are also sometimes accepted.

The second one—Bartlett's test of sphericity, tests the hypothesis that our "correlation matrix is an identity matrix, which would indicate that your variables are unrelated and therefore unsuitable for structure detection. Small values (less than 0.05) of the significance level indicate that a factor analysis may be useful with our data" (*KMO and Bartlett's Test*).

The next important step is to choose a factor extraction method. In this case, the analyst can choose between principal components, maximum likelihood, unweighted squares, generalized least squares, principal axis factoring, alpha factoring, and image factoring. In practice, the first two methods are most often used. It should be added that the "principal components method" is used when we expect that the structure will consist of several unrelated (uncorrelated) factors. It is also applied when we expect the variables to be correlated with each other. On the other hand, the "maximum likelihood" method is used only when we expect that the examined factors will only be correlated with each other. It is not suitable for studying such structures where we do not predict correlation between factors.

In the process of determining the number of factors, various criteria including the Kaiser eigenvalue criterion, the Jolliffe criterion, the Guttman criterion (Jolliffe, 1986), are used.

The next step is the rotation of factors, the aim of which is to obtain such a set that is suitable for their better interpretation in comparison with the primary factors (so that that each variable has a high load in only one factor). In this case, the most frequently used methods include "varimax" rotation, which is used assuming that the dimensions will not be correlated with each other, or "direct oblimin", where it is assumed that the structure will be correlated.

The variance explained analysis enables us to assess whether the separated structure is appropriate. It allows for the determination of differences between the respondents in terms of their attitudes towards the subject of the study, e.g., in terms of sustainable consumer behaviour. All the studied elements (statements) explain 100% of the information, but obviously in further analyses we want to work on reduced data, and we do not want to consider all dimensions. Therefore, the cumulative percentage of the variance should be analysed. It would be good if this value exceeded 50%.

The interpretation of the obtained results is the final stage of factor analysis. By assessing their factor loadings, it is analysed which of the statements are part of a given dimension. If the load for a given statement (questionnaire item) is the highest, it is assigned to a given dimension. If it is also a part of another dimension, it can be decided that is should be removed or assigned to the dimension in which this load is greater. In the case of negative factor loadings, the meaning of this statement should be analysed inversely.

3.3. Sustainable consumption behaviour—an example of application of factor analysis using the IBM SPSS Statistics version 26.0

3.3.1. Model assumptions and selection of variables

In order to construct a model of attitudes towards sustainable consumption, the initial step include extracting the elements that build up these attitudes in cognitive, emotional, and behavioural areas. You can try to create these elements yourself, but it is worth using the scales that have already been built and validated.

The following example uses the scale presented in the article written by Quoquab, Mohammad and Sukari (2019). It contains the statements presented in Table 3.1.

Table 3.1. Scale to measure sustainable consumption behaviour: example

I always try hard to reduce misuse of goods and services (e.g., I switch off light and fan when I am not in the room)
I recycle daily newspaper (e.g., use as pet's litter box)
I avoid being extravagant in my purchase
I avoid overuse / consumption of goods and services (e.g., take print only when needed)
I reuse paper to write on the other side
While dining in restaurant, I order food(s) only in the amount that I can eat in order to avoid wasting food
I choose to buy product(s) with biodegradable container or packaging
I do not like to waste food or beverages
I recycle my old stuffs in every possible way (e.g., distribute old clothes among needy people)
I reuse shopping bag(s) every time I go shopping
I plan carefully before I purchase product or service
I do care for the natural environment
I use eco-friendly products and services

I purchase and use products which are environmentally friendly
I often pay extra money to purchase environmentally friendly product (e.g., organic food)
I am concerned about the shortage of the natural resources
I prefer to use paper bag since it is biodegradable
I love our planet
I always remember that my excess consumption can create hindrance for the future generation to meet up their basic needs
I care for the need fulfilment of the next generation
I often think about future generation's quality of life
I try to control my desire of excessive purchase for the sake of future generation
I am concerned about the future generation
I try to minimize the excess consumption for the sake of preserving environmental resources for the future generation

Source: (Quoquab et al., 2019).

If we want to apply factor analysis with the use of the SPSS package, the "Dimension reduction" and then "Factor" module (Figure 3.2) is applied.

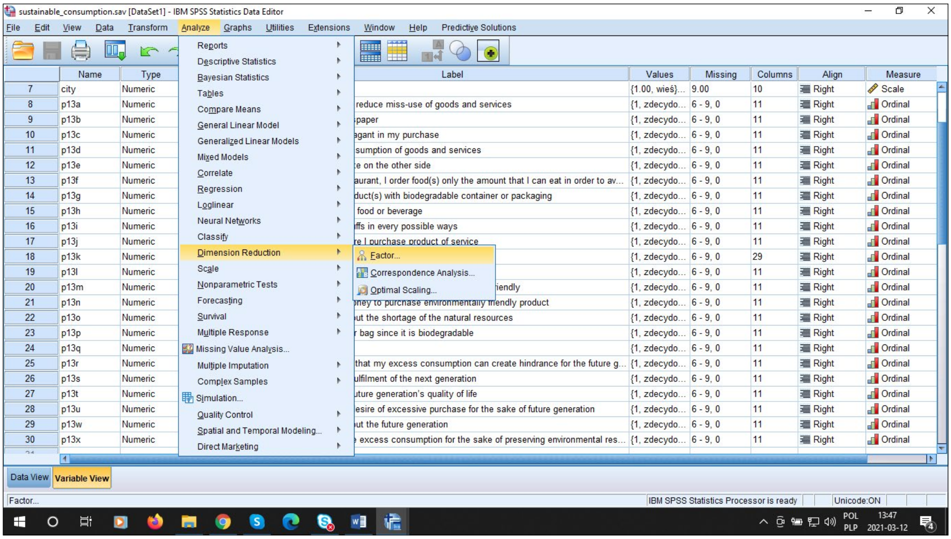


Figure 3.2. Factor analysis in SPSS program—the first step

Source: Own elaboration.

All the variables of interest ("Label") are introduced to the analysis, provided that these statements refer to a given phenomenon examining a certain property or feature. Factor analysis will allow for determining whether there are any relationships between these statements, i.e., if a structure can be created within them (Figure 3.3).

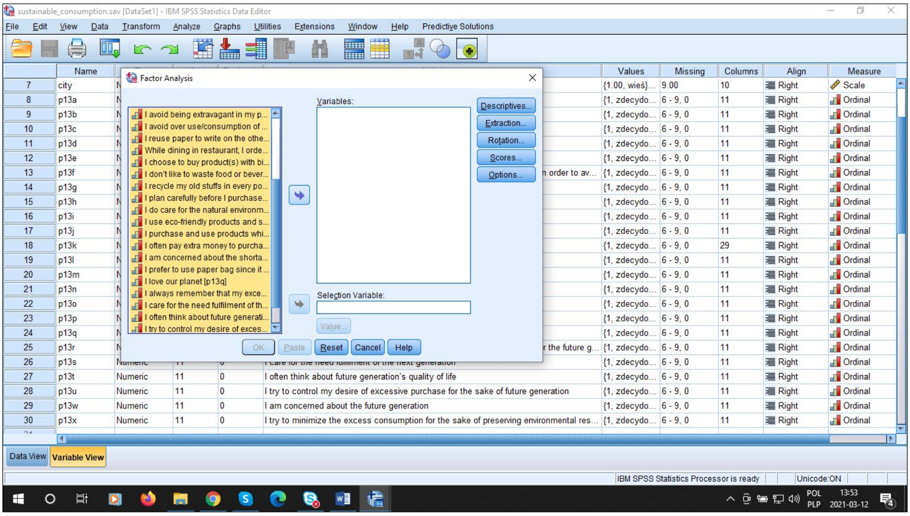


Figure 3.3. Selection of variables for analysis

Source: Own elaboration.

Then "KMO and Bartlett's test of sphericity" and "Anti-image" are selected in "descriptives" (Figure 3.4). It should be reminded that the KMO statistics allow for assessing to what extent the data collected in the research can be used for their FA. In general, it would be best if this test were as large as possible. As previously mentioned, the ideal situation for this test is 0.8, but slightly lower values may also be accepted.

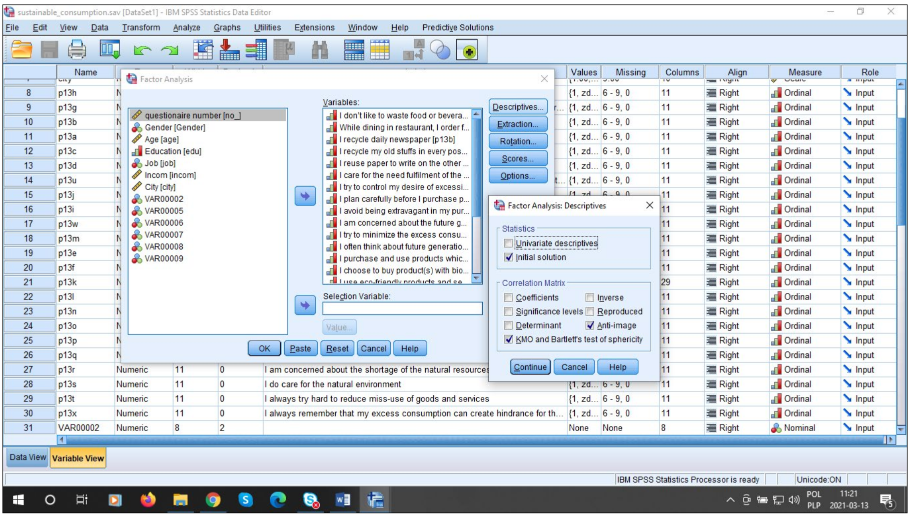


Figure 3.4. Descriptives

Source: Own elaboration.

In turn, by selecting "Anti-image" information on the mutual correlation of the output variables will be obtained. "Anti-image matrices" is a huge table with "Anti-image correlation" found in the second half. If any variable found on its diagonal (marked with letters "a") has a value lower than 0.5, then it should be eliminated from further analysis.

Then we move to the next option, i.e., "Extraction" (Figure 3.5). In our example, the "Principal components" method is selected because we are not sure whether the expected structure will consist of correlated or uncorrelated factors. In addition, we also select the "Scree plot" option, which is used to define how many factors should be distinguished. "Based on Eigenvalue" method is another method that will also allow for determining this. By default, the eigenvalue is 1.

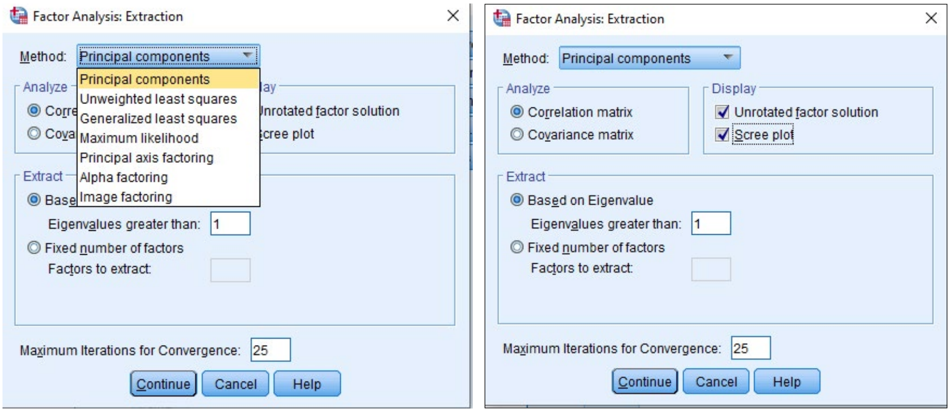


Figure 3.5. Extraction in FA

Source: Own elaboration.

The next step is the selection of "Rotation" (Figure 3.6), i.e., checking whether it is there, and if so, arranging (with the use of certain mathematical algorithms) its factor structure. Since we expect the dimensions not to be correlated with each other, we choose "Varimax" (otherwise we would choose "Direct Oblimin").

The last element, "Options" (Figure 3.7), is used to correctly display the factor loadings, i.e., a certain statistic (correlation) between a given structure element (statement) and a given dimension, i.e., a hypothetical structure that can be observed in the data. To obtain some order, the factor loadings must be sorted by size first. Secondly, it is also worth suppressing the so-called small coefficients to separate items that do not form a factor structure. This is because we want to have such a factor structure whose individual elements are closely related to each other. This means that, e.g., if in the research we work on a method that is well located and recognized in theory, like another version of the sustainable behaviour model, then we assume that the "absolute value below" is 0.3 (i.e., the amount of correlation between the given statement and factor structure); whereas if we test a new tool based on a new theory then we should be more conservative and this value should be 0.4.

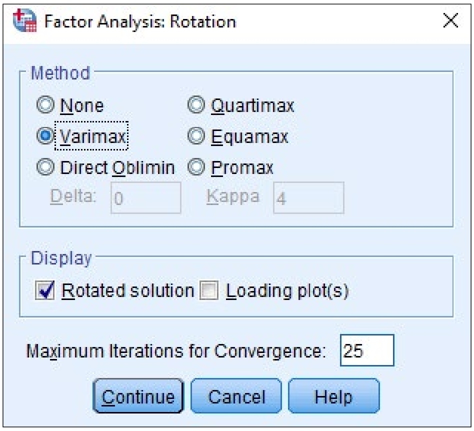


Figure 3.6. Rotation in FA

Source: Own elaboration.

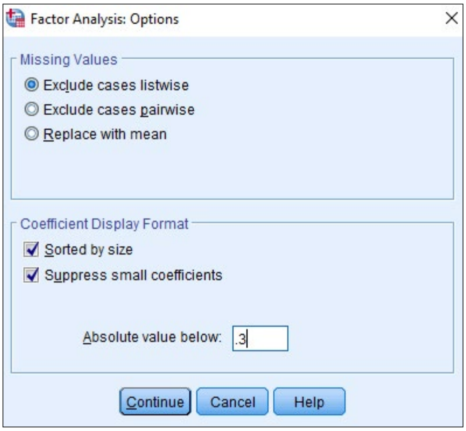


Figure 3.7. Options in FA

Source: Own elaboration.

3.3.2. Model estimation and analysis

After accepting the above assumptions, the tables presented below are obtained (Tables 3.2–3.5). The first one contains the KMO test statistics, which in our example is 0.866. This is a good result, that indicates a great potential of our data. Therefore, they will probably be a good source for distinguishing the factor structure. The Bartlett's test also shows that factor analysis can be a useful method in our research.

Table 3.2. Kaiser-Meyer-Olkin and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.866
Bartlett's Test of Sphericity	approx. chi-square	4795.288
	df	253
	Sig.	0.000

Source: Own elaboration.

The analysis of the "Anti-image correlation" table (the table is too large to be presented here) allowed for including all the statements in further analyses (all correlation coefficients exceeded the value of 0.5).

Table 3.3 presents the resource of common volatility. The "Extraction" column is of primary interest for us. It includes coefficients of correlation between the result related to a given statement and the factor structure. The values in this column should therefore be as large as possible. "I love our planet" statement has the highest value.

Table 3.3. Results of extraction

Communalities	Initial	Extraction
I recycle my old stuffs in every possible way	1.000	0.694
I recycle daily newspaper	1.000	0.686
I reuse paper to write on the other side	1.000	0.448
I care for the need fulfilment of the next generation	1.000	0.587
I often think about future generation's quality of life	1.000	0.444
I purchase and use products which are environmentally friendly	1.000	0.457
While dining in restaurant, I order food(s) only in the amount that I can eat in order to avoid wasting food	1.000	0.683
I do not like to waste food or beverage	1.000	0.575
I avoid being extravagant in my purchase	1.000	0.609
I plan carefully before I purchase product or service	1.000	0.667
I choose to buy product(s) with biodegradable container or packaging	1.000	0.582
I use eco-friendly products and services	1.000	0.674
I try to minimize the excess consumption for the sake of preserving environmental resources for the future generation	1.000	0.497
I often pay extra money to purchase environmentally friendly product	1.000	0.594
I avoid overuse / consumption of goods and services	1.000	0.550
I prefer to use paper bag since it is biodegradable	1.000	0.514
I love our planet	1.000	0.724
I am concerned about the shortage of the natural resources	1.000	0.595
I do care for the natural environment	1.000	0.504
I always try hard to reduce miss-use of goods and services	1.000	0.665
I try to control my desire of excessive purchase for the sake of future generation	1.000	0.564
I am concerned about the future generation	1.000	0.319
I always remember that my excess consumption can create hindrance for the future generation to meet up their basic needs	1.000	0.587

Source: Own elaboration.

Table 3.4 is the first table directly related to the purpose of factor analysis, i.e., identification of the hidden structure. On the left side of the table there are 23 components, i.e., the number of initial statements considered in this example. Such a situation is not very favourable, because it is difficult to describe the phenomenon using so many dimensions, and therefore we want to reduce this value. So how many of these factors should there be? This is determined by the eigenvalue criterion (previously, the standard value was 1—see Figure 3.5). In the discussed example, 6 components exceed the eigenvalue of 1.

Table 3.4. Total variance explained

Com-po-nent	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumula-tive %	total	% of variance	cumula-tive %	total	% of variance	cumula-tive %
1	5.741	24.960	24.960	5.741	24.960	24.960	2.859	12.429	12.429
2	2.180	9.477	34.437	2.180	9.477	34.437	2.515	10.935	23.364
3	1.649	7.170	41.607	1.649	7.170	41.607	2.449	10.650	34.014
4	1.480	6.434	48.041	1.480	6.434	48.041	1.965	8.546	42.559
5	1.135	4.936	52.978	1.135	4.936	52.978	1.893	8.231	50.791
6	1.033	4.491	57.469	1.033	4.491	57.469	1.536	6.678	57.469
7	0.970	4.219	61.688						
8	0.884	3.842	65.530						
9	0.732	3.182	68.712						
10	0.722	3.138	71.850						
11	0.687	2.985	74.836						
12	0.637	2.769	77.605						
13	0.603	2.620	80.225						
14	0.569	2.476	82.701						
15	0.533	2.317	85.018						
16	0.522	2.268	87.285						
17	0.482	2.096	89.382						
18	0.475	2.065	91.447						
19	0.450	1.959	93.405						
20	0.415	1.804	95.209						
21	0.410	1.784	96.993						
22	0.376	1.637	98.630						
23	0.315	1.370	100.000						

Source: Own elaboration.

The scree plot, which is a graphical representation of the data contained in the table above is an alternative to eigenvalue in determining the number of factors. In the discussed example, the scree plot is presented in Figure 3.8. A clear flattening of the line occurs with the 5–6 components, so there should be just 5 or 6 factors. How many of these factors should there finally be? It depends on the researcher's substantive interpretation and the decision which of these choices will be the better solution. In our example, the adoption of 6-factor model was initially decided, because 6 exceeds the eigenvalue of 1. Moreover, when analysing the cumulative percentage (percentage of the variance explained by individual factors, see Table 3.4), 57.5% of the explained variance we will obtained when distinguishing 6 factors. This means that with the help of 6 separate dimensions, we can explain about 58% of the differences between consumers in terms of their sustainable behaviour (with a 23-element factor structure, we would explain 100% of the information).

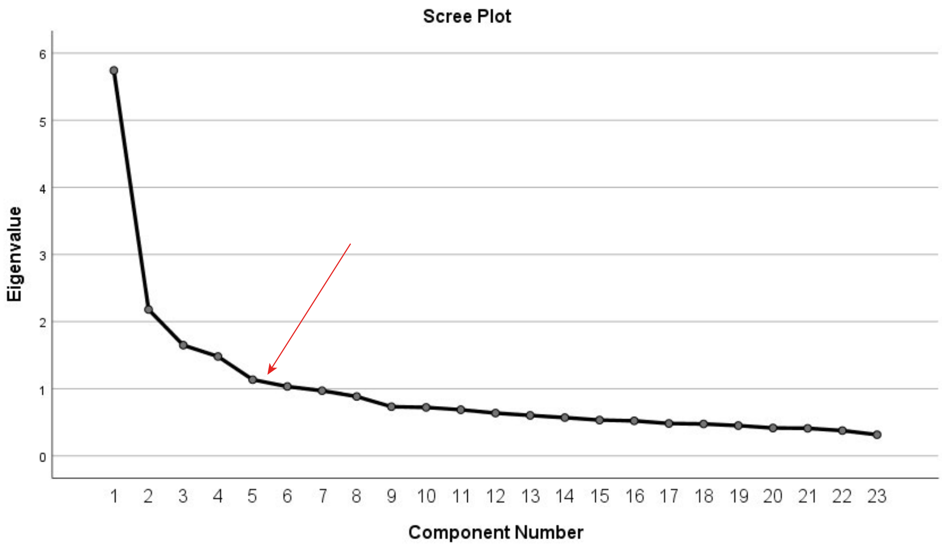


Figure 3.8. Scree plot

Source: Own elaboration.

The detailed factor structure is presented in the rotated component matrix (Table 5), containing the factor loadings that make each of the 6 distinguished factors. In some places it is empty, because we previously specified that the program should not show us loads below 0.3 (Figure 3.7).

How to assign a given statement to individual components? For example, it can be seen that there are 6 statements in the first dimension, including 3 which are also elements of other components. If the factor load in a given dimension (for a given questionnaire item) is the highest, then this item is included in that dimension. For example, the statement "I avoid overuse / consumption of goods and services" has the highest factor load in the first dimension (0.599) and will be included in it, but it also belongs to the second dimension (0.376). As previously mentioned, in this case, the researcher may decide to delete the statements that form more than one factor.

Table 3.5. Rotated Component Matrix

	Component (factor)					
	1	2	3	4	5	6
I use eco-friendly products and services	0.790					
I often pay extra money to purchase environmentally friendly product	0.734					
I choose to buy product(s) with biodegradable container or packaging	0.663					
I avoid overuse / consumption of goods and services	0.599	0.376				
I prefer to use paper bag since it is biodegradable	0.515		0.417			
I purchase and use products which are environmentally friendly	0.400		0.377		0.356	
I always try hard to reduce misuse of goods and services		0.803				
I love our planet		0.796				
I am concerned about the shortage of the natural resources		0.705				
I do care for the natural environment		0.632				
I always remember that my excess consumption can create hindrance for the future generation to meet up their basic needs			0.724			
I often think about future generation's quality of life			0.621			
I care for the need fulfilment of the next generation			0.601			0.347
I try to minimize the excess consumption for the sake of preserving environmental resources for the future generation	0.373		0.519			
I am concerned about the future generation			0.493			
I plan carefully before I purchase product or service				0.787		
I try to control my desire of excessive purchase for the sake of future generation				0.734		
I avoid being extravagant in my purchase				0.705		
I recycle daily newspaper					0.802	
I recycle my old stuffs in every possible way					0.778	
I reuse paper to write on the other side					0.611	
While dining in restaurant, I order food(s) only in the amount that I can eat in order to avoid wasting food						0.803
I do not like to waste food or beverages						0.664

Source: Own elaboration.

After analysing the elements that form individual factors, it is worth naming each of these factors in the next stage. For example, the first factor could be called "friendly environment", the third "care for the future generation", and the last "not wasting food".

3.4. Testing managers' ethics in retail industry: case study no. 1

Implementation of sustainable business polices and application of corporate social responsibility in various industries relies on the level of business ethics of managers. There are numerous scales in contemporary business literature created for assessing managers' business ethics and their general perceptions towards business ethics. Examples of such scales are PRESOR, EPQ and ATBEQ.

PRESOR scale (abbreviation for: The Perceived Role or Ethics and Social Responsibility) is created to measure how an individual perceives the role of ethics and social responsibility in achieving organizational effectiveness. The scale consists of 13 statements and it was designed by Singhapakdi, Vitell, Rallapalli and Kraft (1996).

EPQ (abbreviation for: The Ethics Position Questionnaire) is an instrument designed by Forsyth (1980) to assess individual differences in relativism and idealism. It consists of 20 items divided into 2 scales: 20 items, two scales: (1) idealism and (2) relativism.

ATBEQ (abbreviation for: The Attitudes towards Business Ethics Questionnaire) was created by Preble and Reichel (1988) to measure attitudes on option selection regarding exact business ethics situations. It comprises 30 statements on business ethics.

The above-mentioned scales are applied over time in various industries in original form or with slight modifications. They are proven by numerous authors as reliable tools in the field of corporate social responsibility and business ethics (see Davis, Andersen, & Curtis, 2001; Promislo, Giacalone, & Welch, 2012; Kurnoga, Knežević, & Šimurina, 2017).

In advance, we will show how factor analysis can be used to analyse data collected by applying PRESOR scale on a sample of managers in retail industry.

In the questionnaire distributed to 1000 managers in retail industry in Croatia, there was a question based on a PRESOR original 13-items scale (Singhapakdi et al., 1996). The question was formulated as a Likert scale evaluation table (see Table 3.6).

Table 3.6. Survey question in form of the PRESOR scale

Statement code	Statement	Indicate level of agreement*				
		1	2	3	4	5
Q1	being ethical and socially responsible is the most important thing a firm can do					
Q2	bending and breaking the rules is acceptable if a firm is making a profit					

Q3	the ethics and social responsibility of a firm is essential to its long-term profitability					
Q4	overall effectiveness of a business can be determined, to a great extent, by the degree to which it is ethical and socially responsible					
Q5	to remain competitive in a global environment, business firm will have to disregard ethics and social responsibility					
Q6	social responsibility and profitability can be compatible					
Q7	business ethics and social responsibility are critical to the survival of a business enterprise					
Q8	a firm's first priority should be employee morale					
Q9	business has a social responsibility beyond making profit					
Q10	if a survival of a business enterprise is at stake, then you must forget about ethics and social responsibility					
Q11	efficiency is much more important to a firm than whether or not the firm is seen as ethical or socially responsible					
Q12	good ethics is often good business					
Q13	if the stockholders are unhappy, nothing else matters					

* Note: 1—fully disagree; 2—disagree; 3—neutral; 4—agree; 5—fully agree.

Source: Own elaboration.

Out of 1000 companies, 220 managers responded to our survey, but there were 215 correctly filled questionnaires. Response rate based on valid questionnaires was 21.5%. For factor analysis some authors suggest that a minimum size should be the number of items in questionnaire times ten (Costello & Osborne, 2005; Field 2009). In case of PRESOR we have 13 items (statements), which means that we should have at least 130 valid responses. In our case 215 > 130, therefore, we can perform factor analysis.

Dataset was entered to SPSS and all steps as explained earlier were performed in order to isolate relevant factors for retail industry managers out of a generally proposed scale. In Table 3.7 data for Kaiser-Meyer-Olkin Measure and Bartlett's Test are shown. In Table 3.8 communalities for all statements are shown. Then variance analysis and rotation matrix are given for two factors (Tables 3.9 and 3.10) and three factors potential solution (Tables 3.11 and 3.12). At the end, there are questions for discussion and analysis.

Table 3.7. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.926
Bartlett's Test of Sphericity	approx. chi-square	1876.534
	df	78
	Sig.	0.000

Source: Own elaboration.

Table 3.8. Communalities

	Initial	Extraction
Q1	1.000	0.753
Q2	1.000	0.801
Q3	1.000	0.695
Q4	1.000	0.356
Q5	1.000	0.778
Q6	1.000	0.322
Q7	1.000	0.569
Q8	1.000	0.573
Q9	1.000	0.773
Q10	1.000	0.770
Q11	1.000	0.554
Q12	1.000	0.766
Q13	1.000	0.806

Source: Own elaboration.

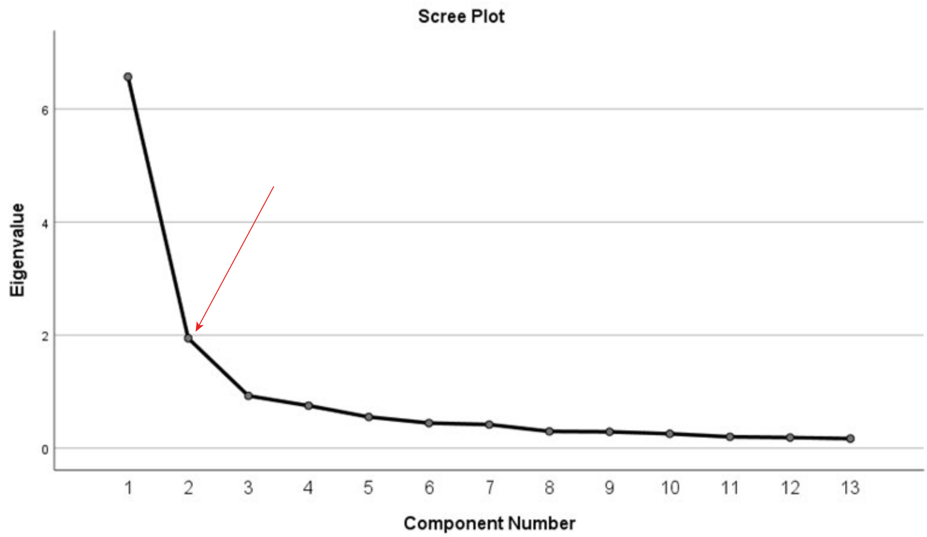


Figure 3.9. Scree plot—PRESOR scale

Source: Own elaboration.

Table 3.9. Total variance explained—solution with two factors

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
1	6.570	50.541	50.541	6.570	50.541	50.541	6.532	50.243	50.243
2	1.945	14.961	65.503	1.945	14.961	65.503	1.984	15.259	65.503
3	0.925	7.118	72.620						
4	0.751	5.777	78.398						
5	0.551	4.241	82.638						
6	0.444	3.414	86.053						
7	0.417	3.209	89.261						
8	0.297	2.282	91.543						
9	0.289	2.221	93.764						
10	0.254	1.952	95.715						
11	0.200	1.541	97.256						
12	0.187	1.439	98.696						
13	0.170	1.304	100.000						

Source: Own elaboration.

Table 3.10. Rotated Component Matrix[a] for two factors solution

	Component	
	1	2
Q13	0.898	
Q2	–0.886	
Q5	–0.877	
Q10	–0.876	
Q12	0.875	
Q9	0.874	
Q1	0.863	
Q8	0.756	
Q11	–0.741	
Q3		0.830
Q7		0.746
Q4		0.572
Q6		0.566

Extraction Method: Principal Component Analysis; Rotation Method: Varimax with Kaiser Normalization
a. 2 components extracted

Source: Own elaboration.

Table 3.11. Total variance explained—solution with three factors

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
1	6.570	50.541	50.541	6.570	50.541	50.541	6.524	50.186	50.186
2	1.945	14.961	65.503	1.945	14.961	65.503	1.805	13.885	64.071
3	0.925	7.118	72.620	0.925	7.118	72.620	1.111	8.550	72.620
4	0.751	5.777	78.398						
5	0.551	4.241	82.638						
6	0.444	3.414	86.053						
7	0.417	3.209	89.261						
8	0.297	2.282	91.543						
9	0.289	2.221	93.764						
10	0.254	1.952	95.715						
11	0.200	1.541	97.256						
12	0.187	1.439	98.696						
13	0.170	1.304	100.000						

Source: Own elaboration.

Table 3.12. Rotated Component Matrix[a] for three factors solution

	Component		
	1	2	3
Q13	0.898		
Q2	–0.886		
Q5	–0.877		
Q10	–0.876		
Q12	0.875		
Q9	0.874		
Q1	0.863		
Q8	0.756		
Q11	–0.741		
Q3		0.830	
Q7		0.746	
Q4		0.572	0.460
Q6		0.566	–0.764

Extraction Method: Principal Component Analysis
a. 3 components extracted

Source: Own elaboration.

3.5. Local government representatives about retailers—from the CSR perspective: case study no. 2

Retail stores have a significant impact on the lives of consumers—they provide products and services, educate consumers, generally they satisfy a wide range of consumer needs. They also are—citizens of local communities, who pay taxes, invest, participate or organize events, support local communities. It seems reasonable to present how local governments representatives perceive retail chains in the context of CSR and their role as members of local communities. According to corporate social responsibility concept, these entities should take responsibility for the effects of their activities—both positive and negative. However, the question occurs: whether their presence and activities are noticed and appreciated by local governments?

The purpose of the analysis was to establish the factors and their importance for the perception of retailers' social involvement.

A survey was conducted among the representatives of local governments with the use of an internet questionnaire. The choice of the research method was dictated by several reasons. First, the pilot study showed that respondents are reluctant to answer over the phone and ask for a list of questions. The territorial scope of the study eliminated the direct interview method—due to the high costs of reaching interlocutors. The research covered the entire population. As a result, the contact database was used and an electronic invitation to participate in the survey was sent to all local governments. A questionnaire was developed with a link provided in the cover letter. In total, 2,800 invitations to participate in the study were sent by e-mail. 431 questionnaires were collected, which gives a return of 15%. Out of them, 104 respondents did not complete the certificate. The Likert scale was used, where 1: fully disagree; 2: disagree; 3: neutral; 4: agree; 5: fully agree. The research tool was constructed based on the areas of retailers' CSR initiatives indicated in the literature. 18 items included into analysis. You will find more information about the research in Stefańska (2014b).

Table 3.13. Reliability statistics

Cronbach's alpha	Number of items
0.924	18

Source: Own elaboration.

Table 3.14. KMO and Bartlett's Test

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.933
Bartlett's Test of Sphericity	approx. chi-square	4282.622
	df	153
	Sig.	0.000

Source: Own elaboration.

Table 3.15. Communalities

	Initial	Extraction
Retailers act with integrity and in accordance with the principles of fair play	1.000	0.580
Retailers organize charity events for residents in need	1.000	0.606
Retailers sponsor people, organizations and entities by providing financial or material resources	1.000	0.610
Retailers identify problems in the local community and help solve them	1.000	0.740
Retailers educate residents about ecology (e.g., waste sorting)	1.000	0.665
Retailers teach consumers how to eat healthily	1.000	0.683
Retailers use their example to teach what it means to counteract discrimination	1.000	0.704
Retailers find talented people among residents and support their development	1.000	0.704
Retailers organize training for unemployed residents	1.000	0.650
Retailers treat residents as partners	1.000	0.538
Retailers are an important investor for us (they develop the area, build roads, pavements, parking lots, etc.).	1.000	0.635
Retailers are actively involved in environmental protection (e.g., they encourage residents to use eco-bags, set up bins for various types of rubbish, organize tree planting campaigns, etc.)	1.000	0.696
Retailers make a significant contribution to the city's budget	1.000	0.615
Retailers engage residents in joint actions to help those in need (charity campaigns)	1.000	0.682
Retailers organize events, the so-called events integrating residents to achieve goals important for the local community (e.g., competitions, shows)	1.000	0.737
Retailers equip shops with facilities for residents with disabilities to an extent greater than is required by law	1.000	0.562
Retailers operating in our town employ people discriminated against, e.g., because of age, gender, disability	1.000	0.574
Retailers focus solely on selling goods and services	1.000	0.530

Source: Own elaboration.

Table 3.16. Total variance explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
1	9.015	50.085	50.085	9.015	50.085	50.085	4.331	24.064	24.064
2	1.276	7.091	57.177	1.276	7.091	57.177	4.060	22.557	46.620
3	1.220	6.776	63.953	1.220	6.776	63.953	3.120	17.333	63.953
4	0.915	5.083	69.036						

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
5	0.761	4.228	73.264						
6	0.623	3.460	76.725						
7	0.558	3.100	79.825						
8	0.524	2.910	82.735						
9	0.510	2.836	85.571						
10	0.482	2.676	88.247						
11	0.366	2.031	90.278						
12	0.338	1.880	92.158						
13	0.317	1.758	93.917						
14	0.248	1.380	95.297						
15	0.243	1.347	96.645						
16	0.217	1.208	97.852						
17	0.210	1.165	99.017						
18	0.177	.983	100.000						

Source: Own elaboration.

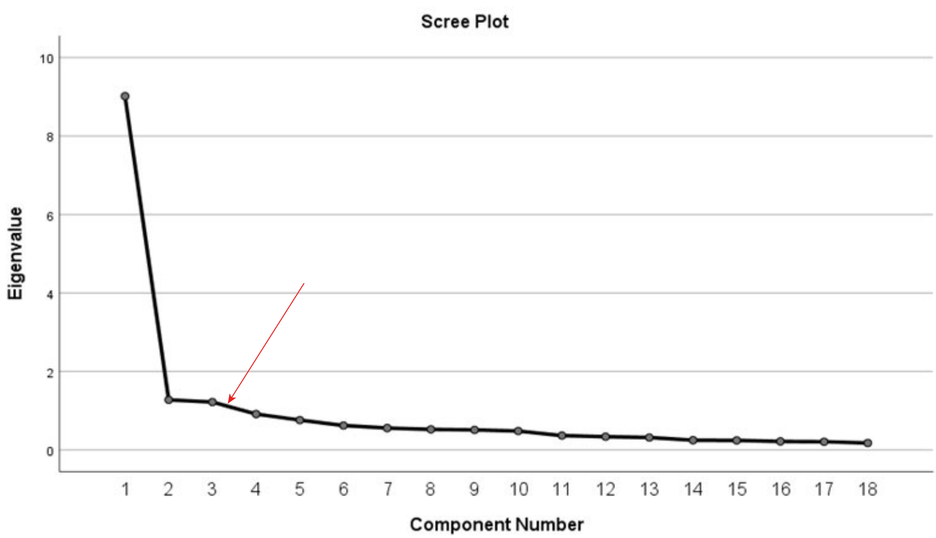


Figure 3.10. Scree plot

Source: Own elaboration.

Table 3.17. Rotated Component Matrix[a]

	Component		
	1	2	3
Retailers identify problems in the local community and help solve them	0.765		
Retailers sponsor people, organizations and entities by providing financial or material resources	0.735		
Retailers act with integrity and in accordance with the principles of fair play	0.734		
Retailers organize charity events for residents in need	0.643		
Retailers treat residents as partners	0.609		
Retailers use their example to teach what it means to counteract discrimination	0.589		0.545
Retailers educate residents about ecology (e.g., waste sorting)	0.568		0.538
Retailers teach consumers how to eat healthily	0.551		0.562
Retailers are an important investor for us (they develop the area, build roads, pavements, parking lots, etc.)		0.737	
Retailers make a significant contribution to the city's budget		0.734	
Retailers are actively involved in environmental protection (e.g., they encourage residents to use eco-bags, set up bins for various types of rubbish, organize tree planting campaigns, etc.)		0.715	
Retailers organize events, the so-called events integrating residents to achieve goals important for the local community (e.g., competitions, shows)		0.683	
Retailers equip shops with facilities for residents with disabilities to an extent greater than is required by law		0.638	
Retailers engage residents in joint actions to help those in need (charity campaigns)		0.636	
Retailers operating in our town employ people discriminated against, e.g., because of age, gender, disability		0.605	
Retailers organize training for unemployed residents			0.673
Retailers focus solely on selling goods and services			–0.713
Retailers find talented people among residents and support their development	0.508		0.601

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 6 iterations

Source: Own elaboration.

The extracted factors were used in *k*-means cluster analysis to identify segments of local governments—according to their opinion about retailer operating in their region. Two segments were identified.

Table 3.18. Final cluster centres

	Cluster	
	1	2
Component 1	0.74097	–0.52443
Component 2	–0.44762	0.31681
Component 3	–0.24051	0.17022

Source: Own elaboration.

Table 3.19. Number of cases in each cluster

Cluster	1	155.000
	2	219.000
Valid		374.000
Missing		57.000

Source: Own elaboration.

3.6. Testing attitude of Socially Responsible Employee: case study no. 3

The aim of the research was to identify the attitude of employees toward environment—both in place of living and in working place, and then to identify whether that fact influences the services at work. The research was conducted among store employees. The survey was quantitative and limited only to the workers directly involved in customer service. The selection of the sample was purposive. The respondents for the survey were selected among shop-assistants from shops selling FMCG, clothes and cosmetics. 272 people participated in the survey in a selected group of Polish cities. The analytical tool for the quantitative survey was based on the author's own detailed collective interviews with employees of shops. Likert scale was applied (from 1 to 5, where 1—means strongly disagree and 5—strongly agree). You will find more results from the research in Stefańska (2014a) and Stefańska (2018).

Table 3.20. Scale to measure attitude of employees toward environment: example

No	Items
a12_1	I take part in charity events organized outside my job
a12_2	I try to segregate rubbish at home (plastic, waste paper, glass, etc.)
a12_3	I try to save water and electricity at home
a12_4	I buy organic products to satisfy my own or my family's needs
a12_5	I help others for myself, not to please my employer
a12_6	I do not hesitate to submit suggestions for improving work so that customers are more satisfied
a12_7	I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money, etc.)
a12_8	I participate in CSR activities more out of fear of getting a job than out of my own free will
a12_9	I make other employees aware if they do something to the disadvantage of clients
a12_10	I would be more involved in helping those in need if it mattered to my boss
a12_11	I admonish someone who forgot to turn off the light or turn off the tap in the store
a12_12	In the case of a product that is put away anywhere, I take it to the right place by the customer

Source: Own elaboration.

Results of research

Table 3.21. Reliability statistics

Cronbach's alpha	*N* of items
0.797	12

Source: Own elaboration.

Table 3.22. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.776
Bartlett's Test of Sphericity	approx. chi-square	845.886
	df	66
	Sig.	0.000

Source: Own elaboration.

Table 3.23. Communalities

	Initial	Extraction
I take part in charity events organized outside of my job	1.000	0.360
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)	1.000	0.704
I try to save water and electricity at home	1.000	0.708
I buy organic products to satisfy my own or my family's needs	1.000	0.541
I help others for myself, not to please my employer	1.000	0.474
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	1.000	0.513
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)	1.000	0.600
I participate in CSR activities more out of fear of getting a job than out of my own free will	1.000	0.655
I make other employees aware if they do something to the disadvantage of clients	1.000	0.609
I would be more involved in helping those in need if it mattered to my boss	1.000	0.536
I admonish someone who forgot to turn off the light or turn off the tap in the store	1.000	0.634
In the case of a product that is put away anywhere, I take it to the right place by the customer	1.000	0.468

Source: Own elaboration.

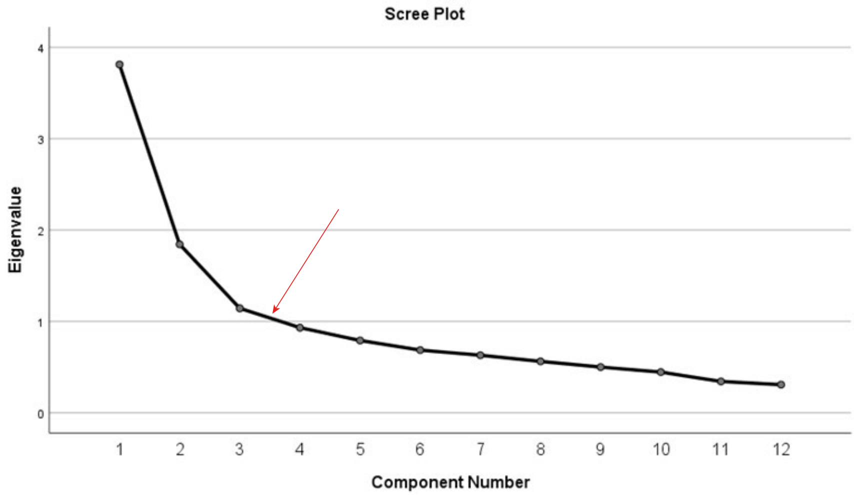


Figure 3.11. Scree plot

Source: Own elaboration.

Table 3.24. Total variance explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
1	3.813	31.774	31.774	3.813	31.774	31.774	2.363	19.690	19.690
2	1.844	15.367	47.141	1.844	15.367	47.141	2.357	19.641	39.331
3	1.144	9.531	56.672	1.144	9.531	56.672	2.081	17.342	56.672
4	0.932	7.763	64.436						
5	0.792	6.596	71.032						
6	0.686	5.718	76.750						
7	0.630	5.249	82.000						
8	0.563	4.689	86.689						
9	0.500	4.170	90.859						
10	0.446	3.715	94.574						
11	0.343	2.859	97.434						
12	0.308	2.566	100.000						

Source: Own elaboration.

Table 3.25. Rotated Component Matrix[a]

	Component		
	1	2	3
I take part in charity events organized outside my job			
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)		0.832	
I try to save water and electricity at home		0.833	
I buy organic products to satisfy my own or my family's needs		0.617	
I help others for myself, not to please my employer		0.540	
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	0.693		
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)			0.645
I participate in CSR activities more out of fear of getting a job than out of my own free will			0.802
I make other employees aware if they do something to the disadvantage of clients	0.714		
I would be more involved in helping those in need if it mattered to my boss			0.728
I admonish someone who forgot to turn off the light or turn off the tap in the store	0.752		
In the case of a product that is put away anywhere, I take it to the right place by the customer	0.592		

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 5 iterations

Source: Own elaboration.

After analysis of table 3.25 the decision was made to remove item a12_1., because the value of the loading was below 0.5 (*I take part in charity events organized outside my job*). New updated analysis was conducted.

Table 3.26. Reliability statistics

Cronbach's alpha	Number of items
0.782	11

Source: Own elaboration.

Table 3.27. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.761
Bartlett's Test of Sphericity	approx. chi-square	775.310
	df	55
	Sig.	0.000

Source: Own elaboration.

Table 3.28. Communalities

	Initial	Extraction
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)	1.000	0.720
I try to save water and electricity at home	1.000	0.727
I buy organic products to satisfy my own or my family's needs	1.000	0.586
I help others for myself, not to please my employer	1.000	0.469
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	1.000	0.529
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)	1.000	0.594
I participate in CSR activities more out of fear of getting a job than out of my own free will	1.000	0.674
I make other employees aware if they do something to the disadvantage of clients	1.000	0.611
I would be more involved in helping those in need if it mattered to my boss	1.000	0.552
I admonish someone who forgot to turn off the light or turn off the tap in the store	1.000	0.625
In the case of a product that is put away anywhere, I take it to the right place by the customer	1.000	0.449

Source: Own elaboration.

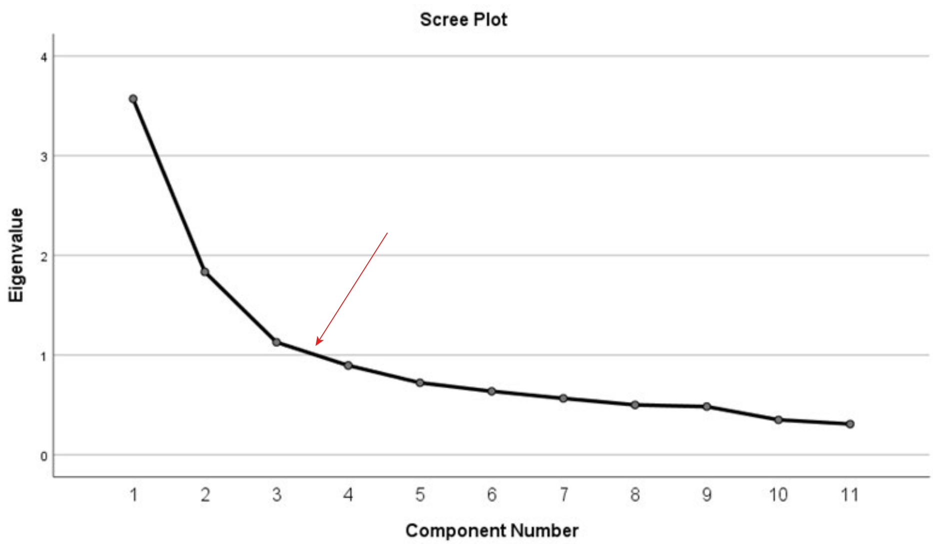


Figure 3.12. Scree plot

Source: Own elaboration.

Table 3.29. Rotated Component Matrix[a]

	Component		
	1	2	3
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)		0.840	
I try to save water and electricity at home		0.842	
I buy organic products to satisfy my own or my family's needs		0.639	
I help others for myself, not to please my employer			
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	0.714		
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)			0.626
I participate in CSR activities more out of fear of getting a job than out of my own free will			0.816
I make other employees aware if they do something to the disadvantage of clients	0.711		
I would be more involved in helping those in need if it mattered to my boss			0.737
I admonish someone who forgot to turn off the light or turn off the tap in the store	0.748		
In the case of a product that is put away anywhere, I take it to the right place by the customer	0.576		

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 5 iterations

Source: Own elaboration.

Again, one of the items loading was below 0.5, so the researcher decided to remove item a12_5. (*I help others for myself, not to please my employer*).

Table 3.30. Reliability statistics

Cronbach's alpha	Number of items
0.767	10

Source: Own elaboration.

Table 3.31. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.738
Bartlett's Test of Sphericity	approx. chi-square	684.353
	df	45
	Sig.	0.000

Source: Own elaboration.

Table 3.32. Communalities

	Initial	Extraction
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)	1.000	0.753
I try to save water and electricity at home	1.000	0.736
I buy organic products to satisfy my own or my family's needs	1.000	0.588
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	1.000	0.492
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)	1.000	0.596
I participate in CSR activities more out of fear of getting a job than out of my own free will	1.000	0.685
I make other employees aware if they do something to the disadvantage of clients	1.000	0.621
I would be more involved in helping those in need if it mattered to my boss	1.000	0.543
I admonish someone who forgot to turn off the light or turn off the tap in the store	1.000	0.667
In the case of a product that is put away anywhere, I take it to the right place by the customer	1.000	0.494

Source: Own elaboration.

Table 3.33. Total variance explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	total	% of variance	cumulative %	total	% of variance	cumulative %	total	% of variance	cumulative %
1	3.309	33.090	33.090	3.309	33.090	33.090	2.265	22.651	22.651
2	1.743	17.428	50.518	1.743	17.428	50.518	2.015	20.148	42.799
3	1.125	11.248	61.766	1.125	11.248	61.766	1.897	18.967	61.766
4	0.861	8.610	70.376						
5	0.672	6.718	77.095						
6	0.570	5.705	82.799						
7	0.566	5.658	88.457						
8	0.489	4.889	93.346						
9	0.352	3.515	96.861						
10	0.314	3.139	100.000						

Source: Own elaboration.

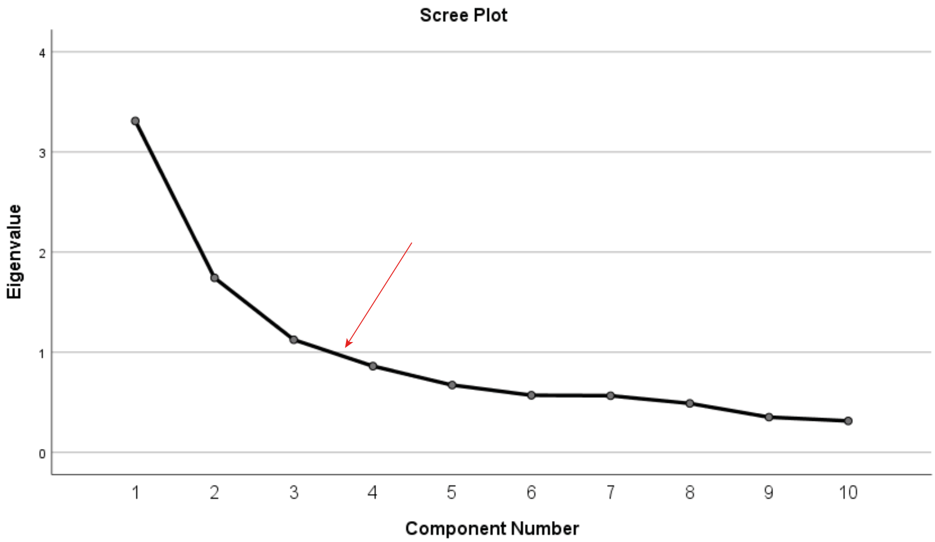


Figure 3.13. Scree plot

Source: Own elaboration.

Table 3.34. Rotated Component Matrix[a]

	Component		
	1	2	3
I try to segregate rubbish at home (plastic, waste paper, glass, etc.)		0.855	
I try to save water and electricity at home		0.845	
I buy organic products to satisfy my own or my family's needs		0.644	
I do not hesitate to submit suggestions for improving work so that customers are more satisfied	0.689		
I try to initiate social, charity or ecological activities in the store for a specific person or family (e.g., collection of clothes, food, caps, money)			0.620
I participate in CSR activities more out of fear of getting a job than out of my own free will			0.824
I make other employees aware if they do something to the disadvantage of clients	0.729		
I would be more involved in helping those in need if it mattered to my boss			0.728
I admonish someone who forgot to turn off the light or turn off the tap in the store	0.783		
In the case of a product that is put away anywhere, I take it to the right place by the customer	0.599		

Extraction Method: Principal Component Analysis
Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 4 iterations

Source: Own elaboration.

The extracted factors were used in regression linear model to identify the influence of those factors on services provided by employees working directly with customers (in accordance with diffusion model). The following results were obtained (Tables 3.35–3.38).

Dependable variable Y consists of 3 items transformed into index: *I advise clients on choosing the right products for them, even if it is not the most profitable for the store (chain of stores); I encourage my customers to buy ecological or fair-trade products, although they are sometimes more expensive; I recommend the goods to customers as if I were buying for myself.*

Independent variables are factors extracted in previous analysis: pro-ecological attitude (1), work engagement (2), extrinsic motivation (3).

Table 3.35. Model summary

Model	*R*	*R* square	Adjusted *R* square	Standard error of the estimate
1	0.583[a]	0.340	0.332	0.64010
2	0.583[b]	0.340	0.334	0.63883

a. Predictors: (Constant), REGR factor score 3 for analysis 1, REGR factor score 2 for analysis 1, REGR factor score 1 for analysis 1
b. Predictors: (Constant), REGR factor score 2 for analysis 1, REGR factor score 1 for analysis 1

Source: Own elaboration.

Table 3.36. ANOVA[a]

Model		Sum of squares	df	Mean square	*F*	Sig.
1	Regression	52.891	3	17.630	43.030	0.000[b]
	Residual	102.841	251	0.410		
	Total	155.732	254			
2	Regression	52.890	2	26.445	64.801	0.000[c]
	Residual	102.841	252	0.408		
	Total	155.732	254			

a. Dependent Variable: Y
b. Predictors: (Constant), REGR factor score 3 for analysis 1, REGR factor score 2 for analysis 1, REGR factor score 1 for analysis 1
c. Predictors: (Constant), REGR factor score 2 for analysis 1, REGR factor score 1 for analysis 1

Source: Own elaboration.

Table 3.37. Coefficients[a]

Model		Unstandardized coefficients		Standardized coefficients	t	Sig.
		B	Std. error	Beta		
1	(Constant)	3.823	0.040		95.369	0.000
	pro-ecological attitude	0.312	0.040	0.396	7.722	0.000
	work engagement	0.333	0.040	0.426	8.302	0.000
	extrinsic motivation	−0.001	0.040	−0.002	−0.033	0.974
2	(Constant)	3.823	0.040		95.558	0.000
	pro-ecological attitude	0.312	0.040	0.396	7.737	0.000
	work engagement	0.333	0.040	0.426	8.318	0.000

a. Dependent Variable: Y

Source: Own elaboration.

Table 3.38. Excluded variables[a]

Model		Beta in	*t*	Sig.	Partial correlation	Collinearity statistics
						tolerance
2	extrinsic motivation	−0.002[b]	−0.033	0.974	−0.002	1.000

a. Dependent Variable: Y
b. Predictors in the Model: (Constant), work engagement, pro-ecological attitude

Source: Own elaboration.

Questions / tasks

1. What is the main purpose of factor analysis?
2. What are the differences between factor analysis and principal-components analysis? Find in literature.
3. Based on Kaiser-Meyer-Olkin test (KMO), is our data suited for Factor analysis? Explain your answer.
4. What number sig. 0,000 in Bartlet's test tells us? Can we proceed with factor analysis?
5. Look at the table 3.8 where communalities are given. Interpret the table.
6. Examine shown tables regarding two and three factor possible solution. Which solution is more acceptable and why?
7. Once you have decided which solution is more acceptable, analyse which statements form which factor. Name factors based on components included.
8. Research a literature how to improve factor analysis by cutting off (or eliminating) questions with a problematic data in commonality table. Suggest which question(s) to eliminate in this example to get more robust and reliable results!

9. Given example was based on PRESOR scale. Investigate which statements are consisted in EPQ and ATBEQ methodology. Discuss how would you organize data collection to enable factor analysis according to those two scales. Can you anticipate what statements could form factors according those two scales?
10. Interpret Cronbach's alpha value, KMO and sphericity Bartlett test parameters.
11. Interpret scree plot.
12. What are your recommendations for items with loadings above 0.5 and doubled in two factors?
13. How would you name those factors to reflect their core meaning? Interpret them.
14. Research a literature how to improve factor analysis by cutting off (or eliminating) questions with a problematic data in commonality table. Suggest which question(s) to eliminate in this example to get more robust and reliable results!
15. Extracted components were applied in *k*-means cluster analysis. As a result, two clusters were identified. What can we tell about them?
16. How would you improve the analysis?
17. Interpret Cronbach's alpha value, KMO and sphericity Bartlett test parameters.
18. Interpret scree plot.
19. What are your recommendations for items with loadings above 0.5 and doubled in two factors?
20. How would you name those factors to reflect their core meaning? Interpret them.
21. What is your opinion about the decision to remove items—first 12_1 and then 12_5. Were they correct?
22. Extracted components were applied in regression linear model to identify whether they influence the quality of provided services by employees. As a result, two clusters were identified. What we can say about them? Analyse results step by step from table 3-35–3.38.
23. How would you improve the analysis?

References

Aczel, A. D. (2009). *Complete business statistics* (7th ed.). New York: McGraw-Hill.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, *10*(7), 1-9.

Davis, M. A., Andersen, M. G., & Curtis, M. B. (2001). Measuring ethical ideology in business ethics: A critical analysis of the ethics position questionnaire. *Journal of Business Ethics*, *32*(1), 35-53.

Field, A. (2009). *Discovering statistics using SPSS*. London: Sage.

Forsyth, D. R. (1980). A taxonomy of ethical ideologies. *Journal of Personality and Social Psychology*, *39*(1), 175-184.

Jolliffe, I. T. (1986). Principal component analysis and factor analysis. In I. T. Jolliffe (Ed.), *Principal component analysis. Springer Series in Statistics* (pp. 115-128) New York: Springer. https://doi.org/10.1007/978-1-4757-1904-8_7

KMO and Bartlett's Test. Retrieved March, 2021 from https://www.ibm.com/support/knowledgecenter/SSLVMB_23.0.0/spss/tutorials/fac_telco_kmo_01.html

Kurnoga, N., Knežević, B., & Šimurina, N. (2017) Multivariate analysis of attitudes on financial and other aspects of business ethics of future managers. *Croatian Operational Research Review*, *8*(1), 93-105.

McDaniel, C., & Gates, R. (2018). *Marketing research* (11th ed.). Hoboken: Wiley.

Preble, J. F., & Reichel, A. (1988). Attitudes towards business ethics of future managers in the US and Israel. *Journal of Business Ethics*, *12*, 941-949.

Promislo, M. D., Giacalone, R. A., & Welch, J. (2012). Consequences of concern: Ethics, social responsibility and well-being. *Business Ethics: A European Review*, *21*(2), 209-219.

Singhapakdi, A., Vitell, S. J., Rallapalli, K. C., & Kraft, K. L. (1996) The perceived role of ethics and social responsibility: A scale development. *Journal of Business Ethics*, *15*(11).

Stefańska, M. (2014a). *Why retailers should make the CSR concept a tool of building satisfaction in working place? Store personnel about CSR*. (The Proceedings of 10th International Conference Marketing Trends, pp. 1-9, Paris–Venice). Retrieved from http://archives.marketing-trends-congress.com/2014/pages/PDF/135.pdf

Stefańska, M. (2014b). *Rola społecznej odpowiedzialności w tworzeniu przewagi konkurencyjnej przedsiębiorstw handlu detalicznego*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Stefańska, M. (2018). *The relationship between the perception of the company and employee attitude in the context of CSR*. (The results of an empirical study research papers of Wrocław University of Economics No. 520, pp. 136-149). Retrieved from http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.desklight-0b50876c-64a7-4229-9493-ac597e07b0e4

Quoquab, F., Mohammad, J., & Sukari, N. N. (2019). A multiple-item scale form measuring "sustainable consumption behaviour" construct development and psychometric evaluation". *Asia Pacific Journal of Marketing and Logistics*, *31*(4).

What is the Kaiser-Meyer-Olkin (KMO) Test?. Retrieved March, 2021 from https://www.statisticshowto.com/ kaiser-meyer-olkin/

Qualitative and quantitative methods in sustainable development, pp. 117-178
https://doi.org/10.18559/978-83-8211-072-2/04

4.

STRUCTURAL EQUATION MODELLING IN SUSTAINABLE DEVELOPMENT RESEARCH

Todor Krastevich
D. A. Tsenov Academy of Economics, Svishtov

Atanaska Reshetkova
D. A. Tsenov Academy of Economics, Svishtov

Abstract: This chapter is dedicated to the structural equation modelling methods applied to solve sustainable development research problems. A structural equation model is an abstraction of reality, and the researcher's job is to build a model that approximates that reality as closely as possible. This task can be difficult if we do not have a clear understanding of what the reality of the studied phenomena is. Sometimes there is a sound theory behind the studied phenomena, and we can use variables that other researchers have already pointed out as valid indicators. In other situations, we have to start with a set of variables and test many hypothetical relationships based only on theoretical work. In this chapter, we focus on providing researchers with the knowledge needed to specify, evaluate, and interpret structural equation models (SEMs) in any field of social sciences, but most and foremost—in research related to the concept of sustainable development.

Keywords: CB-SEM, PLS-SEM, structural equation modelling, sustainable development.

4.1. What is Structural Equation Modelling (SEM)?

Structural equation models represent an a-priori formulated and theoretically and / or logically justified complex relationships between variables in a linear system of equations. These models serve to estimate the effects (as coefficients) between the considered variables, as well as the measurement errors. SEMs are advanced statistical procedures for testing measurement models, predictive, and causal hypotheses. These multivariate statistical tools are very useful to conduct basic or applied research in the behavioural and social sciences (Bagozzi & Yi, 2012, p. 8). The SEM analytical framework represents a generalization of both multiple regression and factor analysis and subsumes most linear modelling methods as special cases. SEM makes it easier to specify and test models that include latent variables, multiple indicators, measurement errors, and complex structural relationships such as reciprocal causation (Heck & Thomas, 2015, p. 13). The emergence and development of SEMs trace back to three different scientific fields: (1) path analysis, originally developed in genetics and later in sociology (2) simultaneous-equation models, as developed in economics, and (3) factor analysis from psychology (Rosseel, 2012, p. 2). The three traditions were ultimately merged and popularized at the application level in the early 1970s by Karl Jöreskog (1970). In recent decades, a number of software applications, such as commercial programs LISREL (Jöreskog, Olsson, & Wallentin, 2016), Amos (Arbuckle, 2019), EQS, (Bentler, 2006), Mplus (Muthén & Muthén, 2017), STATA, (StataCorp, 2017) XLSTAT, JMP, SAS PROC CALIS, as well as non-commercial open source packages `lavaan`, `sem`, `semPLS` and `OpenMx` in R environment, or the Python package `semopy`, have been developed, which further contributed to, and initiated a methodological revolution in the field of consumer research[1].

4.1.1. SEM in a nutshell: basic concepts

Many questions in the field of social sciences are concerned with investigating causal dependencies between certain variables. If causalities are checked and proved with a data set, this is generally referred to as a *causal* analysis (Backhaus, Erichson, & Weiber, 2015, p. 67). In the context of causal analysis, it is particularly important that the researcher makes intensive logical considerations about the relationships between the variables before using the statistical method. Based on a theoretically justified hypothesis system, SEM is used to check whether the theoretically established relationships match the empirically obtained data. SEM therefore has

[1] A comprehensive overview of modern SEM software packages is provided by Westland (2019, pp. 26, 44–45).

a confirmatory nature and belongs to the statistical methods for testing hypotheses. SEM is able to consider interactions between variables and can include both directly *observable* (manifest) and *non-directly observable* (latent) variables in the analysis. If all variables are manifest, path analysis is used, whereas analyses with latent variables are mostly referred to in the literature as causal analyses (Weiber & Mühlhaus, 2014, p. 36). The special feature of SEM in the context of causal analysis can be seen in the fact that it can be used to check interactions between latent variables.

Let us start from the hypothesis that consumer attitude towards a brand determine the consumer buying behaviour. If we note the attitude with letter ξ, and the buying behaviour with letter η, then the causal dependence underlying this hypothesis could be presented visually as:

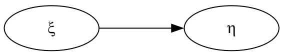


Figure 4.1. Structural model with two latent variables

Figure 4.1 graphically depicts the relationship between two unobservable (latent) variables, which are usually represented by ellipses and lowercase Greek letters. Assuming that the two variables are linearly related to each other, the hypothesis could be expressed mathematically as follows:

$$\eta = \gamma_0 + \gamma_1 \cdot \xi$$

Latent variables are referred to as hypothetical constructs, characterized by an abstract content. It is not possible to decide immediately whether the intended structural models are present in the reality or not. According to Bagozzi and Phillips, a hypothetical (theoretical) construct is an abstract entity which represents the 'true', unobservable state or nature of a phenomena. They achieve their meaning through formal connections to empirical concepts (Bagozzi & Phillips, 1982, p. 465).

In the context of SEM methodology, we can distinguish two types of latent variables: *exogenous* and *endogenous*. Exogenous latent variables are synonymous with independent variables because they "cause" fluctuations in the values of other latent variables in the model. They are considered to be influenced by directly observable variables that are external to the model. These variables serve as *indicators* of the underlying construct they represent. Endogenous latent variables are synonymous with dependent variables as they are influenced by the exogenous variables in the model, either directly or indirectly. Fluctuation in the values of endogenous variables is said to be explained by the model because all latent variables that influence them are included in the model specification (Byrne, 2016, p. 5). Endogenous latent variables are also considered to be influenced by other directly observable *indicator* variables. In Figure 4.1 consumers' attitude toward the brand is an exogenous latent variable and buying behaviour is an endogenous latent variable. Since we cannot directly measure the two latent variables in the model, it is necessary to operationalize them, i.e., to define an appropriate set of indicators that we need to measure for each construct. These indicators must depict the empirical representation of the unobservable, latent variables. SEM helps us verify a theoretically based hypothesis system that the relationships you have hypothesized among the latent variables and between the latent variables and the manifest indicators are indeed consistent with the empirical data at hand (Diamantopoulos & Siguaw, 2000, p. 4).

Structural equation models consist of two parts: a structural model and a measurement model. The **structural model** describes the relationships between the latent variables. Using our example, the structural model would explain whether there is a significant relationship between consumer attitudes towards the brand (exogenous latent variable ξ) and expected buying behaviour (endogenous latent variable η). The **measurement model** describes the set of manifest indicators that correspond to each latent variable. Let us assume that "Attitudes toward a brand" is described by three indicator variables: X_1 "This brand is eco-friendly", X_2 "This brand has a name you can trust", and X_3 "This is a high quality brand", while „Buying behaviour" is operationalized by two observable indicator variables: Y_1 "Number of items purchased" and Y_2 "Frequency of purchase". Based on these assumptions, the full structural equation mode can be represented as in Figure 4.2. Schematic representations of models are termed path diagrams because they provide a visual portrayal of relations that are assumed to hold among the variables under study. Essentially, a path diagram depicting a particular SEM model is actually the graphical equivalent of its mathematical representation whereby a set of equations relates dependent variables to their explanatory variables (Byrne, 2016, p. 10).

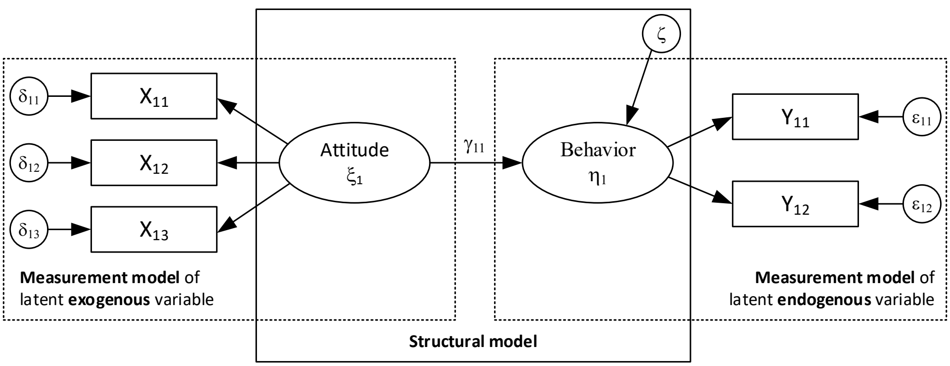


Figure 4.2. An example of full structural equation model

This model is termed "full" (or "complete") because it comprises both a measurement model and a structural model: the measurement model is depicting the links between the latent variables and their observed measures, and the structural model is depicting the links among the latent variables. A full model that specifies the direction of cause from one direction only is termed a recursive model; one that allows for reciprocal or feedback effects is termed a non-recursive model (Byrne, 2016, p. 7).

The notation of variables in Figure 4.2 has become largely unified in the literature.

Table 4.1 gives an overview of used abbreviations and their meanings. Statistical evaluation results in the so-called "path coefficients" that express the dependencies between variables in model. Typically, path coefficients are denoted with lowercase Greek letters, e.g.: β_{ij}, γ_{ij}, φ_{ij}, and ψ_{ij}.

Table 4.1. Variables in a complete structural equation model

Abbreviation	Meaning
η	Latent endogenous variable, which is explained in the model
ξ	Latent exogenous variable, which is not explained in the model
Y	Observable (measurable) indicator variable for a latent endogenous variable η
X	Observable (measurable) indicator variable for a latent exogenous variable ξ
ε	Disturbance (measurement error) for an indicator variable Y
δ	Disturbance (measurement error) for an indicator variable X
ζ	Disturbance for a latent endogenous variable η

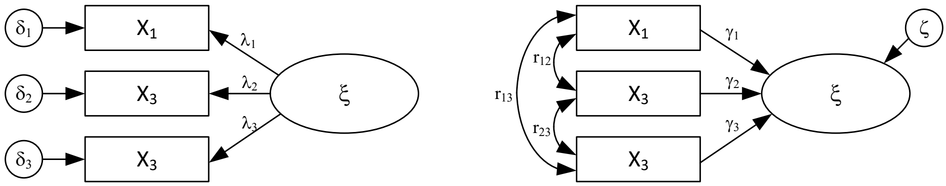
Source: (Weiber & Mühlhaus, 2014, p. 39).

It is important to distinguish between two basic types of *measurement* models—**reflective and formative models** (Edwards & Bagozzi, 2000, pp. 161–164). An indicator is reflective if it is caused by the construct expression and therefore the causal relationship goes from the construct to the indicator as depicted in Figure 4.3a, where each of the X_1 Xiindicator variables is influenced by the construct ξ_1. This model might be appropriate when a researcher wants to test a theoretical explanation of a latent construct. It is obvious that both measurement models in Figure 4.2 are reflective. The reflective measurement model in Figure 4.3a can be represented by a set of regression equations as follows (Diamantopoulos, 1999, p. 446):

$$X_1 = \lambda_1 + \xi \cdot \delta_1 \qquad X_2 = \lambda_2 + \xi \cdot \delta_2 \qquad X_3 = \lambda_3 + \xi \cdot \delta_3$$

where: λ_1 is the expected effect of ξ on X_1, and δ_1 is the measurement error for the i^{th} indicator ($i = 1, 2, 3$).

Reflective measurement is consistent with the confirmatory factor analysis (CFA) model as the variance in each indicator is as a linear function of the underlying latent variable. It is appropriate to use CFA when we have some knowledge of the underlying latent variable structure. Based on theoretical knowledge, empirical research, or both, the researcher postulates relations between the observed measures and the underlying factors a priori and then tests this hypothesized structure. CFA focuses solely on how the observed variables are linked to their underlying latent factors. More specifically, it is concerned with the extent to which the observed variables are generated by the underlying latent constructs and thus the strength of the regression paths from the factors to the observed variables (i.e., the factor loadings λ_i) is of primary interest. Although relations between latent factors are also of interest, any regression structure among them is not considered in the factor analytic model (Byrne, 2016, p. 7).



(a) Reflective measurement (b) Formative measurement

Figure 4.3. Reflective and formative measurement models

Source: (Diamantopoulos, 1999, p. 446).

However, in social-oriented studies we want to identify, e.g., what most important variables that ultimately influence the latent constructs are (e.g., attitudes toward a brand). In this case, we have a formative measurement model (see Figure 4.3b), because the indicators are assumed to be the cause for a latent variable to manifest. The causal relationship is going from the indicators to the constructs. In other words, under a formative perspective, a concept is assumed to be defined by, or to be a function of its measurements. The formal specification of the formative measurement model in Figure 4.3b is as follows (Diamantopoulos, 1999, p. 447):

$$\xi = y_1 X_1 + y_2 X_2 + y_3 X_3 + \zeta$$

where y_1 is the expected effect of X_1 on ξ and ζ is a disturbance term, with $cov(X_1, \zeta) = 0$ and $E(\zeta) = 0$.

In both formative and reflective measurement models, the measurement equation(s) follow a regression approach. The basic equation of a formative measurement model can be formulated as follows (Weiber & Mühlhaus, 2014, p. 257):

$$\boldsymbol{X_T} = X_0 + (X_S + X_R)$$

where $\boldsymbol{X_T}$ is the true construct value (not observable), X_0 is observed value (empirically measurable), X_S is systematic error, and X_R is random error.

In contrast, the following formulated relationship applies to the basic equation of reflective measurement models:

$$X_0 = \boldsymbol{X_T} + (X_S + X_R)$$

The decisive difference between the two measurement approaches is that in the reflective approach the measurement variable X_0 is the dependent variable and the construct is the independent variable. With the formative approach, this is exactly the opposite. One full structural equation model can contain only reflective, only formative or both reflective and formative indicators in its measurement models. It is obvious that in the reflective specification, the explanatory variable is latent, and the dependent variables are visible. In contrast, in the formative specification, the explanatory variables are manifest and the dependent variable is latent. This key difference between the two types of measurement models implies the use of different approaches to the statistical evaluation of parameters of SEMs (Jarvis, MacKenzie, & Podsakoff, 2003).

4.1.2. The model estimation

The relationships between latent and manifest variables cannot be exactly determined since the latent variables do not 'exist' in the same way as indicator variables do. It is even more difficult to determine causalities between two latent variables since both constructs can only be determined indirectly by their indicators. Measurement errors (also called residuals or disturbance) complicate the determination of the relationships. This is why all relationships between variables in SEMs can only be statistically estimated. There are two different approaches that can be used to estimate structural equation models: covariance-based structural equation modelling (CB-SEM) and variance-based partial least squares (PLS) path modelling, also referred to as PLS-SEM. Most notably, the first one estimates all model parameters by minimizing a global optimization criterion, whereas partial least squares path modelling does not involve such a global optimization procedure.[2] This seemingly small methodological difference leads to important practical implications (Hwang & Takane, 2014, p. xi). According to Hair, Ringle, and Sarstedt (2012, p. 312) „CB-SEM is a confirmatory approach that focuses on the model's theoretically established relationships and aims at minimizing the difference between the model implied covariance matrix and the sample covariance matrix. In contrast, PLS-SEM is a prediction-oriented variance-based approach that focuses on endogenous target

[2] In recent years another novel approach for structural equation modelling has also gained popularity. This approach, named *generalized structured component analysis*, was developed by Hwang and Takane (2004, 2014). and remains out of scope of this chapter.

constructs in the model and aims at maximizing their explained variance.“ There are some rules of thumbs that can be applied when deciding whether to use CB-SEM or PLS-SEM (Hair, Hult, Ringle, & Sarstedt, 2017, p. 18).

In general, if the purpose of the study is to verify and confirm some hypothetical theoretical concept (respectively theoretically defined psychological construction), the appropriate method is CB-SEM. However, if the aim of the study is to predict or develop an already known theory (containing some latent constructs in its description), the appropriate method is PLS-SEM. Statistically, PLS-SEM is similar to multiple regression analysis. The main goal of this approach is to maximize the explained part of the dispersion in the dependent (endogenous) constructs. PLS-SEM is less pretentious about the sample size and the requirement for normal statistical distribution of indicator variables. In other words, data requirements are less restrictive and, in general, should be applied in exploratory rather than confirmatory studies (Hair, Ringle, & Sarstedt, 2011, p. 140).

4.1.2.1. Model estimation using CB-SEM approach

Covariance-based approach, proposed by Jöreskog is the most often used approach to estimate and assess structural equation models (1973, 1970). The aim is to study the structure of the observed variables resulting from the variance-covariance matrix. In this way, it is possible to consider relationships between latent variables measured by manifest variables.

Since the variance-covariance matrix includes relationships between the variables, we can estimate the strength and direction of the links between variables in the model. However, there are some important assumption and features in CB-SEM. The latent variables are not defined by linear combination of manifest variables. They are true latent variables. Furthermore, the vector of error variables of measurement models is assumed to be pairwise uncorrelated and uncorrelated with all the latent variables (see Figure 4.2). This assumption turns CB-SEM approach into a “hard” model, in contrast to the “soft” PLS-SEM model (Schneeweiss, 1991, p. 152).

The most common parameter estimation algorithm for CB-SEM is the maximum likelihood method (ML). It is robust and consistent, and provides extensive and good quality measurements, which is why it is used most frequently. The disadvantage of this method is the assumed multivariate normal distribution of the data. It is problematic in this context that data normally distributed in marketing are rarely found (Jahn, 2007, p. 12). Another normal theory estimator is generalized least squares (GLS). When the assumption of multivariate normality is met, ML and GLS estimates are asymptotically equal. However, ML estimation has been shown to outperform GLS estimation under model misspecification conditions (Olsson, Foss, Troye, & Howell, 2000; Pituch & Stevens, 2016, p. 649).

The quality of estimation is an important topic of the covariance analysis. Different criteria and goodness-of-fit measures are applied for this purpose. Some of the more important ones will be discussed later. In addition to assessing the proposed model, it is also possible to predefine and compare different variants of the model.

4.1.2.2. Model estimation using PLS-SEM approach

Partial least square methods (sometimes referred to as component-based SEM or simply PLS-SEM) is an alternative statistical approach for casual modelling with latent variables. This method was first proposed by Wold (1975), but later Lohmöller (1989) contributed to the significant expansion and upgrading of algorithm for latent variable path modelling with partial least squares. PLS-SEM estimates model structures by combining principal components analysis with ordinary least squares regression and is typically viewed as an alternative that overcomes the very restrictive assumption of CB-SEM (Hair, Risher, Sarstedt, & Ringle, 2019, p. 4). It aims to maximize explained variance of dependent latent variables in SEM (i.e., maximizes the values). PLS optimizes locally, i.e., it maximizes the prediction of each dependent variable. It is therefore forecast-oriented (Jahn, 2007, pp. 14–15). In recent years, there has been a growing interest in the use of PLS-SEM in the field of empirical marketing and management studies (Hair, Hult et al., 2017, p. xiv). The increased recent interest is primarily due to the attention given to formative indicators in the construction of latent variables. An essential advantage of using this approach is that PLS makes no distribution assumptions due to the nature of the least squares estimate. This means that models can be estimated using PLS without the data having to be multivariate normal distributed. For example, if there is a large skewness of the sample, this does not affect the parameter estimation in PLS. Another key characteristics and advantages of PLS-SEM is that this method has no identification issue with small sample size and works fine with metric data, quasi-metric (ordinal) scaled data, and binary coded variables (for comparison, CB-SEM works only with metric data). This approach also easily incorporates reflective and formative measurement models (Tenenhaus, Vinzi, Chatelin, & Lauro, 2005, p. 165). At the same time, however, it has some limitations concerning model evaluation. The most important of them is that there is no established global goodness-of-fit criterion for model adequacy.[3]

[3] For more details see (Hair, Hult, et al., 2017, pp. 19–20).

4.1.2.3. Choosing the right approach

There are several fundamental differences between CB-SEM and PLS-SEM approaches (Scholderer & Balderjahn, 2005). The first one lies in the understanding of latent variables that underlie the respective approaches. Latent variables in the first approach follow the tradition of psychometric theory, i.e., that is, they satisfy the conditions of local stochastic independence. If there are one or more latent variables that cause the relationships between the observed variables, these relationships should disappear if the latent variables are kept constant. However, the latent variables in PLS-SEM do not have such properties. They follow the econometric theory tradition, which defines latent variables as unobserved, but does not exclude deterministic functions of the observed variables from the class of latent variables. The second fundamental difference between CB-SEM and PLS-SEM lies in the assumptions about the sample distribution of the variables. CB-SEM approach requires a multivariate normal distribution of the observed and latent variables, which in many cases is a serious limitation when working with empirical data. PLS-SEM, on the other hand, makes no assumptions regarding the distribution of the model variables, but precisely because of this, it cannot offer the inferential statistical possibilities of CB-SEM. Furthermore, it can be seen that the advantages of PLS arise precisely from the weak points of the maximum likelihood estimation. On the other hand, CB-SEM have some strengths over PLS-SEM. These mainly relate to the quality measures, but also the efficiency, consistency and robustness of the estimation results. The central weak point of PLS-SEM is that there are no criteria for assessing the overall model; there is still no convincing goodness-of-fit index for the joint assessment of the measurement and structural model in the PLS approach (Hulland, 1999, p. 202). Some authors recommended the application of PLS-SEM for sample size of $n < 100$ and less than four indicators per latent variable or if there are uncertain model hypotheses. Otherwise, they strongly recommend the use of CB-SEM approach due to the higher performance and application potential (Scholderer & Balderjahn, 2005). According to other views, they should not be seen as competitive alternatives, but rather as complementary approaches (Jöreskog & Wold, 1982, p. 270). Several years ago, Sarstedt and others published detailed results of simulation studies and derived on their basis valuable recommendations for choosing the appropriate approach for SEM assessment, in cases of reflective and formative models (Sarstedt, Hair, Ringle, Thiele, & Gudergan, 2016, p. 4007).

In conclusion it can be summarized that CB-SEM and PLS-SEM estimates of the same models can only be compared to a limited extent. A CB-PLS model always contains a number of additional restrictions compared to PLS, even if the path diagrams appear identical at first glance. However, we strongly suggest you have a look at Hair and others "rules of thumb" for selecting CB-SEM or

PLS-SEM (Hair, Ringle, & Sarstedt, 2011, p. 144), according to which CB-SEM and PLS-SEM results should be similar, but if CB-SEM requirements cannot be met (e.g., model specification, identification, nonconvergence, data distributional assumptions), use PLS-SEM as a good approximation of CB-SEM results. In general, PLS-SEM is the best approach, if you need to use latent variable scores in subsequent analyses. CM-PLS is recommended when the goal is theory/conformation, or comparison of alternative theories. PLS-SEM would also be recommended when we conduct exploratory research and if there are formative measurement models.

4.1.3. Identification issues and model adequacy

Not every theoretically justified structural equation model with latent variables is susceptible to statistical identification using the CB-SEM approach.[4] The problem with "identifiability" stems from the fact that in order to unambiguously solve a system of linear equations (through which each SEM is described mathematically), it is necessary that the number of equations contained in it be greater than the number of estimating model parameters. The number of equations always corresponds to the number of elements of the correlation matrix of the theoretical model. The number of elements of this correlation matrix is determined by the number of indicator variables that are subject to empirical observation.[5] The difference between the number of equations and the number of unknown parameters to be evaluated represents the degree of freedom of the model. A necessary condition for the identification of any SEM is that the degree of freedom of the model is greater than, or equal to zero (Backhaus et al., 2015, p. 86). Formally, if we denote the number of endogenous variables in a model by p, respectively the number of exogenous variables by q and the number of parameters to be evaluated by t, to achieve identifiability (i.e., to be able to statistically evaluate the parameters of SEM), the following condition must be valid:

$$t \leq \frac{1}{2}(p+q)\cdot(p+q+1)$$

However, the fulfilment of the above condition is not sufficient to conclusively prove the identifiability of the SEM. An additional criterion for checking the possibility for statistical estimation of the model parameters is the absence of a linear relationship between the equations that describe it. If there are linear interactions,

[4] PLS-SEM is not constrained by identification issues, even if the model becomes complex—a situation that typically restricts CB-SEM use (Hair, Sarstedt, Ringle, & Gudergan, 2017, p. 34).

[5] The number of elements (i.e., the correlation coefficients) of the correlation matrix is always equal to $n(n + 1)/2$, where n is the number of indicator variables included in the model.

the initial empirically observed correlation matrix must be positively defined (i.e., to be invertible). A necessary condition for this is that the number of estimated parameters of the model is greater than the number of indicator variables (Backhaus et al., 2015, p. 87). If the model is identifiable, the CB-SEM approach can be applied to evaluate its parameters. As the alternative PLS-SEM approach is not limited by the problem of identifiability, it can also be applied in cases of non-fulfilment of the described conditions for identifiability.

However, regardless of the approach used, the interpretation of the SEM should always begin by examining the logical validity of the sign and the magnitude of the estimated parameters. Typically, the dependencies between variables whose standardized estimates have an absolute value above 0.2 and ideally above 0.3 are usually interpreted and considered meaningful. In the presence of reflective indicators, it is recommended that the factor weights are above 0.6–0.7, ensuring that at least 50 % of the information in the indicator variable is explained by the relevant latent variable (Chin, 1998a, p. xiii). When working with formative indicators and the PLS-SEM approach, the estimated parameters no longer have the meaning of factor loadings, but their interpretation as the "strength" of the relationship is the same.

Of course, each parameter of an estimated SEM could also be evaluated in terms of its statistical significance. The basic rule here is that at a significance level below 0.05, the parameter is considered significant and can be interpreted. However, the interpretation of the statistical significance of coefficients estimated by PLS is problematic, as it is presumed that there is no explicit requirement for PLS to have a normal statistical distribution of the data of the indicator variables. In such a scenario, it is necessary to use bootstrap resampling to simulate a sufficiently large number of replicates (usually over 1,000) and to calculate the level of statistical significance p-value for each parameter of the model.

After this preliminary examination of the individual assessments of the parameters of the model, it is possible to proceed to an in-depth analysis of its quality using indices of goodness-of-fit.

So, what exactly are the indices of model fit? When we run a structural equation modelling, we are building or proposing a model that is based on some conceptualization, some past work by other researchers, which is called a review of literature. So, based on all this work, you conceptualize a model in which different variables are related to each other. Some variables may act like causes and other variables may act like effects. Then we also have the error term in covariances. Now, how do we know that your model is a good model, or it is an acceptable model? We assume that if the role played by all variables in our model is more or less the same as the actual role played by these variables in real life, we say that we are approximating to the reality. Our model captures the actual reality in a sufficient way. These indices of model fit tell us, to what extent our model is a good model and to what extent the model fits the sample behaviour. Perhaps, we are not only restricted to the sample behaviour but many indices which also tell us, what the chances of this model being replicated are, in the entire population. So, the indices of model fit are very important and very useful and without using indices of model fit, we cannot be very sure about whether our model is a good fitting model or not.

Different indices (assessment criteria) for the "quality" of the model as a whole are used. The quality of the assessed model can be assessed by two different perspectives—in terms of its reliability and its validity. Statistical reliability in this case means the absence of random errors in the evaluated model, i.e., in repeated empirical observations with other random samples, the estimates of the model will remain stable, i.e., their interpretation should not change. Where the measurement is reliable, it allows a summary relating to a wide variety of circumstances to be derived from one particular use of the model. One of the forms of reliability proof is to carry out repeated tests on the same respondents twice or more. The aim is to ensure that respondents' responses do not vary significantly over time, so that measurement leads to stable results. The second, more commonly used form to demonstrate reliability is the calculation of an internal consistency criterion between the indicator variables associated with a particular latent structure. The rationale for internal consistency is that individual indicators must measure the same construct and therefore correlate strongly with the latent variable being explained. The most popular and widely-used measure of internal consistency are the alpha coefficient (better known as Cronbach's alpha) and the omega coefficient, also known as composite reliability, whose values are recommended to exceed the threshold of 0.7 (Bacon, Sauer, & Young, 1995, p. 400).

Statistical validity means the circumstance in which the differences in the observed estimates reflect only the actual differences of the studied characteristic, which is the object of measurement (i.e., the model to really measure and reflect what needs to be measured and reflected). Validity therefore means conceptual correctness of measurement. In other words, SEM is considered reliable when it does not contain systematic errors and is respectively defined as valid when it does not contain random errors.

Reliability is a necessary but not a sufficient condition for validity (Peter, 1979, p. 6). The validity of any SEM could be considered content, convergent, discriminant and nomological (Homburg & Giering, 1996, S. 7). Content validity is determined by the degree to which the variables in a given measurement model belong to the semantic area of the studied construct, and the constructed elements fully reflect all content aspects of the construct. Convergent validity is explained by the extent to which there are strong associative (correlation) relationships between the indicator variables describing a latent variable in SEM. Discriminant validity means that the associative relationships between indicator variables attributed to different latent variables should generally be weaker than those that measure the same latent variable. Ensuring nomological validity requires that the construction of a construct be integrated and empirically justified in a higher-level theoretical framework. To ensure content and nomological validity, careful selection and arrangement of the indicator variables is needed, as well as precise definition and interpretation of working hypotheses in the model.

However, proving convergent and discriminatory validity in SEM can be performed using two types of statistical criteria—local and global goodness-of-fit evaluation statistics.

4.1.3.1. Local criteria for model evaluation

When evaluating the reliability and validity of reflective measuring models with local goodness-of-fit statistical criteria, the aim is to check the adequacy of the measuring models. Here, on the one hand, the indicators are evaluated, and on the other hand—their relationship with the latent variables.

It is usually started by assessing the structural reliability by examining the factor loadings, the values of which must exceed 0.707 to make sure that more than half of the variance in the indicator variables is related to the latent construct.

The omega index is then checked for composite reliability, the values of which must be above 0.7 (Fornell & Larcker, 1981, p. 45). It is also possible to use the AVE indicator (AVE = average variance extracted). Fornel and Larcker claim, that „if AVE is less than 0,5, the variance due to measurement error is larger than the variance captured by the construct, and the validity of the individual indicators, as well as the construct, is questionable" (Fornell & Larcker, 1981, p. 46). The construct reliability and the average variance are suitable as test variables for the convergence validity of the indicators assigned to a factor. In order to complete the reliability and validity considerations, the discriminant validity is assessed using the Fornell / Larcker criterion (Fornell & Larcker, 1981, p. 46). It says that the average variance of a factor must be greater than any squared correlation between it and another construct. In addition, the coefficient of determination R^2, describing the share of the variance of an endogenous construct explained by the relationships in the model could always be calculated. Here the recommended minimum threshold is around 0,3 (Drengner, Gaus, & Jahn, 2008, p. 143). In summary, the evaluation of reflective measurement models using local statistical criteria should be in line with the following recommended scheme in Table 4.2.

Table 4.2. Recommended thresholds for assessment of reflective measurement models with local fit evaluation criteria

Construct / Indicators	Convergent validity			Discriminant validity	
	factor loadings	composite reliability	AVE[a]	Fornell / Larcker	R^2
(Requirement)	(≥0.707)	(≥0.7)	(≥0.5)	(AVE > $Corr^2$)[b]	(>0.3)
Construct 1		…	…	… > …	…
Indicator 1–1	…				
Indicator 1–2	…				
Construct 2		…	…	… > …	…
Indicator 2–1	…				
Indicator 2–2	…				

[a] AVE = average variance extracted
[b] $Corr^2$ = highest squared correlation between the model constructs

Source: Based on (Drengner et al., 2008, p. 143).

When working with formative measurement models there is no need to assess internal composite reliability as well as convergent (and therefore discriminant) validity. For this reason, only the estimates of the path coefficients should be statistically significant (Jahn, 2007, S. 23). Since the evaluation of formative measurement models uses multiple regression, one of the main requirements is the absence of multicollinearity between the independent variables. Essentially, if estimated factor loadings, composite reliability, AVE, and Fornell / Larcker criterion meet the requirements for reflective measurement models (see Table 4.2), a reliable and valid measurement could be assumed. Afterwards it is possible to check the structural model. If the assumption that the model withstands the empirical examination is supported, then the hypotheses can be checked. The investigation of whether the data collected contradicts the relationships expressed by the model or not is conducted with the help of the global quality measures. These measures are called by different authors by different names, including "fit indices" (Marsh, Balla, & Hau, 1996, p. 315), "goodness-of-fit indices" (Jöreskog et al., 2016, p. 500) or simply "fit Statistics" (Bollen & Long, 1992, pp. 1–9).

4.1.3.2. Global criteria for model evaluation

Due to the different logic and statistical nature of CB-SEM and PLS-SEM assessment, different fit indices are applied to assess the adequacy of the defined models. As arguments for the 'best' global fit criterion cannot be argued, the following will outline some of the most commonly used ones.

CB-SEM evaluation

Two types of fit indices are used as global criteria for assessing the adequacy of CB-SEM—***absolute*** and ***incremental***. *Absolute* fit indices assess how well an a priori model reproduces the sample data. In order to test this agreement statistically, a chi-square test can be performed. The perfect representation of reality through the model mean that the null hypothesis is likely to be rejected as soon as the number of samples is large enough. In order to neutralize the sample size effect on the test result, it is also recommended to compute a relative chi-square (χ^2/df). A ratio of approximately 3 or less is considered 'beginning to be reasonable' (Arbuckle, 2019, p. 641). Furthermore, the probability p is calculated that the rejection of the null hypothesis would represent a wrong decision. Some authors recommend rejecting the model if p is less than 0,1 (Weiber & Mühlhaus, 2014, p. 204).

The Goodness of Fit Index (**GFI**) and the Adjusted Goodness of Fit Index (**AGFI**), which considers the degrees of freedom, are also commonly used. The possible range of GFI values is 0 to 1, with higher values indicating better fit (Hair, Black, Babin, & Anderson, 2019, p. 637). These indicators should be used with caution as they are sensitive to sample size as well as the size of the model (Anderson & Gerbing, 1984, p. 172).

Root Mean Square Residual (**RMR**), which corresponds to the standard error in the regression analysis, can also be used as a measure of the covariance that is not explained on average in a model. There is also a standardized variant of this index (**SRMR**), that is useful for comparing fit across models. Lower RMR and SRMR values represent better fit and higher values represent worse fits. Hair et al. claim that an SRMR over 0,1 suggests a problem with the fit (Hair, Black et al., 2019, p. 638).

Browne and Cudeck (Browne & Cudeck, 1992, pp. 238–239) recommend another absolute index for global assessing the adequacy of CB-SEM—Root Mean Squared Error of Approximation (**RMSEA**). According to them, a value of RMSEA of about 0,05 or less would indicate a "close" fit of the model. A value greater than 0,1 indicates an unacceptable approximation. In addition, some software programs (such as AMOS) also calculate the probability value for testing the null hypothesis that the population RMSEA is no greater than 0,05 (Arbuckle, 2019, p. 645).

Incremental indices differ from absolute ones in the fact that they are resulted from the comparison of the estimated model with a null model that is more restricted than the target model because its variables must not correlate with each other. The most commonly used incremental index is Normed Fit Index (**NFI**). It ranges between 0 and 1, and a model with perfect fit would produce an NFI of 1. An alternative is Tucker Lewis Index (**TLI**), that is conceptually similar, but considers model complexity to some extent. Other popular indices are Comparative Fit Index (**CFI**) and Relative Non-centrality Index (**RNI**). Like the other incremental fit indices, higher values represent better fit, and the possible values generally range between 0 and 1 (Hair, Black et al., 2019, p. 638).

In general, the simultaneous use of the mentioned indices is recommended for the reliable assessment of CB-SEM, and the cut-off values given in

Table 4.3 can be used as a guide.

Table 4.3. Cut-off values for fit indices for global evaluation CB-SEM

RMSEA	RMR	SMRM	CFI/RNI/NFI/ TLI	χ^2/df	AGFI
< 0,08	< 0.05 (but not > 0.1)	< 0.05	> 0.90	< 3	> 0.90

Source: (Jahn, 2007, p. 27).

The fit indices discussed are also suitable for comparing different variants of a CB-SEM model. However, these alternative models should be similar in their complexity to ensure comparability.

PLS-SEM evaluation

The PLS approach has no reasonable global criterion for assessing the model quality, so that it cannot be assessed comprehensively. Some quasi-global quality measures applicable to PLS-SEM estimates are somewhat similar to those of a linear regression. For example, to assess the explanatory power of a PLS model, the coefficient of determination R^2 can be used for each latent endogenous variable. The interpretation is identical to that of tradition regression—describe proportion of the variance explained. According to Wynne Chin, values for R^2 of 0,19 are interpreted as "weak", of 0,33 as "moderate", and of 0.66 as "substantial" (Chin, 1998b, p. 323).

Suitable criterion is the so-called "effect size" f^2. Specifically, the effect size f^2 checks whether a particular exogenous latent variable has substantive impact on an endogenous variable by using or omitting independent variables in the structural equation, respectively. High values indicate that the exclusion of the corresponding exogenous variables causes a significant drop in R^2, which in turn means a high relevance in explaining the endogenous variables. Effect size of 0,02, can be viewed as small, 0.15 as medium, and 0,35 as large effect at the structural level (Chin, 1998b, p. 317).

An important criterion for assessing global predictive relevance of PLS models is Stone-Geisser Q_n^2 criterion (Chin, 1998b, p. 317). The Stone-Geisser Q_n^2 criterion evaluates the predictive relevance, i.e., how well the dependent (endogenous) is described on their independent (exogenous) variables in the structural model. If the value of this criterion is above zero, the model is predictive significant. A value of zero means that the model does not predict the original data better than an average estimate. Values less than zero speak against the quality prediction of the model structure. It should be noted that this measure can only be used sensibly for reflective measurement models (Herrmann, Huber, & Kressmann, 2006, p. 58; Weiber & Mühlhaus, 2014, p. 329).

Finally, the robustness of the results of a PLS-SEM estimate can be assessed using the bootstrapping method, in which different samples are used to estimate the PLS model (Chin, 1998b, p. 320). If the parameter estimators vary widely across the different samples, this speaks against the robustness of the estimation results. Usually, the first two moments of the distribution of the individual estimators, the mean value and the standard deviation over the samples are primarily considered.

However, in order to check the quality or "appropriateness" of a PLS solution, it is recommended to consider all available individual criteria for assessing the measurement models and the structural model in a kind of "synopsis".

Table 4.4 shows the criteria discussed above and their recommended cut-off levels.

If individual measurement models have deficits, it makes sense to perform modifications in order to achieve statistically significant results in all elements of the causal model, at least for partial structures or exploratory modifications (Ringle, 2004, p. 23).

Table 4.4. Recommended values for global evaluation criteria of PLS-SEM

Coefficient of determination (R^2)	Effect size (f^2)	Stone-Geisser criterion (Q^2_η)	Path coefficients (after bootstrapping)
~ 0.66 (substantial)	~ 0.35 (large)	> 0	> 0.2
~ 0.33 (moderate)	~ 0.15 (medium)		
~ 0.19 (weak)	~ 0.02 (small)		

Source: Based on (Chin 1998b, pp. 317–323; Ringle, 2004, p. 22; Jahn, 2007, p. 28).

If the local and global evaluation result is an acceptable mode (i.e., reliable and valid), the hypotheses can be tested. If the relationship between two constructs is significantly different from zero and runs in the positive direction (positive or negative relationship), the hypothesis expressed by them can be confirmed or rejected. Furthermore, the consideration of total effects helps to better understand the interdependencies in a complex model. A total effect is the total influence of one variable on another across all conceivable relationships with other constructs (Jahn, 2007, S. 30).

4.2. Comparing the performance of SEM approaches with simulated data

Following the research framework for evaluating customer satisfaction and loyalty proposed by Tenenhaus and others (2005) and Fornel, Johnson, Anderson, Cha, and Bryant (1996), we simulated data from a hypothetical survey (n = 799)[6]. The data consisted of 22 items (indicators) of corporate image, customer expectations, per-

[6] Similar models with artificially simulated other data are publicly available and found in Addinsoft (2020) and Ringle, Wende and Becker (2015).

ceived product and services quality, perceived value of product, customer satisfaction, customer loyalty and complaints. The proposed structural model for the associations among the hypothetical reflective constructs is illustrated in Figure 4.4. Obviously, the model contains six exogenous (ξ_j) and one endogenous (η) latent variables.

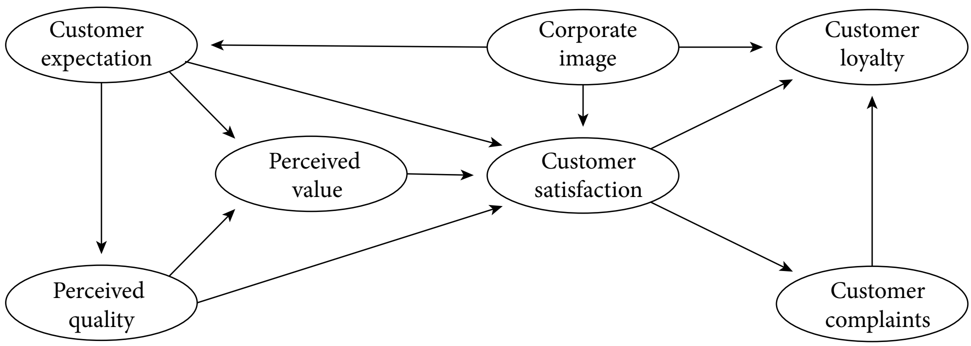


Figure 4.4. A structural model of customer satisfaction index

Source: (Tenenhaus et al., 2005, p. 161).

For clarity, on the path diagram, individual items for each factor are omitted. However, the observable indicators and variable names of latent factors are shown in Table 4.5 (Tenenhaus et al., 2005, p. 162; O'Loughlin & Coenders, 2004, p. 1236).

Table 4.5. Conceivable names of latent variables and their indicators

Latent variables (LV) {name / description}	Manifest variables (MV) name		Description (Likert-type statements) {All the items are scaled from 1 to 7. Scale 1 expresses a very negative, while scale 7 a very positive opinion}
CUEX (ξ_1) Customer expectation of the overall quality	Cuex1	(x_1)	Expectations for the overall quality of product
	Cuex2	(x_2)	Expectations for product to meet your personal need
	Cuex3	(x_3)	Expectation that things could go wrong at your product
PERQ (ξ_2) Perceived quality	Perq1	(x_4)	Please rate the overall quality of the product
	Perq2	(x_5)	Please rate the technical product features
	Perq3	(x_6)	Please rate the customer service and personal advice offered
	Perq4	(x_7)	Please rate the reliability and accuracy of the product
	Perq5	(x_8)	Please rate the clarity of information provided
PERV (ξ_3) Perceived value	Perv1	(x_9)	Please rate the quality of the product given the prices you pay
	Perv2	(x_{10})	Please rate the prices of product given the quality
IMAG (ξ_4) Corporate image	Imag1	(x_{11})	The product provider is a reliable and trustworthy company
	Imag2	(x_{12})	The product provider is a customer-centric company
	Imag3	(x_{13})	The product provider is innovative and forward looking
	Imag4	(x_{14})	The product provider has a social contribution for the society

CUSA (ξ_5) Customer satisfaction	Cusa1	(x_{15})	Overall, how satisfied are you with the product?
	Cusa2	(x_{16})	How close is this product to your ideal product?
	Cusa3	(x_{17})	Considering your expectations, to what extent has the product fallen short of, or exceeded your expectations?
CUSCO (ξ_6) Customer complaints	Cusco1	(x_{18})	How many times have you complained (either formally or informally) to sales or support personnel?
	Cusco2	(x_{19})	To what extent do you think that your product provider will / would care about your complaint?
CUSL (η_1) Customer loyalty	Cusl1	(y_1)	If you needed to choose a new product how likely it is that you would choose the same provider again?
	Cusl2	(y_2)	Let us now suppose that other providers decide to lower their prices, but your provider stays at the same level as today. At which level of difference (in %) would you choose another provider? [requires transformation into a seven-point scale]
	Cusl3	(y_3)	If a friend or colleague asked you for advice, how likely is it that you would recommend your product provider?

Source: Own work.

Based on the structural model presented in Figure 4.4, it is possible to formulate several hypotheses:

H1: *Customer expectation and perceived quality have positive impact on perceived value.*

H2: *Corporate image, customer expectation, perceived quality and perceived value have positive effects on customer satisfaction.*

H3: *Customer satisfaction has negative effect on customer complains*

H4: *Customer satisfaction and corporate image have positive effects on customer loyalty*

H5: *Customer complaints have negative effects on customer loyalty.*

If we collect customer survey data, we wish to ascertain whether our proposed model of hypothetic influence is an adequate model for the data. Since our task is related to explore performance and implementation issues of the two approaches considered for SEM, we created simulated data set[7] using the factor loadings similar to those reported by Tenenhaus and others (2005, p. 180), Liu, Ren, and Liu (2013, p. 780) and Askariazad and Babakhani (2015).

Following a five-step Bolen's procedure (Bollen & Long, 1992, p. 123), namely (a) model specification, (b) identification, (c) estimation, (d) testing fit, and (e) re-specification, we first try to identify the model already defined on the basis of past research by Tenenhaus and others (2005, p. 161) and illustrated in Figure 4.4. We use a covariance-based SEM and try to fit the model with three popular software packages—a free open source package 'lavaan' (Rosseel, 2012), IBM SPSS Amos (Arbuckle, 2019), as well as alternatively with the very robust procedure for SEM, developed by

[7] The data set is available for download here https://bit.ly/3mxzo8W.

STATA (StataCorp, 2017). We do this to find out whether the choice of software tool would affect the reliability of the results. Later, we will try to evaluate the instrumental validity of outputs, comparing the results obtained vie CB-SEM and PLS-SEM.

In order for a model to be identifiable, we need to compare the number of data points to the number of parameters to be estimated. Since the input data set is the variance/covariance matrix, the number of data points is equal to the number of elements of this matrix. Our simulated data set has 799 observations and 22 variables (corresponding to the manifest variables in Table 4.5). Therefore, the number of data point is $\{22 \cdot (22 + 1) / 2 = 253\}$. Because the number of parameters $\{0{,}5 \cdot (1 + 6) \cdot (1+ 6 + 1) = 28\}$ to be estimated is less than the number of data points, the model is "over identified", and the analysis can proceed (Byrne, 2016, p. 41).

4.2.1. CB-SEM approach

The subsequent estimation of the parameters of the model within the CB approach is possible with different numerical methods. In this case, we use the maximum likelihood (ML) function, to minimize the difference between the sample covariance and those predicted by the theoretical model. The method was chosen because we assumed and generated simulations data set with multivariate normal statistical distribution.[8]

4.2.1.1. Fit a model to data using 'lavaan' package in R/RStudio

To assess the model represented in Figure 4, its specification is needed first. After loading `lavaan` package in R/RStudio (Rosseel, 2012), each latent factor is described by its name on the left, followed by the 'is manifested by' symbol "=~" with the latent variables that it influences and its observed manifest variables. In our case, we call the model "ECSI" and use the following program code to build it:

```
ECSI <–"
 # Measurement model
 CUEX    =~ Cuex1 + Cuex2 + Cuex3
 PERQ    =~ Perq1 + Perq2 + Perq3 + Perq4 + Perq5
 PERV    =~ Perv1 + Perv2
 IMAG    =~ Imag1 + Imag2 + Imag3 + Imag4
 CUSA    =~ Cusa1 + Cusa2 + Cusa3
 CUSCO   =~ Cusco1 + Cusco2
 CUSL    =~ Cusl1 + Cusl2 + Cusl3
 # Structural model (defining latent variables)
 CUEX    ~ IMAG
```

[8] The data quality checks here are omitted, but it is plausible to inspect those with some descriptive statistics and tests for normality.

```
  PERQ     ~ CUEX
  PERV     ~ CUEX + PERQ
  CUSA     ~ CUEX + PERV + PERQ + IMAG
  CUSCO    ~ CUSA
  CUSL     ~ CUSA + CUSCO + IMAG “
```

To fit the model, we use the `sem()` command with the following syntax:

```
ECSI.fit <-sem(ECSI, data=ECSISimData, std.lv=TRUE)
```

We add an argument `std.lv=TRUE` to standardize the latent variables, which help us to compare relative influence strength. Below there is a part of the abbreviated output.

```
> summary(ECSI.fit, standardized=TRUE, fit.measures=TRUE )
lavaan 0.6-6 ended normally after 32 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of free parameters                         56

  Number of observations                           799

 Model Test User Model:

  Test statistic                               230.917
  Degrees of freedom                               197
  P-value (Chi-square)                           0.049

… …

 User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    0.992
  Tucker-Lewis Index (TLI)                       0.990

… …

 Root Mean Square Error of Approximation:

  RMSEA 0.015
  90 Percent confidence interval-lower           0.001
  90 Percent confidence interval-upper           0.022
  P-value RMSEA <= 0.05                          1.000

 Standardized Root Mean Square Residual:

  SRMR                                           0.025
```

We begin with the interpretation of chi-square test statistic, which in our case is equal to 230.917. It measures the difference between the sample covariance (correlation) matrix and the fitted covariance (correlation) matrix. According to Jöreskog and others (2016, p. 499), chi-square should be used as goodness-of-fit measure rather than a statistical test. To be used as a test statistic, all observed variables must have a multivariate normal distribution. A small chi-square corresponds to good fit and a large chi-square to bad fit. Chi-square tends to be large in large samples, because it is calculated as N times the minimum value of the fit function, where N is the sample size (number of observations). If chi-square test statistic is not significant ($p > 0.05$), we accepted null hypothesis, i.e., H0: There is no significant difference between sample covariance matrix and population covariance matrix. Hence the default model is almost on the verge of acceptable.

Because chi-square test statistics is very often significant in samples of large size, suggesting rejection of the proposed model, the relative chi-square could also be calculated by dividing the chi-squared test statistic by degrees of freedom of the model (Wheaton, Muthen, Alwin, & Summers, 1977, p. 99). Accepted range of values are between 1 and 3 (Carmines & McIver, 1983, p. 64). However, some researchers have recommended relative χ^2/df value between 2 to 5 indicating a reasonable fit (Marsh & Hocevar, 1985, p. 567). In our case, the value of relative chi-square is 1.17 (= 230.917/197), which indicates an acceptable fit between the hypothetical model and the sample data.

Next in the output we see the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI). Both incremental indices have values higher than 0.95, which indicate a strong model fit. This conclusion is also confirmed by the low residuals (RMSEA = 0.015 and SRMR = 0.025, where 0 means perfect fit).

The resulting structural coefficients for the proposed model can be plotted using `semPlot` package (Epskamp et al., 2019) with the following code:

```
semPaths(ECSI.fit , what="est", fade=FALSE , residuals=FALSE, rotation = 2,
 structural=FALSE , nCharNodes =6, edge.label.cex =0.6,
 sizeMan = 5, sizeLat = 6)
```

Now we could draw a few conclusions. Because the model shows good fit to the data, we are able to interpret the results. We can use the coefficient estimates to answer questions about the association of the latent factors with the outcomes of interest. However, before proceeding, it is necessary to check whether the coefficient estimates are statistically significant. The function summary(ECSI.fit) also displays this result (only the essential part of the result is shown below).

```
… …
                Estimate    Std.Err    z-value    P(>|z|)
 CUEX ~
```

```
  IMAG                 0.488      0.064      7.635     0.000
  PERQ ~
  CUEX                 0.547      0.059      9.261     0.000
  PERV ~
  CUEX                 0.647      0.102      6.363     0.000
  PERQ                -0.008      0.072     -0.108     0.914
  CUSA ~
  CUEX                 0.028      0.092      0.306     0.759
  PERV                 0.273      0.074      3.696     0.000
  PERQ                 0.485      0.063      7.714     0.000
  IMAG                 0.073      0.066      1.111     0.267
  CUSCO ~
  CUSA                -0.618      0.063     -9.803     0.000
  CUSL ~
  CUSA                 0.676      0.078      8.641     0.000
  CUSCO               -0.177      0.069     -2.550     0.011
  IMAG                 0.156      0.063      2.481     0.013
… …
```

It is obvious that `perceived quality` (PERQ) does not have a statistically significant effect on `perceived value` (PERV). The same conclusion can be drawn for the latent variables `customer expectation of the overall quality` (CUEX) and `corporate image` (IMAG) and their effect on `customer satisfaction` (CUSA). All other coefficient estimates (not shown) are statistically significant and can be interpreted.

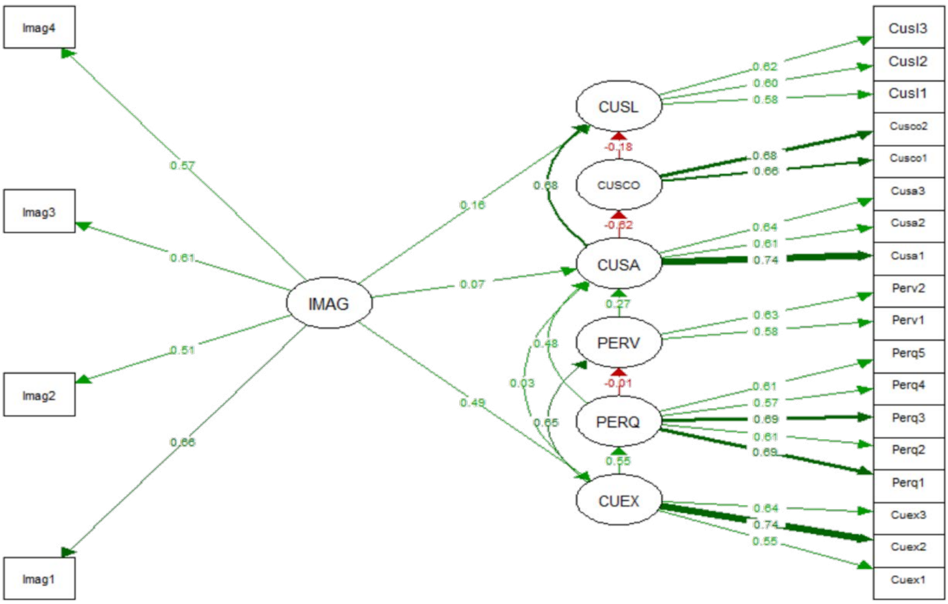


Figure 4.5. Path diagram with coefficient estimates for the ECSI model using `lavaan`

Source: Own work.

However, it should be noted that a good fit in CB-SEM is not enough for reliable interpretation and valid conclusions. It is recommended to compare the proposed model to one or more plausible alternative models, in order to prove that our proposal is superior to other reasonable models (Chapman & Feit, 2019, p. 283). The specification of an alternative model depends on the research objectives and / or the theory it is based on. When defining it, one can also start from the so-called "weak" and / or statistically insignificant relationships observed in the assessment of the baseline model. It is also possible to make a comparison with an already existing model from the literature, the result of past research.

In the present example, we could "clear" the basic model of "weak" relationships, such as the dependence of `Customer satisfaction` on` Corporate image` and `Customer expectation`, as well as the dependence of `Perceived value` on `Perceived quality`. This alternative model ECSIalt can be described as follows:

```
# Specification of ECSCalt model----
ECSIalt <-"
 # Measurement model
 CUEX     =~ Cuex1   + Cuex2    + Cuex3
 PERQ     =~ Perq1   + Perq2    + Perq3   + Perq4   + Perq5
 PERV     =~ Perv1   + Perv2
 IMAG     =~ Imag1   + Imag2    + Imag3   + Imag4
 CUSA     =~ Cusa1   + Cusa2    + Cusa3
 CUSCO    =~ Cusco1  + Cusco2
 CUSL     =~ Cusl1   + Cusl2    + Cusl3
 # Structural mode
 CUEX     ~ IMAG
 PERQ     ~ CUEX
 PERV     ~ CUEX
 CUSA     ~ PERV + PERQ
 CUSCO    ~ CUSA
 CUSL     ~ CUSA + CUSCO + IMAG "
```

After fitting and plotting (Figure 4.6) the alternative model to the initial data, it is necessary to compare the obtained results with those of the basic model. Using the function `compareFit()` (available after installing and loading the `semTools` package) we can directly compare the performance of the two models:

```
# Fit the ECSIalt model with CB-SEM
ECSIalt.fit <-sem(ECSIalt, data=ECSISimData, std.lv=TRUE)

# Creating a path diagram of ECSIalt
semPaths(ECSIalt.fit , what="est", fade=FALSE , residuals=FALSE, rota-
tion = 2,
      structural=FALSE , nCharNodes =6, edge.label.cex =0.6,
      sizeMan = 5, sizeLat = 6)
```

```
# Compare the proposed model with the alternative model&
> summary(compareFit(ECSI.fit , ECSIalt.fit , nested=TRUE))

################### Nested Model Comparison #####################
Chi-Squared Difference Test

              Df   AIC    BIC   Chisq   diff Df diff Pr(>Chisq)
ECSI.fit     197  51358  51620  230.92
ECSIalt.fit  200  51354  51602  232.94       2.027     3     0.5668

####################### Model Fit Indices #######################
 chisq df pvalue cfi tli aic bic rmsea srmr
ECSI.fit   230.917† 197 .049  .992  .990  51357.820  51620.089  .015
.025†
ECSIalt.fit 232.944 200 .055 .992† .991† 51353.847† 51602.066† .014†
.026

################## Differences in Fit Indices ##################
                     df   cfi   tli    aic     bic rmsea    srmr
ECSIalt.fit–ECSI.fit   3     0     0 -3.973 -18.023     0   0.001
```

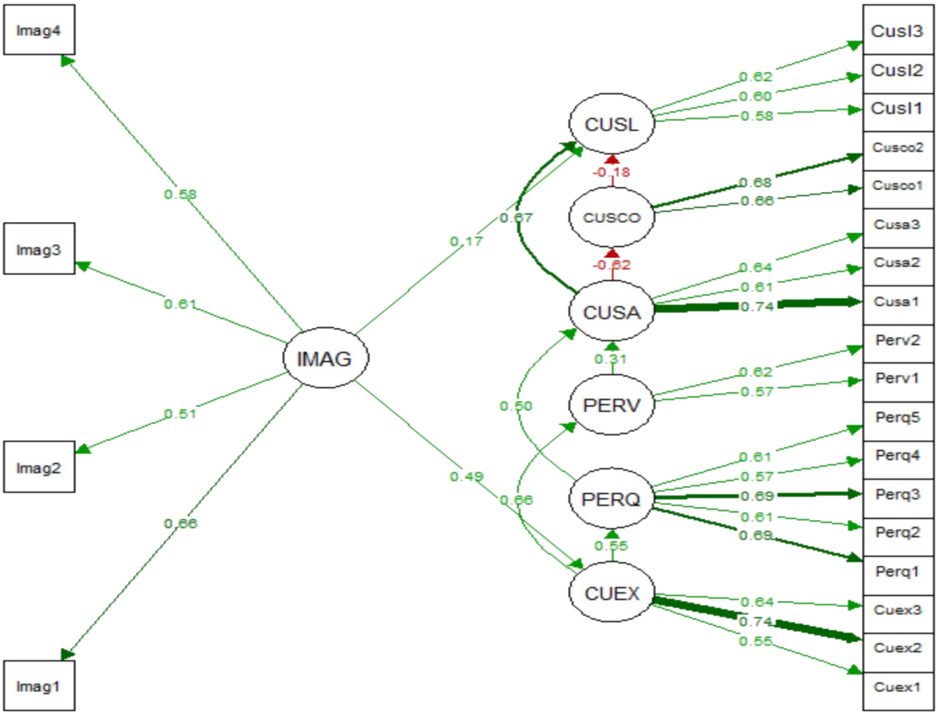


Figure 4.6. Path diagram with coefficient estimates for the ECSIalt model using `lavaan`

Source: Own work.

From these results it can be concluded that the alternative model is slightly better than the basic model. This can be seen from the direct comparison of test statistics, incremental indices and information criteria AIC and BIC (lower is better). However, this superiority is not statistically significant because chi-square difference between the two models (2.027) is not statistically significant ($p = 0.5668$). In these circumstances, it is reasonable to interpret the alternative model.

4.2.1.2. Fit a model to data using IBM SPSS AMOS

Probably one of the most popular software programs for evaluating CB-SEM is IBM SPSS AMOS. AMOS is distributed commercially, and its use is associated with significant costs. A trial version could be obtained from https://ibm.co/35VjjEe and a free downloadable user guide is also available. In this part we will demonstrate briefly how to use this package to evaluate the proposed linear structure model depicted in Figure 4.4. We use the same data set as in the previous demonstration.

With the AMOS program it is very easy to build a model using simple graphical tools. This is undoubtedly an advantage, especially when the researcher is feeling uncomfortable with programming, and is very useful when one is taking first steps in SEM. Typical graphic elements of path diagrams are used to draw the model: rectangles for observed variables; ellipses for unobserved variables; single-headed arrows for causal relationships; double-headed arrows for covariance; circle for error terms. A data file has to be selected, then AMOS will perform all necessary computations for evaluating the model and display an output with the results. However, other approaches to specify the desired model are also available in AMOS, but they involve describing it with equation statements. Choosing "AMOS graphic" option when starting the program will let the researcher draw the theoretical model graphically and there will be no need to express the relationships with manually written equation statements. We will not describe in detail how to build the model using the graphic interface of AMOS as it is well described in the user manual provided by the developer. Instead, we will focus on the evaluation results and compare it with the ones obtained with `lavaan` package.

When the model is specified and the data file is selected, the calculation of the estimates can be easily done with one click. AMOS will provide us with both graphical and textual view on the results. The graphical output shows the estimates next to each arrow on the model path diagram, and fit indexes are displayed under the path diagram (see Figure 4.7). This way of reviewing the results is comprehensive and allows the researcher to take a quick glance at the estimates and the evaluation criteria.

It is evident from the fitness indexes under the path diagram that the model fits the data. The χ^2 test and its degrees of freedom and p-value indicate that (since we can reject the null hypothesis) our predicted model matches the data. Alternatively, we can rely on the relative $\chi^2 = 1.171$, which also confirms the good model fit. Amos provides the benchmark values of all reported fitness indexes in brackets, so it is easier to compare each measure to its threshold. In this case GFI, CFI, and RMSEA all indicate that the model fits the data. However, detailed results are presented in the textual output. It includes several sections which are shown in the upper left corner in Figure 4.8.

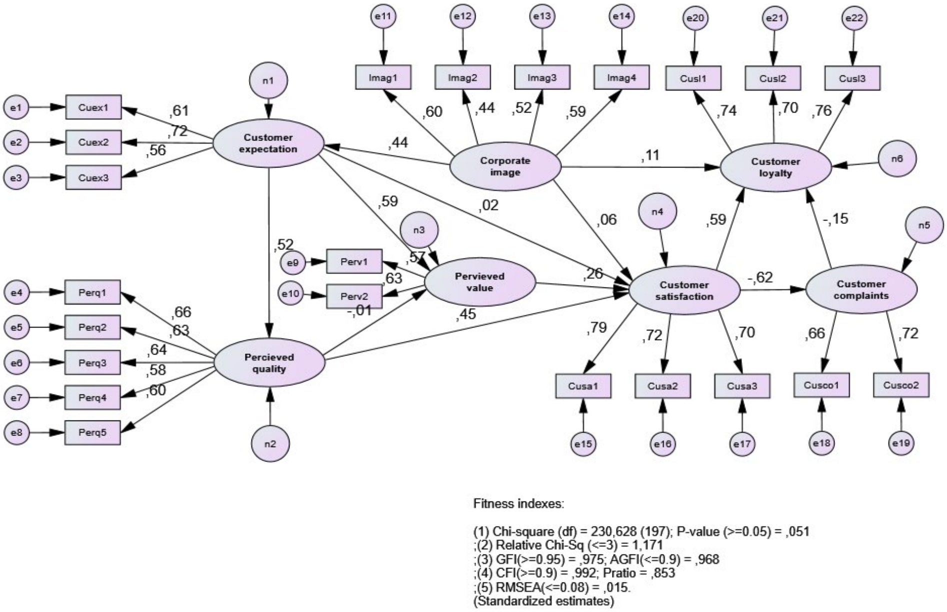
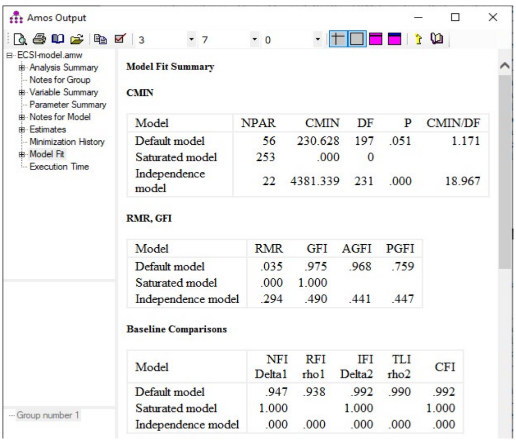


Figure 4.7. Path diagram with coefficient estimates for the ECSI model using IBM SPSS Amos

Source: Own work.



Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	56	230.628	197	.051	1.171
Saturated model	253	.000	0		
Independence model	22	4381.339	231	.000	18.967

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.035	.975	.968	.759
Saturated model	.000	1.000		
Independence model	.294	.490	.441	.447

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.947	.938	.992	.990	.992
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Figure 4.8. Textual output from ECSI model evaluation with Amos

Source: Own work.

Since the χ^2 test is sensitive to the sample size we cannot rely solely on this test to make a conclusion about the model fit. We need to look at the goodness-of-fit statistics, reported in `Model fit` section, and decide whether it is correctly specified and thus—fits the data well. Amos reports plenty of goodness-of-fit measures and sometimes they may lead to different conclusions. It is important to choose to rely on those measures that are shown to be appropriate in situations similar to the one at hand in terms of sample size, applied estimation procedure, model complexity, and the presence/absence of multivariate normality and variable independence (Byrne, 2016, p. 101). In this example, we will use the same criteria that were reported in the `lavaan` package output and explained in part "1.3.2. Global criteria for model evaluation" of this chapter. In the output we can see that both CFI and TLI have values higher than 0.95 (CFI = 0.992 and TLI = 0.990), which indicates a good model fit. This conclusion is also confirmed by the low residuals: RMSEA = 0.015 and SRMR = 0.026, where 0 means perfect fit. We can proceed the assessment by observing the model estimates.

Selecting the 'Estimates' section will show all parameters estimates and their *p*-values (Table 4.6). These estimates are almost identical with those we obtained with the `lavaan` package. The relationships between CUSA and both CUEX and IMAG, and between PERV and PERQ are not significant, and we should remove them from the initial model and re-evaluate an alternative model, as we did with `lavaan` package. However, all regression weights of each measurement model are statistically significant and there is no need to remove any indicator variables.

Table 4.6. Regression weights of the initial ECSI model

			Estimate	S.E.	C.R.	P
CUEX	<---	IMAG	0.473	0.069	6.865	***
PERQ	<---	CUEX	0.519	0.062	8.416	***
PERV	<---	CUEX	0.638	0.093	6.850	***
PERV	<---	PERQ	−0.008	0.075	–.108	**0.914**
CUSA	<---	PERV	0.276	0.079	3.510	***
CUSA	<---	CUEX	0.028	0.092	.306	**0.760**
CUSA	<---	PERQ	0.508	0.068	7.523	***
CUSA	<---	IMAG	0.071	0.064	1.109	**0.267**
CUSCO	<---	CUSA	−0.662	0.059	−11.261	***
CUSL	<---	CUSA	0.615	0.067	9.227	***
CUSL	<---	CUSCO	−0.150	0.059	−2.543	0.011
CUSL	<---	IMAG	0.137	0.056	2.459	0.014
Cuex3	<---	CUEX	1.000			
Cuex2	<---	CUEX	1.154	.095	12.090	***
Cuex1	<---	CUEX	.859	.075	11.418	***

			Estimate	S.E.	C.R.	P
Perq5	<---	PERQ	1.000			
Perq4	<---	PERQ	0.944	0.077	12.305	***
Perq3	<---	PERQ	1.142	0.087	13.172	***
Perq2	<---	PERQ	1.011	0.078	13.016	***
Perq1	<---	PERQ	1.130	0.084	13.476	***
Perv2	<---	PERV	1.000			
Perv1	<---	PERV	0.918	0.115	8.014	***
Imag1	<---	IMAG	1.000			
Imag2	<---	IMAG	0.779	0.092	8.505	***
Imag3	<---	IMAG	0.928	0.099	9.400	***
Imag4	<---	IMAG	0.870	0.088	9.913	***
Cusa3	<---	CUSA	1.000			
Cusa2	<---	CUSA	0.955	0.054	17.537	***
Cusa1	<---	CUSA	1.159	0.062	18.619	***
Cusl1	<---	CUSL	1.000			
Cusl2	<---	CUSL	1.033	0.060	17.119	***
Cusl3	<---	CUSL	1.076	0.059	18.099	***
Cusco2	<---	CUSCO	1.000			
Cusco1	<---	CUSCO	0.970	0.087	11.107	***

Source: Own work.

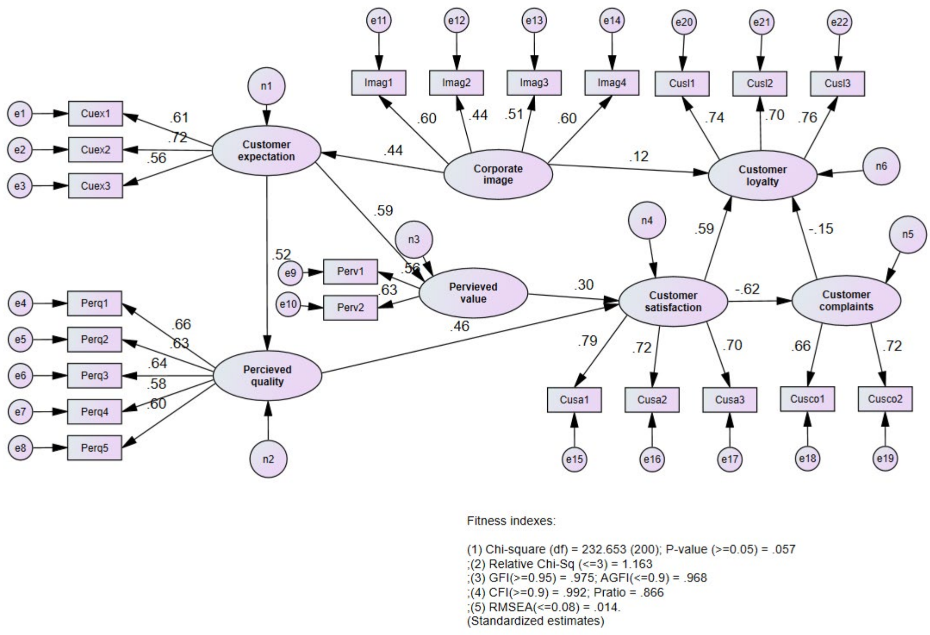


Figure 4.9. Path diagram with coefficient estimates for the ECSIalt model using IBM SPSS Amos

Source: Own work.

The evaluated alternative model is almost indistinguishable from that obtained with the `lavaan` package and we can make the same conclusions about the adequacy of its fit to the data.

4.2.1.3. Comparing and interpreting the results

Comparing the values of the model fit measures that resulted from using the two different software solutions confirms the good model fit (see Table 4.7). The obtained parameter estimates for the ESCIalt model are all statistically significant. Once again, no indicator variables need to be removed from the measurement model, as all regression weights are also significant.

Table 4.7. Fit indices of the ECSIalt model obtained from `lavaan` and IBM SPSS Amos

	Model fit indexes	
	`lavaan`	IBM SPSS Amos
χ^2 (df)	232.944 (200)	232.653 (200)
χ^2/ ratio	1.165	1.163
	0.992	0.992
	0.991	0.991
	0.014	0.014
	0.026	0.026*

* Amos does not produce this statistic as one of its regular fit indices. It can be calculated by selecting `Plugins—> Standardized RMR` from the drop-down menus, then clicking on the `Calculate Estimates` button again.

Source: Own work.

We can also compare the standardized parameter estimates of the structural model, which are presented in Table 4.8. Amos reports these values in the `Estimates` section of the output. In order to obtain standardized estimates with `lavaan`, you can use `parameterEstimates()` function and add `standardized=TRUE` as an argument:

```
# Obtaining the standardized parameter estimates for ECSIalt model
Stand.par <-parameterEstimates(ECSIalt.fit, standardized = TRUE)
```

The standardized estimates are located in the column "std.all" in the output. It is obvious that the two software solutions produce identical estimates.

Table 4.8. Standardized estimates of the ECSIalt model obtained from `lavaan` and IBM SPSS Amos

	Standardized Weights	
	`lavaan`	*IBM SPSS Amos*
CUEX <--—IMAG	0.442	0.442
PERQ <--—CUEX	0.522	0.522
PERV <--—CUEX	0.590	0.590
CUSA <--—PERV	0.297	0.463
CUSA <--—PERQ	0.463	0.297
CUSCO <--—CUSA	–0.619	–0.619
CUSL <--—CUSA	0.590	0.590
CUSL <--—CUSCO	–0.155	–0.155
CUSL <--—IMAG	0.115	0.115

Source: Own work.

We would use the coefficient estimates in the alternative model to answer substantive questions about the associations of the latent factors with the outcomes of interest and to prove defined five hypotheses. In the general case, all statistically significant path coefficients whose estimates are greater than 0.5 can be interpreted as strong impact, between 0.3 and 0.5 as moderate, and below 0.3 as weak impact. The sign in front of the respective coefficient on the other hand shows the direction of this influence (positive or negative). The summary of the predefined hypotheses in this example is presented in Table 4.9.

Table 4.9. Summarized results of the defined hypotheses testing

Hypothesis		Fulfilment of hypothesis
H1:	Customer expectation and perceived quality have positive impact on perceived value	partially supported due to lack of statistical significance of PERQ
H2:	Corporate image, customer expectation, perceived quality and perceived value have positive effects on customer satisfaction	partially supported due to lack of statistical significance of CUEX and IMAG
H3:	Customer satisfaction has negative effect on customer complains	supported
H4:	Customer satisfaction and corporate image have positive effects on customer loyalty	supported
H5:	Customer complaints have negative effects on customer loyalty	supported

Source: Own work.

4.2.2. PLS-SEM approach

In the previous section, we considered the possibility for estimating a linear structural model based on a covariance-based approach. We note that in this approach the aim is to consider as much of the total covariance in the data as possible, among all observed and latent variables. CB-SEM requires relatively strict assumptions about data sets, such as continuous and multivariate normally distributed data, normally distributed residuals, relatively large sample size, generally three or more indicators per latent construct, reliability of indicators (Hair, Hult et al., 2017, p. 3). Although CB-SEM is a powerful analytical tool that test a model rigorously and allows for model comparison (Chapman & Feit, 2019, p. 285), all of these requirements at the same time are relatively rare in empirical marketing research. Therefore, in practical cases where the structural model is complex (many constructs and many indicators), the sample size is small, and the research goal is to predict key target constructs or identify key "driver" constructs, the use of the PLS-SEM approach is recommended (Hair, Risher et al., 2019).

However, we would like to emphasize once again that PLS-SEM approach in general does not outperform CB-SEM, as it does not allow to assess a global "goodness of fit" that is comparable across models. We recommend its use in cases where CB-SEM fails when estimating the model.

In the next section we will demonstrate three alternative software solutions for linear structural modelling based on PLS-SEM approach.[9] To ensure comparability, we will use the same data set and try to evaluate the same model presented in Figure 4.4.

4.2.2.1. Fit a model to data using `semPLS` package in R/RStudio

The R package `semPLS` is probably the most popular free open-source software for estimating complex structural equation models (Monecke & Leisch, 2012; Monecke, 2015; Ravand & Baghaei, 2016). After the package has been installed in R/RStudio, the following program code can be used to fit the presented ECSI model to the same data set ECSISimData. What is special about compiling the program code is that, unlike the procedure in the `lavaan` package, here the specification of the complete model requires two separate steps. The first step consists in defining the measurement model, which describes the relationships between the latent variables and their observable indicator variables.

The second step specifies the structural model that describes the relationships between the latent variables.

[9] There are many other open-source and commercial software statistical package for fitting PLS-SEM. Venturini and Mehmetoglu (2019, pp. 3–4) give a concise overview of the most popular of them.

Referring to the model in Figure 4.4, the measurement model could be defined as follows:

```
# Step ONE: Defining a measurement model
ECSIPLSmm <—matrix(c(
 “CUEX”, “Cuex1”,
 “CUEX”, “Cuex2”,
 “CUEX”, “Cuex3”,
 “PERQ”, “Perq1”,
 “PERQ”, “Perq2”,
 “PERQ”, “Perq3”,
 “PERQ”, “Perq4”,
 “PERQ”, “Perq5”,
 “PERV”, “Perv1”,
 “PERV”, “Perv2”,
 “IMAG”, “Imag1”,
 “IMAG”, “Imag2”,
 “IMAG”, “Imag3”,
 “IMAG”, “Imag4”,
 “CUSA”, “Cusa1”,
 “CUSA”, “Cusa2”,
 “CUSA”, “Cusa3”,
 “CUSCO”,”Cusco1”,
 “CUSCO”,”Cusco2”,
 “CUSL”, “Cusl1”,
 “CUSL”, “Cusl2”,
 “CUSL”, “Cusl3” ), ncol=2, byrow=TRUE)
```

The definition of the structural model from Figure 4.4 is performed in a similar matrix format:

```
# Step TWO: Defining a structural model
ECSIPLSsm <—matrix(c(
 “CUEX”, “PERQ”,
 “CUEX”, “PERV”,
 “CUEX”, “CUSA”,
 “PERQ”, “PERV”,
 “PERQ”, “CUSA”,
 “PERV”, “CUSA”,
 “CUSA”, “CUSL”,
 “CUSA”, “CUSCO”,
 “CUSCO”,”CUSL”,
 “IMAG”, “CUEX”,
 “IMAG”, “CUSA”,
 “IMAG”, “CUSL” ), ncol=2, byrow=TRUE)
```

To fit the PLS model to the simulated 799-respondent data set, we use `plsm(data, strucmod, measuremod)` command line from semPLS package, based on matrices that previously were defined on step one and step two. Next, using `sempls(model, data)` command line we can estimate model parameters:

```
# Defining the whole model and fit it to data
library(semPLS)

ECSIPLS.mod <-plsm(data=ECSISimData , strucmod=ECSIPLSsm , measuremod
=ECSIPLSmm)
ECSIPLS.fit <-sempls(model=ECSIPLS.mod , data=ECSISimData)
```

The evaluation results are contained in the `ECSIPLS.fit` object. To estimate the factor structure (so-called factor loadings) of the measurement model, it is necessary to use the command plsLoadings (MODEL):

```
> plsLoadings(ECSIPLS.fit)
  IMAG CUEX PERQ PERV CUSA CUSCO CUSL
Imag1 0.71      .       .       .       .       .       .
Imag2 0.59      .       .       .       .       .       .
Imag3 0.70      .       .       .       .       .       .
Imag4 0.72      .       .       .       .       .       .
Cuex1   .    0.76       .       .       .       .       .
Cuex2   .    0.83       .       .       .       .       .
Cuex3   .    0.72       .       .       .       .       .
Perq1   .       .    0.75       .       .       .       .
Perq2   .       .    0.71       .       .       .       .
Perq3   .       .    0.72       .       .       .       .
Perq4   .       .    0.68       .       .       .       .
Perq5   .       .    0.70       .       .       .       .
Perv1   .       .       .    0.78       .       .       .
Perv2   .       .       .    0.86       .       .       .
Cusa1   .       .       .       .    0.86       .       .
Cusa2   .       .       .       .    0.83       .       .
Cusa3   .       .       .       .    0.81       .       .
Cusco1  .       .       .       .       .    0.85       .
Cusco2  .       .       .       .       .    0.87       .
Cusl1   .       .       .       .       .       .    0.83
Cusl2   .       .       .       .       .       .    0.81
Cusl3   .       .       .       .       .       .    0.85
```

In this example, all latent variables have strong loadings with its manifest indicator variables. If a latent variable has a factor loading below 0.3 for any indicator (which in this case is not present), or below 0.5 for all of its indicators—then the reliability of the measures will be debatable and further investigation and / or re-specification of the model structure is needed (Chapman & Feit, 2019, p. 288).

To obtain the values of the coefficients of the relationships between the latent variables (i.e., the structural model) it is necessary to use the command `pathCoeff(MODEL)`. The results are presented below:

```
> pathCoeff(ECSIPLS.fit)
       IMAG   CUEX   PERQ   PERV    CUSA  CUSCO   CUSL
IMAG      .  0.287      .      .   0.062      .  0.096
CUEX      .      .  0.370  0.334   0.114      .      .
PERQ      .      .      .  0.065   0.354      .      .
PERV      .      .      .      .   0.157      .      .
CUSA      .      .      .      .       . -0.443  0.475
CUSCO     .      .      .      .       .      . -0.161
CUSL      .      .      .      .       .      .      .
```

We see that 'Customer complaints' has a negative influence on 'Customer loyalty', and 'Customer satisfaction' has also negative influence on 'Customer complaints', which corresponds to the expectations (face validity).

In order to visualize the results of the evaluation through the familiar path diagram, it is necessary to take several additional technical steps, using the free available third-party software component Graphviz.[10] Using the program command `pathDiagram(MODEL, FILE, full-TRUE, …)` we can create a plot with the structural coefficients and loadings as an output as *.dot file. Once Graphiz software is installed, we can process this file as input and produces the corresponding image as a PDF or PNG file. For this purpose, first it is necessary to execute the following program line:

```
# Creating object for path diagram for Graphviz
pathDiagram(ECSIPLS.fit , file = "ECSIPLSfull", full = TRUE , digits = 2,
 edge.labels = "values", output.type = "graphics", graphics.fmt = "pdf")
```

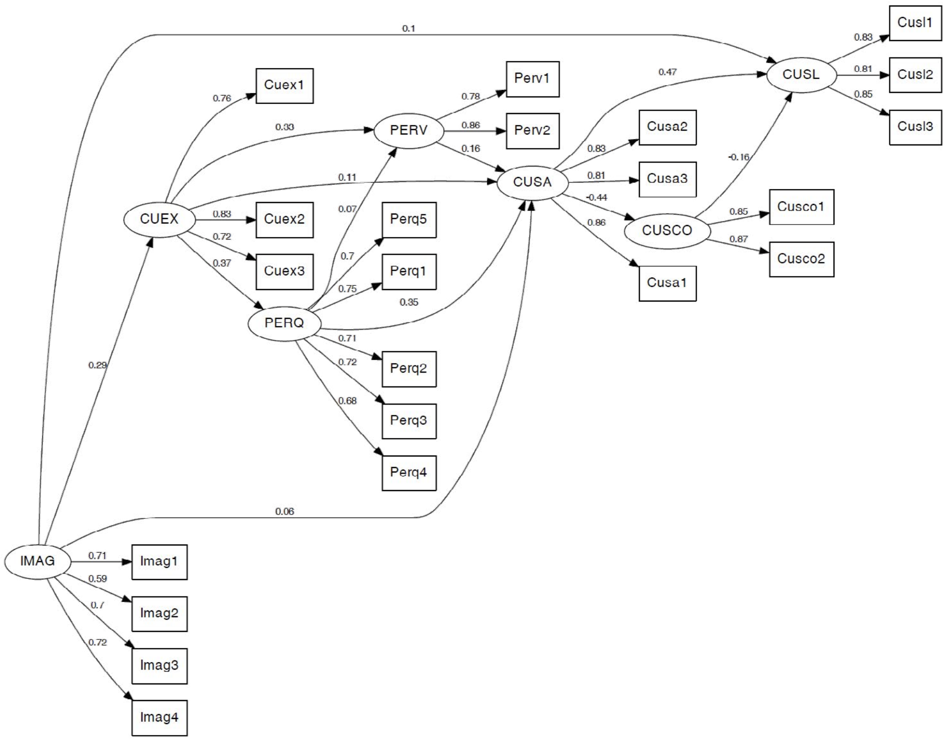
Because PLS models do not assess global model fit, according to Chapman and Feit (2019, p. 289), „(…) there is not a general way to compare CB-SEM and PLS-SEM results apart from interpreting the models and their implications, so it is not advisable to compare the coefficients directly".

After fitting the model to the data, however, the question of its statistical evaluation remains open. Because by PLS-SEM there is no summary metric for global model assessment and comparison, there are three possible approaches to its validation. First, we can examine the model's coefficients for their face validity. *Face validity* is the extent to which a coefficient estimated corresponds to our subjective expectation (in size and sign). Second, we can calculate the overall coefficient of determination (R^2) for the model, which is a measure of overall variance explained

[10] Graphviz can be downloaded for free from: http://www.graphviz.org/download/

within each part of the model. Thirdly, one can think about applying a bootstrap method to examine coefficient stability (Hair, Sarsted et al., 2011, pp. 423–424; Chapman & Feit, 2019, p. 289).

Figure 4.10. Path diagram with PLS coefficient estimates for the ECSI model using `semPLS`



Source: Own work.

The calculation of the coefficient of determination R^2 for each of the latent variable is possible by the `rsquared()` function:

```
> rSquared(ECSIPLS.fit)
 R-squared
IMAG .
CUEX 0.082
PERQ 0.137
PERV 0.132
CUSA 0.244
CUSCO 0.196
CUSL 0.345
```

There is no general "rule of thumb" for interpreting this measure, but since the R^2 values range is between 0 and 1, the closer the empirical score of R^2 to 1, the better the model.

A more general approach to access coefficient stability is to use a bootstrapping procedure. We can perform it with `bootsempls()` command. A key point in performing the procedure is setting a sufficiently large value for resample sets of observation. In this case, we set 1000:

```
# Bootstrapping
set.seed (5250)
ECSIPLS.boot <-bootsempls(ECSIPLS.fit , nboot =1000, start="ones")
```

After 1000 resample through command summary() we get the following results:[11]

```
> summary(ECSIPLS.boot)
Call: bootsempls(object = ECSIPLS.fit, nboot = 1000, start = "ones")

Lower and upper limits are for the 95 percent perc confidence interval

         Estimate        Bias  Std. Error       Lower     Upper
lam_1_1    0.7125   -4.38e-03     0.03923    0.620324     0.779
lam_1_2    0.5876   -2.54e-03     0.05264    0.464342     0.677
lam_1_3    0.7003   -2.02e-03     0.04066    0.613629     0.771
lam_1_4    0.7246   -2.90e-03     0.03804    0.643485     0.789
lam_2_1    0.7570   -5.85e-04     0.02085    0.712676     0.793
lam_2_2    0.8300   -4.54e-04     0.01547    0.797551     0.857
lam_2_3    0.7235   -2.77e-04     0.02533    0.670878     0.768
lam_3_1    0.7504   -8.70e-04     0.01963    0.708437     0.786
lam_3_2    0.7102   -1.42e-03     0.02290    0.659964     0.753
lam_3_3    0.7207   -7.41e-04     0.02166    0.674986     0.761
lam_3_4    0.6786    4.17e-05     0.02510    0.629882     0.723
lam_3_5    0.6991   -5.56e-04     0.02452    0.645462     0.745
lam_4_1    0.7815   -3.36e-03     0.03328    0.701759     0.832
lam_4_2    0.8624    9.87e-04     0.02183    0.817765     0.906
lam_5_1    0.8602   -2.73e-04     0.00911    0.840928     0.876
lam_5_2    0.8269   -1.02e-03     0.01329    0.798288     0.851
lam_5_3    0.8122   -7.55e-04     0.01537    0.780249     0.840
lam_6_1    0.8481   -9.06e-04     0.01594    0.813603     0.876
lam_6_2    0.8690    5.01e-05     0.01334    0.841386     0.894
lam_7_1    0.8345   -7.65e-04     0.01297    0.806461     0.857
lam_7_2    0.8090   -5.58e-04     0.01502    0.776243     0.836
```

[11] Note that the use of bootstrapping procedure with relatively small samples (e.g., < 100) is sometimes problematic and the model could be unstable because bootstrap iterations failed to converge.

```
 lam_7_3     0.8502     -3.05e-04    0.01032       0.829233     0.869
 beta_1_2    0.2870      2.70e-03    0.02939       0.229344     0.348
 beta_2_3    0.3697      3.80e-03    0.02903       0.313689     0.430
 beta_2_4    0.3337      5.64e-04    0.03285       0.265878     0.395
 beta_3_4    0.0654      3.65e-04    0.03545      -0.003022     0.133
 beta_1_5    0.0619      3.02e-03    0.03275       0.000189     0.130
 beta_2_5    0.1135     -1.47e-03    0.03321       0.047941     0.177
 beta_3_5    0.3544     -1.38e-03    0.03247       0.285964     0.417
 beta_4_5    0.1567      1.28e-03    0.03407       0.090795     0.225
 beta_5_6   -0.4433     -1.75e-03    0.03018      -0.505684    -0.385
 beta_1_7    0.0961      1.11e-03    0.02896       0.037563     0.155
 beta_5_7    0.4747     -1.63e-03    0.02759       0.418293     0.525
 beta_6_7   -0.1606     -1.35e-03    0.02969      -0.220249    -0.102
```

In the output above, the labels lam_x_x correspond to the factor loadings of the measurement model, while the labels beta_x_x correspond to the coefficients of the structural model. Problematic coefficients are those in which upper and lower bounds include 0, and for which we therefore do not have even directional confidence. We observe only two cases, at beta_3_4 and beta_6_7. They correspond to the influence of IMAG on CUSA, i.e., of ` Corporate image` on `Consumer satisfaction`, and PERQ on PERV, i.e., of `Perceived quality` on `Perceived value`. The interpretation should be that such an influence is unlike. The results of the simplified alternative CB-SEM models prompted us to a similar conclusion.

The results of the bootstrapping procedure can be visualized as a parallel plot. This plot can be created using the following program code:

```
parallelplot(ECSIPLS.boot , reflinesAt = 0, alpha =0.8,
 varnames=attr(ECSIPLS.boot$t, "path") [23:34],
 main="Path coefficients in 1000 PLS bootstrap iterations (N=799)")
```

The resulting plot is shown in Figure 4.11. We can see all bootstraps estimates of the coefficients of structural model and their spread between lower and upper limits for the 95% confidence interval. We should read this by looking at the spread of estimates along each of the horizontal grid lines representing one structural model coefficient. For example, the influence of `Perceived quality` on `Perceived value`, as well as `Corporate image` on `Customer satisfaction` are generally estimated to be positive, but several of the estimates hold the relationship to be slightly negative. Therefore, the interpretation of this coefficient could not be reliable, and we could not use them confidently. We made a similar conclusion after the application of CB-SEM in the previous subsection.

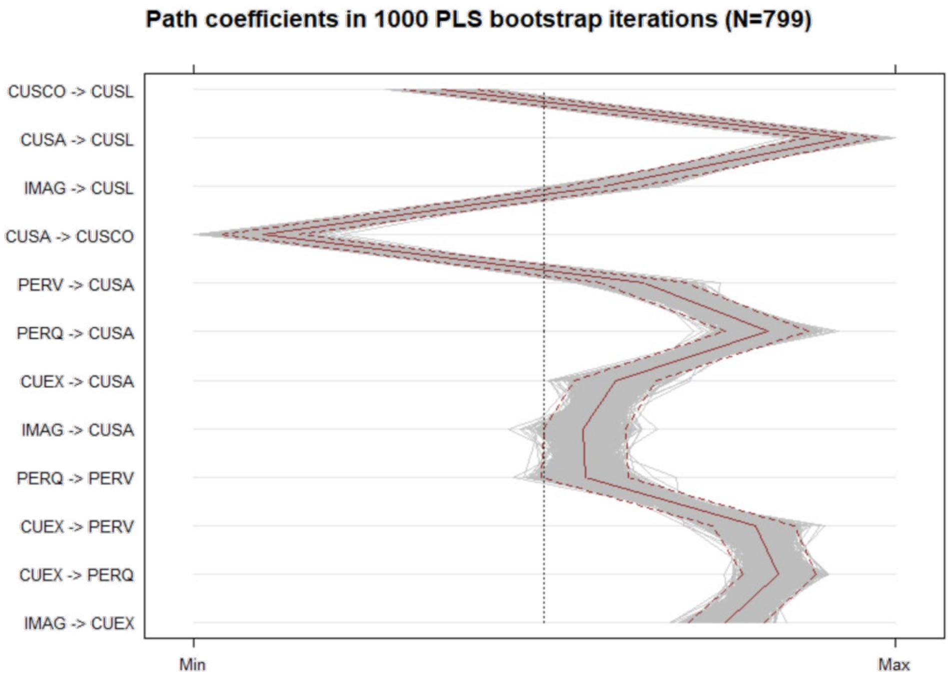


Figure 4.11. Bootstrapped coefficients for the PLS model. Each line plots the twelve estimated coefficients of the structural for one complete bootstrap iteration

Source: Own work.

4.2.2.2. Fit a model to data using SmartPLS

Currently, SmartPLS[12] is probably the most popular commercially available software to use the PLS-SEM method. It is more convenient than any other open-source or commercial software solution. The graphical interface is very easy and intuitive to work with and it allows us to focus on the research itself rather than to struggle with the software. In this sub-section, we will illustrate the use of SmartPLS to evaluate the theoretical model we already tested with `lavaan` package, IBM SPSS Amos, and `semPLS` package, using the same simulated dataset.

Drawing the conceptual model with the use of graphical tools is easy. You can use the guidelines in the official recourses available here. After a data file is imported, you can easily drag and drop indicator variables from a list to a particular construct in the model you have drawn. Once you have built the structural and measurement models, you can proceed to the model evaluation.

There are two types of PLS algorithms available in SmartPLS—a regular PLS algorithm and a consistent PLS algorithm (PLSc). The latter is appropriate when

[12] For a trial version, visit: www.smartpls.com.

some or all constructs have reflective measurement models because it performs a correction of these constructs' correlations. However, one can run both algorithms and compare the results. The model in Figure 4.12 shows the estimated ECSI model applying the regular PLS algorithm.

SmartPLS contains many options for evaluating and verifying analytical results. Detailed instructions, guidelines and recommendations for reporting the results of the PLS-SEM approach using this software can be found in Hair, Sarsted and others (2011, 2019). We will briefly comment on the model estimation procedure.

The results show the outer weights and/or loadings of each measurement model. If the path diagram at hand has a construct with formative measurement model, then we are interested at the outer weights for this construct. For reflective measurement models, we need to look at the outer loadings. The ECSI model has only reflective measurement models, thus the parameters for all indicator variables displayed in Figure 4.12 are the outer loadings.[13] There are few outer loadings that fall under the recommended threshold value of 0.7 (see Table 4.2). We can remove each indicator that loads poorly on the construct and check whether this improves the reliability and validity of the measurement model. Before we do this, we should assess the current measurement models.

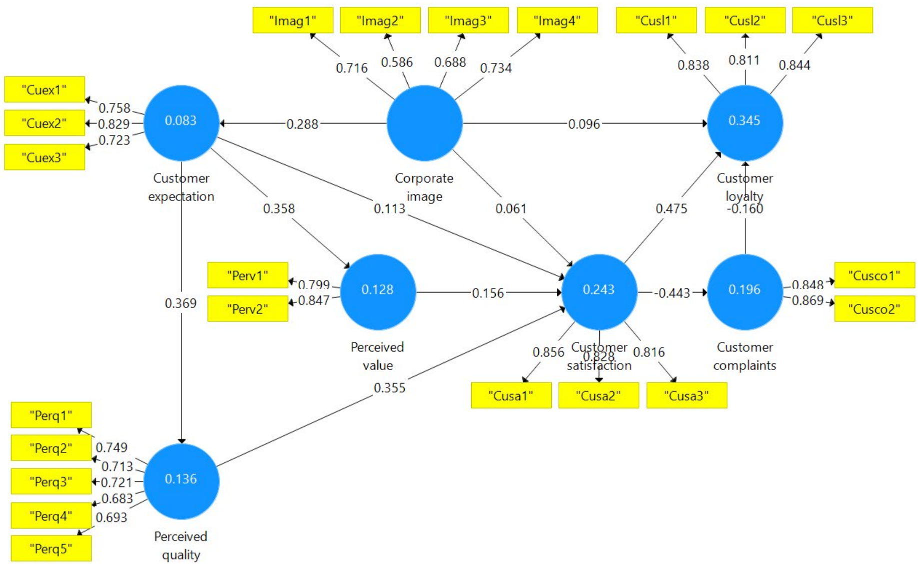


Figure 4.12. Path diagram with PLS coefficient estimates for the ECSI model using SmartPLS

Source: Own work.

[13] In SmartPLS you can choose to display the outer weights, loadings or both on the path diagram depending on the type of the measurement models you have.

There are several criteria reported in the results that can be used to evaluate the measurement models (see Table 4.10). Both Cronbach's alpha and rho (ρ) are used to assess the internal consistency of each construct. In our example, most of the latent variables have values of these two measures above 0.6, except for "Perceived value".

We can use the average variance extracted to assess the convergent validity of the constructs. In our example, all constructs have AVE > 0.5, except for IMAG. However, since this value is very close to the threshold, we will not delete any indictors at this point.

Table 4.10. Evaluation of reliability and validity of the measurement models

	Cronbach's alpha	rho_A	Composite reliability	Average variance extracted (AVE)
Perceived value	0.53	0.53	0.81	0.68
Perceived quality	0.76	0.76	0.84	0.51
Corporate image	0.62	0.63	0.78	0.47
Customer loyalty	0.78	0.78	0.87	0.69
Customer complaints	0.64	0.65	0.85	0.74
Customer satisfaction	0.78	0.79	0.87	0.69
Customer expectations	0.66	0.68	0.81	0.60

Source: Own work.

The discriminant validity of each measurement model can be judged by the Fornell / Larcker criterion. In Table 4.11 we can see the square root of AVE for each construct, which should be higher than the correlations with other constructs. In our example, this is true for all constructs, thus their measurement models have a good discriminant validity.

Table 4.11. Square root of AVEs and correlations between constructs (Fornell/ Larcker criterion)

	CUEX	CUSA	CUSCO	CUSL	IMAG	PERQ	PERV
Customer expectations	0.77						
Customer satisfaction	0.32	0.83					
Customer complaints	–0.15	–0.44	0.86				
Customer loyalty	0.25	0.56	–0.38	0.83			
Corporate image	0.29	0.17	–0.07	0.19	0.68		
Perceived quality	0.37	0.44	–0.24	0.26	0.16	0.71	
Perceived value	0.36	0.27	–0.14	0.22	0.11	0.19	0.82

Source: Own work.

The next step is to evaluate the relationships between latent variables and if they explain each construct to a satisfactory extent. Estimated R^2 values of the latent variables appear in the circles representing each construct (see Figure 4.12).

They are identical to those obtained using the `semPLS` package (see Figure 4.10). SmartPLS contains a powerful bootstrapping algorithm, with the help of which it is possible to estimate the statistical significance of each of the estimated coefficients of the model (see Table 4.12). From the table it is clear that only the influence of `Corporate image` on `Customer satisfaction` is not significant (*p* value > 0.05).

Table 4.12. Statistical significance of coefficients after bootstrapping

	Original sample (O)	**Sample mean (M)**	**Standard deviation (STDEV)**	**T-Statistics (|O/STDEV|)**	***p*-values**
Corporate image—> Customer expectation	0.288	0.291	0.032	9.117	0.000
Corporate image—> Customer loyalty	0.096	0.095	0.030	3.171	0.002
Corporate image—> Customer satisfaction	0.061	0.065	0.033	1.830	**0.067**
Customer complaints—> Customer loyalty	–0.160	–0.160	0.031	5.243	0.000
Customer expectation—> Customer satisfaction	0.113	0.111	0.035	3.224	0.001
Customer expectation—> Perceived quality	0.369	0.371	0.031	11.864	0.000
Customer expectation—> Perceived value	0.358	0.359	0.030	12.018	0.000
Customer satisfaction—> Customer complaints	-0.443	-0.445	0.028	15.754	0.000
Customer satisfaction—> Customer loyalty	0.475	0.475	0.029	16.649	0.000
Perceived quality—> Customer satisfaction	0.355	0.356	0.032	11.069	0.000
Perceived value—> Customer satisfaction	0.156	0.157	0.035	4.474	0.000

Source: Own work.

4.3. Solving sustainability research problems with SEM

Sustainable development is a concept that recommends a set of ethical-oriented goals for nations who are aspiring to make economic growth widespread, to encourage social welfare, and to protect the environment from human-induced degradation (Sachs, 2015, p. 3). We can apply structural equation modelling in various sustainable development research areas when the problem at hand requires testing complex relationships between latent variables. In a recent meta-analysis of scientific articles related to sustainable development (in which some form of SEM was used), the authors conclude that 61% of all research teams used SmartPls, 26% used Amos, 8% used LISREL, and 5% used other software solutions when evaluating conceptual models (Mardani et al., 2017). Survey data was used for model evaluation and the respondents were either clients (end-users of products and services) or companies' representatives (usually senior or junior managers, but also other types of employees).

In this sub-section, we will focus on examples of research problems solved with the use of SEM. While this is just a summary of possible areas in SD where this type of analysis is appropriate for solving specific problems, applying SEM can answer many other research questions.

4.3.1. Sustainable development as a concept and strategy

Whenever we want to study general perceptions of sustainability as a concept or as a strategy implied by firms or public entities, we can use SEM to test our theoretical models. The three facets of SD can be understood differently by different communities or organizations. For example, a researcher may want to know how residents of one area perceive local ecological initiatives, economic measures, and social inclusion practices and whether these constructs influence the perceptions of sustainable development of this area. A simple structural model of this type is depicted in Figure 4.13. Each construct in this model could represent the complex understanding of people for each aspect of sustainable development or could reflect people's perceptions of specific measures undertaken by the government to encourage sustainable development. Previous research is a good starting point when selecting appropriate indicator variables, as is a preliminary qualitative study.

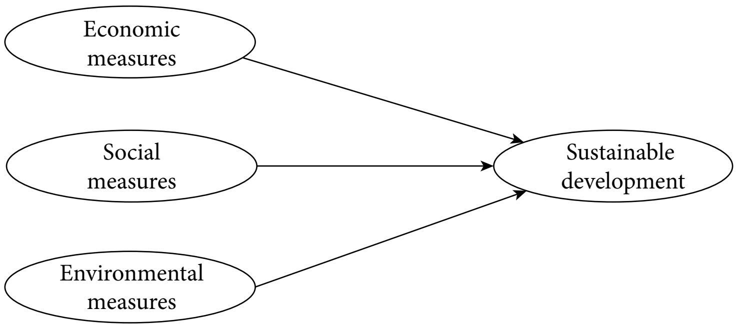


Figure 4.13. Structural model of the influence of economic, social and environmental measures undertaken by the government to the perception of sustainable development

Source: Own work.

4.3.2. Supply chain management

Practices related to supply chain management have a considerable environmental effect. Researchers in this field might be interested in studying the supplier's perspective on shifting to more sustainable transportation practices. The willingness to adopt a sustainable strategy can depend on many factors.

The theoretical model in Figure 4.14 shows the relationships between a company's environmental performance and different aspects of its purchasing behaviour (Large & Thomsen, 2011). Working with green suppliers and collaborating on environmental issues hypothetically influence 'Environmental performance improvement' and 'Purchasing performance'. The model was validated with survey data obtained from a sample of 109 purchasing and supply managers, using PLS-SEM methodology.

The relationships between latent variables in the structural model depicted in Figure 4.14 was initially considered positive. All measurement models are reflective, having Cronbach's alpha > 0.7 (cited in previous research), average variance extracted > 0.6, and composite reliability > 0.7. The latent variable "Strategic level of purchasing" had three items aimed at measuring the perceived actions that the purchase department took on a strategic level. The next construct—"Environmental commitment"—was also measured with three indicators: policy statement, value and understanding, all designed to measure the extent to which supply managers perceive environmental concerns are part of their company's policy, values and efforts to make employees also acknowledge the importance of environmental management. The construct "Purchasing's environmental capabilities" is measured with three indicators that captured the respondent's perceptions about the environmental goals of the purchase department and environmental training as well as activities performed by the purchasing personnel. The "Green supplier assessment" is measured through three items, i.e., reflecting the extent to which companies perform environmental assessment of potential suppliers, giving them feedback and performing environmental audits. "Green collaboration with suppliers" is supposed to reflect the respondents' perceptions of the extent to which their companies make joint efforts with the suppliers to reduce waste, help suppliers improve their performance and provide them with training. "Environmental performance improvement" is measured by six indicator variables—waste reduction, improved compliance with environmental laws, increased level of recycling, environmental protection, improved environmental reputation, and improved overall environmental performance. The last construct in the structural model is measured through five items: specifications, on time delivery, quantities, and internal satisfaction with the purchasing department.

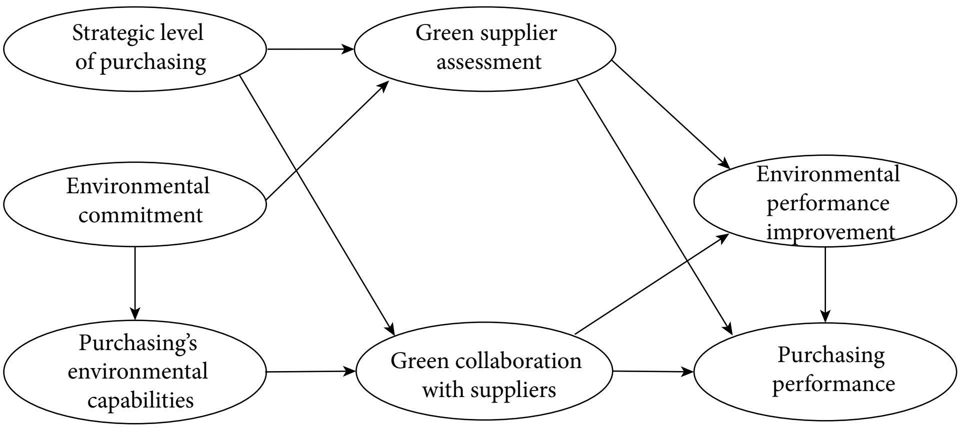


Figure 4.14. Structural model of environmental and purchase performance of firms buying from 'green' suppliers

Source: (Large & Thomsen, 2011).

All hypothetical relationships in the model in Figure 4.14 was confirmed, except the relationship between "Green collaboration with suppliers" and "Purchasing performance". The research contributed to the understanding of how green supplier assessment and collaboration with suppliers on environmental issues can positively influence firm's environmental performance. On the other hand, the improvement of environmental performance had a positive influence on the purchasing performance of companies.

4.3.3. Corporate social responsibility

A research dedicated to the link between corporate social responsibility (CSR) and environmental supplier development (ESD) has concluded that the latter influence positively the financial performance and competitive advantage of firms (Agan, Kuzey, Acar, & Açıkgoz, 2016).

The model in Figure 4.15 includes four latent variables: "Corporate social responsibility", "Environmental supplier development", "Financial performance", and "Competitive advantage". Four multi-item scales were used to measure these constructs, which included 45 different items. However, due to low factor loadings, some of the items were propped, which resulted in 35 items in the final measurement model. These indicators loaded on three factors (sub-dimensions) describing the CSR: "CSR to employees", "CSR to customers", "CSR to environment", "CSR to media", and "Partnership with NGOs". Items related to ESD were loaded on three factors, namely "Supplier evaluation", "Incentives", and "Direct involvement". Performance questions were loaded on two factors named 'Financial Performance' and 'Competitive Advantage'.

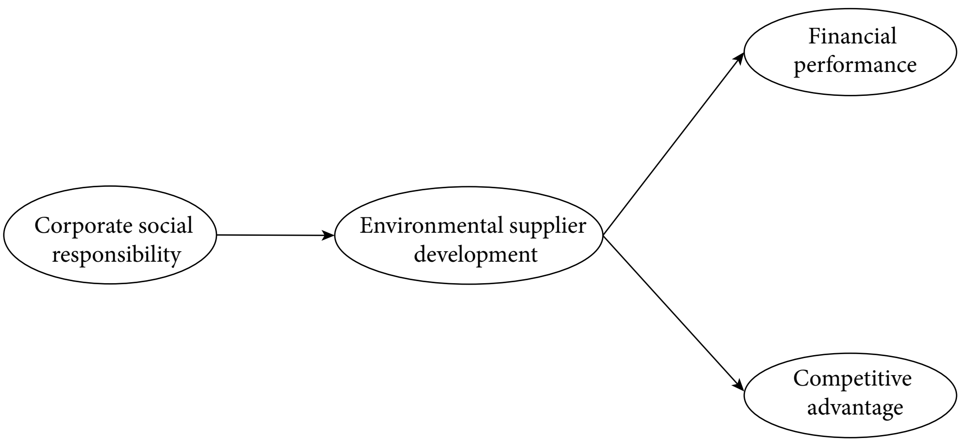


Figure 4.15. Structural model of the relationship between CSR, ESD and company performance

Source: (Agan et al., 2016).

All three relationships in the model in Figure 4.15 are hypothesized as being positive. This was confirmed by applying PLS-SEM. The data needed to evaluate the conceptual model was gathered through a survey and included 314 responses from mid- or high-level managers, directors or engineers in manufacturing firms. The research showed that CSR have a positive influence on ESD, although the latter is poorly explained solely by CSR and there are probably other relevant determinants. However, ESD also impacts positively the financial performance and competitive advantage of companies.

4.3.4. Innovations linked to sustainability

Green innovation can consist of either creating 'green' products or adopting 'green' processes. Green innovation comprises innovation in technologies for energy saving, pollution prevention, waste recycling, green product designing, and corporate environmental management.

The theoretical model in Figure 4.16 shows the hypothetical relationships between green supply chain management (GSCM) practices and technological innovation in manufacturing firms (Lee, Ooi, Yee-Loong, & Seow, 2014).

- Six indicators measure "Internal environmental management": commitment of GSCM from senior managers; support for GSCM from mid-level managers; cross-functional cooperation for environmental improvements; total quality environment management; environmental compliance and auditing programs; environmental management system exists.
- Three indicators measure "Eco-design": design of products for reduced consumption of material / energy; design of products for reuse, recycle, recovery of material, component parts; design of products to avoid or reduce use of hazardous of products and / or their manufacturing process.
- Three indicators measure "Investment recovery": investment recovery (sale) or excess inventories / materials; sale of scrap and used materials; sale of excess capital equipment.
- Four indicators measure "Green purchasing": cooperation with suppliers for environmental objectives; environmental audit for supplier's internal management; suppliers' ISO 14000 certification; second-tier supplier environmentally friendly practice evaluation.
- Three indicators measure "Cooperation with customers": cooperation with customer for eco-design; cooperation with customer for cleaner production; cooperation with customer for green packaging.
- "Technological innovation" is measured with nine indicators: "We are able to produce products with novelty features"; "We use the latest technology for new product development"; "The speed of new product development is fast enough / competitive"; "We have enough new products introduced to the market"; "We

have new products which are first-in-market (early market entrants)"; "We are technologically competitive"; "We use up-to-date/new technology in the process"; "We are fast in adopting process with the latest technological innovations"; "The process, techniques and technology change rapidly in our company."

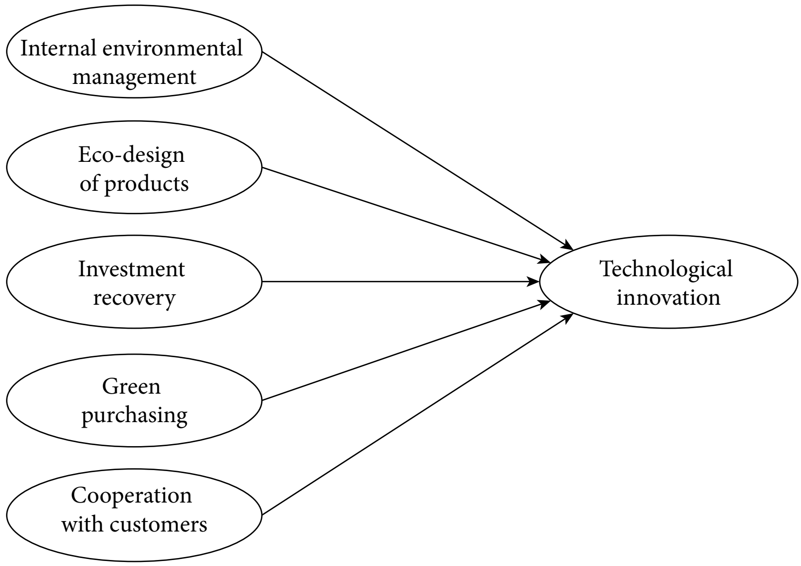


Figure 4.16. Structural model of the impact of green supply chain management practices on the technological innovation in firms

Source: (Lee et al., 2014).

The model was tested with survey data collected from environmental and operation managers in manufacturing companies of different sizes. The results showed that latent variables 'Internal environmental management', 'Eco-design of products' and 'Investment recovery' have a significant positive effect on technological innovation in firms.

4.3.5. Consumer behaviour and sustainable consumption

Human psychology is everything but simple. Most of the processes that make us think, feel or act in a certain way are not straightforward and easy to measure or interpret. It is quite the opposite: there can be many possible variables that have an effect on the way we perceive information and derive meaning from it, then form feelings and decide how to behave. Studying consumer behaviour almost always includes efforts to measure multi-dimensional psychological phenomena, such as consumer satisfaction, consumer loyalty, expectations and experience with the product, attitude and many others. In other words, the researcher is interested in analysing complex structures of latent nature, that are not a subject of direct observation and measurement. This is where SEM comes at hand, and by applying this type of analysis, we can answer some common questions regarding consumer behaviour linked to sustainability.

Globally, consumers are focusing their preferences on products and brands that are implementing different innovations to promote sustainability. Environmental awareness is increasing, and this makes consumers look for, and choose eco-friendly products, avoid waste, and reuse products and materials (Euromonitor International, 2020). Very often, the researcher or practitioner wants to know, e.g.:

– What factors contribute to the adoption of sustainability-related behaviours?
– Why do consumers choose to buy a brand positioned as sustainable?
– What factors contribute the most to the satisfaction with a sustainable product alternative?
– How do consumers perceive sustainable product attributes?
– Do perceptions of sustainable product attributes affect the product choice?
– What influences loyal behaviour towards sustainable companies and products?

The structural model in Figure 4.17 shows the hypothetical influence of different aspects of sustainable development on customer satisfaction, loyalty and willingness to pay (Xu & Gursoy, 2015). This model was tested in the field of hospitality supply chain management, with survey data obtained from 499 consumers who stayed in a hotel within the last 6 months. The data is reported to fit the measurement model well, with acceptable levels of reliability and discriminant validity for each construct.

The results show that all three dimensions of sustainable development adopted by hotels have a positive effect (direct or indirect) on consumer behaviour. The more hotels invest in sustainable practices, the more satisfied and loyal customers are, and the higher willingness to pay for the service they express.

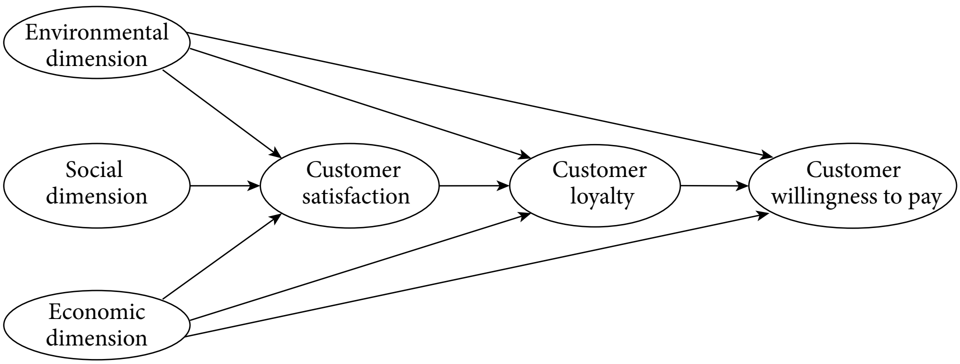


Figure 4.17. Structural model of the impact of SD dimensions on customer satisfaction, loyalty, and willingness to pay

Source: (Xu & Gursoy, 2015).

Another model explaining the link between product information, perceived value and specific buying decision process on consumer's willingness to buy 'green' products is given in Figure 4.18. This conceptual model was tested with survey data gathered in 27 member states of the EU, with a sample size $n = 26573$ (Couto, Tiago, Gil, Tiago, & Faria, 2016). The available product information, perceived value and considerations linked to the buying decision process hypothetically influence positively consumers' willingness to pay for a green product. Six indicators are used to measure the "Product information" construct: information on the shelf, information in advertisements, on the internet, in a leaflet at the shop, on a bar code that can be scanned by a smartphone, and on the label on the product. Participants are asked to rate each of these sources of information as more or less preferable when looking for environmental information about a product. The next construct, "Perceived value", is also measured with six indicators, aimed at capturing the different aspects of value related to environmentally friendly products: good value for money; as effective as other products; using them being the right thing to do; setting a good example with the purchase; making a difference to the environment; positive opinion of friends and family about using environmentally friendly products.

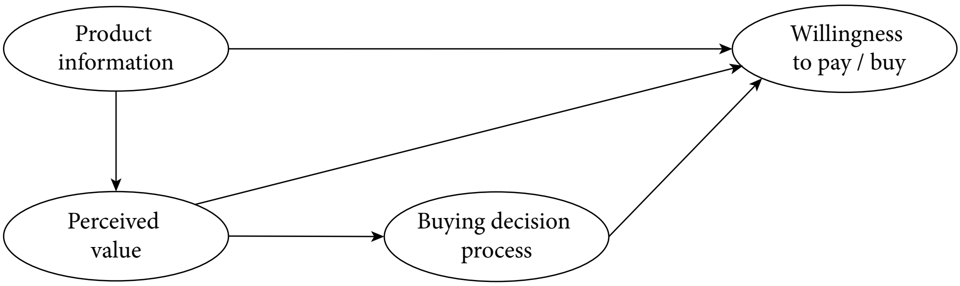


Figure 4.18. Structural model of willingness to pay / buy as a result of product information

Source: (Couto et al., 2016).

"Buying decision process" is measured with four indicators, reflecting different dimensions used to evaluate products before making a purchase decision. These are: product's impact on the environment; price; quality; brand name. The last construct represents the consumers' willingness to pay more for a product if they consider it as environmentally friendly. There are five indicators for this construct, all representing an increasing extra payment rate for environmentally friendly products. Only two out of five research hypothesizes were supported after the analysis. Willingness to pay was positively influenced by product information, but not by perceived information and buying decision process. Perceived product value did influence positively the buying decision process.

The willingness to pay / (WTP) for renewable energy is in the focus of another research (Lin & Syrgabayeva, 2016). Four constructs are tested for their direct or indirect effects on the willingness to pay more for renewable energy. The single construct that hypothetically have a direct positive effect on WTP more for renewable energy is “Attitude toward renewable energy”. The latter is positively influenced by “Environmental concern”, “Environmental belief” and “Knowledge about renewable energy”. All other hypothetical relationships between the constructs are also positive (see Figure 4.19).

Multiple items are used to reflect the constructs, all measured on a 5-point Likert scale. “Environmental concern” construct has three indicators representing the concern about pollution, air pollution, and water usage. “Knowledge about renewable energy” reflects on three indicators about knowledge and familiarity with renewable energy sources and with wind-generated energy. “Environmental belief” is measured with three indicators representing the extent to which respondents believe that the environment, reliability, and environmental safety are important when considering renewable energy. “Attitude toward renewable energy” reflects on three statements indicating the extent to which respondents like renewable energy more than traditional one, their preference to buy and to use renewable energy. Finally, “Willingness to pay more for renewable energy” reflects on three variables, measuring the intention to pay more for renewable energy.

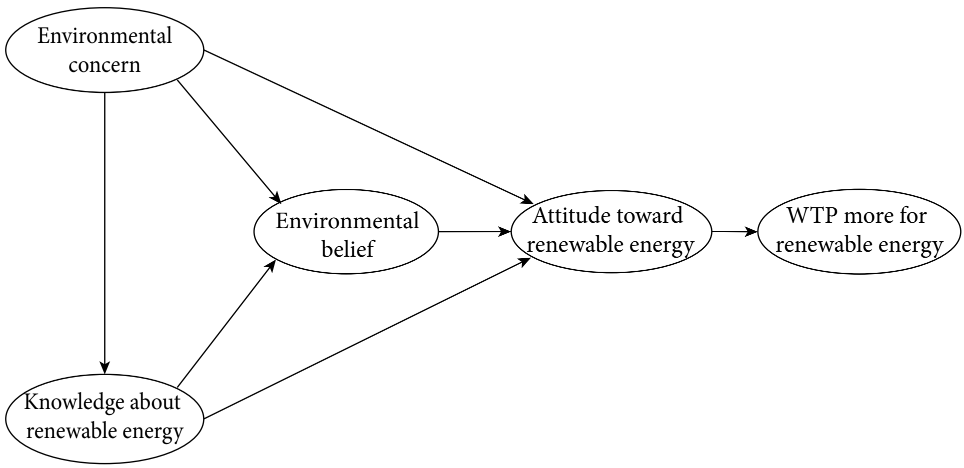


Figure 4.19. Structural model of the impact of the attitude toward renewable energy on the willingness to pay more.

Source: (Lin & Syrgabayeva, 2016).

PLS-SEM was used to assess the described conceptual model. The measurement model was found to be reliable, and all constructs had good convergent validity. The structural model was found to fit the data relatively well. Two of the hypothetical relationships were not statistically significant: the environmental concern did not have a direct impact on the knowledge about renewable energy, and the latter did not influence the attitude toward renewable energy. All other hypotheses used to build the conceptual model were supported. Consumers were found to have a better attitude toward renewable energy when they rated higher the importance of their environmental beliefs and when their concerns about the environment were also higher. A better attitude toward renewable energy also had a positive effect on the WTP more for renewable energy.

4.3.6. Human resource management

Promoting social welfare is one of the pillars of sustainable development. Public policies in this regard are aimed at ensuring healthy lives and well-being; providing access to quality education; ensuring gender equality; creating inclusive societies and institutions. In an effort to study each goal within the social aspect of sustainable development, researchers inevitably have to deal with complex phenomena of latent nature. As social constructs, most of these categories have many nuances and very often are interrelated, thus their studying require application of SEM methods.

Some social sustainability problems addressed in public policies are also relevant for private companies. Businesses that aim to be socially responsible are making efforts not only to minimize their environmental footprints but also to establish fair treatment of their employees and guarantee equality and inclusion. Sometimes these efforts are imposed by legal norms, and other times companies just want to stand out, emphasizing their fair working conditions and social care for employees. There is no doubt, however, that human resource management is a part of the sustainability strategies of companies.

Another noteworthy aspect of sustainability strategy, implemented by different human resource activities, is promoting sustainable behaviours to employees on every hierarchy level. This is also linked with the idea that even though companies are taking actions to become environmentally responsible, it is the motivation of employees that boosts the environmental performance of firms.

In the next example, human resource management practices are hypothesized to have an impact on the adoption of environmental practices by companies, which also have a direct effect on the operational performance (Jabbour, Jabbour, Govindan, Teixeira, & Freitas, 2013). Another latent variable in the theoretical model is the lean manufacturing that is hypothesized to influence the adoption of environmental practices (see Figure 4.20). The theoretical model was tested using survey data. The questionnaire included 28 items reflecting the latent variables and was distributed among automotive sector companies. The final sample included 75 respondents on managerial positions in production / operations areas.

- Indicator variables of 'Environmental management' construct include: clear policy of valorising environmental management; environmental training for all employees; 3Rs (reduction, reuse, and recycling); development of products with smaller environmental impact; development of production processes with smaller environmental impact; supplier selection based on environmental criteria; ISO 14001 or other environmental management system; voluntary promotion of information on environmental performance.
- Indicator variables of 'Human resource' construct include recruiting and selection; training; performance evaluation; rewards; benefits.
- Indicator variables of 'Operational performance' construct include: cost; time-to-market; new products; quality; flexibility; delivery.
- Indicator variables of 'Lean manufacturing' construct include: multifunctional involvement in the process; continuous improvement; 5S; total productive maintenance; Kanban; just-in-time; lot reduction; improvement circles; vendor development.

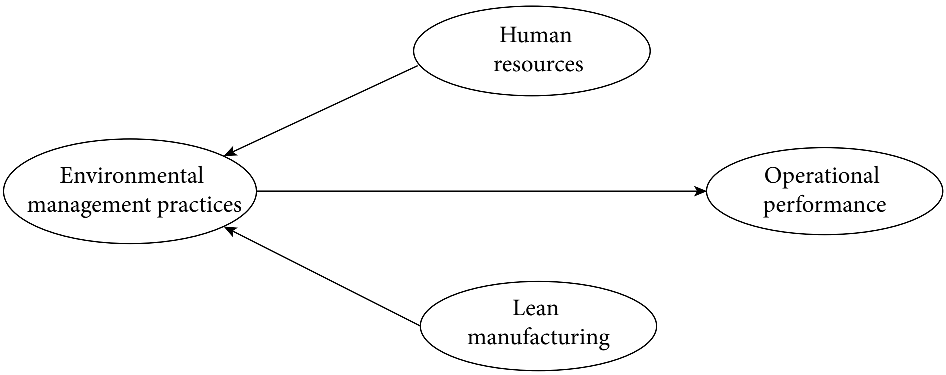


Figure 4.20. Structural model of environmental practices adoption impact on operational performance

Source: (Jabbour, Jabbour, Govindan, Teixeira, & Freitas, 2013).

The final measurement model included 25 indicator variables. The results showed that human resource and lean manufacturing practices positively influence the adoption of environmental practices, although this influence is weak to moderate. Environmental practices also positively influence operational performance.

*

Structural equation modelling methods are a powerful toolbox for researchers in any field to study complex phenomena of latent nature. A structural equation model relates observed manifest variables (indicators) to underlying constructs. It estimates the strength of the associations in a proposed model, as well as the degree to which the model fits the data.

There are two general SEM approaches—the covariance-based approach (CB-SEM) and the partial least squares approach (PLS-SEM). Both CB-SEM and PLS-SEM approaches represent an analysis at the aggregation level and suggest a set of statistical constraints. These two approaches should not be seen as an alternative to each other but rather as complementary.

CB-SEM is more restrictive in terms of statistical requirements, the fulfilment of which the researcher should start with it. PLS-SEM is less restrictive, but it has other shortcomings, such as its inability to use a global criterion to assess the structural model.

In addition to the statistical requirements, structural equation modelling methods also have requirements regarding the data to be analysed. This is why SEM should be applied if there is a reliable theory about the relationships between the variables, and as much information as possible is included in the analysis, whereby one can obtain this information from theoretical sources or previous exploratory analysis. Building a "good" model that is most likely to reflect the studied phenomena and can be validated with empirical data requires us to have solid theoretical knowledge of the studied phenomena.

Questions / tasks

1. What is the main idea of the SEM method?
2. What are the main differences between CB-SEM and PLS-SEM methodologies?
3. Can you name some of the advantages of PLS-SEM?
4. What is the difference between formative and reflective measurement models?
5. What kind of sustainable development research problems can be solved using SEM?
6. Is the PLS-SEM approach to model testing appropriate for the model depicted in Figure 4.14? Why?
7. What are the possible formulations of items used to measure the "Environmental commitment" construct, which was part of the structural model depicted in Figure 4.14?
8. What indicators can you suggest being used for the latent variables in Figure 4.17?
9. The structural model depicted in Figure 4.20 is based on three research hypotheses. Can you formulate them?

Task 1

Allnature Cream Cheese is a product made with natural ingredients. It is intended to satisfy the need of consumers for healthy alternative of processed cheese. The product can be directly consumed or used as an ingredient in various recipes. The marketing and sales department of the manufacturing company wanted to know what influences the buying decision of consumers and to use this knowledge in their strategy. They decided to gather qualitative data about the product attributes that influence the purchase intentions of potential buyers. Three parallel focus groups were organised with participants representing targeted consumer segments. The following product attributes were lined out as most important when deciding to buy cream cheese:

- **Health**
  Consumers are generally concerned about the health effects of the products they consume. They perceive cream cheese as healthier than the processed cheese, but they still pay attention to product ingredients when choosing a product. Consumers prefer products made from natural ingredients, although sometimes they do not trust the labels "eco", "green", or "bio".
- **Sustainability**
  While health benefits are a more personal reason why consumers choose to buy *Allnature Cream Cheese*, perceived sustainability of the production process is related to their concerns about environmental changes and nature preservation. Generally, consumers perceive favourably any company that adopt sustainability innovation processes, and this is expected to influence the perception of the company's products.
- **Price**
  Consumers expect prices of healthier product alternatives to be higher than those of conventional products. They are somehow ready to pay more for healthier products. However, the price level is still a concern when choosing a product.

These results were somehow satisfying for the marketing and sales team. The all-natural ingredients of the product and the ongoing transformation of production practices in an effort to apply a sustainability strategy are good premises for market success of the new product. The research team created a conceptual model of the potential influence of cream cheese product attributes to the purchase intentions of consumers (Figure 4.21). This influence is assumed to have both direct and indirect aspects. On the one hand, when consumers perceive the product as more healthy, sustainable and cheaper, their intentions to purchase it are expected to be higher. On the other hand, these perceived product attributes could influence the overall perception of the product as attractive, which in turn can affect consumers' purchase intentions. After a thorough discussion, the research team defined the following hypotheses for the next stage of the study:

H$_1$: The healthier the product is perceived by the consumer, the more attractive it is.
H$_2$: The healthier the product is perceived by the consumer, the stronger is the intention to buy.
H$_3$: The more sustainable the production process is, the more attractive the product is to consumers.
H$_4$: The more sustainable the production process is, the higher the probability that it will be purchased.
H$_5$: The more attractive the product is perceived by consumers, the more likely it is to be purchased.
H$_6$: The higher the price level of the product, the lower the probability that it will be purchased.

In the discussion on the above hypotheses, it was pointed out that there might be correlations between the latent variables "Health", "Sustainability" and "Price".

The next step in the study was to gather quantitative data to test the stated hypotheses. The team made a list of suitable indicator variables that should be included in the measurement models of each latent variables (see Table 4.13). It took 3 months to gather 1000 responses to an online-distributed survey. 262 cases were removed due to incomplete answers and the remaining 738 are available for downloading here: The respondents were asked: "How would you rate *Allnature Cream Cheese* on each of the following (…)". A 6-point Likert scale, where 1 = "low" and 6 = "high", was used to measure all indicator variables.

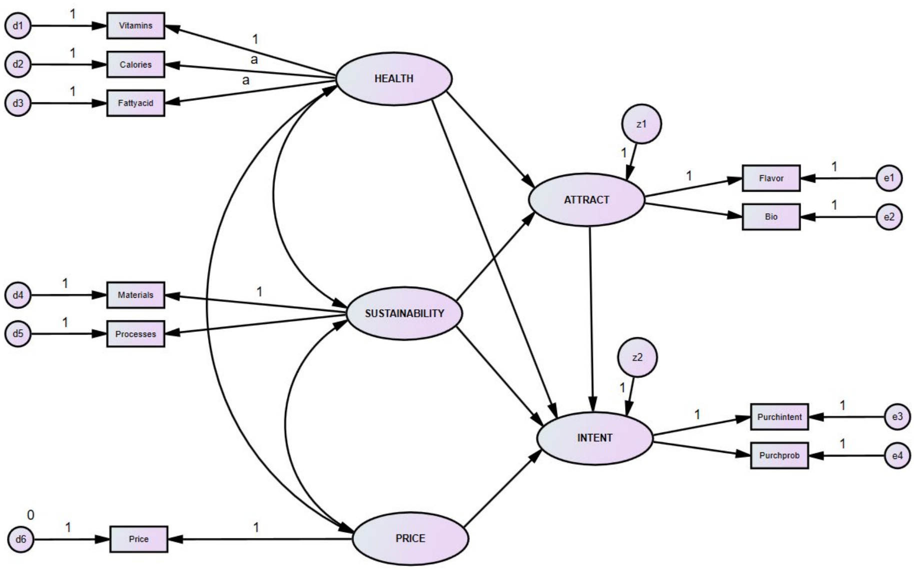


Figure 4.21. Path diagram of cream cheese purchase behaviour

Source: Own work.

The number of parameter estimates in this model is as follows:

- In the structural model there are:
  $\beta_{21}; \gamma_{11}; \gamma_{12}; \gamma_{21}; \gamma_{22}; \gamma_{23}; \zeta_{11}; \zeta_{11} \rightarrow$ 8 parameters.
- In measurement models of the endogenous latent variables, there are:
  $\lambda_{21}; \lambda_{41}; \varepsilon_{21}; \varepsilon_{22}; \varepsilon_{33}; \varepsilon_{44} \rightarrow$ 6 parameters.
- In measurement models of the exogeneous latent variables, there are:
  λ_{21} $(= \lambda_{31}); \lambda_{52}; \delta_{11}; \delta_{22}; \delta_{33}; \delta_{44}; \delta_{55} \rightarrow$ 7 parameters.
- Correlations between the latent exogeneous variables and their variances will also be estimated:
  $\varphi_{11}; \varphi_{21}; \varphi_{22}; \varphi_{31}; \varphi_{32}; \varphi_{33} \rightarrow$ 6 parameters.

The overall number of parameters to be estimated is 27. We have four y – indicator variables and six x – indicator variables, which means that there are $\frac{1}{2}(4+6).(4+6+1)=55$ empirical correlations, variances and covariances. Hence, the degrees of freedom of the model are and this model is identifiable.

Table 4.13. Variables in the model

Latent variable	Indicator variable
Exogeneous variables (ξ)	
ξ_1 : HEALTH	x_1 : Vitamins
	x_2 : Calories
	x_3 : Fattyacids
ξ_2 : SUSTAINABILITY	x_4 : Materials
	x_5 : Processes
ξ_3 : PRICE	x_6 : Price
Endogenous variables (η)	
η_1 : ATTRACT	y_3 : Flavour
	y_3 : Bio
η_1 : INTENT	y_3 : Purchintent
	y_3 : Purchprob

Source: Own work.

Answer the following questions:

1. Which latent variables are part of the structural model?
2. Which variables are part of the measurement model of ‘USAGE’?
3. Which latent variables are endogenous? Why?
4. Which latent variables are exogenous? Why?
5. Can you find any potential problems in the specification of the model in Figure 4.21

6. What approach to the model evaluation would you use and why?

Work individually on the following tasks:

1. Assess the measurement models using Cronbach's alpha.
2. Use an appropriate SEM approach to test the research hypothesis.
3. Apply another SEM approach and compare the results.

Some practical guidelines:

- Latent variable PRICE is explained by only one indicator variable. This is why we assume that the factor loading λ_{63} equals 1. This means that this parameter is fixed and will not be a subject of evaluation. The error term δ_{63} must also be fixed to 0. This is needed in order to apply CB-SEM approach to a model featuring latent variable with only one indicator.
- The latent variable HEALTH reflects equally indicator variables *calories* and *fattyacids*. This is why, in the model in Figure 4.21, λ_{21} equals λ_{31} (both arrows are marked with letter 'a').
- Note that the expected relationship between PRICE and INTENT is negative, while all other relationships between latent variables are expected to be positive.

The data file in .csv format is available here: http://bit.ly/3kMpqQD

References

Addinsoft. (2020, March 1). *XLSTAT PLS-PM* [Computer software]. Retrieved July 27, 2020 from https://help.xlstat.com/s/article/consumer-satisfaction-analysis-in-excel-with-plspm?language=en_US

Agan, Y., Kuzey, C., Acar, M. F., & Açıkgoz, A. (2016). The relationships between corporate social responsibility, environmental supplier development, and firm performance. *Journal of Cleaner Production*, *112*, 1872-1881.

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*(2), 155-173.

Arbuckle, J. L. (2019). *IBM® SPSS® Amos™ 26 user's guide*. Retrieved July 27, 2020 from https://ibm.co/2CRD6bF

Askariazad, M. H., & Babakhani, N. (2015). An application of European Customer Satisfaction Index (ECSI) in business to business (B2B) context. *Journal of Business & Industrial Marketing*, *30*(1), 17-31.

Backhaus, K., Erichson, B., & Weiber, R. (2015). *Fortgeschrittene Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (3. Ausg.). Berlin: Springer Gabler.

Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite reliability in Structural Equations Modelling. *Educational and Psychological Measurement*, *55*(3), 394-406.

Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly*, *27*(3), 459-489.

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, *40*(1), 8-34. https://doi.org/10.1007/s11747-011-0278-x

Bentler, P. M. (2006). *EQS 6 Structural equations program manual*. Encino: Multivariate Software.

Bollen, K. A., & Long, J. S. (1992). Tests for Structural Equation Models. *Sociological Methods & Research*, *21*(2), 123-131. https://doi.org/10.1177/0049124192021002001

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230-258.

Byrne, B. M. (2016). *Structural Equation Modelling with Amos* (3rd ed.). New York: Routledge.

Carmines, E. G., & McIver, J. P. (1983). An introduction to the analysis of models with unobserved variables. *Political Methodology*, *9*(1) (Special Issue on Modelling), 51-102.

Chapman, C., & Feit, E. M. (2019). *R for marketing research and analytics* (2nd ed.). Cham: Springer.

Chin, W. W. (1998a). Commentary: Issues and opinion on Structural Equation Modelling. *MIS Quarterly*, *22*(1), vii-xvi.

Chin, W. W. (1998b). The Partial Least Squares approach to Structural Equation Modelling. In G. A. Marcoulides (Ed.), *Modern business research methods* (pp. 295-336). Mahwah: Lawrence Erlbaum Associates.

Couto, J., Tiago, T., Gil, A., Tiago, F., & Faria, S. (2016). It's hard to be green: Reverse green value chain. *Environmental Research*, *149*, 302-313.

Diamantopoulos, A. (1999). Export performance measurement: Reflective versus formative indicators. *International Marketing Review*, *16*(6), 444-457.

Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL: A guide for the uninitiated*. Los Angeles: Sage.

Drengner, J., Gaus, H., & Jahn, S. (2008, March). Image effects of marketing events: The impact of flow experiences. *Journal of Advertising Research*, *48*(1), 138-147.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, *5*(2), 155-174.

Epskamp, S., Stuber, S., Nak, J., Veenman, M., & Jordensen, T. D. (2019). *semPlot*. Retrieved September 19, 2020 from https://github.com/SachaEpskamp/semPlot

Euromonitor International. (2020). *Top 10 Global Consumer Trends 2020*. Euromonitor International.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 38-50.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996, October). The American Customer Satisfaction Index: Nature, purpose, and findings. *Journal of Marketing*, *60*(4), 7-18.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8 ed.). Andover: Cengage.

Hair, J. F., Hult, G. T., Ringle, C., & Sarstedt, M. (2017). *A primer on Partial Least Squares Structural Equation Modelling (PLS-SEM)* (2nd ed.). Los Angeles: Sage.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139-152.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2012). Partial Least Squares: The better approach to Structural Equation Modelling? *Long Range Planning*, *45*(5-6), 312-319.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced issues in Partial Least Squares Structural Equation Modelling*. Los Angeles: Sage.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2011). An assessment of the use of Partial Least Squares Structural Equation Modelling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414-433.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modelling techniques: MLM and SEM approaches using Mplus* (3rd ed.). New York: Routledge.

Herrmann, A., Huber, F., & Kressmann, F. (2006). Varianz- und Kovarianzbasierte Strukturgleichungsmodelle—Ein Leitfaden zu deren Spezifikation, Schatzung und Beurteilung. *Schmalenbachs Zeitschrift Für Betriebswirtschaftliche Forschung*, *58*(1), 34-66.

Homburg, C., & Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte: Ein Leitfaden für die Marketingforschung. *Marketing: Zeitschrift für Forschung und Praxis*, *18*(1), 5-24.

Hulland, J. (1999). Use of Partial Least Squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, *20*(2), 195-204.

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika*, *69*(1), 81-99.

Hwang, H., & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to Structural Equation Modelling*. New York: Chapman and Hall/CRC.

Jabbour, C. J. C., Jabbour, A. B., Govindan, K., Teixeira, A. A., & Freitas, W. R. (2013). Environmental management and operational performance in automotive companies in Brazil: The role of human resource management and lean manufacturing. *Journal of Cleaner Production*, *47*, 129-140.

Jahn, S. (2007). Strukturgleichungsmodellierung mit LISREL, AMOS and SmartPLS: Eine Einführung. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2729658

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, *30*(2), 199-218.

Jöreskog, K. G. (1970). *A general method for estimating a linear structural equation system*. Princeton: Educational Testing Service. Retrieved from https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.1970.tb00783.x

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberg & O. D. Dunkan (Eds.), *Structural Equation Models in the social sciences* (pp. 255-284). New York: Seminar Press.

Jöreskog, K. G., Olsson, U. H., & Wallentin, F. Y. (2016). *Multivariate analysis with LISREL*. Springer International Publishing. https://doi.org/10.1007/978-3-319-33153-9

Jöreskog, K. G., & Wold, H. (1982). The ML and PLS techniques for modelling with latent variables: Historical and comparative aspects. In K. G. Jöreskog, & H. Wold (Eds.), *Systems under direct observations: Causality, structure, prediction, part 1* (pp. 263-270). Amsterdam: North-Holland.

Large, R. O., & Thomsen, C. G. (2011). Drivers of green supply management performance: Evidence from Germany. *Journal of Purchasing & Supply Management*, 176-184.

Lee, V.-.H, Ooi, K.-B., Yee-Loong, C. A., & Seow, C. (2014). Creating technological innovation via green supply chain management: An empirical analysis. *Expert Systems with Applications*, *41*(16), 6983–6994

Lin, C.-Y., & Syrgabayeva, D. (2016). Mechanism of environmental concern on intention to pay more for renewable energy: Application to a developing country. *Asia Pacific Management Review*, *21*(3), 125-134.

Liu, L., Ren, J., & Liu, X. (2013). *Constructing and analyzing Customer Satisfaction Index of mobile phone video service*. (2013 International Conference on Management Science & Engineering, pp. 777-782). Harbin. https://doi.org/10.1109/icmse.2013.6586367

Lohmöller, J.-B. (1989). *Latent variable path modelling with Partial Least Squares*. Heidelberg: Physica--Verlag.

Mardani, A., Streimikiene, D., Zavadskas, E. K., Cavallaro, F., Nilashi, M., Jusoh, A., & Zare, H. (2017). Application of Structural Equation Modelling (SEM) to solve environmental sustainability problems: A comprehensive review and meta-Analysis. *Sustainability*, *9*(10). https://doi.org/10.3390/su9101814

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. In G. A. Marcoulides, & R. E. Schuhmacher, *Advanced Structural Equation Modelling: Issues and techniques* (pp. 315-353). Mahwah: Lawrence Erlbaum Associates.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self--concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, *97*(3), 562-582.

Monecke, A. (2015). *'semPLS' Structural Equation Modelling using Partial Least Squares. R package version 1.0-10*. Retrieved from https://cran.r-project.org/web/packages/semPLS/semPLS.pdf

Monecke, A., & Leisch, F. (2012). semPLS: Structural Equation Modelling using Partial Least Squares. *Journal of Statistical Software*, *48*(3), 1-32. Retrieved from https://www.jstatsoft.org/article/view/v048i03

Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical analysis with latent variables (user's guide)* (8th ed.). Los Angeles: Muthén & Muthén.

O'Loughlin, C., & Coenders, G. (2002). *Application of the European Customer Satisfaction Index to postal services. Structural Equation Models versus Partial Least Squares*. Girona: Universitat de Girona. Retrieved from https://core.ac.uk/download/pdf/6548589.pdf

O'Loughlin, C., & Coenders, G. (2004). Estimation of the European Customer Satisfaction Index: Maximum likelihood versus Partial Least Squares. Application to postal services. *Total Quality Management & Business Excellence*, *15*(9-10), 1231-1255. https://doi.org/10.1080/1478336042000255604

Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in Structural Equation Modelling under conditions of misspecification and nonnormality. *Structural Equation Modelling: A Multidisciplinary Journal*, *7*(4), 557-595. https://doi.org/10.1207/s15328007sem0704_3

Peter, J. P. (1979). Reliability: A review of psychometric basics and recent marketing practices. *Journal of Marketing Research*, *16*(1), 6-17.

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences* (6th ed.). New York: Routledge.

Ravand, H., & Baghaei, P. (2016). Partial Least Squares Structural Equation Modelling with R. *Practical Assessment, Research, and Evaluation*, *21*(11). https://doi.org/10.7275/d2fa-qv48

Ringle, C. M. (2004). *Gütemaße* für den Partial Least *Squares Ansatz zur Bestimmung von Kausalmodellen*. Universität Hamburg, Instituts für Industriebetriebslehre und Organisation. Hamburg: K.-W. Hansmann.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3*. Retrieved from www.smartpls.com

Rosseel, Y. (2012, May). Iavaan: An R package for Structural Equation Modelling. *Journal of Statistical Software*, *48*(2). https://doi.org/10.18637/jss.v048.i02

Sachs, J. D. (2015). *The age of sustainable development*. New York: Columbia University Press.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies!. *Journal of Business Research*, *69*(10), 3998-4010.

Schneeweiss, H. (1991). Models with latent variables: LISREL versus PLS. *Statistica Neerlandica*, *45*(2), 145-157.

Scholderer, J., & Balderjahn, I. (2005). PLS versus LISREL: Ein Methodenvergleich. In F. Bliemel, A. Eggert, G. Fassott, & J. Henseler (Eds.), *Handbuch PLS-Pfadmodellierung. Methode, Anwendung, Praxisbeispiele* (pp. 87-98). Stuttgart: Schäffer-Poeschel.

StataCorp. (2017). *STATA: Structural Equation Modelling* (15th ed.). College Station: A Stata Press Publishing. Retrieved from https://www.stata.com/manuals15/sem.pdf

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modelling. *Computational Statistics & Data Analysis*, *48*(1), 159-205.

Venturini, S., & Mehmetoglu, M. (2019). plssem: A stata package for Structural Equation Modelling with Partial Least Squares. *Journal of Statistical Software*, *88*(8). https://doi.org/10.18637/jss.v088.i08

Weiber, R., & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung: Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* (2. Aufl. Ausg.). Berlin: Springer Gabler.

Westland, J. C. (2019). *Structural equation models—from paths to networks* (2nd ed.). Cham: Springer.

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociological Methodology*, *8*, 84-136.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & R. Capecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modelling* (pp. 307-359). New York: Academic Press.

Xu, X., & Gursoy, D. (2015). Influence of sustainable hospitality supply chain management on customers' attitudes and behaviors. *International Journal of Hospitality Management*, *49*, 105-116.

Qualitative and quantitative methods in sustainable development, pp. 179-205
https://doi.org/10.18559/978-83-8211-072-2/05

5.

DATA ENVELOPMENT ANALYSIS METHODS IN SUSTAINABLE AGRICULTURAL DEVELOPMENT RESEARCH

Katarzyna Smędzik-Ambroży
Poznań University of Economics and Business

Agnieszka Sapa
Poznań University of Economics and Business

Abstract: Sustainable development of business entities can be analysed in terms of three dimensions, i.e., economic, social and environmental ones. The economic dimension of sustainable development can be assessed, *inter alia*, by entities' technical efficiency defined as the relation of outputs to inputs. One of the methods that is used to assess the technical efficiency of business entities compared to other entities is the Data Envelopment Analysis (DEA) method.

The aim of the chapter is to determine the relative technical efficiency of representative agricultural farms from the individual European Union countries in 2018. Moreover, the scale efficiency indexes and the area of scale effects (increasing or decreasing) of the analysed farms were also determined. In the study the data from the Farm Accountancy Data Network (FADN) for 2018 were applied.

In order to achieve the assumed research goals, the input-oriented DEA model was used, and the technical efficiency indexes of farms were estimated with the assumption of constant return to scale (CRS) and variable return to scale (VRS). This allowed, among others, for indicating the countries with farms achieving the highest technical efficiency (Belgium, Spain, Italy, Malta and Netherlands assuming CRS, and Belgium, Spain, Italy, Malta and Netherlands, Greece, Ireland, Romania and Slovenia assuming VRS), the lowest technical efficiency (the Czech Republic and Slovakia) within surveyed group of farms. All relatively inefficient farms (except Slovakia) functioned in the area of increasing economies of scale.

Keywords: DEA method, economics sustainability, effect of scale, farms, technical efficiency.

5.1. DEA—theoretical background

In production process inputs are converted into effects. Efficiency describes how effectively the company transforms inputs into effects. The measurement of efficiency is very important for the company because it informs, e.g., if the assumed goals were achieved or not. Moreover, it allows to compare the achieved efficiency level of a particular company with the results of other similar units. One of the popular methods used to determine the relative efficiency of the examined units is the Data Envelopment Analysis (DEA) method. The DEA method was proposed by Charnes, Cooper and Rhodes in 1978, the so-called CCR model (abbreviation of the first letters of the authors' surnames) (Charnes, Cooper, & Rhodes, 1978). It is an extension of Farrell's (1957) work on technical efficiency estimation.

The DEA method is used to measure the relative efficiency of the selected objects (units) in a situation where the units use many inputs simultaneously and achieve many effects. Before proceeding with further analysis, the notion of relative efficiency and the difference between efficiency (i.e., the relation of effects to inputs) and relative efficiency should be introduced. Table 5.1 presents four objects and each of them is described by one effect (y) and one input (x). The highest efficiency is attributed to unit A, as it transforms a given input into an effect in the best way. To determine the relative efficiency index, all examined units should be compared to the best units, in this case to object A.

Table 5.1. Efficiency and related efficiency of the selected objects A–D

Unit	Input (x)	Output (y)	Efficiency (y/x)	Related efficiency indicator (%)
A	18	125	6,94	100
B	16	44	2,75	40
C	17	80	4,71	68
D	11	23	2,09	30

Source: (Domagała, 2007, p. 24–25).

The efficient unit, i.e., the one whose relative efficiency index is 100% (unit A), determines the efficiency frontier (also the production possibilities curve of the examined objects, Figure 5.1). All units on this curve achieve a relative efficiency equal to 100%, i.e., they are efficient units. The remaining units are inefficient units, e.g., units C, D, E (see Figure 5.1). When the efficiency frontier is indicated, it can be determined how the inefficient units can approach the efficiency frontier. The improvement of the relative efficiency index of inefficient units can be reached by:

– increasing (maximizing) the effects with unchanged inputs (the so-called outputs orientation; unit C’),
– reducing (minimizing) inputs for given effects (the so-called inputs orientation; unit C”),
– the lack of orientation on effects or inputs (the so-called mixed approach; units lying between units C” and C’).

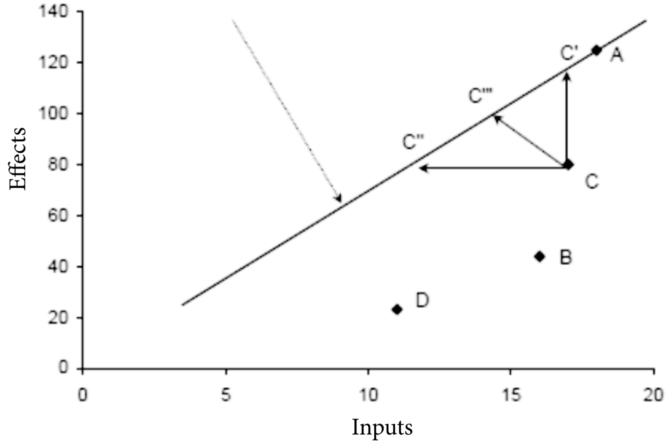


Figure 5.1. The production possibilities curve

Source: (Domagała, 2007, pp. 24–25).

In practice, units are rarely described by one input and one effect. This means that it becomes impossible to assess the efficiency of the examined units, as in the example above (Table 5.1, Figure 5.1). Then the DEA method is useful, as in this method it is assumed that the measure of relative efficiency is the relation of the weighted sum of the outputs and the weighted sum of the inputs. This is a relative efficiency determined for a specific set of objects called DMUs (Decision Making Units).

$$\textit{relative efficiency} = \frac{\text{weighted sum of the outputs}}{\text{weighted sum of the inputs}}$$

The DEA method is a frontier non-parametric method of estimating efficiency. This method allows for identification of the efficiency frontier (the production possibilities curve). The efficiency frontier is constructed in a non-parametric manner using linear programming techniques. This means that the efficiency frontier is spread over the best units (efficient DMUs, called frontiers) in a given studied group of units. Thus, the efficiency frontier is not spread on units determined by a specific production function or by the adopted specific values of inputs and outputs. It is difficult to indicate both an efficiency benchmark level and an objective (functional) relationship between inputs and outputs. So, identification of the units' efficiency only within the group of units and setting out the efficiency benchmarks based on the best units in the examined group can be a useful solution. Since the DEA method does not require the adoption of many predetermined limitations and assumptions, it has become popular and is widely used in social science research, both at the microeconomic and macroeconomic levels.

The DEA method enables to rank the surveyed units according to their efficiency indexes and to determine the relative differences between them in this respect. The object with the highest efficiency index (the so-called frontier) is selected within an examined group of units. The remaining units from the analysed group are considered inefficient units. The efficient unit is the reference (the benchmark) for the other objects, the efficiency indicators of which were evaluated in relation to this efficient unit.

The basic DEA-CCR model (the input-oriented model) can be presented as:
where:

$$\max h_0 = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

assumed that:

$$\mathrm{x} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1; \ j = 1, \ldots, n,$$

$$v_r v_i \geq 0; \quad r = 1, \ldots, s; \quad i = 1, \ldots, m.$$

n—number of decision-making units DMUs (j = 1, …, n),
s—number of outputs (r = 1, …, s),
m—number of inputs (i = 1, …, m),
x_{ij}—value of i input of decision-making unit j,
y_{rj}—value of r output of decision-making unit j,
v_i—decision variable; weight related to i input,
u_r—decision variable; weight related to r input,
o—index of examined decision making unit, $1 \leq o \leq n$,
h_o—efficiency index of object o.

Technical efficiency indicators, estimated using the DEA method, range from 0 to 1. Depending on whether the input-oriented or the output-oriented DEA model is used, the difference between the value of the estimated indicator for a given object and unity means:
- how much a given object should proportionally reduce its inputs without changing its outputs in order to achieve full efficiency (input-oriented model),

– how much a given object should proportionally increase its outputs without changing the level of used inputs in order to achieve full efficiency (output-oriented model).

For example, if in input-oriented DEA model the estimated value of the technical efficiency index for a given object is 0.7, it means that the object should reduce its inputs by 30%. In other words, the object should reduce their inputs to 70% of the current inputs level to manage its inputs in a fully efficient way and to achieve the same relation of outputs to inputs as the best units (frontiers).[1] In the case of output-oriented DEA model, the estimated technical efficiency index at the level of 0.7 for a given object means that it should increase its effects by 30% in order to achieve full efficiency with the given inputs.

It is worth underlining that regardless of the model orientation (input or output), the higher the efficiency index of a given unit, the higher its efficiency. It should also be emphasized that the achieved efficiency indicators in DEA methods are always relative, so adding a new unit to the analysis may change their values.

Analysis with DEA can assume constant or variable returns to scale. The basic version of DEA model assumes constant returns to scale (DEA–CRS). CRS implies linearity between inputs and outputs, meaning that doubling the inputs used, will double the outputs, which is rare in practice. Banker, Charnes and Cooper modified the basic DEA model by introducing the assumption of variable returns to scale (VRS). This is BCC model (DEA–BCC). Thus, the adoption of this approach allows to estimate the technical efficiency without the CRS assumption.

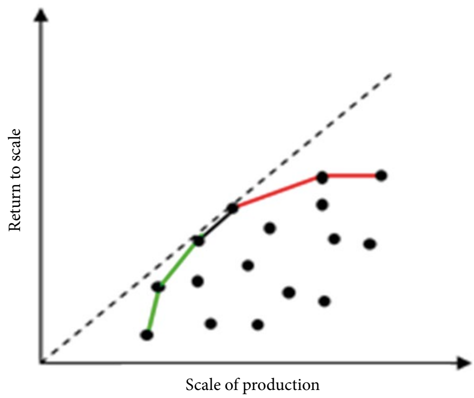


Figure 5.2. Relation between the scale of production and return to scale

Source: Own elaboration based on (Guzik, 2009; Czyżewski, Smędzik-Ambroży & Mrówczyńska-Kamińska, 2020).

[1] It should be emphasized that in the case of the input-oriented model, the main assumption is full complementarity of inputs and zero degree of their substitution (Guzik, 2009).

Considering VRS assumption, it should be stated that a company may experience increasing return to scale (IRS) or decreasing return to scale (DRS). IRS means a situation where increasing the inputs by *t*-times affects the outputs increase by more than *t*-times (green line, Figure 5.2). In some companies, especially with a very huge scale of production, the decreasing return to scale (DRS) can be observed. It results from the law of diminishing marginal productivity and is an unfavourable situation (red line, Figure 5.2). It means that an increase in inputs by *t*-times is accompanied by outputs increase by less than *t*-times. The company's goal is to achieve the optimal production scale and after reaching that the company can operate within decreasing return to scale conditions. The type of return to scale (CRS or DRS) in which a given unit operates can also be determined using the DEA method. In this case, the model with non-increasing returns to scale (NIRS) can be applied.

The next index the DEA enables to estimate is the production scale efficiency index. It is the relation between the unit efficiency index assuming constant return to scale and the unit efficiency index assuming variable return to scale. The production scale efficiency index takes values from 0 to 1. The index informs how much less inputs could be used if the outputs volume were optimal. The index equal to 1 means the optimal production (output) scale in a given group of units.

Summarizing the above, it can be concluded that the essence of the DEA model, is, among others:

– searching for the best units in the examined group of objects (the best ones create the efficiency frontier) in a situation where there are many outputs and inputs, which can be expressed in non-monetary units,
– identifying inefficient units (in Farrel's efficiency sense) and creating their ranking,
– indicating the inefficient objects the distance from the efficiency frontier (i.e., the size of their inefficiency; output surplus or input deficit),
– indicating units with an optimal production scale in the examined group of units (the production scale efficiency),
– determining the area of economies of scale within a given unit.

The basic DEA model (DEA–CCR) became the basis for modifying and developing this method. It allowed for the creation of many varieties of DEA, e.g., the aforementioned BCC model (Bankers, Charnes, & Cooper, 1984), but also the SBM model (slack-based model) (Charnes, Cooper, Golany, Seiford, & Stutz, 1985), CEM model (cross-efficiency model) (Sexton, Silkman, & Hogan, 1986) and many others.

The DEA procedure is presented in the second section. It allows for imagining the main steps you have to take to achieve the assumed goals of the research. The third part of this chapter includes the case study where the DEA method is employed. The application of the DEA input-oriented model in its classic basic variant including CCR model (assuming constant return to scale—CRS) and the BBC model (assuming the variable return to scale—VRS) is presented. In the last part of this chapter, you can find some tasks and questions related to the DEA method.

5.2. DEA procedure: main steps

5.2.1. Aims of research and data (inputs and outputs) selection

Step 1
At the beginning, you have to define the aim of you research. If your goal is to assess efficiency of units in comparison to other units, the DEA method is appropriate. You can evaluate efficiency of units at micro or macro level, so you can compare efficiency of particular firms, farms, institutions, sectors or countries.

Step 2
When you have the aim of your research, in the next step you have to define and prepare the input and output variables. Because of DEA features, you are allowed to take multiple inputs and outputs. You can consider, e.g., three inputs and two outputs in the efficiency estimation instead of the single efficiency calculation (only one output divided by only one input).

The number of inputs and outputs are not limitless and depends on the number of analysed entities. One of the rough rules of numbs in the DEA method is to take the number of units equal or greater than three times the sum of the total number of inputs and outputs (3 · (inputs + outputs)).

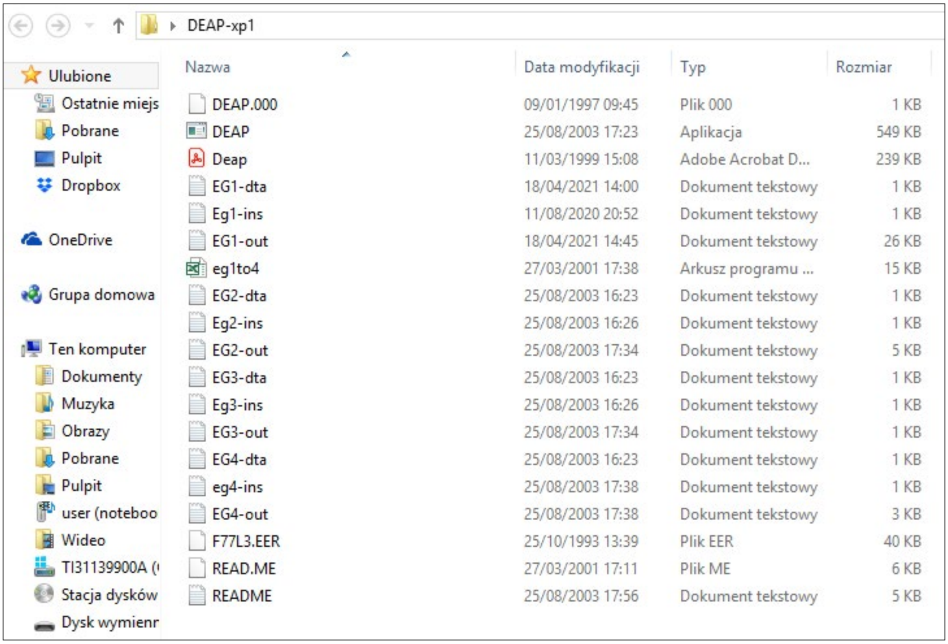
It is important to know, that DEA does not accept negative or zero values for inputs and outputs. In this case, you can employ technics to avoid such a situation, e.g., zero value can be replaced by very low values such as 0.01.

When preparing the data set, you have to remember that each analysed unit must have the same number of inputs and outputs in order to be compared.

You can use primary or secondary data. The first group of data can be gathered during your own primary survey. The second group of data can be retrieved, e.g., from public databases that embrace different range of data at local, country, international or sectoral level, etc. (World Bank, UNCTAD, FADN).

Next, your selected outputs and inputs should be prepared in a table form in Excel file. And considering the requirements of the DEA software, the selected outputs should be put in columns before columns with inputs.

Step 3



Picture 5.1. Folder DEAP-xp1

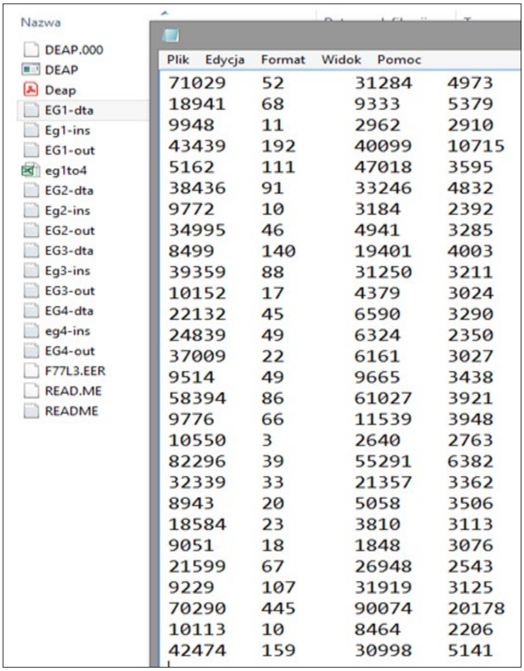
To engage DEA model and make all calculations you have to install the software. The core calculations are made using the DEAP computer program, used by the Centre for Efficiency and Productivity Analyses at the University of Queensland. DEAP is a free software constructed by Tim Coelli and can be downloaded from: https://economics.uq.edu.au/cepa/software (Centre for Efficiency). After the installing process, the DEAP-xp1 folder (Picture 5.1) is available.

Of course, you can download DEAP software at the very beginning of the research process. Here is the last time you have to do it, if you want to employ DEA model.

Step 4

In the next stage, the prepared data file should be imported to DEA program. It will be saved in the EG1-dta file (Picture 5.2).

Remember that you cannot put text data, but numbers in this file. Therefore, the names of variables (inputs, outputs) and names of DMUs (analysed units) cannot be included in the table. So, in the final file, the analysed units will be numbered, i.e., 1, 2, 3, etc. A properly prepared data file looks like the one in Picture 5.2.


Nazwa
DEAP.000
DEAP
Deap
EG1-dta
Eg1-ins
EG1-out
eg1to4
EG2-dta
Eg2-ins
EG2-out
EG3-dta
Eg3-ins
EG3-out
EG4-dta
eg4-ins
EG4-out
F77L3.EER
READ.ME
README

Plik Edycja Format Widok Pomoc

71029	52	31284	4973
18941	68	9333	5379
9948	11	2962	2910
43439	192	40099	10715
5162	111	47018	3595
38436	91	33246	4832
9772	10	3184	2392
34995	46	4941	3285
8499	140	19401	4003
39359	88	31250	3211
10152	17	4379	3024
22132	45	6590	3290
24839	49	6324	2350
37009	22	6161	3027
9514	49	9665	3438
58394	86	61027	3921
9776	66	11539	3948
10550	3	2640	2763
82296	39	55291	6382
32339	33	21357	3362
8943	20	5058	3506
18584	23	3810	3113
9051	18	1848	3076
21599	67	26948	2543
9229	107	31919	3125
70290	445	90074	20178
10113	10	8464	2206
42474	159	30998	5141

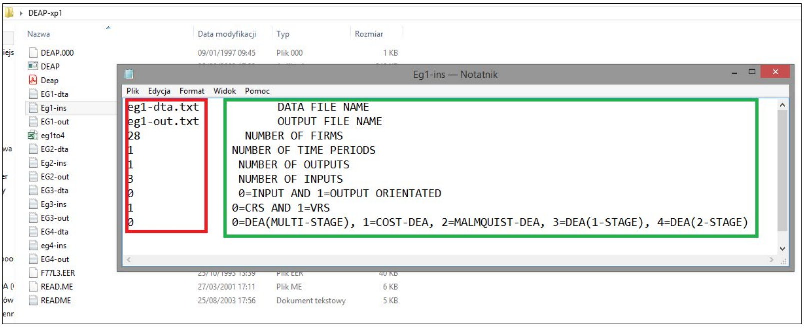
Picture 5.2. A ready-to-use spreadsheet in EG1-dta.txt file

5.2.2. Model calibration and calculation

Step 5
To run a DEA model, the user has to calibrate the model. There are some parameters the researcher should consider. These parameters can be found in the Eg1-ins file (Picture 5.3). In the right column there are the names of parameters (green frame in Picture 5.3) and in the left column there is a place [space] to be filled with the values of parameters according to the particular data set (red frame in the Picture 5.3):

– number of firms (number of analysed objects),
– number of time periods (numbers of analysed years),
– number of outputs (numbers of engaged outputs),
– number of inputs (numbers of engaged inputs),
– 0 = input and 1 = output oriented (you can choose between input-oriented model or output-oriented. 0 means the assumption of input-oriented model and 1 means the assumption of the output-oriented model),
– 0 = CRS and 1 = VRS (you can choose the model with constant returns to scale (CRS) or the model with variables returns to scale (VRS); 0 means the assumption of CRS and 1 means the assumption VRS. If the user puts 1, the program

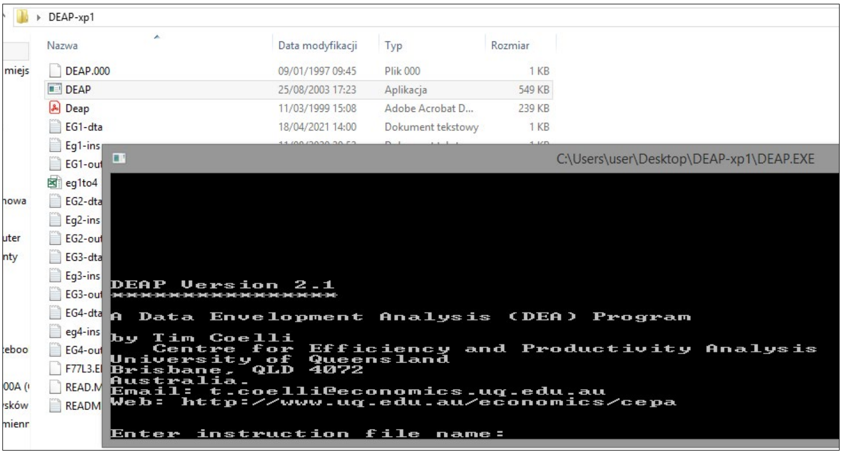
gives the efficiency indicators for both the CRS and VRS models, which is necessary to get the production scale efficiency indicators. If the user puts 0, the program calculates the efficiency indicators only for CRS model),
- the description of the econometric procedure for determining effectiveness. A more detailed description of that part of the procedure can be found in the DEAP manual included with the software folder (Coelli, 1996).



Picture 5.3. Model calibration in Eg1-ins file

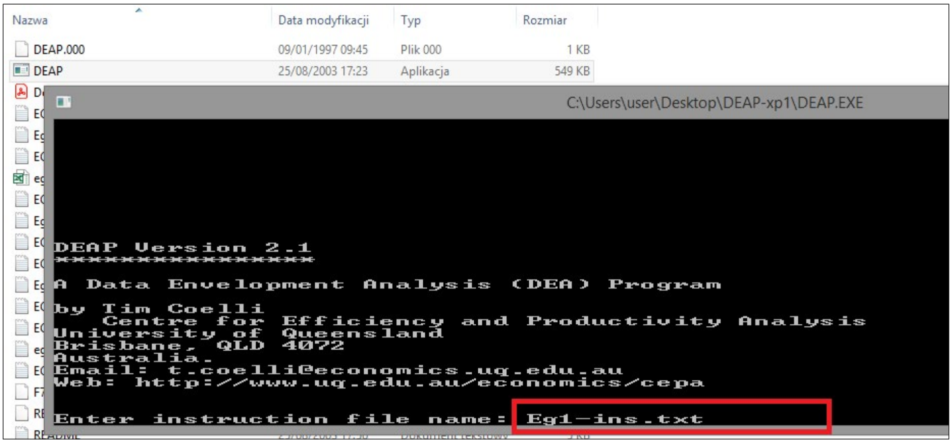
Step 6

The next step is to let the program make calculations. To obtain the final results, the DEAP file should be opened, and the black frame will appear (Picture 5.4).



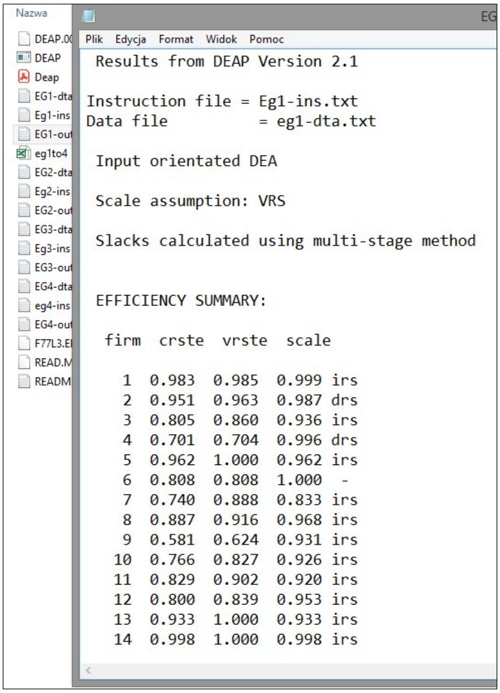
Picture 5.4. The DEAP file opened (i)

Enter the name of the configuration file, i.e., Eg1-ins.txt in the place of the blinking cursor (red frame in the Picture 5.5). Next, press enter, and the window will close automatically.



Picture 5.5. The DEAP file opened (ii)

The researcher can find the results of the calculation in the EG1-out file (Picture 5.6).



Picture 5.6. Results in the EG1-out file (i)

In Table 5.2. there are some abbreviations that help you to calibrate the model and read the results.

Table 5.2. Abbreviations used in DEA calculation

CRS	Constant Returns to Scale
VRS	Variable Returns to Scale
CRSTE	Constant Returns to Scale Technical Efficiency
VRSTE	Variable Returns to Scale Technical Efficiency
Scale	Scale Efficiency
IRS	Increasing Returns to Scale
DRS	Decreasing Returns to Scale

Own elaboration.

5.2.3. Results interpretation

Step 7

The next step is the interpretation of the results that are obtained in the calculation process. The table with the results is presented in Picture 5.7.

The first column (firms) contains the numbers of the analysed units. The second column (crste) includes technical efficiency indicators for selected units assuming the constant returns to scale (CRS). The third column (vrste) presents the technical efficiency indexes of the analysed units assuming variable returns to scale (VRS). In the next column (scale) there are the scale efficiency indicators, while in the last column there are the scale effects areas (IRS or DRS).

If the production scale of the analysed objects is optimal, then the scale effects area is not marked in the last column (e.g., no abbreviation in the fifth column for firm no. 1, Picture 5.7). The scale effects area (IRS or DRS) can only be determined in units with a non-optimal production scale, i.e., those where the scale efficiency index is lower than 1 (present abbreviation in the fifth column for firms no. 2–7, Picture 5.7).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

    1  1.000  1.000  1.000  -
    2  0.311  0.473  0.656 irs
    3  0.557  0.917  0.607 irs
    4  0.311  0.315  0.985 irs
    5  0.097  0.614  0.159 irs
    6  0.555  0.627  0.886 irs
    7  0.543  1.000  0.543 irs
    8  1.000  1.000  1.000  -
    9  0.156  0.551  0.284 irs
   10  0.845  0.944  0.895 irs
   11  0.382  0.788  0.485 irs
   12  0.555  0.771  0.719 irs
   13  0.843  1.000  0.843 irs
   14  1.000  1.000  1.000  -
   15  0.220  0.663  0.331 irs
   16  1.000  1.000  1.000  -
   17  0.196  0.575  0.340 irs
   18  1.000  1.000  1.000  -
   19  1.000  1.000  1.000  -
   20  0.686  0.836  0.820 irs
   21  0.291  0.680  0.428 irs
   22  0.754  0.921  0.819 irs
   23  0.692  1.000  0.692 irs
   24  0.583  0.912  0.640 irs
   25  0.203  0.706  0.288 irs
   26  0.260  0.341  0.762 drs
   27  0.495  1.000  0.495 irs
   28  0.584  0.634  0.920 irs

 mean  0.576  0.795  0.700

Note: crste = technical efficiency from CRS DEA
      vrste = technical efficiency from VRS DEA
      scale = scale efficiency = crste/vrste
```

Picture 5.7. Results in the EG1-out file (ii)

When interpreting the results, it is worth remembering that all the obtained indicators are relative, i.e., their values are estimated in comparison to other entities. Adding or subtracting another unit could change the obtained results of all analysed units. This applies to both the efficiency indicators (assuming CRS or VRS) and the scale efficiency indicators. It should also be emphasised that results equal to 1 are always present for units achieving the highest technical efficiency (the highest relation of outputs to inputs in the analysed sample) and / or the scale efficiency. Adding or subtracting one unit can cause that fully efficient units (efficiency score equals to 1) could not be such units in a new research sample.

There are some main elements / aspects that should be indicated and interpreted:

- efficiency scores assuming constant returns to scale technology; units which are granted 100% efficiency scores (efficiency equals to 1) and units which are inefficient (column 2, Picture 5.7);
- efficiency scores assuming variable returns to scale technology; units which are granted 100% efficiency scores and units which are inefficient (column 3, Picture 5.7);
- scale of production in relation to efficient and inefficient units assuming constant or variable returns to scale (column 4, Picture 5.7);
- the level of input reduction (or augment of outputs) in order to become efficient maintaining output (or input) level;
- mean values for the analysed group of units in relation to efficiency score.

You can find a more detailed interpretation in two examples in part 3.3. The first example is devoted to the analysis of the relative economic efficiency of farms in the European Union. Then the analysis is in-depth at more disaggregated level as the second example is concentrated only on relative economic efficiency of crops farms in European Union.

5.3. Comparison of farms' efficiency in the European Union: case study no. 1

5.3.1. Aims of research and data selection from FADN

Step 1

In our case study: the main aim of the research is to determine the relative economic efficiency of representative agriculture farms from the European Union countries in 2018. In other words, we want to compare agricultural farms in the EU considering their efficiency. The production scale efficiency and areas of scale effects (IRS or DRS) will also be estimated.

Step 2
In the next step, the input and output variables should be defined. In our case, to estimate efficiency index for agricultural farm and compare them, the FADN database is used. FADN (Farm Accountancy Data Network) is a system for collecting accountancy data from agricultural farms in each country of the European Union.

The data collected under FADN structure are used for the annual description of the income of farms operating in the individual EU countries, the analysis of the activity of farms and the assessment of the effects of implemented and planned changes in the EU agricultural policy. The FADN was developed as a harmonized system of sample surveys, using precisely defined terms with a precisely developed method of selecting a sample of farms and transparent control procedures. As a result, the data obtained by the FADN are reliable and representative, what determines the possibility of reflecting the actual results of farms operating in the EU countries (Goraj & Olewnik, 2011). The data collected under the FADN is publicly available and published on the website: https://ec.europa.eu/agriculture/rica/database/database_en.cfm.

To achieve comparability between the variables used in the FADN from individual EU countries, each variable that occurs in the FADN database is precisely defined. These variables are described by symbols SE with a specific number, e.g., SE011. To explain the individual abbreviations, take a look at the diagrams presenting the creation of individual variables of FADN available on the website: https://ec.europa.eu/agriculture/rica/annex003_en.cfm.

Technical efficiency can express the economic dimension of the farm sustainability. For the measurement of economics sustainability of farms, the researchers use e.g.: the value of income per person or farm, the number of the holding's expenses, less often the farm wage level. Among other measures of economic dimension of farm sustainability, the employment and professional activity indicators, workforce productivity, fixed asset capital intensity, and energy intensity indicators, investment level, outlays on research and development activity are the most commonly used in the literature.

In our case study, on the input side, we use annual values of three variables: 1) depreciation in EUR (SE 360), 2) labour input in hours (SE 011), and 3) land inputs expressed in Utilized Agricultural Area in hectares (SE 025).

On the output side the annual farm net income in EUR (SE 420) is applied.

As mentioned before, the selected input and output variables for individual EU countries from FADN database should be downloaded and saved in Excel file. Considering the requirements of the software DEA, the selected outputs should be put before inputs in data table (Table 5.3).

Table 5.3. Output and input values of farms in selected EU countries in 2018

No.	Country	Farm Net Income	Utilised Agricultural Area	Depreciation	Labour input
1	Belgium	71029	52	31284	4973
2	Bulgaria	18941	68	9333	5379
3	Cyprus	9948	11	2962	2910
4	Czech Republic	43439	192	40099	10715
5	Denmark	5162	111	47018	3595
6	Germany	38436	91	33246	4832
7	Greece	9772	10	3184	2392
8	Spain	34995	46	4941	3285
9	Estonia	8499	140	19401	4003
10	France	39359	88	31250	3211
11	Croatia	10152	17	4379	3024
12	Hungary	22132	45	6590	3290
13	Ireland	24839	49	6324	2350
14	Italy	37009	22	6161	3027
15	Lithuania	9514	49	9665	3438
16	Luxembourg	58394	86	61027	3921
17	Latvia	9776	66	11539	3948
18	Malta	10550	3	2640	2763
19	Netherlands	82296	39	55291	6382
20	Austria	32339	33	21357	3362
21	Poland	8943	20	5058	3506
22	Portugal	18584	23	3810	3113
23	Romania	9051	18	1848	3076
24	Finland	21599	67	26948	2543
25	Sweden	9229	107	31919	3125
26	Slovakia	70290	445	90074	20178
27	Slovenia	10113	10	8464	2206
28	United Kingdom	42474	159	30998	5141

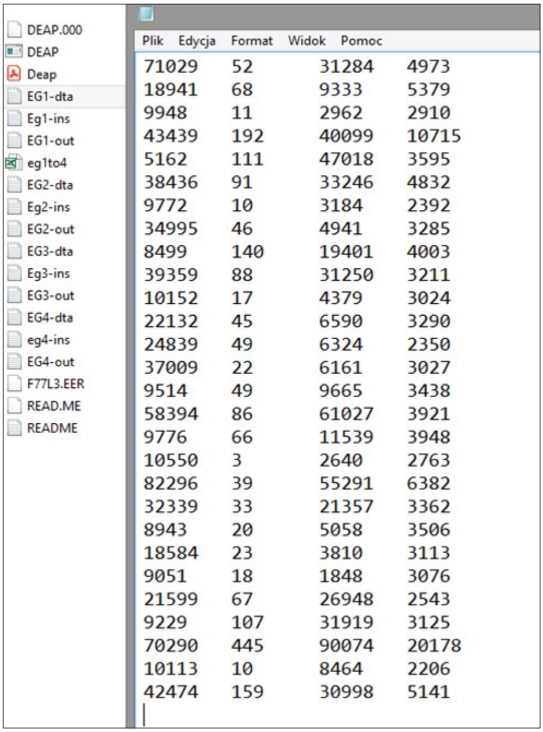
Source: Own survey based on FADN database.

Step 3

The next step: install DEAP software following steps presented in point 5.2.1. *Step 3*. You can skip that step if you already have DEAP software.

Step 4
In the next step, the downloaded data from FADN prepared in a suitable table (see Table 5.3) should be imported to DEA program and saved in the EG1-dta file (Picture 5.8).

As it was said, you have to put only numbers not text data in this file. Therefore, the names of inputs (Utilised Agricultural Area; Depreciation, Labour Input) and output (Farm Net Income), and names of DMUs (Belgium, Bulgaria, Cyprus, etc.) cannot be included in the table. In our case the prepared data file looks like in Picture 5.8.



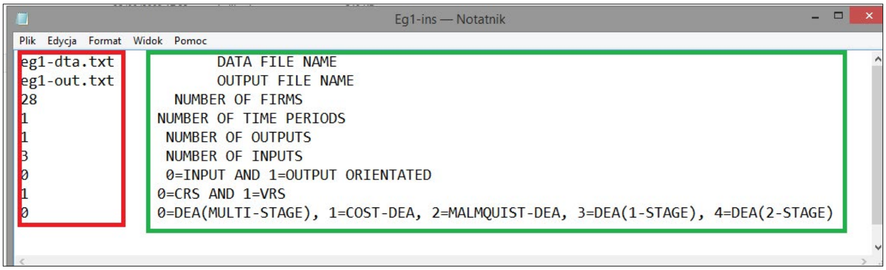
71029	52	31284	4973
18941	68	9333	5379
9948	11	2962	2910
43439	192	40099	10715
5162	111	47018	3595
38436	91	33246	4832
9772	10	3184	2392
34995	46	4941	3285
8499	140	19401	4003
39359	88	31250	3211
10152	17	4379	3024
22132	45	6590	3290
24839	49	6324	2350
37009	22	6161	3027
9514	49	9665	3438
58394	86	61027	3921
9776	66	11539	3948
10550	3	2640	2763
82296	39	55291	6382
32339	33	21357	3362
8943	20	5058	3506
18584	23	3810	3113
9051	18	1848	3076
21599	67	26948	2543
9229	107	31919	3125
70290	445	90074	20178
10113	10	8464	2206
42474	159	30998	5141

Picture 5.8. A ready-to-use spreadsheet in EG1-dta.txt file

5.3.2. Model calibration and calculation

Step 5
In the next step the model is calibrated according to our assumption. The calibration is made in the Eg1-ins file. Considering our research goal and data set, the left column should be fulfilled (red frame in Picture 5.9):

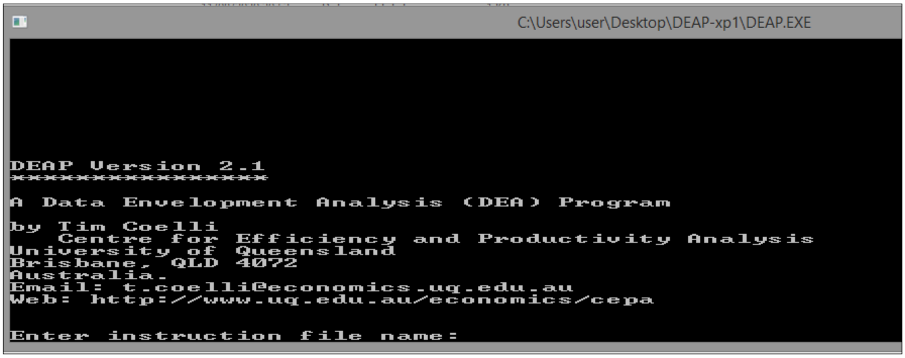
- 28 (number of firms; number of analysed objects, in the presented case study 28 countries)
- 1 (number of times; 1 because the case study concerns only 2018),
- 1 (number of outputs; 1 because there is only one output Farm Net Income in the case study),
- 3 (number of inputs; there are three inputs in our case study: Utilized Agricultural Area, Depreciation and Labour Input),
- 0 (model orientation; in our case study, the input-oriented model was assumed),
- 1 (model with variables return to scale is adopted),
- 0 (the multistage econometric procedure was applied).



Picture 5.9. Model calibration in Eg1-ins file

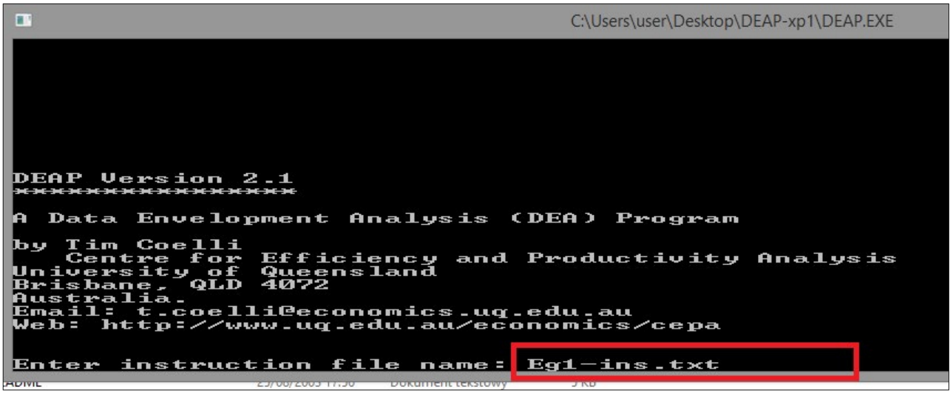
Step 6

To calculate and to obtain final result, the DEAP file should be opened, and the black frame will appear (Picture 5.10).



Picture 5.10. The DEAP file opened (i)

In the place of the blinking cursor (red frame in Picture 5.11), enter the name of the configuration file, i.e., Eg1-ins.txt. Next, press enter, and the window will close automatically.



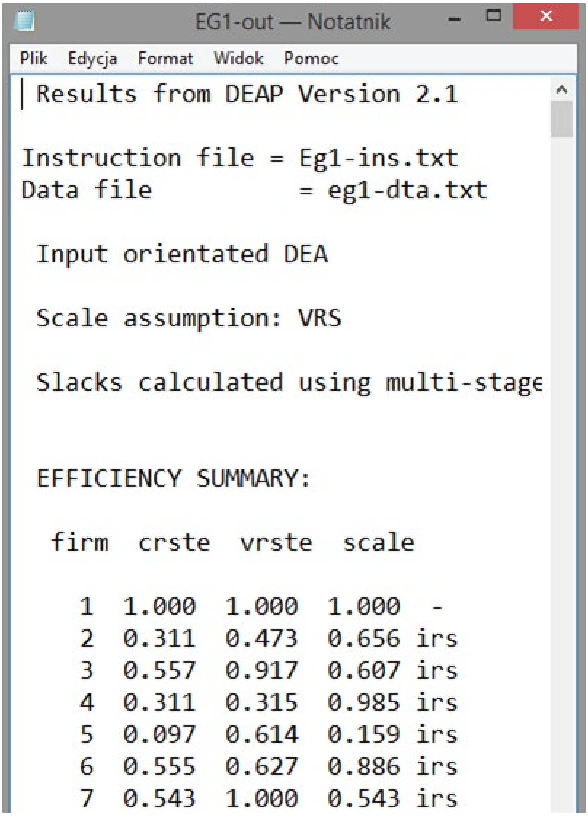
C:\Users\user\Desktop\DEAP-xp1\DEAP.EXE

```
DEAP Version 2.1
*******************
A Data Envelopment Analysis (DEA) Program
by Tim Coelli
   Centre for Efficiency and Productivity Analysis
University of Queensland
Brisbane, QLD 4072
Australia.
Email: t.coelli@economics.uq.edu.au
Web: http://www.uq.edu.au/economics/cepa

Enter instruction file name: Eg1-ins.txt
```

Picture 5.11. The DEAP file opened (ii)

The final results of calculation can be found in the EG1-out file (Picture 5.12). Having it the researcher should interpret the obtained values.



EG1-out — Notatnik

Plik Edycja Format Widok Pomoc

```
 Results from DEAP Version 2.1

Instruction file = Eg1-ins.txt
Data file         = eg1-dta.txt

 Input orientated DEA

 Scale assumption: VRS

 Slacks calculated using multi-stage


 EFFICIENCY SUMMARY:

  firm  crste  vrste  scale

    1  1.000  1.000  1.000  -
    2  0.311  0.473  0.656 irs
    3  0.557  0.917  0.607 irs
    4  0.311  0.315  0.985 irs
    5  0.097  0.614  0.159 irs
    6  0.555  0.627  0.886 irs
    7  0.543  1.000  0.543 irs
```

Picture 5.12. Results in the EG1-out file (i)

5.3.3. Results interpretation

Step 7

The results of our case study are presented in Picture 5.13. The first column (firms) contains the numbers of the EU countries. And as mentioned before, the objects are numbered, in our case from 1 to 28 (Member countries of the European Union in 2018, Picture 5.13 and Table 5.3). The second one (crste) includes technical efficiency indicators for representative farms in these selected countries assuming the constant returns to scale (CRS). The third column (vrste) presents the technical efficiency indexes of farms in the individual EU countries assuming variable returns to scale (VRS). In the next column (scale) there are scale efficiency indicators, while in the last one there are the scale effect areas (IRS or DRS).

In our sample the most efficient farms were in Belgium (no. 1), Spain (no. 8), Italy (no. 14), Malta (no. 16) and Netherlands (no. 19) in 2018 (Picture 5.13 and Table 5.3). The representative farms from these countries, regardless of the assumption about the returns to scale (CRS or VRS), achieved the efficiency indicators equal to 1. It means that they were located on the so-called frontier curve and constituted benchmarks for farms from other EU countries; they are granted 100% efficiency score. In other words, farms from these countries (Belgium, Spain, Italy, Malta and Netherlands) fully efficiently use land, labour and capital inputs to achieve assumed output (expressed in Farm Net Income in the case study). The farms from these countries have also an optimal scale of production, as evidenced by the scale efficiency index equal to 1 (column 4 in Picture 5.13).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

   1   1.000  1.000  1.000   -
   2   0.311  0.473  0.656  irs
   3   0.557  0.917  0.607  irs
   4   0.311  0.315  0.985  irs
   5   0.097  0.614  0.159  irs
   6   0.555  0.627  0.886  irs
   7   0.543  1.000  0.543  irs
   8   1.000  1.000  1.000   -
   9   0.156  0.551  0.284  irs
  10   0.845  0.944  0.895  irs
  11   0.382  0.788  0.485  irs
  12   0.555  0.771  0.719  irs
  13   0.843  1.000  0.843  irs
  14   1.000  1.000  1.000   -
  15   0.220  0.663  0.331  irs
  16   1.000  1.000  1.000   -
  17   0.196  0.575  0.340  irs
  18   1.000  1.000  1.000   -
  19   1.000  1.000  1.000   -
  20   0.686  0.836  0.820  irs
  21   0.291  0.680  0.428  irs
  22   0.754  0.921  0.819  irs
  23   0.692  1.000  0.692  irs
  24   0.583  0.912  0.640  irs
  25   0.203  0.706  0.288  irs
  26   0.260  0.341  0.762  drs
  27   0.495  1.000  0.495  irs
  28   0.584  0.634  0.920  irs

 mean  0.576  0.795  0.700

Note: crste = technical efficiency from CRS DEA
      vrste = technical efficiency from VRS DEA
      scale = scale efficiency = crste/vrste
```

Picture 5.13. Results in the EG1-out file (ii)

If the analysis is based on variable returns to scale effects (VRS), the farms from Greece (no. 7), Ireland (no. 13), Romania (no. 23) and Slovenia (no. 27) are also among farms with the highest technical efficiency (Picture 5.13 and Table 5.3). However, farms from these countries do not achieve the optimal production scale, because the adequate scale efficiency indexes differ from one (column 4, Picture 5.13). All these farms operate in the area of increasing economies to scale (column 5 in Picture 5.13). That allows to conclude that increasing the inputs (and further production volume as a consequence) would result in an over-proportional increase in the farm output (Farm Net Income in this case study).

In our case study, only the farms from Slovakia (no. 26) operate in decreasing returns to scale (column 5, Picture 5.13). It means, that the increase in inputs (and agricultural production as a consequence) causes less than proportional increase in output (Farm Net Income in our case study). Therefore, it is not economically justified to further increase (expand) production by these farms.

In our research sample, assuming VRS, the farms from the Czech Republic (no. 4) and Slovakia (no. 26) had the lowest technical efficiency in 2018 (Picture 5.13). For the Czech Republic, the technical efficiency index of farms was equal to 0.315. Assuming input-oriented model, it means that the same output (Farm Net Income in the case study) can be achieved by input reduction by 68.5% (1 – 0.315 = 0.685). The input reduction by 68.5% would cause these farms to achieve technical efficiency equal to 1. In Slovakia, a 65.9% (1 – 0.341 = 65.9) reduction in farm inputs would make farms achieve efficiency index at the level of 1 while maintaining output level. Thus, the higher the technical efficiency index, the lesser the need to reduce inputs to achieve full technical efficiency at a given output. The other efficiency indicators for the surveyed objects (in the case study for representative agricultural holdings from the individual EU countries) should be interpreted in a similar way.

The lowest scale efficiency of representative farms, amounting to only 0.159, was recorded in Denmark (no. 5) (Picture 5.13 and Table 5.3). This means that adjusting the production volume to the optimal level would allow them to save as much as 84.10% of inputs (1–0.159).

The average values for the studied group of units can also be assessed (red frame in Picture 13). Depending on the adopted assumption, the efficiency ratio was 0.576 and 0.795 for CRS and VRS, respectively. Assuming the VRS, in the EU countries' farms, it was necessary to reduce inputs on farms by 20.5% (1 – –0.795 = 0.205) to achieve full technical efficiency (equal to 1) at a given output level (Farm Net Income in the case study) in 2018. Adjusting the production volume to the optimal scale in surveyed EU farms, would allow them to save as much as 30% of the current inputs (1 – 0.70 = 0.30).

5.4. Comparison of crops farm efficiency in the European Union: case study no. 2

5.4.1. Aims of research and data selection from FADN

Step 1

The main aim of the research is to determine the relative technical efficiency of representative farms specialized in field crops from the European Union in 2018. The production scale efficiency and areas of scale effects (IRS or DRS) will be also estimated.

Step 2

To calculate efficiency index, the input and output variables should be defined first. The data are retrieved from FADN database. On the input side, we use annual values of three variables: 1) intermediate consumption value in EUR (SE 275), 2) labour input in hours (SE 011), and 3) land inputs expressed in Utilized Agricultural Area in hectares (SE 025). On the output side the total production value of farm in EUR (SE 131) is applied. The data should be downloaded and saved in Excel in the form as presented in Table 5.4.

Table 5.4. Output and input values of crop farms in selected EU countries in 2018 prepared for DEA analysis

No.	Country	Production	Utilised Agricultural Area	Intermediate consumption	Labour input
1	Belgium	168822	58	85512	3059
2	Bulgaria	116525	141	59960	6192
3	Cyprus	26763	22	16047	2881
4	Czech Republic	249883	192	177192	7728
5	Denmark	217927	111	145081	2295
6	Germany	211486	126	130570	4139
7	Greece	21067	14	13753	2160
8	Spain	58195	65	32298	2611
9	Estonia	96454	163	82684	3047
10	France	175561	116	114547	2559
11	Croatia	28011	23	16304	2590
12	Hungary	65462	59	40447	2602
13	Ireland	141296	84	75668	1853
14	Italy	52861	26	25906	2710
15	Lithuania	48983	77	34477	3341
16	Luxembourg	116663	77	70880	2782
17	Latvia	62516	95	48387	3496
18	Malta	15624	3	7989	2694

No.	Country	Production	Utilised Agricultural Area	Intermediate consumption	Labour input
19	Netherlands	351170	59	175664	3559
20	Austria	90487	52	52059	2565
21	Poland	21468	22	13377	2984
22	Portugal	41815	18	20187	3105
23	Romania	41774	50	20358	3443
24	Finland	49475	63	46824	1174
25	Sweden	147086	118	103154	2446
26	Slovakia	432884	379	311825	13722
27	Slovenia	20514	9	12260	1763
28	United Kingdom	302020	175	183398	5005

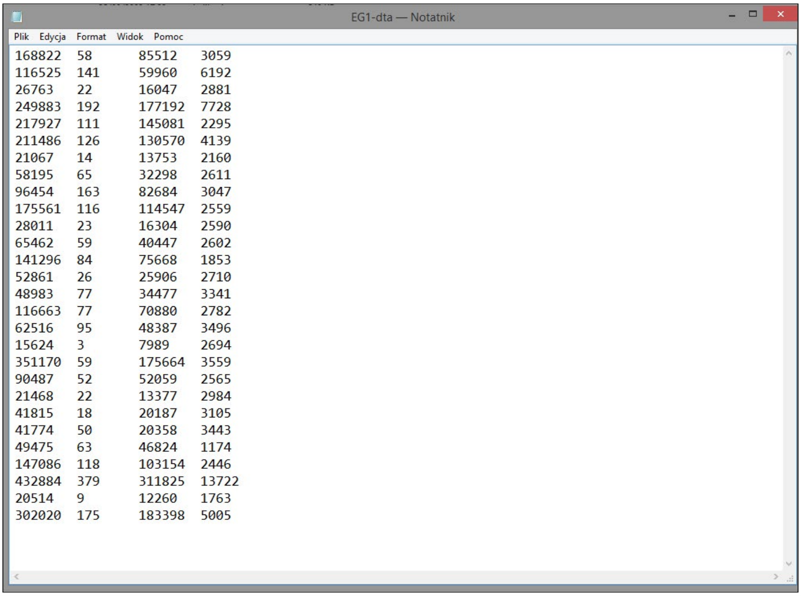
Source: Own survey based on FADN database.

Step 3

Install DEAP software following point 5.2.1 *Step 3*.

Step 4

Import data prepared to for DEA and saved in the EG1-dta file (Picture 5.14). The names of variables (inputs, outputs) and names of countries (analysed units) were numbered, i.e., 1, 2, 3, etc. A properly prepared data file looks like the one in Picture 5.14.



EG1-dta — Notatnik

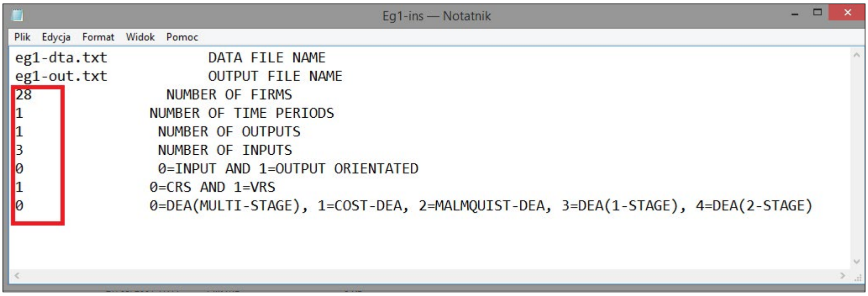
Plik Edycja Format Widok Pomoc

168822	58	85512	3059
116525	141	59960	6192
26763	22	16047	2881
249883	192	177192	7728
217927	111	145081	2295
211486	126	130570	4139
21067	14	13753	2160
58195	65	32298	2611
96454	163	82684	3047
175561	116	114547	2559
28011	23	16304	2590
65462	59	40447	2602
141296	84	75668	1853
52861	26	25906	2710
48983	77	34477	3341
116663	77	70880	2782
62516	95	48387	3496
15624	3	7989	2694
351170	59	175664	3559
90487	52	52059	2565
21468	22	13377	2984
41815	18	20187	3105
41774	50	20358	3443
49475	63	46824	1174
147086	118	103154	2446
432884	379	311825	13722
20514	9	12260	1763
302020	175	183398	5005

Picture 5.14. A ready-to-use spreadsheet in EG1-dta.txt file

5.4.2. Model calibration and calculation

Step 5
Calibrate model according to the assumption. The calibration is made in the Eg1-ins file.



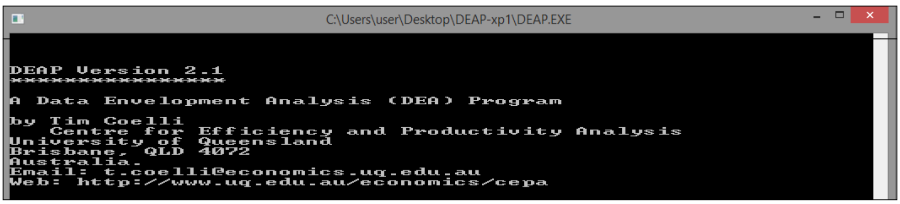
```
eg1-dta.txt          DATA FILE NAME
eg1-out.txt          OUTPUT FILE NAME
28              NUMBER OF FIRMS
1             NUMBER OF TIME PERIODS
1              NUMBER OF OUTPUTS
3              NUMBER OF INPUTS
0              0=INPUT AND 1=OUTPUT ORIENTATED
1             0=CRS AND 1=VRS
0             0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-STAGE), 4=DEA(2-STAGE)
```

Picture 5.15. Model calibration in Eg1-ins file

Considering the research goal and data set, the left column should be fulfilled (red frame in Picture 5.15):

- 28 (number of firms; number of analysed objects, in the presented case study 28 countries),
- 1 (number of times; 1 because the case study concerns only 2018),
- 1 (number of outputs; 1 because there is only one output production value in this example),
- 3 (number of inputs; there are three inputs in the example: Utilized Agricultural Area, Intermediate Consumption, Labour Input),
- 0 (model orientation; in our case study, the input-oriented model was assumed),
- 1 (model with variables return to scale is adopted),
- 0 (the multistage econometric procedure was applied).

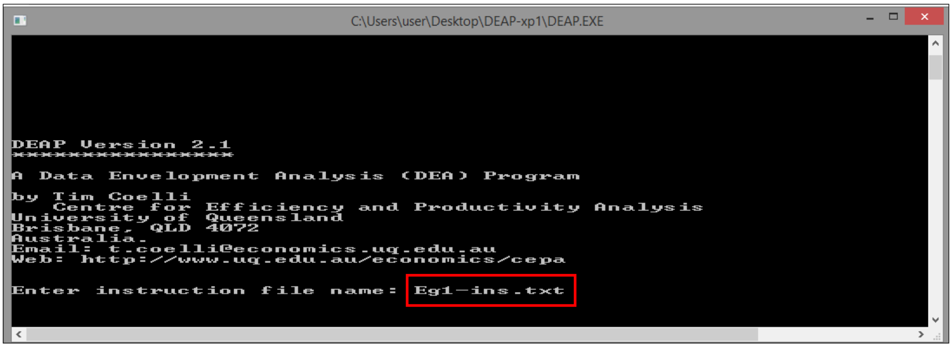
Step 6
To calculate and to obtain final result the DEAP file should be opened, and the black frame will appear (Picture 5.16).



```
DEAP Version 2.1
****************
A Data Envelopment Analysis (DEA) Program
by Tim Coelli
   Centre for Efficiency and Productivity Analysis
University of Queensland
Brisbane, QLD 4072
Australia.
Email: t.coelli@economics.uq.edu.au
Web: http://www.uq.edu.au/economics/cepa
```

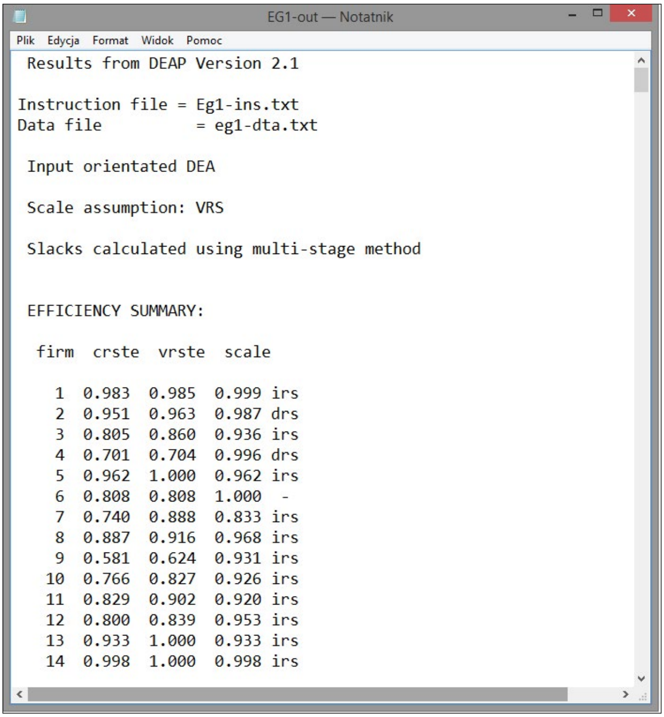
Picture 5.16. The DEAP file opened (i)

In the place of the blinking cursor (red frame in the Picture 5.17), enter the name of the configuration file, i.e., Eg1-ins.txt. Next, press enter, and the window will close automatically.



Picture 5.17. The DEAP file opened (ii)

The final results of calculation can be found in the EG1-out file (Picture 5.18). Having it the researcher should interpret the obtained values.



Results from DEAP Version 2.1

Instruction file = Eg1-ins.txt
Data file = eg1-dta.txt

Input orientated DEA

Scale assumption: VRS

Slacks calculated using multi-stage method

EFFICIENCY SUMMARY:

firm	crste	vrste	scale	
1	0.983	0.985	0.999	irs
2	0.951	0.963	0.987	drs
3	0.805	0.860	0.936	irs
4	0.701	0.704	0.996	drs
5	0.962	1.000	0.962	irs
6	0.808	0.808	1.000	-
7	0.740	0.888	0.833	irs
8	0.887	0.916	0.968	irs
9	0.581	0.624	0.931	irs
10	0.766	0.827	0.926	irs
11	0.829	0.902	0.920	irs
12	0.800	0.839	0.953	irs
13	0.933	1.000	0.933	irs
14	0.998	1.000	0.998	irs

Picture 5.18. Results in the EG1-out file (i)

5.4.3. Results interpretation

Step 7

The most efficient farms specialized in field crops in 218 were in Netherlands (no. 19), and Portugal (no. 22) (Picture 5.14). The representative farms from these countries, regardless of the assumption about the returns to scale (CRS or VRS), achieved the efficiency indicators equal to 1. It means that they were located on the so-called frontier curve and constituted benchmarks for farms from the other EU countries; they are granted 100% efficiency score. In other words, farms from these countries (Netherlands, Portugal) fully efficiently use land, labour and capital inputs to achieve assumed output (expressed in total production value). The farms from these countries have also an optimal scale of production, as evidenced by the scale efficiency index equal to 1 (column 4 in Picture 5.19).

If the analysis is based on variable returns to scale effects (VRS), the farms from Denmark (no. 5), Ireland (no. 13), Italy (no. 14), Malta (no. 18), Finland (no. 24), Slovakia (no. 26), Slovenia (no. 27) are also among farms with the highest technical efficiency (Picture 5.19). However, farms from these countries do not achieve the optimal production scale, because the adequate scale efficiency indexes differ from one (column 4, Picture 5.19). These farms without farms from Slovakia (no. 26) operate in the area of increasing economies to scale (column 5 in Picture 5.19). That allows to conclude that increasing the inputs (and further production volume as a consequence) would result in an over-proportional increase in the farms output (total production value). In this case study, only the farms from Slovakia (no. 26), Bulgaria (no. 2), Czech Republic (no. 4) and United Kingdom (no. 28) operate in decreasing returns to scale (column 5, Picture 5.19). It means, that the increase in input (and agricultural production as a consequence) causes less than proportional increase in output (total production value). Therefore, it is not economically justified to further increase (expand) production by these farms. In our research sample, assuming VRS, the farms from Estonia (no. 9) and Latvia (no. 17) had the lowest technical efficiency in 2018 (Picture 5.19). For Estonia, the technical efficiency index of farms was equal to 0.624. Assuming input-oriented model, it means that the same output (total value of production) can be achieved by input reduction by 37.6% (1 – 0.624 = 0.376). The input reduction by 37.6% would cause these farms to achieve technical efficiency equals to 1. In Latvia, a 33.8% (1 – 0.662 = 0.338) reduction in farm inputs would make farms achieve efficiency index at the level of 1 while maintaining output level. The lowest scale efficiency of representative farms, amounting to 0.528, was recorded in Finland (no. 24) (Picture 5.14). This means that adjusting the production volume to the optimal level would allow them to save as much as 47.2% of inputs (1 – 0.528).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

    1  0.983  0.985  0.999 irs
    2  0.951  0.963  0.987 drs
    3  0.805  0.860  0.936 irs
    4  0.701  0.704  0.996 drs
    5  0.962  1.000  0.962 irs
    6  0.808  0.808  1.000  -
    7  0.740  0.888  0.833 irs
    8  0.887  0.916  0.968 irs
    9  0.581  0.624  0.931 irs
   10  0.766  0.827  0.926 irs
   11  0.829  0.902  0.920 irs
   12  0.800  0.839  0.953 irs
   13  0.933  1.000  0.933 irs
   14  0.998  1.000  0.998 irs
   15  0.696  0.721  0.966 irs
   16  0.819  0.841  0.974 irs
   17  0.637  0.662  0.963 irs
   18  0.976  1.000  0.976 irs
   19  1.000  1.000  1.000  -
   20  0.863  0.891  0.968 irs
   21  0.775  0.838  0.925 irs
   22  1.000  1.000  1.000  -
   23  0.991  0.991  1.000  -
   24  0.528  1.000  0.528 irs
   25  0.713  0.774  0.921 irs
   26  0.690  1.000  0.690 drs
   27  0.816  1.000  0.816 irs
   28  0.822  0.823  0.999 drs

mean   0.824  0.888  0.931
```

Picture 5.19. Results in the EG1-out file (ii)

The average values for the studied group of units can also be assessed. Depending on the adopted assumption, the efficiency ratio was 0.824 and 0.888 for CRS and VRS, respectively. Assuming the VRS, in the EU countries' farms, it was necessary to reduce inputs on farms by 11.2% (1 – 0.888 = 0.112) to achieve full technical efficiency (equal to 1) at a given output level (total value of production) in 2018. Adjusting the production volume to the optimal scale in surveyed EU farms, would allow them to save 6.9% of the current inputs (1 – 0.931 = 0.069).

Questions / tasks

The DEA method can also be successfully used to assess the performance of non-agricultural units such as banks, hospitals, commercial enterprises, etc. However, it should be remembered that DMUs should operate in similar conditions and be comparable in terms of technology used, scale of operation and specialization. Using publicly available statistical databases and the DEA method, try to do the following exercises:

1. Compare efficiency of the agricultural farms in two EU selected countries in 2004–2007 (use average data for 2004–2017 from the EUFADN database).
2. Evaluate and compare efficiency of farms specialized in field crops and in horticulture within the EU-28 countries in 2017 (use averaged data from the EU-FADN database).
3. Which of the EU countries has the highest efficiency of the food industry measured by the relation of income and the number of employees and the value of fixed capital involved in production process (use Eurostat database)?

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078-1092.

Centre for Efficiency and Productivity Analysis (CEPA). Page which describes the computer program DEAP Version 2.1. Retrieved July 15, 2018 from https://economics.uq.edu.au/cepa/software

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). *Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical productions functions.* (Texas University at Austin Center for Cybernetic Studies No. CCS-504).

Coelli T. J. (1996). *A guide to DEAP version 2.1.: A Data Envelopment Analysis (computer) program*. (Centre for Efficiency and Productivity Analysis (CEPA) Working Papers No. 8/96).

Czyżewski, B., Smędzik-Ambroży, K., & Mrówczyńska-Kamińska, A. (2020). Impact of environmental policy on eco-efficiency in country districts in Poland: How does the decreasing return to scale change perspectives?. *Environmental Impact Assessment Review*, *84*, 106431.

Domagała, A. (2007). Metoda Data Envelopment Analysis jako narzędzie badania względnej efektywności technicznej. *Badania Operacyjne i Decyzje*, (3-4), 21-34.

Farm Accountancy Data Network (FADN). *FADN Public Database*. Retrieved July 5, 2018 from https://ec.europa.eu/agriculture/rica/database/database_en.cfm

Farm Accountancy Data Network (FADN). *Annex: Standard results indicators*. Retrieved July 5, 2018 from https://ec.europa.eu/agriculture/rica/annex003_en.cfm

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, *120*(3), 253-281.

Goraj, L., & Olewnik, E. (2011). *FADN i Polski FADN*. Warszawa: Wydawnictwo Instytutu Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy. Retrieved July 20, 2018 from http://fadn.pl/wp-content/uploads/2011/06/FADN_polski_FADN.pdf

Guzik, B. (2009). *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, (32), 73-105.

eISBN: 978-83-8211-072-2