Springer Proceedings in Business and Economics

Kesra Nermend
Małgorzata Łatuszyńska *Editors*

Experimental and Quantitative Methods in Contemporary Economics

Computational Methods in Experimental Economics (CMEE) 2018 Conference

Springer

Springer Proceedings in Business and Economics

More information about this series at http://www.springer.com/series/11960

Kesra Nermend · Małgorzata Łatuszyńska
Editors

Experimental and Quantitative Methods in Contemporary Economics

Computational Methods in Experimental Economics (CMEE) 2018 Conference

Springer

Editors
Kesra Nermend
Faculty of Economics and Management
Institute of IT in Management
University of Szczecin
Szczecin, Poland

Małgorzata Łatuszyńska
Faculty of Economics and Management
Institute of IT in Management
University of Szczecin
Szczecin, Poland

ISSN 2198-7246 ISSN 2198-7254 (electronic)
Springer Proceedings in Business and Economics
ISBN 978-3-030-30250-4 ISBN 978-3-030-30251-1 (eBook)
https://doi.org/10.1007/978-3-030-30251-1

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Contemporary economics, when analyzing economic behavior of people, increasingly takes into account non-market and non-economic factors, such as cognitive, behavioral, and emotional. It uses insights from humanities and natural sciences, becoming more and more an interdisciplinary science. Nowadays, economists realize the necessity to develop a new methodology and investigation instruments, and the need to use the diversity of research methods and modern ways of discovering knowledge. The increasing popularity of using economic experiments requires the use of IT tools and quantitative methods that facilitate the analysis of the research material obtained as a result of the experiments and the formulation of correct conclusions.

This book presents experiences and innovative solutions showing issues related to contemporary economics and what and how quantitative and experimental tools can be applied to solve modern economic problems.

In particular, it refers, among others, to important issues, such as growth theory, sustainability, public policy, real estate market, unemployment, silver economy, behavioral finance, consumer preferences, social campaigns, and neuromarketing. It shows the possibility of using tools and methods from the above-mentioned areas for solving problems, such as multicriteria analysis, multidimensional analysis, statistical methods, computer simulation, and computer games or techniques of cognitive neuroscience.

This book does not provide a comprehensive overview but focuses on selected topics. It includes papers of researchers which represent a certain level of experience in mentioned fields above. Its main purpose is to illustrate the links between various domains of contemporary economics and its methodological apparatus.

The book is divided into three parts:

I Theoretical Aspects of Contemporary Economics
II Application of Quantitative and Experimental Methods and Tools in Contemporary Economics
III Practical Issues—Case Studies

As the title suggests, the first part of the book presents the theoretical foundations of selected aspects in the contemporary economics. It raises issues related to methodological foundations of unified growth theory, assumptions of the silver economy and the main achievements of innovative technology supporting the independent functioning of seniors. It also refers to the theoretical basis for measuring sustainable development, the problems of using the achievements from the field of behavioral economics in the public policy and using experimental methods in behavioral economic. It finally shows the attempt of handling knowledge in experimental economics.

The second part of the book contains a general outline of possible methods and tools that support scientists in solving the current economic problems in different fields. It particularly shows the application of econometric and statistical methods in the mass appraisal process, the ESG risk analysis, the analysis of unemployment, measuring investors' optimism or in solving contemporary problems of green management from the perspective of modern economics. It also covers issues related to the use of such tools as IT, neuroscience techniques and simulation models in various domains of economics and business.

The last part of the volume presents examples of quantitative and experimental research. There are only selected case studies shown, but they outline a wide range of topics connected to contemporary economics, for instance, investigating the effectiveness of the product placement, assessing the level of sustainable energy provision in EU countries, assessment of the quality characteristics of real estate, mass land appraisal or analyzing intra-Community supplies. Moreover, an important place in this part is occupied by case studies related to the use of experimental methods.

The issues raised in the monograph do not exhaust the subject of contemporary economics. Yet, in the opinion of the editors, it does show the diversity of areas, current problems, methods, techniques and domains concerning this subject.

Szczecin, Poland

Małgorzata Łatuszyńska
Kesra Nermend

Contents

Part I
Theoretical Aspects of Contemporary Economics

Chapter 1
Some Remarks on Methodological Foundations of Unified Growth Theory

Jerzy Boehlke

Abstract The purpose of this chapter is to present and evaluate methodological foundations of unified growth theory. In a context of the modern theory of economic growth, Oded Galor, the protoplast of unified growth theory, provides a fresh theoretical and methodological approach in comparison with Keynesian and neoclassical economics. From the methodological perspective, he emphasizes an importance a context of the economic history of growth and development process observed in market type economies. This chapter shows some links with ideas coming from Ferdinand Braudel, institutional economics, and convergence theory. An interdisciplinary approach is the main achievement in unified growth theory.

Keywords History of economic development · Unified growth theory · Methodology

1.1 Introduction

In studies on the historical process of socioeconomic transformations, for many years attempts have been made to describe, explain, and interpret it in the context of historical sciences, sociology, geoeconomics, theory of economic growth and development, cultural anthropology, and many other scientific disciplines. Regardless of the historical conditions of that process, its causes, mechanism, and consequences are relevant for the present and future of human civilization. According to Braudel [10], the founder of "Annales" school, conducting research in this field requires references to the so-called near history, whose topic of interest are events from the nearest past, distant history, known as "tele-history," which encompasses the history of different civilizations and distinguishes between what is important and what is of secondary importance. It is worth mentioning that the theory of pluralism of social time had been formulated earlier by F. Braudel, in a famous essay entitled "History and social

J. Boehlke (✉)
Department of Economics, The Faculty of Economic Sciences and Management, Nicolaus Copernicus University, Toruń, Poland
e-mail: jurekbo@umk.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_1

sciences: The Long Duration" [9], where he distinguished between three categories: short time, time of cycles and intercycles, and time of secular tendencies. Those three categories were reflected in the history of events, history of business cycles, and structural history, which was related to long duration. According to F. Braudel, the result of such an approach is the multitude of ways of understanding history, a variety of points of view and interpretations which is made uniform due to dialectics relevant to the history. A synthesis of studies on the historical process of socioeconomic development is a consequence of an analysis of the so-called long duration, which is a historical process in which all the changes have a renewable and cyclical nature. Thanks to such an approach, social sciences gain a historical perspective, whereas historical sciences are capable of formulating various types of generalizations, avoiding the randomness and uniqueness of the studied events and processes. F. Braudel's methodological proposal is accepted and successfully implemented in the studies by I. Wallerstein [9]. Making use of the three categories of time proposed by F. Braudel, he makes an attempt to interpret and understand both the dynamics of changes of the analyzed historical social systems and the factors and mechanisms causing them to last. For I. Wallerstein, the primary category of analysis is the term of historical social systems, so-called world-systems, which are autonomous structures of long duration with relatively wide geographical range. Among world-systems, there is also "world-capitalist economy," which just like every social system has its time boundaries established within the long duration (hence, it has a beginning and an end), as well as spatial boundaries, which undergo various transformations in the course of its history. Economic theory cannot, obviously, ignore Braudel's methodological proposals but, on the other hand, it must not resign from studying the economic growth and development within own essential and methodological concepts which dwell upon it.

1.2 Economic Growth and Development in the Theory of Economy

In the economic theory, it is commonly accepted that there are three main groups of factors which influence the course of historical economic growth, namely the size and structure of means of production assets, institutional order and its transformations, and the model and quality of the economic policy implemented. The answer to the question concerning the force and character of the impact of particular groups of factors in the course of the historical process of economic growth remains open. The search for the answer to that question is very often associated with a comparative analysis of long-term economic growth rates in particular countries. They served as a basis for the dynamic development of economic growth theory in the first half of the nineteenth century, both within the classical, neoclassical, and Keynesian tradition. However, at the turn of the twentieth and twenty-first centuries, there was a visibly increased interest among the scholars in the impact of institutional order and

its transformations on the course of economic growth and development [17]. It was caused, among other factors, by disillusionment with modernity, and in particular, by the phenomenon of divergence, observed since the nineteenth century both between the studied countries and within particular national economies. For example, the research findings by L. Pritchett and A. Madison quoted by Wojtyna [66, p. 15] indicate a significant growth of disparities in the level of wealth around the world during the years 1870–2000. According to L. Pritchett, they grew fivefold, whereas A. Madison claims that they increased during that period from three- to nineteen-fold.

The abovementioned development of research concerning the impact of institutional changes on the development and functioning of the economy, both in the macro- and microeconomic scale is mostly related to expansion of institutional economics and especially new institutional economics. However, one cannot at this point ignore the accomplishments of evolutionary economics [14, 15]. Already plentiful and synthetic elaborations from that field [21] demonstrate, on the one hand, the vastness of the field of scientific research conducted within that current of contemporary economic thought, while on the other, their still insufficient applicability.

An analysis of institutional changes is also a canon of research on development economics [48]. They emphasize the impact of differences in the institutional order of individual countries, as well as their constituent regions, on the course of processes of socioeconomic development, whose result is usually a particular pace of economic growth or a change in the value of HDI indicator [3, 42]. A special example here may be the research recently conducted within the framework of endogenic theory of economic growth [63] concerning the impact of institutional changes on the pace and stability of economic growth [66]. Significant development was also observed in the field of research concerning the determination of cultural conditions for development processes [64, 29]. A theoretical background for a majority of analyses of the relationships between the system of institutions and economic development is the theory of economic growth and various theories of economic and political development [4, 19, 54]. Recently, increasing importance is attributed to research on the conditions for regional development conducted upon the foundation of new economic geography, whose ideological and methodological foundations were formulated by P. Krugman [33, 41, 65]. In the theory of economic development, the currents which are traditionally distinguished are research on the growth of highly developed economies and the analysis of the possibilities as well as conditions, strategies, and policies of dealing with economic underdevelopment. For research on underdevelopment, a useful and justified distinction turned out to be the one between economic growth and development. Known examples from the literature which confirm that opinion are, among others, the "big push" theory by Rosenstein-Rodan [57], the "vicious circle of poverty" theory by Nurkse [46], theory of economic development by Schumpeter [60], growth poles theory by Perroux [50], and the theory of structural changes by Chenery [12], all known from subject literature. A separate group is constituted by theories of economic growth of highly developed countries.

As it is known, economic literature contains a large number of economic growth models for highly developed countries which refer to Keynesian tradition, as well

as neoclassical one [6]. It should be emphasized that the Keynesian and neoclassical models are characterized by significant differences concerning institutional assumptions regarding the role of state, the nature and importance of competition, the breakdown structure of the economy, the impact of globalization, etc. However, it is assumed that for analysis of growth in highly developed countries, the best tool is growth models based on neoclassical tradition, particularly the Solow–Swan model [61] and its various mutations. It is on the basis of that model that a conclusion was formulated, according to which accumulation of physical capital is incapable of explaining very high growth of GDP per capita and its enormous diversity among countries, which are the reasons for the phenomenon of divergence observed worldwide, especially among highly developed and undeveloped countries. As it is known, the model is based on the analysis of production function in its general form; $Y(t) = f(K(t), L(t))$, where t means time, K—capital, L—labor force, and Y—product. Constant income in relation to scale in terms of capital and labor is assumed. In the Solow–Swan model, economic growth is not determined by the policy of the state. In particular, it does not depend on savings, since over a longer period the market mechanism causes the growth to be balanced, with capital and production growing at the pace determined by growth in labor assets. Technological progress is of exogenous nature and serves as a factor which compensates for dwindling income of capital and labor. The phenomena of divergence observed worldwide have been a reason for modification of that model. In alternative models, endogenous nature of technological progress was assumed [55], and the assumption of dwindling income from the capital was revoked. However, the assumption of falling income from labor was preserved, as a rejection of the dwindling income assumption, both for capital and labor, would result in endogenous growth rates, which would be constantly growing. That, in turn, would be against the empirical practice. A change in assumptions has allowed Mankiw, Romer, and Weil to explain the phenomena of divergence in the contemporary global economy with the differences in the population growth and capital accumulation rate, using the concept of human capital [39]. A generalization of that model was made by Nonneman and Vanhoudt [43]. It is commonly assumed that endogenous growth models allow for a more thorough analysis of the role of capital, labor, and technology in the process of growth. Unlike the models which assume the exogenous nature of growth and institutions, in the context of which it is difficult to explain and understand the enormous diversity of institutional order among countries, they make a much broader attempt to analyze the impact of institutions in the course of growth. They even introduce a factor known as "public order," created by the state as a provider of public goods [18]. Especially interesting in this field are social impact models [5, 19] and creative destruction models [2].

In exogenous growth models, one can only analyze the influence of institutional order on economic growth in the long-term perspective. Such an approach, however, is highly insufficient in the context of socioeconomic reforms attempted in many countries, which accelerate or stabilize its pace over short and medium periods. Theories which assume exogenous changes of production factors and institutions may, first of all, explain sudden and unexpected changes in the pace of growth [13, p. 232]. On the other hand, theories of that type are unable to answer the question why the

Industrial Revolution and the emergence of capitalism took place, why economic reforms in China caused a rapid acceleration of growth process, whereas attempts to replicate those reforms in Russia ended in failure. According to Rodrik (2011, p. 68), effective growth strategies should be always based on two-way changes. The first direction is related to short-term strategy, whose aim is to accelerate growth, whereas the second direction should include medium- and long-term strategies, oriented toward sustaining the growth. The necessity to combine the policy of economic growth and institutional changes which facilitate development and are related to so-called inclusive institutions is emphasized by Acemoglu and Robinson [1]. It is currently the most widely acknowledged belief in economics.

An important achievement of economic thought in recent years in research on the historical process of economic growth was undoubtedly the formulation of unified economic growth theory. Its creator was Oded Galor, and it is undoubtedly an attempt to synthesize the use of the concept of long duration according to F. Braudel with the development of the economic theory of endogenous growth. For that reason, it can be viewed as a theoretical concept which belongs to the theory of socioeconomic development. It definitely goes beyond the well-established framework of mainstream economics. In comparison with development economics and the theory of economic development and economic history, the unified growth theory by Oded Galor is based on more sophisticated methodological foundations.

That obviously makes it more novel, as well as more effective when describing, explaining, and interpreting findings of research on the course of historical transformation process.

1.3 Studies on Development and Growth in Economic History

Studies on economic development and growth in the context of socioeconomic history are development and supplementation of analyses concerning those issues conducted within the framework of economics, especially the theory of economic development and growth, evolutionary economics, and institutional economics. Analyses of economic development and growth from the historical perspective most commonly include the issues related to modernization and emergence of capitalism in Western Europe and its expansion to other geographical areas. The subjects of analysis are also the conditions and consequences of industrialization. The scope of that analysis may be viewed as not very precise due to existing controversies concerning the understanding of terms such as "modernity," "modernization," or "contemporaneity," which is, for instance, mentioned by Golka [26], Osińska [47]. Regardless of the aforementioned interpretation problems, the starting point for historians of the economy is usually the question regarding the origins, results, and mechanisms of transformations to subsequent periods and stages of development in the particular historiosophical concepts they acknowledge and analyze. An example which confirms

the correctness of the abovementioned ascertainment is the analysis of economic underdevelopment conducted by Gerschenkron [25]. He is the author of a concept well-known in subject literature which presents the relationships between institutional forms and the types of ideology and the degree of relative underdevelopment. As a result of the conducted studies, he defined the conditions for entering a path of sustainable and relatively fast economic growth and industrialization. According to Gerscherkron, the observed differences in the course of industrialization processes result from the degree of underdevelopment of individual economies. He also formulated the thesis of developmental pluralism, according to which one should allow for differentiation of development and industrialization in various countries. It should be noted, however, that the interpretation of the idea of developmental pluralism is not unambiguous.

For example, whereas according to Gerschenkron there is a major problem in the context of historical experience of the West concerning the degree to which the processes of industrialization of European countries may be considered as series of repetitions of the history of England's industrialization or deviations from that path, Marcin Kula suggests that one should completely abandon the perception of their economic development paths with regard to repetitions or deviations from the English pattern [34, pp. 67–70] and acknowledge the relevance of studying the so-called historical path of dependence. Otherwise, despite Gerschenkron's intentions, one should adopt some version of one-way interpretation of the vision of development. In practical terms, it means engaging in a discussion on the prerequisites of modern industrialization, which are related, among other factors, to the role of original accumulation of capital.

Another important issue was touched upon by Berger (1995). Commenting on Rostov's known concept of growth stages [58], he formulated the thesis that acceptance of the idea of gradual development does not mean the necessity of mechanical transition to subsequent stages. The view was shared also by Marks [40] and Hobsbawn [27]. Step changes are also possible [32], since development can be treated as a process of progressing economic growth which leads to an increase in the society's wealth. As it is known, Rostov's concept was viewed as Eurocentric and overly optimistic. As experience shows, in reality we deal with coexistence of phenomena of capitalist wealth and poverty, which since Prebish's [53] formulation of the "center" and "periphery" theory, which was developed within Wallerstein's [65] well-known concept of global economy system and the dependent development theory [49, 20], and in recent years within the research current called colonial studies [38, 59], are considered as mutually conditioned and functional. However, history does not confirm their continuity and durability, the examples of which are the countries that managed to overcome barriers of underdevelopment or dropped out of the "center" into the "periphery."

Another important issue in the studies on economic growth and development is the shape and role of the state. As emphasized by Polanyi [52], regardless of the achieved level of development, industrialization and strong acceleration of economic growth usually require good organization and high effectiveness of that institution, although

according to aforementioned Gerschenkron the case of England proves that the development could occur on its own, being reinforced by *laissez-faire* ideology. It should thus be concluded that, generally speaking, the methodological proposal concerning the thesis of developmental pluralism is fully justified. For example, the thesis is accepted by Landes [36], yet according to him the historical approach is merely a necessary condition, but definitely not a sufficient one for finding the answer to the question concerning the causes of poverty and wealth. He claims that when it comes, for instance, to the wealth and dominance of the West there are generally two explanations, the first of which maintains that it is a result of better organization, intelligence, hard work, and innovations of all kinds, whereas the second one points to aggression, greed, and ruthlessness of the Europeans. In light of the conducted analyses, factors of great interest and significance are ascertainments by Eisenstadt [16], Levy [37] regarding a faster pace of spreading capitalism in the form of modernization process than in the form of changes in the production methods and social structures. Progress and underdevelopment are a result of accelerated accumulation of capital which took place in Europe at the turn of the eighteenth and nineteenth centuries [7]. However, the understanding of lasting underdevelopment requires referring to the mechanism of functioning of the global economy according to Prebish and Wallerstein, as an effect of the emergence of industrial capitalism. Thus, whereas according to the aforementioned authors, economic growth and development depends mostly on institutional factors, in the opinion of Rosenberg and Birdzel [56] the wealth of the West should be treated as a coincidence which occurred in particular historical conditions. It resulted in economic, military, technological, and political global supremacy of Western Europe after the year 1500, which was the period of modernization and emergence of capitalism. It also caused the shaping of a bipolar world model, in which the "center" dominated the "periphery" [28, 30]. The opinion is also shared by Kennedy [31]. He notices, however, that it was a dynamic process fueled mostly by technological and economic progress in particular geographical and social conditions.

Within the current analyses of the course of the historical process of growth and development, it seems important to notice the concepts concerning its structure. An example here may be an interesting work by Gelei [24]. She introduced a division of all processes into differentiative, substitutive, and shaping the absolute growth. The phenomenon of differentiation, as a complex, multi-level one, involves changes in the manner of using means of production which influence division of labor, increased labor efficiency and capital, and uneven development and processes of specialization, integration, and concentration. As far as the substitutive processes are concerned, they are related to changes in the systems of relations between old and new types of activity and combinations of means of production, as well as transformations of their structure. Finally, the term of absolute growths expresses the essence of the economic growth process, manifested, for example, by an increase in the factors of production and the volume of goods and services produced. Differentiation, substitution, and absolute growths constitute and inseparable unity. According to Gelei, economic growth processes occur thanks to the creation of absolute growths through

differentiation, which results from substitution. In her opinion, the process of long-distance economic growth includes two periods, namely stagnation and accelerated growth. In the accelerated growth period, an economy ceases to be a one-sector economy. The ultimate effect of those changes is faster development and dominance of money–goods relations. The emergence of a global market increases the economic interdependence of individual countries.

In the literature concerning the topic of economic growth and development, a worthwhile element is the theoretically and empirically interesting discussion on the legitimacy of accepting the logistic curve as a proper description of the historical process of economic growth. The curve is even called the growth curve and is used for the analysis of numerous social phenomena. An example here may be the remarks by Cameron [11], who claims that whereas it is possible to use the logistic curve while analyzing economic growth due to the reversibility of the process, it is highly questionable whether it is possible to reverse development, since organizations of social, political, and economic structures rarely have the same form at particular stages. Kurowski [35], on the other hand, while describing the growth mechanism according to the logistic curve, claims that it has a dialectic nature. According to him, the thesis of logistic nature of growth may be deductively derived from the geometric progress rule and the theory which assumes the finiteness of the material world.

According to Best [8], on the other hand, economic growth processes depend on the so-called triad, which means a particular size and structure of factors of production, efficient institutions for coordination of actions taken within the economy, and policy based on experience and knowledge, especially institutional one.

Summing up, the research conducted upon the foundation of economic history concerning development and growth processes demonstrate their evolutionary character. The processes are subject to certain regularities, derived inductively from empirical studies, or deductively from the accepted presumptions concerning their nature and the ontological nature of the material world. Their pace grew significantly along with the emergence of structures of capitalism and the industrial economy. According to North [45], growth periods occurred throughout history when the economy generated strong stimuli for productivity growth. He formulates the idea that focusing attention on the studies of institutional and organizational structure allows for identification of the relations between a society's economic and political organization in the context of changing beliefs which drive human actions. The best example here is the history of the emergence of capitalism in Western Europe. The key to success in this case turned out to be a huge variety of the opportunities taken and increasing the probability of entering the path of economic growth thanks to shaping a competitive economic order which facilitated the extension of individual liberties. As claimed by North [44], the analysis and comprehension of the processes of long-term economic changes are possible on the grounds of the concept of historical path of dependence. It should, however, be remembered that understanding the processes of long-term growth and development is not equivalent to explaining them. A review of the literature concerning the course of economic growth and development processes within the economic history indicates wide variability of their interpretation without an unambiguous and ultimate explanation.

1.4 General Characteristics of the Unified Economic Growth Theory by Oded Galor

The unified economic growth theory by Oded Galor is a concept which describes the transformation of societies and economies from the period of stagnation, defined as the Malthusian period to the regime of modern growth, which means the period of sustainable growth. Its author assumes in his analysis that the course of the historical economic process is significantly influenced by geographical factors (when analyzing the growth paths one must thus not neglect the importance of the spatial criterion), institutional and cultural factors, and the characteristics of human capital. They determine not only the course but also the diversity of the pace and outcomes of growth, manifested mostly by the diversity in the level of income and life standard of the populations of individual countries. According to Galor [23], the starting point for the studies on the historical growth process should be the Malthusian period, with its distinctive stagnation, low standard of life of the population, low labor efficiency, and primitive technologies, accompanied by the growing population and territorial expansion. Improvement of life standards is not of a universal nature. Experience shows the existence of strong divergence processes in that area. However, many societies have succeeded in escaping the trap of poverty and underdevelopment and subsequently enter the path of economic growth, associated with sustainable growth. According to Galor, the transformation from the Malthusian period to the era of sustainable growth and the accompanying processes of divergence in the level of per capita income appear to be a problem of crucial significance in the studies on economic growth. The unified growth theory concentrates of recognizing and describing the fundamental factors which facilitate the ability to escape the Malthusian trap and progress to the post-Malthusian period, and then to the period of sustainable growth. The process of transition from the post-Malthusian period into the period of modern growth is characterized by technological changes which are more rapid in comparison to the ones observed during the Malthusian period, a relatively high pace of human capital accumulation and a significant decrease in the population growth rate, which has become the reason for economic surplus, allowing for an increase in per capita income, and consequently contributing to an improvement in the population's life standard. Within the unified theory of economic growth, Galor makes an attempt to answer the questions of why a major part of the history of human civilization is marked with stagnation and other characteristic features of the Malthusian period and what factors and mechanisms trigger the growth process. He concludes that providing an answer to those questions require an analysis of the mutual influence of technological changes on the changes in the population size and the changes in the level of per capita income which are related to that. Galor's intention was to depart from the instrumental treatment of the theory and to make an attempt to describe, explain, and interpret the growth process with reference to the historical and cultural context. That is why the starting point for him is characterizing three aforementioned growth regimes. Thus, as mentioned earlier, the Malthusian period, due to the principles of the Malthusian Law, is marked with lack of changes in the population

size, technological changes and changes in the level of per capita income. During the post-Malthusian period, due to an industrial revolution, technological changes take place, along with a population growth caused by an increase in the per capita income, due to increased labor efficiency. During the modern growth period, on the other hand, in comparison with the Malthusian and post-Malthusian period, one can observe rapid changes in the level of human capital and per capita income. In the unified growth theory, it is assumed that from the methodological point of view, the process of growth is considered to by the basic unit for theoretical and practical analysis [22]. The process is of endogenous and evolutionary nature. It includes the already mentioned changes in the population, technology, human capital, and per capita income in a spontaneous process of transition from the Malthusian stagnation to sustainable growth, with the accompanying phenomenon of divergence. Galor's theory involves studies of the dynamics of interaction between human evolution and the process of economic development. A hypothesis is tested concerning a significant role of the mechanism of natural selection in the evolution processes of the global economy and its transformation from stagnation to growth. According to Galor, the working of Malthusian law is crucial to the changes in population. At the same time, he expresses the belief that understanding and a proper interpretation of the global differences in the course of economic development processes will be incomplete as long as an analysis is conducted concerning the changes in the standard of life of particular communities, both in highly developed and undeveloped countries. Within the field of unified growth theory, it is also assumed that some differences in the observed growth processes are conditioned by biogeographical factors, as well as historical factors, related to, for instance, colonization, or political and military ones. Thus, on the one hand, the unified growth theory is an analysis of the prehistorical and historical factors which influence the course of contemporary processes of development, while, on the other hand, it is an example of an analysis of human evolution and the growth process. Philosophical and epistemological bases of the unified growth theory refer to the ideas of historicism, empiricism, hermeneutics, and the already mentioned evolutionism. Their embedment in history provides the analysis with an empirical basis. On the other hand, the time–space sequence of changes adopted by Galor, as well as acknowledging the key role of changes related to technology and human capital, as well as accumulation of capital, whose acceleration took place in the period of Industrial Revolution in England clearly indicate implicit acceptance for Braudel's idea of long duration and the idea of progress from the Enlightenment period. The reason to explain the reasons for wealth has a long history, especially in the domain of economy. From that point of view, Galor's concept is not anything unique. In his model, development, despite being marked with strong divergence, is ultimately one-directional, progressing from less advanced to more advanced stages. In this sense, it can be perceived as a realization of the Enlightenment idea of progress, although in a slightly modified form, which is manifested by associating development with divergence, which is uneven development simultaneously occurring both in time and space. Galor's original proposal is an acknowledged periodization of development. Similarly to the case of Rostov's growth stages theory, the former

stage is related to the Malthusian and encompasses the longest period in the history of human civilization, whereas the contemporary growth originates from the Industrial Revolution. As it is known, the Industrial Revolution instigated the period over 200 years of economic growth. However, whereas until the half of the twentieth century the average worldwide pace of growth had oscillated around 3.4%, after that period its changes demonstrate a declining trend, which is the fact emphasized by Pinker [51]. He claims that the phenomenon has become the basis for the formulation of new long-term stagnation hypothesis by Lawrence Summers. On that basis, the necessity appears to search for the answer to the question concerning the relevance of the periodization of history adopted by Galor. Another question is the manner of referring to and using the idea of evolutionism. It seems that assuming the historical growth process to be an analytical unit on the one hand and studying the influence of changes in the relations between human behaviors and between the level of human capital and diversity in terms of growth, and as a result also the observed phenomenon of divergence on the other, locates Galor's concept within the current of holistic evolutionism, which characterizes certain theories in contemporary social sciences. In commentaries concerning the unified growth theory, it is commonly acknowledged that it has an interdisciplinary character. In view of the current state of knowledge of contemporary scientific methodology, hermeneutics in particular, such an approach significantly increases the number of possible interpretations of the course and outcomes of the economic growth process. References to history locate Galor's theory among the concepts of long duration in the sense postulated by Braudel, without at the same time rejecting the manner of interpretation of the time appropriate for the endogenous and exogenous growth. What makes it similar to the latter ones is the use of the language of mathematics and statistics, which undoubtedly increases the level of accuracy of the conducted analyses and improves the possibilities for empirical verification.

1.5 Conclusion

Summing up, the concept of economic growth analysis devised by Galor is an important development of the contemporary economic growth theory, being at the same time strongly embedded in earlier experiences and traditions of studies of this process. His theoretical and methodological approach is a link between economic history, modern methodology proposed by hermeneutics, evolutionary economics, institutional economics, and mainstream economics.

Oded Galor proposes an innovative methodology of development and growth processes investigation.

Acknowledgements Financial support from research project no. 2015/17/B/HS4/01000 supported by the National Science Centre in Poland is gratefully acknowledged.

References

1. Acemoglu, D., Robinson, J.A.: Why Nations Fail. The Origins of Power, Prosperity and Poverty. Profile Books, London (2013)
2. Aghion, P., Hewitt, P.: A model of growth through creative destruction. Econometrica **60**(2), 323–351 (1992)
3. Amin, A.: An institutional perspective on regional economic development. Blackwell Publishers Ltd, Oxford (1999)
4. Balcerowicz, L., Rzońca, A.: Puzzles of Economic Growth. The World Bank, Washington, DC (2014)
5. Barro, R.: Economic growth in a cross section of countries. Quart. J. Econ. **106**, 407–443 (1991)
6. Barro, R., Sala-i-Martin, X.: Economic Growth. McGraw-Hill, New York (1995)
7. Berger, P.L.: The Capitalist Revolution. Fifty Propositions about Prosperity, Equality and Liberty. Basic Books, a Division of Harper Collins Publisher, New York (1986)
8. Best, M.H.: How Growth Really Happens. The Making of Economic Miracles Through Production, Governance, and Skills. Princeton University Press, Princeton & Oxford (2018)
9. Braudel, F.: Historia i trwanie. Czytelnik, Warszawa (1999)
10. Braudel, F.: Gramatyka cywilizacji. Oficyna Naukowa, Warszawa (2006)
11. Cameron, R.: A Concise Economic History of the Word: From Paleolithic Times to the Present. Oxford University Press, Oxford (Polish edition) (1993)
12. Chenery, H.: Structural Change and Development Policy. John Hopkins University Press, Baltimore (1979)
13. Clark, G.: Farewell to Alms: A Brief Economic History of the World. Princeton University Press, Princeton (2008)
14. Dopfer, K.: Evolutionary Foundations of Economics. Cambridge University Press, Cambridge (2006)
15. Dopfer, K., Potts, J.: The General Theory of Economic Evolution. Routledge, London, New York (2008)
16. Eisenstadt, S.: Modernisation: Protest and Change. Prentice Hall, New Jersey (1966)
17. Elsner, W.: The theory of institutional change revisited: the institutional dichotomy, its dynamic, and its policy implications in a more formal analysis. J. Econ. Issues **XLVI**(1) (2012)
18. Findlay, R., Wilson, J.D.: The political economy of Lewiatan. In: Razin, A., Sadka, E. (eds.) Economic Policy in Theory and Practice. St. Martin's Press, New York (1987)
19. Fukuyama, F.: Trust: The Social Virtues and the Creation of Prosperity. Free Press, New York (1995)
20. Furtado, C.: Development and Underdevelopment. U.C. Press, California (1964)
21. Furubotn, E.G., Richter, R.: Institutions and Economic Theory. The Contribution of the New Institutional Economics. The University of Michigan Press, An Arbor (1997)
22. Galor, O.: From Stagnation to growth: Unified groeth theory. In: Aghion, P., Durlauf S. N. (eds.) Handbook of Economiv Growth, vol. IA. Elsevier North-Holland, Amsterdam (2005)
23. Galor, O.: Unified Growth Theory. Princeton University Press, Princeton, New Jersey (2011)
24. Gelei, A.: Trendy wzrostu gospodarczego. Państwowe Wydawnictwo Ekonomiczne, Warszawa (1974)
25. Gerschenkron, A.: Economic Backwardness in Historical Perspective (A Book of Essays). The Belknap Press of Harvard, Cambridge (1962)
26. Golka, M.: Cywilizacja współczesna i globalne problemy. Oficyna Naukowa, Warszawa (2012)
27. Hobsbawn, E.J.: Labouring Man. Weidenfeld & Nicolson, London (1964)
28. Hobson, J.: The Eastern Origins of Western Civilisation. Cambridge University Press, Cambridge (2004)
29. Huntington: Zderzenie cywilizacji. Warszawskie Wydawnictwo Literackie Muza, Warszawa (2000)
30. Jones, E.L.: The European Miracle: Environments, Economies and Geopolitics in the History of Europe and Asia. Cambridge University Press, Cambridge (1981)

31. Kennedy, P.: The Rise and Fall of the Great Powers. Paul Kennedy (1988)
32. Kler, J., Perczyński, M., Dobosiewicz, Z.: Z doświadczeń rozwoju skokowego. Państwowe Wydawnictwo Ekonomiczne, Warszawa (1989)
33. Krugman, P., Venables, A.J.: Globalization and the inequality of nations. Quart. J. Econ. **110**(4), 857–880 (1995)
34. Kula, M.: Aleksandra Gerschenkrona uwagi o zacofaniu gospodarczym. In: O gospodarce mieszanej w krajach „Trzeciego Świata". Państwowe Wydawnictwo Naukowe, Warszawa (1967)
35. Kurowski, S.: Historyczny proces wzrostu gospodarczego. Państwowe wydawnictwo Naukowe, Warszawa (1963)
36. Landes, D.S.: The Wealth and Poverty of Nations. W. W. Norton & Company, New York (2000)
37. Levy Jr., F.D.: Economic Planning in Venezuela. Frederick A. Praeger, New York (1968)
38. Loomba, A.: Colonialism/Postcolonialism. Routledge, London, New York (2005)
39. Mankiw, N.G, Romer, D., Weil, D.: A contribution to the empirics of economic growth theory. Quart. J. Econ. **107**, 407–437 (1992)
40. Marks, K.: Kapitał. Krytyka ekonomii politycznej, vol. 1. Książka i Wiedza, Warszawa (1951)
41. Martin, R., Sunley, P.: Paul Krugman's geographical economics and its implication for regional development theory: a critical assessment. Econ. Geogr. **2** (1996)
42. Martin, R., Sunley, P.: Slow convergence? The new endobenous growth theory and regional development. J. Econ. Geo. **74**(3) (1998)
43. Nonneman, W., Vanhoudt, P.: A further argumentation of the Solow model and the empirics of economic growth for OECD countries. Quart. J. Econ. **111**(3), 943–953 (1996)
44. North, D.C.: Institutions, Institutional Change, and Economic Performance. Cambridge University Press, Cambridge (1990)
45. North, D.: Understanding the Process of Economic Change. Princeton University Press, Princeton, New Jersey (2005)
46. Nurkse, R.: Problems of Capital Formation in Underdeveloped Countries. Oxford University Press, Oxford (1953)
47. Osińska, M. (ed.): Economic Miracles in the European Economies. Springer Nature Switzerland AG, Switzerland (2019)
48. Oström, E.: Governing the Commons; The Evolution of Institutions for Collective Action. Cambridge University Press, Cambridge (1990)
49. Palma, G.: Dependency: a formal theory of underdevelopment or a methodology for the analysis of concrete situations of underdevelopment? World Dev **6**(7–8) (1978)
50. Perroux, F.: Economic Space, Theory and Applications. Quart. J. Econ. (1950)
51. Pinker, S.: Nowe oświecenie. Argumenty za rozumem, nauką, humanizmem i postępem. Wydawnictwo Zysk i S-ka, Poznań (2018)
52. Polanyi, K.: The Great Transformation: The Political and Economic Origins of Our Time. Beacon Press (1947)
53. Prebisch, R.: Towards a Dynamic Development Policy for Latin America. U.N. Publication (1963)
54. Rodrik, D.: One Economy, Many Recipes. Princeton University Press, Princeton, NJ (2007)
55. Romer, P.M.: Advanced Macroeconomics. McGrow-Hill, New York (1996)
56. Rosenberg, N., Birdzell Jr., L.E.: How The West Grew Rich: The Economic Transformation of the Industrial World. Basic Books, Inc (1986)
57. Rosenstein-Rodan, P.N.: Notes on the Theory of the 'Big Push". MIT Center for International Studies, Cambridge, Mass (1957)
58. Rostow, W.W.: The Stages of Economic Growth. A Non-communist Manifesto. Cambridge University Press, Cambridge (1960)
59. Said, E.W.: Orientalism. New York (2003)
60. Schumpeter, J.: The Theory of Economic Development. Taylor & Francis, Milton Park (2017)

61. Solow, R.: A contribution to the theory of economic growth. Quart. J. Econ. **70**, 65–94 (1956)
62. Swan, T.W.: Economic growth and capital accumulation. Econ. Rec. **32**, 334–361 (1956)
63. Tokarski, T.: Ekonomia matematyczna. Modele makroekonomiczne. Państwowe Wydawnictwo Ekonomiczne, Warszawa (2009)
64. Trompenaars, F., Hampden-Turner, C.: Riding the Waves of Culture. Understanding Cultural Diversity in Business. Intercultural Management Publishers NV (1997)
65. Wallerstein, I.: The Modern World-System: Capitalist Agriculture and Origins of the European World-Economy in the Sixteenth Century. Academic Press, New York (1976)
66. Wojtyna, A. (ed.): Instytucje a polityka ekonomiczna w krajach na średnim poziomie rozwoju. Polskie Wydawnictwo Ekonomiczne, Warszawa (2008)

Chapter 2
Innovative Technologies in the Process of Development of the Silver Economy

Grażyna Krzyminiewska

Abstract This chapter discusses the issue of the role of innovative technology in the process of development of the silver economy. It is recognized here that the new paradigm of development constitutes a response to the global processes of population ageing, and that computer science in management is an indispensable element in this process, allowing for the development of issues in the field of information technology and management sciences, as it provides a perfect combination of technological knowledge from the scope of IT and business skills. The chapter indicates that we should incorporate the values associated with the silver economy into our thinking about the development of advanced technologies, and that we should treat it as an opportunity for technological progress, development of business and the implementation of important social values associated with the goal of providing a good life to an ageing population. The author refers to extensive scientific research and the application of these studies to the socio-economic practice within the adopted paradigm of the silver economy. It is also pointed out that the new paradigm of development does not focus exclusively on the silver economy as a new economic sector, but develops the thinking about the older generation, which we are obliged to respect and to provide with a dignified life. These two points of view are mutually integrated, and the purpose of the article is to present the assumptions of the silver economy and the main achievements of innovative technology supporting the independent functioning of seniors and improving the quality of their lives. The contents contained in the chapter fall within the category of economic and social analyses, utilize the achievements of sociological and economic studies and were based both on existing sources and the results of own research. The methods of desk research and descriptive analysis were used in the presentation of the studied issues. In conclusion, the chapter emphasizes the need to overcome cultural, ethical and mental barriers, so that the successful implementation of advanced technologies to the silver economy translates into general economic and social success.

Keywords Silver economy · Gerontechnology

G. Krzyminiewska (✉)
Poznań University of Economics and Business, Poznań, Poland
e-mail: grazyna.krzyminiewska@ue.poznan.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_2

2.1 Introduction

An analysis of the demographic trends clearly indicates that the phenomenon of ageing of the world's population is occurring on a previously unprecedented scale. The global population aged 60 years or over numbered 962 million in 2017, more than twice as large as in 1980 when there were 382 million older persons worldwide. The number of older persons is expected to double again by 2050, when it is projected to reach nearly 2.1 billion. The process of population ageing is most advanced in Europe and in Northern America. In 2050, older persons are expected to account for 35% of the population in Europe, 28% in Northern America, 25% in Latin America and the Caribbean, 24% in Asia, 23% in Oceania and 9% in Africa [12].

According to the forecasts of the Polish Central Statistical Office, in 2050 the share of people aged 65 and above in the total population in Poland will amount to 30.2%, while people aged 80 and above will account for 10.4% of the total population (in 2013—3.9%) [9].

These figures undoubtedly prompt us to pursue an analysis of the phenomenon, to conduct a broad discussion and to develop effective strategies taking into account its nature and consequences. Public attention frequently focuses on the negative effects of population ageing, which result from the fear that the demographic burdens will negatively affect the developmental capabilities of the given society (including Poland). On the one hand, emphasis is primarily placed on the inefficiency of the existing eldercare systems and the inefficiency of the social security system. On the other hand—we are dealing with a view through the prism of new development opportunities, consistent with the concept of silver economy and exploring the economic potential inherent in the silver economy [5].

The concept of the silver economy ties into the concept of social development based on the belief that economic growth must be associated with social progress in all spheres of life. The economy of ageing societies is both a challenge and an opportunity for development. On the one hand, the nature of the silver economy could be beneficial to the economy for economic reasons, and on the other hand, it is a manifestation of care for the seniors and their needs. The silver economy could positively affect the quality of life of older people, but could also positively influence the labour market, allowing for the friendly treatment of older workers, who better understand the needs and problems of seniors, and could also stimulate the creative sector and lead to an increase in employment in industries utilizing innovative technologies. According to the OECD estimates, the ageing marketplace will be huge: by 2020, the +60 consumer spending power will total US$15 trillion, up from US$8trillion in 2010 [8].

Based on the assumption that the silver economy could become a developmental impulse, this article focuses on:

1. The assumptions of the silver economy, the implementation of which will allow for the transformation of demographic problems into development opportunities and the development strategy, which takes into account the needs of an ageing

population and is conducive to an improvement in the quality of life, applying not only to the elderly, but also to other members of the population.
2. Innovative technology supporting the independent functioning of seniors and improving the quality of their lives.

The contents contained in the article fall within the category of economic and social analyses, utilize the achievements of sociological and economic studies and were based both on existing sources and the results of own research. The methods of desk research and descriptive analysis were used in the presentation of the studied issues.

2.2 The Essence of the Silver Economy

In the last decade, the silver economy has become the subject of many studies. This is primarily due to the fact that an unprecedented phenomenon of population ageing is currently taking place, primarily (but not only) in highly developed countries. This phenomenon leads people to change their views on economic processes and to create a development paradigm utilizing the process of population ageing. We are frequently dealing with an approach which points to the great threat to economic growth resulting from the fact of population ageing, and also with a contrary approach, pointing out, that these processes may lead to the formation of a "new path of development". In OECD's document [8] entitled *The Silver Economy as a Pathway for Growth*, the experts clearly state that "several efforts are underway to drive a paradigm shift that views the ageing as assets not liabilities".

This point of view has become characteristic for many institutions that include the senior economy in their development strategies. For example, the outline of the European concept of the silver economy appeared many years ago in the documents of the European Union, where it was pointed out, that the ageing of the European population could be an opportunity to increase the competitiveness of the economy through the production of goods and services targeted at older people [2], and, as was noted in later documents, that the development of innovative enterprises will also allow for the expansion of development opportunities based on global markets, due to the fact that the issue of population ageing is also affecting many non-European countries. This creates the possibility for the developed solutions, products and services to be exported to other countries [3].

The meaning of the concept of silver economy is not uniform—it is a conceptual construct engendered by the various needs of older people. As indicated by Szukalski [10], a mature economy utilizing the potential of silver economy is an economy that exhibits a proactive—and not a reactive—approach. This means actions ensuring the longest possible independence of older people through the development of services enabling them to stay at home longer (cleaning, shopping, care services) and the provision of knowledge about the existence and the use of the so-called civilizational prostheses which compensate for the lost fitness. It is the organization of

the transportation services, the activity of companies preparing the so-called smart apartments (premises with systems of devices supporting independent living), as well as the management of senior citizens' free time (education, entertainment, tourism and recreation), services related to the breeding of pets and "age sensitive" financial services.

What is crucial here is the fact that a mature senior economy works towards ensuring social integration, as well as the creation of products and services enabling the elderly to continue the professional activity, which ultimately benefits the entire society.

There is no single universally accepted definition of the silver economy, but all the existing definitions have certain common elements. For example, we can recall the definitions presented by two major institutions:

- the European Commission, which defines the silver economy as the existing and emerging economic opportunities associated with the growing public and consumer expenditures related to population ageing and the specific needs of the population over 50 [6];
- and the definition used by OECD, which in turn equates the silver economy with silver production, i.e. one that produces and supplies products and services targeted at older people, shaping an environment in which people over 60 years of age interact and thrive in the workplace, engage in innovative enterprise, help drive the marketplace as consumers and lead healthy, active and productive lives [8].

A reflection on the essence of the silver economy significantly enriches its understanding and the method of analysis. Figure 2.1 indicates that the essence of the silver economy is not only the production of goods and services dedicated to seniors, but a wide spectrum of activities that can be arranged in silver industries, social innovations and gerontechnology.

It should be emphasized that the silver economy understood in such a way covers an extremely interdisciplinary field of activities in which scientific research should reach the areas of the economy, society, psychology and technology.

In the process of designing activities in the field of the silver economy, it is extremely important to take into consideration that the population of older people is not a homogeneous group. Seniors differ from each other in a variety of ways—both due to the features of social differentiation characteristic of society as a whole (wealth, social status, social and cultural capital, gender, physical and mental condition), as well as the differentiation based on age. The needs of a person aged 60+ are totally different than those of a person aged 90+ . There are more differences than similarities between them. Therefore, in order to adequately respond to the needs of seniors and to achieve economic and social success, it is necessary to take into account the heterogeneity of the senior population. In accordance with the proposals presented by Zsarnoczky [14], seniors can be divided according to the following age criteria:

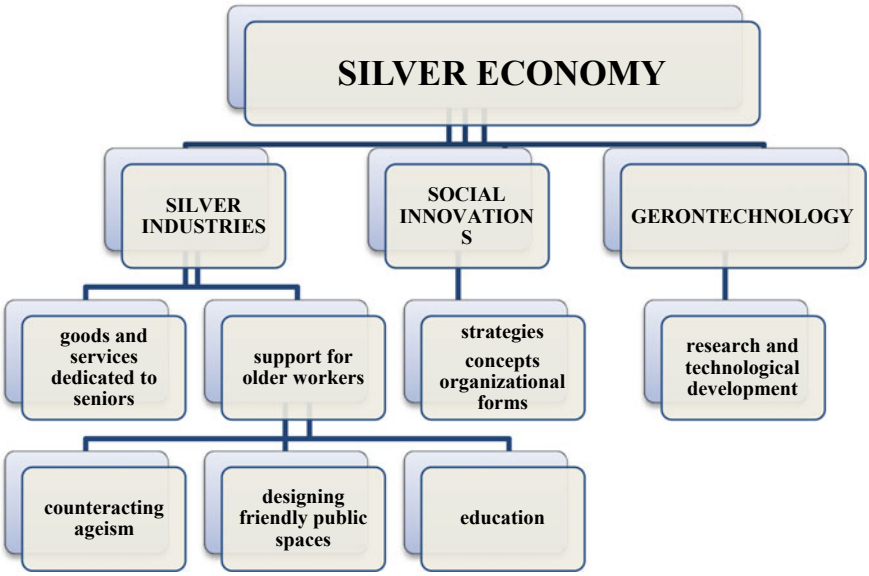


Fig. 2.1 Areas of activities of the silver economy. *Source* Author's own compilation based on the literature of the subject

1. Mature—51 to 64
2. Young-Old—65 to 74
3. Middle-Old—75 to 84
4. Old-Old—85 to 94
5. Very Old—95+.

The varied needs of an ageing population provide a wide range of opportunities for entrepreneurs to create new jobs and opportunities on the labour market and also create various challenges for science.

2.3 Innovative Technology as a Challenge for Science

After conducting interviews with experts in various fields, Zsarnoczky [14] analysed their opinions in relation to the possibilities for the development of the silver economy. It turns out that they are very broad. In this regard, we can distinguish areas such as media, fashion, gerontology, health services, home services, real estate, smart homes, educational system, tourism, medical tourism, nursing home, assistant living, fitness, finance, insurance, cosmetics, mobility, culture, recreation, IT, innovative technology, home delivery, robotics, architecture, design, public transport, local markets (e.g. food).

Being aware that each of these areas unquestionably involves in its activities the participation of technology and innovation, without which no field could function, for the purposes of this text, our attention will be limited to two of the above areas—innovative technology and robotics. They tie into the concept of gerontechnology, shaped at the interface between technology and advanced age. It is one of the five *Key Insights* highlighted by the OECD, which recognizes, that "technology solutions empower the ageing by helping build new markets, delivering new products and services, supporting new work practices and creating connected communities that respond to their needs […] A technology-driven silver economy ecosystem is emerging to support the unique requirements of the ageing. Physical and mental deterioration among the oldest cohorts create demand for new products and services and offer opportunities for innovators to respond to unmet needs" [8].

Gerontechnology relates to the designing of electronic aids to daily living, such as telework solutions, detectors, alarms, equipment to improve hearing, sense of smell, eyesight, bathroom systems, intelligent houses. The processes of technologization are undoubtedly mainly taking place in the field of health care, and one of the most intensively developing areas of innovative technology is telemedicine and telecare. These are various forms of exercising full-time care over an elderly person and round-the-clock monitoring of their health. As indicated by Zapędowska-Kling [13], new technologies in the healthcare sector are not limited to complicated micro devices that are still in the experimental phase. So-called assistive technologies are popular and widely used. Their primary objective is to improve the functioning of older people with locomotor system dysfunctions, as well as hearing or vision impairment. The group of assistive technologies includes wheelchairs/walkers, hearing aids, speech synthesizers, prostheses, as well as devices from the category of home automation [11]. The postulate of increased availability of assistive technologies goes hand in hand with the idea of a public space friendly to seniors, free of architectural barriers, adapted to the capabilities and needs of older people.

Table 2.1 indicates some of the devices supporting the safe functioning of seniors at their homes.

Medical technology is one field that is very advanced, but requires constant new research. "Medical instruments, devices, and equipment, including medical diagnostic machines medical therapeutic devices (drug delivery, surgical instruments, pacemakers, artificial organs); and other health-related products, such as medical monitoring equipment, handicap aids, reading glasses, and contact lenses. Medical technology also includes molecular diagnostic devices and health information technology, such as smartphone and IT applications. This broad range of products goes from simple, non-invasive equipment, such as wheelchairs, to high-tech and highly regulated invasive devices, such as pacemakers and insulin pumps" [7].

The safe functioning of seniors is also associated with the design of so-called smart homes [11]. It includes a range of facilities and user-friendly solutions, which were also indicated above, but it mainly involves devices enabling the automated control of the home environment, as well as devices (Table 2.2) used for recreation (virtual reality, robots) [1].

Table 2.1 Selected assistive devices

Device	Assistive devices	
	Description	Application
Portable GPS systems—in the form of a watch (including waterproof products)—shoes equipped with GPS	Real-time monitoring and directing the user to the right place in case they are lost	Devices dedicated to people in the early stages of dementia diseases
Medication dosing devices	Medications sorted into cups with alarm signals, if the drug is not taken after 90 min of alarm reminders, a message is sent to the carer	Devices dedicated to people in the early stages of dementia diseases
Home alarm systems—necklaces, bracelets	Real-time monitoring. They enable direct contact with the emergency services	Devices dedicated to seniors in order to increase their safety
Smart carpet	Monitoring of steps and activation of an alarm in the case of a fall	Fall detector
Environmental voice controller	Controlling household electronic devices	Hands-free control elements
Emotion controller (in the form of a small item—earring, brooch)	Device for monitoring of a person's emotional state	Devices dedicated to persons with emotional disorders threatening their life
Biomedical sensors	Monitoring of physiological functions	Useful, for example, in the measurement of the glucose levels in the blood, etc.
Sensors and detectors	Control of gas valves and water valves, reminding the user to switch off the oven, etc.	Devices dedicated to people in the early stages of dementia diseases

Source Author's own compilation based on the literature of the subject

The listed technological achievements do not cover the entire spectrum of already existing solutions, but allow for at least a cursory overview of the broad possibilities offered to seniors by new technologies. A huge role in this regard is also attributed to computer programming and designing, along with the introduction of these solutions to economics, project management, marketing and human resource management. Computer science in management is an indispensable element in this process, allowing for the development of issues in the field of information technology and management sciences as it provides a perfect combination of technological knowledge from the scope of IT and business skills. It is important to incorporate the values associated with the silver economy into our thinking about the development of advanced technologies and to treat it as an opportunity for technological progress, development of business and the implementation of important social values associated with the goal of providing a good life to an ageing population.

Table 2.2 Selected robotic technologies

Technologies			
Name	Author	Substitution	Description
Aibo	Sony	Improvement of the quality of life A robot companion, supporting the life of an elderly person	A set of sensors and drives. Programmed for having fun and maintaining contact with people. It is not currently produced
Paro	Intelligent Systems Research Institute (ISRI) belonging to the National Institute of Advanced (AIST)	Robot companion. Therapeutic purposes, mainly in dementia diseases. It fulfils the need for touch and safety	Robot in the shape of a seal programming environment and a set of sensors covering the entire surface of the robot. It is not mobile
iCat	Philips Electronics	Supporting basic activities and monitoring of persons requiring constant attention	Platform for the purposes of man–robot interactions. Mobile robot. Its appearance is similar to a cat. The face expresses emotions
Pearl	Carnegie Mellon University	Cognitive and mobile support	The mobile robot supports the movement of elderly people in care centres

Source Author's own compilation based on the literature of the subject

The essence of such an approach can be found in the work of [4], which is based on cutting-edge research into demographics, technology and the ambitions and aspirations of those in the over-fifty market. The director of the MIT AgeLab shows how businesses can profit from designing for and marketing to people aged fifty and over and from selling them products and services they truly want and need. The aforementioned AgeLab is one of the growing number of more well-known research centres, which "sit at the intersection of infrastructure, information, and institutions […] explore ideas in information-seeking, the power of place and 'things' in the environment, as well as business strategies and government policies to develop solutions for ageing successfully". The AgeLab bases its activities on the search for the answer to four basic questions:

- How do older adults learn, trust, adopt and use new technology and related services?
- How do novel technologies such as robotics, telemedicine and vehicle interfaces extend the safety and independence of older adults?
- What is the optimal design of everyday "things" and places to facilitate use as well as to excite and delight across the lifespan? How does the framing of information, formation of social networks and trust in "experts" impact the decision-making process of a consumer, patient or client?

- What are the implications of disruptive demographics on how businesses engage the consumer, develop products and services, deliver value and see ageing as a source of economic opportunity? (http://agelab.mit.edu/research).

The AgeLab uses its own methods and research tools in the course of the carried out experiments.

The development of innovative technologies understood in such a way translates into technological progress and the development of the economy, but also contributes to an increased quality of life of the older generation and the society as a whole, which also benefits from these achievements. The focus on utilizing the achievements of science and modern technologies is the answer to the question of what it will contribute to our everyday life, especially since what is useful for the older generation is also useful for all other people.

2.4 Barriers to the Development of the Senior Economy

Regardless of the deep conviction concerning the opportunities and benefits coming from the development of the silver economy, it is necessary to reflect on the numerous barriers to its development. Above all, it should be noted that many business entities may not be interested in the production and distribution of products for the elderly. It may result from the internal barriers of a given company, which does not have enough technical knowledge and the ability to quickly implement innovations, as well as the lack of a strategy for the company's long-term development. Companies are often too attached to the existing ways of doing things, not noticing new paths of development.

An important issue here is also the perception that the provision of various senior services is unattractive and the resulting focus on affluent consumers (luxury products and services). We are often dealing with stereotypical thinking about the consumption patterns of older people, presuming that they focus on satisfying their basic needs and are characterized by distrust for new solutions, reluctance towards innovations in their way of life. The poverty of older people is also pointed out, which leads to a lack of faith in the economic effectiveness of activities geared towards production and services dedicated to seniors. It is worth emphasizing, however, that there are many reasons why it is necessary to quickly let go of this stereotype.

First of all—while poverty among the elderly certainly exists, both in the world and in Poland, when referring to Poland or other countries, e.g. in the region of Central and Eastern Europe, it should be remembered that the age groups known as the so-called winners of the economic transformation are now entering old age (and, contrary to popular belief, this is not a small group). This relates to people who succeeded in the conditions of the market economy during the systemic transformation, becoming business owners, high-income employees, as well as owners of real estate and movable property with significant market value.

Secondly—along with the economic transformations, we are observing significant cultural changes in terms of attitudes and lifestyles. Basic mental changes are taking place both in the seniors' social environment and among the elderly themselves. Nowadays, the process of ageing is seen as a certain stage of life, which does not necessarily have to mean giving up on activities and being homebound, i.e. rarely or never leaving one's home. Modern seniors are open to new experiences and challenges, which is especially true for those who are in the so-called early phase of old age and are still full of vitality. Consequently, they expect their life to remain at least at the same level as before, and they are ready to use their accumulated assets in order to pursue their passions and interests. Oftentimes, however, they need products and services of the right quality, as well as appropriate support (including technological support) in order to be able to do so. The traditional attitudes of seniors are undergoing a significant change. They are no longer solely thinking about transferring the accumulated assets to the younger generation—they are increasingly frequently spending their wealth on their own long life. It is worth realizing that the preferences regarding the ways of spending free time will, in the near future to a significant extent, become the preferences of the future group of seniors.

Thirdly, in the countries of Central and Eastern Europe, during the last decades the phenomenon of economic emigration was observed on a significant scale (and was typical not only for Poland). Many of the economic emigrants intend to return to their native country for retirement and are ready to pay for the appropriate comfort of their remaining lives. Additionally, the economic emigration of the young generation and the associated phenomenon of the so-called elder orphans result in a situation where the children who live away from their parents often use their purchasing power to provide them with care and good life and finance the needs of their elderly parents.

While discussing the barriers to the development of the silver economy, especially in the context of the utilization of information technologies, we should not forget about the social responsibility associated with its implementation. Technology, and in particular artificial intelligence, could trigger numerous ethical dilemmas. The main point is that its use should not lead to the stigmatization of seniors. When special products are created for seniors, or interactive devices monitoring their lives are introduced into the lives of seniors (Table 2.1), this could deepen the seniors' sense of dependence on other people and on technologies and create a feeling of marginalization and exclusion. Unfortunately, the ethical side of this phenomenon is not clear, much like the issue of the ethical consequences of the participation of artificial intelligence in our lives and the replacement of personal relationships and social bonds with impersonal technology. According to many researchers, the development of gerontechnology may lead to a dehumanization of life in which the contact with a "living person" will be replaced by a personalized robot.

2.5 Conclusion

Knowledge-based economy, technological innovativeness and entrepreneurship can help us to use demographic trends to create new jobs, new professions and management trends. According to these assumptions, by creating challenges for contemporary economic and technological thought, the silver economy may become a new paradigm of development. The stimulation of the creative sector and the increasing employment in industries using modern technologies is already becoming a fact. Researchers and practitioners of economic life point out that the silver economy is worth billions of dollars, and that there is a great future ahead for the development of IT tools that provide the basis of creative products and senior services.

References

1. Bang, W., Bien, Z., Stefanov, D.H.: The smart house for older persons and persons with physical disabilities: structure, technology, arrangements, and perspectives. IEEE Trans. Neural Syst. Rehabil. Eng. **12**(2), 228–250 (2004)
2. Commission of the European Communities: The demographic future of Europe—from challenge to opportunity (COM(2006) 571 final). EU COM, Brussels (2006)
3. Commission of the European Communities: Dealing with the impact of an ageing population in the EU (2009 Ageing Report) (COM(2009 180 final). EU COM, Brussels (2009)
4. Coughlin, J.F.: The Longevity Economy. Inside the World's Fastest-Growing. Most Misunderstood Market. Ingram Publisher Services, USA (2017)
5. Domonkost, T., Horvát, P., Hvozdíková, V., Lichner, I., Miklošovič, T., Páleník, V., Radvanský, M., Štefánik, M.: Modelling The Economic Potential Of The Silver Economy. Neujobs Working Paper No. D 12.3/20.8.2013. https://www.google.pl/search?source=hp&ei=HL9EW76XEI2LmgXkvJLADA&q=modeling+the+economic+potential+of+the+silver+economy&oq=modeling+the+economic+potential+of+the+silver+economy&gs_l=psy-b.3…2213.24769.0.25975.55.42.0.0.0.0.561.2995.0j4j8j5-1.13.0..2..0…1.1.64.psy-ab..42.9.1883.0..33i160k1j33i21k1.0.Bd2ir-jafb0 (2013). Accessed 10 July 2018
6. European Commission: The 2015 Ageing Report: Economic and budgetary projections for the EU 28 Member States (2013–2060). European Economy 2015/3, Luxembourg. https://ec.europa.eu/info/sites/info/files/about_the_european_commission/eu_budget/draft-budget-2015-statement-of-estimates-sec-2014-357_2014_en.pdf (2015). Accessed 10 July 2018
7. Khaled, M.: Applying ICT & Innovation for the Aged—Enabling Better Health and Enhanced Participation to Create Economic and Societal Multiplier Effects, Tokyo, 12th Sept 2012. http://www.oecd.org/sti/ieconomy/Khaled%20ICT%20and%20Active%20Silver%20Aging_MKHALED_2012.pdf (2012). Accessed 10 July 2018
8. OECD: University of Oxford, Global Coalition for Ageing, The Silver Economy as a Pathway for Growth. Insights from the OECD-GCOA Expert Consultation, 26 June 2014. https://www.oecd.org/sti/the-silver-economy-as-a-pathway-to-growth.pdf (2014). Accessed 10 July 2018
9. Prognoza ludności na lata 2014-2050: GUS, Departament Badań Demograficznych i Rynku Pracy, Warszawa (2014)
10. Szukalski, P.: Demografia i Gerontologia Społeczna. Biuletyn Informacyjny **7**, 108 (2012)

11. Wieczorek, A.: Gerontechnologia w rozwiązywaniu problemów osób starszych. Syst Supporting Prod Eng **2**(14), 358–370 (2016)
12. World Population Ageing Department of Economic and Social Affairs World Population Ageing: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf (2017). Accessed 10 July 2018
13. Zapędowska-Kling, K.A.: Nowe technologie w służbie seniorom. Acta Universitatis Lodziensis Folia Oeconomica **4**(315), 206–207 (2015)
14. Zsarnoczky, M.: The_New_Hope_for_the_EU. http://www.regionalstudies.org/uploads/eng.pdf (2018). Accessed 12 June 2018

Chapter 3
Possibilities of Measuring Sustainable Development—Selected Aspects

Ewa Mazur-Wierzbicka

Abstract A developed system of indicators and measurement methods, both quantitative and qualitative, has fundamental importance for the possibilities of observation and implementation of a concept and pursuing its objectives and principles. The same applies in the case of sustainable development. An assessment of processes involving sustainable development is usually carried out with the use of indicators as well as qualitative and quantitative methods. This paper focuses on theoretical aspects and a general presentation of issues concerning the measurement of sustainable development. The aim of the paper is to show the diversity and complexity of measuring sustainable development (indicators, quantitative and qualitative methods) and to indicate difficulties associated with measuring it. The paper's layout is dictated by the implementation of this aim. The paper presents the theoretical basis of measuring sustainable development, after which various possibilities of measuring sustainable development are characterized and the difficulties associated with measuring thereof are pointed out. Quantitative and qualitative methods for the assessment of the course of sustainable development are pointed out.

Keywords Sustainable development · Measurement · Indicators · Quantitative and qualitative methods

3.1 Introduction

Since the second half of the twentieth century, the idea of sustainable development has become the basis of a new development paradigm involving the creation of a social and economic model based on progress ensuring a better quality of life at the same time taking care of the state of the natural environment.

In the concept of sustainable development, its holistic approach to the development policy needs to be emphasized which takes into account interrelations between the

E. Mazur-Wierzbicka (✉)
Department of Human Capital Management, Faculty of Economics and Management, University of Szczecin, Szczecin, Poland
e-mail: ewa.mazur-wierzbicka@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_3

following dimensions: economic, social and environmental (see [29]).[1] At the same time, it is assumed that these dimensions are to be equivalent. However, some authors believe that the environmental dimension (or more precisely the so-called ecosphere) is crucial as it determines life expectancy on Earth [22]. This is reflected in graphical models of sustainable development, especially in the so-called concentric circles model [18] or the spatial model proposed by Mauerhofer [28].

Lasting satisfaction of human needs achieved through the implementation of economic, social and environmental goals is considered the concept's prime objective.

The concept's essence itself seems to be universally understood, yet defining and a precise analysis of what sustainable development is a subject of discussion both among theorists and practitioners. Thus, the literature lacks one, generally accepted definition of sustainable development.[2] Existing definitions are formulated in national documents, documents of various organizations (e.g. the UN) as well as by individual authors.

The multiplicity and variety of definitions of sustainable development allow one to point out a few common elements. Therefore, sustainable development is (cf. [25]):

- a kind of social and economic development which rejects the concept of "zero growth",
- an intergenerational concept,
- a concept in which a significant role is played by distributive and intergenerational justice,
- a process integrating all actions taken by man,
- a development which strives to sustain three dimensions: economic, social and environmental,
- a concept which points out non-material values, including in particular the quality of life [19].

The approach to sustainable development defining it as action which should be economically viable, ecologically sound and socially acceptable can be assumed as most universal and synthetic (cf. [41]).

Indicators for which no single, universally applicable system has been developed should be considered a basic tool of monitoring sustainable development (cf. [2]). Therefore, one can either apply the systems of indicators created by various organizations or use their proposed modifications featuring in the literature.

An assessment of processes involving sustainable development is usually carried out with the use of indicators as well as qualitative and quantitative methods.

---

[1] Of the three dimensions of sustainable development, the environmental dimension was identified the earliest. Then, the time came for the social and economic dimensions. It was them that were the main subject of discussion in the report "Our Common Future". Over the years, the discussion was clearly reduced to two main themes: environmental and economic. Very little focus was given to the social dimension. In the literature, apart from the above-mentioned division into three basic dimensions of sustainable development, there are also more complex divisions, e.g. one can also encounter, i.e. ethical, technical, legal, political or spatial orders.

[2] Towards the 1980s, environmental economist Jack Pezzey identified over 60 definitions of SD [34], in the next decade Michael Jacobs established as many as 386 definitions [30].

The aim of the paper is to show the diversity and complexity of measuring sustainable development (indicators, quantitative and qualitative methods) and to indicate difficulties associated with measuring it.

The paper focuses on theoretical aspects and a general presentation of issues concerning the measurement of sustainable development. This topic is continued in other studies the subject matter of which concerns problems associated with practical measurement of sustainable development and with the use of quantitative methods for the assessment of the level of sustainable development.

3.2 Measurement of Sustainable Development—Introductory Issues

Sustainable development includes very complex and important issues, which should be measured in an appropriate way. When taking the effort to monitor sustainable development, one needs to pay attention to the dynamic character of this concept (the same applies to forecasting). Economic, social and environmental phenomena function in different rhythms[3] [20].

Indicators[4] have a fundamental importance to the possibilities of monitoring sustainable development. Owing to them, it is possible to approach the essence of this concept of development in a measurable manner. Not only do they allow an assessment of the current state, but also a formulation of measurable objectives, the monitoring of changes and the assessment of the effects of decisions and actions taken.

Sustainable development indicators are tasked with presenting in a measurable way the degree of implementation of the principles and targets assumed in this concept. They ought to, depending on the prevailing conditions and circumstances, first, make it easier for the authorities of a given area (to which they refer) and its residents to assess the degree of implementation of the concept of sustainable development, second, take up activating measures for sustainable development, third, verify the currently applicable policy directions and development aims together with the strategies to achieve them.

When drawing up the list of sustainable development indicators, one should be guided first and foremost by the specifics of the geographic area. A different degree of importance, a different rank may be assigned to specific aspects of sustainable development for different areas (country, region), thus different issues will be subject to monitoring. Due to the above, the most important issue is to specify which phenomena are considered important from the point of view of implementation of the outlined vision of sustainable development, and, what is entailed, which of them should be monitored. Each time the team developing and monitoring the indicators

[3]For example, planning an energy project requires a different approach (where the period of at least 30 years needs to be taken into account) to an approach taken when making financial transactions, when it is seconds that decide about a large loss or profit.

[4]The term indicator was first used by an American sociologist Dodd [15] in his work.

will chose those that to the greatest degree will express the essence of sustainable development of a specific area (cf. [7]).

The indicators should meet a number of various correctness criteria, the fulfilment of which gives meaning to constructing the system of indicators. Of these, the greatest significance is attributed to substantial correctness and formal correctness [9, 36].

The set of indicators which are recommended for monitoring sustainable development should have a uniform description manner, the so-called data sheet, which would include exhaustive information about individual measures. A sample description should include [9]:

- governance quantified in terms of indicators (e.g. environmental governance, social governance) and the field within the given governance,
- the management level the indicator refers to (e.g. national level),
- a definition of the indicator with a precisely specified measurable parameter, range of volatility or range of most favourable values,
- the objectives of the creation of the indicator and its place in the indicators' cause and effect sequence, as well as the interrelation with other indicators of this sequence,
- the indicator's strategic context, that is its reference to strategic objectives and principles of sustainable development,
- the methodology and possibilities of calculating (estimating) the indicator,
- the sources of information (statistical and non-statistical, etc.) necessary to apply the indicator.

A reliable preparation and appropriate use of sustainable development indicators may bring a lot of benefits, for instance [6]:

- reviewing the progress of implementation of policy targets and strategies conducive to the idea of sustainable development (or the assessment of policies and strategies not conducive to sustainable development),
- making reliable comparisons of the degree of implementation of the concept of sustainable development,
- facilitating strategic and operational planning of sustainable development,
- ordering and improving existing databases about the environment and individual spheres of the functioning of social and economic life as well as the system of statistical reporting,
- providing information in a form that will facilitate communication between experts, politicians, decision-makers and regular citizens,
- broad participation in creating the indicators contributes to increased understanding of sustainable development and its social acceptance.

A set of indicators proposed by the World Bank should also be noted. World Development Indicators (WDI) The World Development Indicators is a compilation of relevant, high-quality, and internationally comparable statistics about global development and the fight against poverty. The database contains 1600 time series indicators for 217 economies and more than 40 country groups, with data for many indicators going back more than 50 years (datatopics.worldbank.org).

3.3 International Systems of Sustainable Development Indicators

In the 1990s, the Organisation for Economic Co-operation and Development (OECD) developed the system of short-, medium- and long-term indicators that concerned environmental issues. Their objective was a unified presentation of cross-sectoral data in individual countries. These indicators were ordered according to the cause and effect layout, the so-called Pressure–State–Response.[5] The principle of causality is considered the basis for this division, namely, it is assumed that man, through his action (caused pressure), leads to the occurrence of damage to the natural environment. A response to it includes various undertakings resulting from the environmental policy, thanks to which one strives to minimize the damage caused to the environment. Therefore, a feedback loop occurs (Fig. 3.1) [38].[6]

The D–P–S–I–R sequence is the development of the classical P–S–R model. Next to the indicators of pressure, state and response, there appear additionally indicators of casual factors—D (*Driving Force*) and indicators of the assessment of impact—I (*Impact*).[7] In the D–P–S–I–R analysis, first, the driving force factors (D) are established, then their impact understood as pressure on, e.g. the state of the natural environment (P), which causes specified negative consequences (S) affecting both

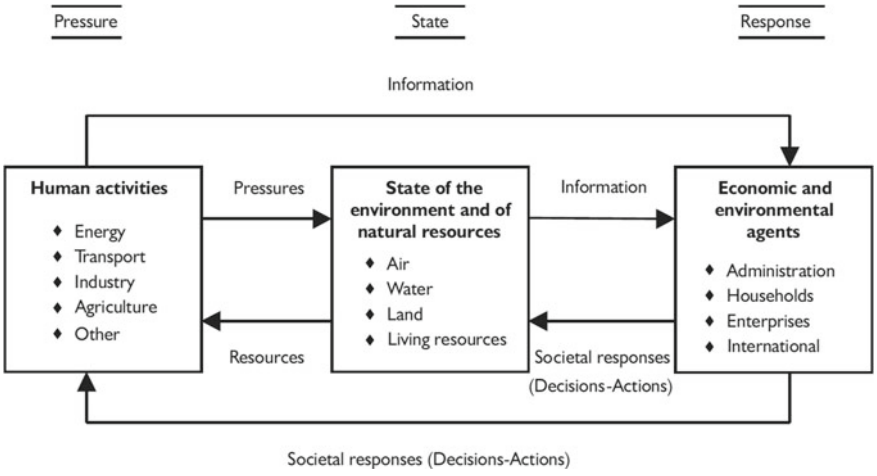


Fig. 3.1 Pressure–State–Response framework. *Source* Indicators of Environment [23]

[5]P (*Causes/Determinants*)—indicators of causes (origins of the state, factors, determinants, premises, correlates) of a specific phenomenon,

S (*States/Syndromes*)—indicators of states or syndromes of a specified phenomenon often expanded by indicators of consequences (R—*Results*),

R (*Responses*)—indicators of response which can also be called indicators of preventive action.

[6]The description of the OECD indicators is included in the OECD work [13].

[7]The D–P–S–I–R model is often used by the European Environment Agency (EEA).

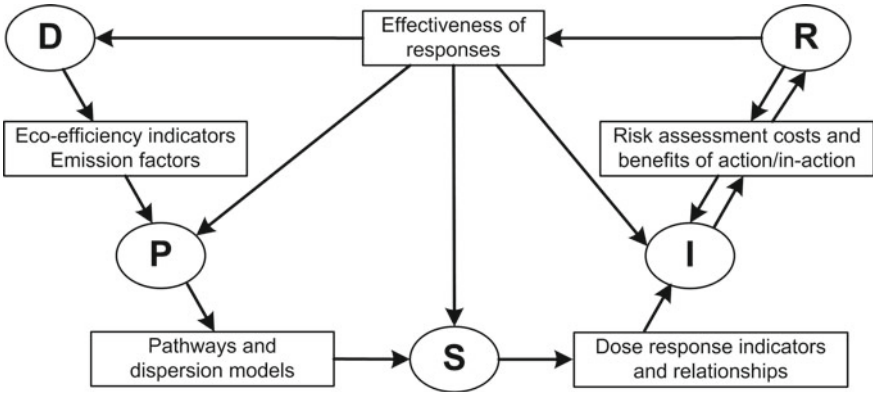


Fig. 3.2 DPSIR framework. Adapted from EEA. *Source* Smeets and Weterings [40]

the state of the environment and also, e.g. the society's quality of life (health, rest). A response to it involves taking action aiming to eliminate the damage occurred or preventing its emergence (R) (cf. [17]) (Fig. 3.2).

The OCED has also brought about grouping of environmental indicators. Therefore, one can identify the so-called core environmental indicators allowing continuous monitoring of changes occurring in the natural environment; a few groups of sectoral indicators in order to promote integration of environmental objectives in sectoral policies, and the group of indicators stemming from statistical accounting with the aim of promoting integration of environmental objectives in economics and sustainable use and management of natural resources [8].

Another important system of sustainable development indicators is the one proposed by the UN. Chapter 40 of Agenda 21 was the starting point for building and implementing sustainable development indicators by the international community. The created system of indicators was supposed to help decision-makers of all levels in the process of building a national development policy in which integration of economic, social and environmental objectives was to be emphasized (see more: [16]).

In April 1995, the Commission on Sustainable Development—CSD (UNCSD) adopted a work programme for the creation of the system of sustainable development indicators. Ultimately, the proposed set comprised 134 indicators (at the outset it was 130 indicators) ordered according to the Driving Force–State–Response scheme, which can be considered an analogy of the P–S–R concept adopted by the OECD. The indicators were grouped into four dimensions, 15 themes and 38 sub-themes (see more: [43]). The created list of indicators was not considered final and generally applicable. The indicators were often interrelated within the dimensions or main themes. Each indicator included a brief definition and description as well as the manner of calculation and interpretation. The sustainable development indicators proposed by the UN Commission on Sustainable Development were formulated somehow as an alternative to the OECD indicators.

After reviewing the indicators, their number decreased to 59. They were grouped into four dimensions: economic, social, environmental and institutional [44]. The next stage of changes concerning the specification of the number of indicators took place in 2006. Then, the number of indicators was established at 98, of which 50 were the so-called core indicators. They were grouped into 14 themes [45]. They referred to Agenda 21, the Johannesburg Plan of Implementation and the Millennium Development Goals [45]. The third change in the number of indicators and in the grouping of sustainable development indicators was dictated by the desire to achieve greater cohesion with the system of monitoring the Millennium Development Goals. The last change was proposed in 2015 and was associated with the adoption of the *2030 Agenda for Sustainable Development (Post-2015* Development *Agenda) including 17* Sustainable Development Goals and the related 169 targets monitored by means of appropriate indicators.

The currently applicable set of sustainable development indicators of the European Union refers directly to the 17 Sustainable Development Goals. The indicators were grouped into 17 themes which reflect the 17 Sustainable Development Goals. The previous set of EU's indicators referred to the *Renewed Sustainable Development Strategy (the indictors were grouped into 10 themes (areas) which reflected key challenges of the Sustainable Development Strategy)* [46].

Along with the modification of sustainable development indicators by the UN's agendas, one can see the adoption of a new way of ordering them. This is due to a departure from grouping them according to the D–S–R model in favour of the division into subject areas, main and specific problems.

3.4 Synthetic Indicators of Sustainable Development

The indicators thanks to which it is possible to enclose important issues in the form of one figure [26] play a significant role in the process of communication and providing information. There are many attempts to introduce synthetic indicators on the international scale. They can also be assigned to individual dimensions of sustainable development. They allow making comparisons in a more or less general manner in selected areas between countries. They allow observation of changes in the development of individual indicators in time. Thanks to this, it is possible to obtain at least approximate information on the directions of changes in concrete areas (economic, social and environmental) of individual countries or in groups of countries.

Synthetic indicators of sustainable development may be assigned to three basic groups, i.e. indicators of a general nature, indicators focusing on social and economic issues, indicators addressing environmental matters (Table 3.1).

Table 3.1 presents those synthetic indicators which are most popular (including the ones referred to in the literature) and positively assessed. The list of synthetic indicators is not closed. Such a thesis is constantly being confirmed by undertaken work for the creation of further synthetic indicators, as well as for modifying the

Table 3.1 Selected synthetic indicators of sustainable development (author's own compilation)

Name	Author(s)/responsible body	Objective
Synthetic indicators of a general nature—Group I		
Human Development Index (HDI)	The United Nations	Measurement of the development of individual countries taking into account three basic areas: life expectancy, level of GDP per capita and education[a]
Commitment to Development Index (CDI)	Center for Global Development	Assessment of wealthy countries in terms of their contribution to the improvement of the economic and environmental situation in poorer countries
Well-being of Nations (WoN)	Indicator authored by many international organizations, most often associated with Robert Prescott-Allen	Promotion of an appropriate condition of societies and ecosystems
Sustainable Society Index (SSI)	Sustainable Society Foundation	Measuring the sustainability of development at the level of national economy by specifying the distance that separates it from the predefined desired state
Happy Planet Index (HPI)	New Economics Foundation	Specifying the length of satisfying life which should remain in harmony with the environment as a source of resources and services necessary to people
Synthetic indicators of an economic nature—Group II		
Index of Sustainable Economic Welfare (ISEW)	H. E. Daly, J. B. Cobb	Non-pecuniary measurement of social welfare which takes into account both mineral resources management and intergenerational social justice
Genuine Progress Indicator (GPI) (results of continued work on improving ISEW)	Redefining Progress	
Index of Economic Well-Being (IEWB)	Centre for the Study of Living Standards—Canada	Referring to those areas of management which create well-being and social satisfaction

(continued)

Table 3.1 (continued)

Name	Author(s)/responsible body	Objective
Genuine Savings (GS)	K. Hamilton, M. Clemens/World Bank	Capturing the problem of sustainability of operation of national economy
Balanced Development Index (BDI)	Kozminski University	Analysis of and forecast for social and economic development
Synthetic indicators of an environmental nature—Group III		
Living Planet Index (LPI)	Word Wildlife Foundation	Identification of biodiversity threatened by human activity
Environmental Sustainability Index (ESI)	Scholars of Columbia University and Yale University	Tracking social and economic, environmental and institutional changes
Environmental Performance Index (EPI)	Scholars of Columbia University and Yale University	Pointing out the lowering of the pressure on the environment that negatively affects human health as well as the protection of ecosystems and vitality of natural resources
Ecological Footprint (EF)	Wackernagel [47, 48], Rees [35], Global Footprint Network	It shows how much highly industrialized areas are dependent on natural resources located in other parts of the world
Total Material Requirement (TMR)	Wuppertal Institute	It measures total material outlays which comprise national and imported material outlays
NCI (Natural Capital Index)	OECD	Investigated area: loss of biodiversity on the basis of changes occurring in ecosystems (changes in their areas and in their degree of naturalness)
RLI (Red List Index)	IUCN	Investigated area: loss of biodiversity on the basis of the number of endangered species (fungi, plants and animals) and the degree to which they are endangered (categories)

[a]Since 1990, there have been statistical attempts to include the issue of the environment in HDI calculations, though with little success (see more: [1, 12])

already existing ones. However, it is not important for the number of synthetic indicators to keep growing, but for the created indicators to provide adequate and concrete information concerning a specific area or phenomenon, etc., and to reflect essential, interesting features of sustainable development [39].

Indicators from the first group are very diverse and represent a great degree of generality. Their task is to show the effect of mutual interactions of the economic, social and environmental spheres—which is very difficult. Therefore, they are made up of a number of partial indicators which to the greatest extent are to provide concrete information from specified areas.

The second as well as third group of synthetic indicators includes domain-oriented indicators. To a large degree, they refer to concrete dimensions of sustainable development.

Synthetic indicators addressing the environmental issues concern the degree of human impact on the natural environment. Very often, their components include indicators characteristic to individual fields of the protection of the environment, e.g. biodiversity, protection of the air and waters.

The last group mentioned here includes indicators referring to broadly understood social and economic issues. One can find in them references both to prosperity economics as well and to the economics of the public sphere. In their case, the essential difficulty involves combining available statistical information with data, in a qualitative approach. Such data may be not registered by official statistics, yet it has essential significance to prosperity and sustainability of managing social and economic phenomena.

3.5 Difficulties in Measuring Sustainable Development

The measurement of sustainable development is a difficult and multidimensional activity, and it raises a lot of controversies and brings a number of difficulties. The various measurement tools presented in the paper show that there is no single, universally acceptable one. Most often each organization addressing the issue of sustainable development, each country and individual researchers adopt different measurement tools, which is mainly determined by different circumstances in which they function, availability of data, the level of analysis or its objective [27].

Of the problems in measuring sustainable development, the following are mentioned most often: the quality of data and the lack of data comparability, a high cost of obtaining data and an ambiguously defined function of the objective. Another question involves the lack of agreement between researchers as to the gravity of individual dimensions of the concept of sustainable development and the fact that some objectives within different dimensions are mutually exclusive. Establishing (in principle often impossible) which variables within individual dimensions should be considered key brings additional difficulties. One should point to the fact that the significance of individual dimensions and their components or selected indicators is

to a large extent determined by, i.e. the level of general civilizational development of a given country or the issue of priorities of the economic policy.

When using in practice, the division of indicators arranged according to the cause and effect order, the so-called Pressure–State–Response model, one can encounter a number of significant difficulties. They include:

- an insufficiently developed theory of social or economic sciences in areas that are to be subject to investigation in order to establish the division of indicators according to the P–S–R model (in such a case, a significant difficulty involves correct assignment of indicators to one of the three groups),
- the lack of appropriate statistical data, that is, data that has attributes necessary for conducting an analysis of the occurring cause and effect relationships—it needs to meet a few criteria, i.e. it must be complete, of a dynamic nature and it must take the form of time series,
- the lack of practical opportunities to define all relationships between the phenomena and their indicators [32].

In turn, in the case of synthetic measures, the main issue addressed is the fact that they are trying to approach a phenomenon of a very complex nature in a relatively simple way [2, 24].

Problems related to the measurement may be identified at each stage of constructing a synthetic indicator. Already at the beginning stage—selection of variables—there occur difficulties in the form of a lack of data presenting different dimensions of sustainability. This applies in particular to developing countries and as far as indicators are concerned—indicators of the social dimension area. A number of requirements have been formulated in terms of variables, yet at the moment finding a set of indicators meeting them fully is problematic. This stems from the fact that the established requirements are rather of a postulate nature.

The next stage concerns the unification of partial indicators. A significant difficulty in this respect involves the establishment of a reference point, often identified with the state of sustainability. Attempts to establish it have been made, certain solutions have been offered, yet none were sufficiently convincing for all and did not get full approval.

In the next stage, namely aggregation, it is the mechanisms of weighing individual components of a synthetic indicator that are mainly problematic. This results from the fact that they are usually of a subjective nature [5].

3.6 Quantitative and Qualitative Methods Used in the Analysis and Forecasting of Sustainable Development

Sustainable development has a multidirectional, dynamic and, at the same time, evolving nature. It is dependent on the course of numerous complex social and economic phenomena occurring in the micro- and macro-economic scale. This brings a

need to examine, analyse, diagnose as well as predict them systematically—be means of appropriate indicators and by the application of specified methods (qualitative and quantitative). They allow, in a longer time perspective, for taking orderly and rational decisions and resulting actions towards effective and stable sustainable development. It needs to be noted that quantitative methods are indispensable in creating synthetic indicators of sustainable development (described earlier in the study). It is necessary to point, e.g. in the case of building synthetic measures, at the initial stage to procedures of selecting variables for the design of the indicators, then to unifying partial indicators (an essential difficulty occurring here involves establishing a reference point), whereas at the final stage, to aggregation—weighing individual components of a synthetic indicator.

It the case of quantitative methods, it is characteristic in conducting an analysis of the occurring phenomena to use solely variables of a measurable nature. The most important quantitative methods that are used to measure and optimize variables which describe the potential and possibilities for development of sustainable development of a given area were included in Table 3.2.

As part of forecasting (apart from methods shown in the table), econometric forecasting is also worth noting. It is composed of two elements, i.e. modelling, where identification and a thorough analysis of development processes as well as conditions and mechanism of its implementation are assumed, as well as proper forecasting, i.e. using the model and independent variables—which emerged at the modelling stage—to develop the forecast for the development [4]. Fundamental benefits of forecasting in the context of sustainable development include greater effectiveness in managing sustainable development processes, possibility to see in advance potential effects of undertaken actions and ways of achieving adopted pro-development goals, opportunity for significant limitation of risk—by selecting appropriate instruments.

Sustainable development is based on a number of processes and components, i.e. economics, social science or natural environment. A significant role in this concept is played by relationships occurring between these elements. Adopting issues related to social disciplines brings about a need to apply qualitative formalisms, not solely quantitative ones. Qualitative studies involve gathering, analysing and use of various empirical materials: case studies, personal experience, introspection, interviews (see more: [49]) or observation materials (cf. [14]). A person carrying out qualitative research applies a number of interrelated practices in order to understand the essence of the investigated phenomenon [cf. 31, 33]. Each of the practices allows for seeing the investigated phenomenon from a different perspective. Qualitative methods include a few research systems, i.e. grounded theory [11, 37, 42] and qualitative case study research [10, 14, 21].

The grounded theory is based on an assumption that the social reality is best understood by actors involved in it. Therefore, it rejects a traditional functionalist approach where the researcher analyses a population using a previously developed theoretical model because he concludes that it only causes self-affirmation of a given theory (the researcher confirms his conviction since he finds what he wants to find) (see more [11, 37, 42]).

Table 3.2 Selected quantitative methods serving to measure and optimize variables describing the potential of sustainable development (author's own compilation)

Quantitative method	Use
Measures of descriptive statistics	Characterizing the research sample; Classification taking into account a differentiated level of development; Specifying regularity and trend that characterize the sample
Correlation analysis	Examining the strength of the relationship that exists between variables $\langle -1;1\rangle$; Determining the direction of correlation dependence
Regression analysis	Identifying variables that affect development
Estimation of frontiers DEA (Data Envelopment Analysis)	Assessment of investment layouts that stimulate the process of sustainable development; Assessment of the effectiveness of pro-development activities carried out by individual entities and of development layouts
Taxonomic analysis	Grouping entities that feature similar level of development
Classification method	Dividing areas due to specific attributes which characterize the level of their development
Linear trend forecasting and multiple criteria linear programming	Optimizing development processes (e.g. national, local) and optimizing variables which determine sustainable development
Multidimensional comparative analysis	Identifying areas with a similar level of development of the phenomenon subject to analysis and then grouping them and identifying phases of development of the investigated areas
Assessment of sustainable development **using aggregation measures**	Assessment of sustainable development, its individual components, governances

A qualitative case study is a research method where the researcher strives to describe a certain population or individual comprehensively taking into account an abundant set of variables, where he is interested both in values of variables as well as relations between them. The subject matter of research is of an individual nature. The research is initiated without preliminary hypotheses with the aim of precise examination of an assumed phenomenon in its real context. The main tasks of a qualitative case study entail determining the case's boundaries and defining its context, formulating research questions (in terms of form, content, addressee), conducting an analysis (using, for example: coding methods, clustering methods, a decision tree and relationship matrices), finding models in the gathered data and

formulating theories concerning the investigated case (cf. [14]). Results obtained by means of qualitative methods are the basis for creating models for quantitative methods.

3.7 Summary/Conclusions

To conclude, the contents included in the paper allowed the implementation of the aim set in the study.

The appearance of the concept of sustainable development in science and development strategies of organizations, countries and regions has brought about the need to measure it. It is because measurement has fundamental significance for the possibilities of observing the introduction of the concept and the implementation of its objectives.

Currently, numerous international organizations are working on creating new indicators in terms of sustainable development or improving the already existing ones [3]. A lot of substantive arguments make a case for it. First, it is essential that precise monitoring of the implementation of objectives that serve sustainable development, especially in the context of reviewing progress, should be possible. Second, the functioning set of indicators accepted and used on an international scale allows comparisons between countries in terms of the level of implementation of sustainable development. Third, constructing, e.g. an international base of indicators, will facilitate easier communication and carrying out discussion on the international forum, and it will allow countries to present their own achievements. Fourth, the work on indicators is a good opportunity for ordering and improving the existing databases and the system of national statistical reporting.

However, one needs to bear in mind a number of difficulties accompanying the measurement of sustainable development and to take them into account, especially when carrying out various analyses or comparisons.

The search for a synthetic measure still remains an open research topic. Current attempts of creating and implementing it include too many controversial elements of sustainable development in order for them to be considered as de facto measuring sustainable development.

One needs to point out the great importance of quantitative methods for the measurement, as well as analysing and diagnosing numerous complex social and economic phenomena affecting sustainable development. They provide as objective solutions as possible in the scope of shaping social and economic phenomena both in the past and in the future. In the most precise way, they facilitate the interpretation of the emerging development processes, as well as their efficient management and effective use. The use of issues related to social disciplines in the concept of sustainable development encourages also the use of qualitative methods and not exclusively quantitative ones.

The author wanted to point out in the study the specific nature and complexity of measuring sustainable development. Investigating the level of sustainable development (mainly in reference to individual countries) will be addressed in author's further research and studies.

References

1. Atkinson, G., Hamilton, K.: Accounting for progress: indicators for sustainable development. Environment **38**(7), 16–44 (1996)
2. Bell, S., Morse, S.: Sustainability Indicators. Measuring the Immeasurable. Earthscan, London (1987)
3. Bell, S., Morse, S.: Sustainability indicators past and present: what next? Sustainability **10**, 1688 (2018)
4. Błaszczuk, D.: Wstęp do prognozowania i symulacji. PWN, Warszawa (2006)
5. Borychowski, M., Staniszewski, J., Zagierski, B.: Problemy pomiaru rozwoju zrównoważonego na przykładzie wybranych wskaźników. Roczniki Ekonomiczne Kujawsko-Pomorskiej Szkoły Wyższej w Bydgoszczy **9**, 28–43 (2016)
6. Borys, T.: Dekalog ekorozwoju. In: Lokalna Agenda 21, czyli gminny program ekorozwoju. t. 2, Jelenia Góra- Wrocław (1996)
7. Borys, T.: Strategie i wskaźniki zrównoważonego rozwoju. In: Kronenberg, J., Bergier, T. (eds.) Wyzwania zrównoważonego rozwoju w Polsce, pp. 217–232. Fundacja Sendzimira, Kraków (2010)
8. Borys, T. (ed.): Wskaźniki ekorozwoju. Wydawnictwo Ekonomia i Środowisko, Białystok (1999)
9. Borys, T., Fiedor, B.: Operacjonalizacja i pomiar kategorii zrównoważonego rozwoju – przyczynek do dyskusji. In: Plich, M. (ed.) Rachunki narodowe, pp. 115–131. Wybrane problemy i przykłady zastosowań, Główny Urząd Statystyczny i Uniwersytet Łódzki, Łódź (2008)
10. Brown, L.: A review of the literature on case study research. Can. J. New Scholars Educ. **1**(1), 1–13 (2008)
11. Charmaz, C.: Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Sage, London (2006)
12. Constantini, V., Monni, S.: Sustainable human development for European countries. J. Hum. Dev. Capabilities **6**(3), 329–351 (2005)
13. Core Set of Indicators for Environmental Performance Reviews: A Synthesis Report by the Group on the State of the Environment Monographs. Paris **83**, 61–163 (1993)
14. Denzin, N.K., Lincoln, Y.S.: Metody badań jakościowych. PWN, Warszawa (2009)
15. Dodd, S.S.: Dimensions of Society. A Quantitative Systematics for the Social Sciences, New York (1942)
16. Dokumenty końcowe konferencji Narodów Zjednoczonych "Środowisko i Rozwój", Rio de Janeiro, 3–14 czerwca 1992 (1998) Wydawnictwo IOŚ, Warszawa
17. Dong, Y., Hauschild, M.Z.: Indicators for environmental sustainability. Procedia CIRP **61**, 697–702 (2017)
18. Elliott, J.A.: An Introduction to Sustainable Development. London, New York (2013)
19. Florczak, W.: W kierunku endogenicznego i zrównoważonego rozwoju – perspektywa makroekonometryczna. Wydawnictwo Uniwersytetu Łódzkiego, Łódź (2011)
20. Gabrusewicz, T.: Pomiar w zrównoważonym rozwoju - w poszukiwaniu wspólnej koncepcji. In: Czternasty, W., (ed.) Gospodarka Polski po akcesji do Unii Europejskiej. Wymiar finansowy. Zeszyt Naukowy nr 207, pp. 7–19 Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań (2011)

21. Harrison, H., Birks, M., Franklin, R., Mills, J.: Case study research: foundations and methodological orientations [34 paragraphs]. Forum Qual. Soc. Res. **18**(1), Art. 19. http://datatopics.worldbank.org (2017). Accessed 10 Apr 2019
22. Imran, S., Alam, K., Beaumont, N.: Reinterpreting the definition of sustainable development for a more ecocentric reorientation. Sustain. Dev. **22**, 134–144 (2014)
23. Indicators of Environment and Sustainable Development: Theories and Practical Experience. Environmental Economics Series, Background Paper no. 89, World Bank, Washington (2002)
24. Jesinghaus, J.: How evil is aggregation? Lessons from the dashboard of sustainability. In: Bell, S., Morse, S. (eds.) Routledge Handbook of Sustainability Indicators, pp. 392–406. Routledge, London, UK (2018)
25. Kistowski, M.: Regionalny model zrównoważonego rozwoju i ochrony środowiska Polski a strategie rozwoju województw. Wydawnictwo Naukowe Bogucki, Poznań (2003)
26. Lawn, P. (ed.): Sustainable Development Indicators in Ecological Economics. Edward Elgar, Cheltenham UK (2006)
27. Matuszczak, A.: Zróżnicowanie rozwoju rolnictwa w regionach Unii Europejskiej w aspekcie jego zrównoważenia. PWN, Warszawa (2013)
28. Mauerhofer, V.: 3-D Sustainability: an approach for priority setting in situation of conflicting interests towards a sustainable development. Ecol. Econ. **64**, 497–501 (2008)
29. Mebratu, D.: Sustainability and sustainable development: historical and conceptual review. Environ. Impact Assess. Rev. **18**, 493–520 (1998)
30. Michael, J.: Sustainable development, capital substitution and economic humility: a response to Beckerman. Environ. Values **4**(1), 57–68 (1995)
31. Mills, J.: Methodology and methods. In: Mills, J., Birks, M. (eds.) Qualitative Methodology: A Practical Guide, pp. 31–47. Sage, Thousand Oaks, CA (2014)
32. Muras, M., Ivanov, A. (eds.): Raport. Wykluczenie i integracja społeczna w Polsce. Ujęcie wskaźnikowe, CeDeWu, Warszawa (2006)
33. Olejnik, I., Kaczmarek, M., Springer, A.: Badania jakościowe metody i zastosowania. CEDEWU, Warszawa (2018)
34. Pezzey, J.: Economic analysis of sustainable growth and sustainable development. Environment Department working paper. no. ENV 15. World Bank, Washington DC (1989)
35. Rees, W.E.: Ecological footprints and appropriated carrying capacity: what urban economics leave out. Environ. Urbanisation **4**(2), 121–130 (1992)
36. Sekerka, B., Obrsalova, I., Bata, R.: Analyse options for relationship between sustainability development indicators. WSEAS Trans. Environ. Dev. **10**, 223–232 (2014)
37. Silverman, D.: Prowadzenie badań jakościowych. PWN, Warszawa (2008)
38. Śleszyński, J.: Rola wskaźników realizacji trwałego rozwoju. In: Graczyk, A., (ed.) Zrównoważony rozwój w teorii i w praktyce. Prace Naukowe AE we Wrocławiu, no 1190, pp 99–114. Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław (2007)
39. Śleszyński, J.: Obrona syntetycznych wskaźników rozwoju trwałego. In: Poskrobko, B. (ed.) Ekonomia zrównoważonego rozwoju w świetle kanonów nauki, pp. 82–97. Wyższa Szkoła Ekonomiczna w Białymstoku, Białystok (2011)
40. Smeets, E., Weterings, R.: Environmental Indicators: Typology and Overview. Copenhagen (1999)
41. Stanny, M., Czarnecki, A.: Zrównoważony rozwój obszarów wiejskich Zielonych Płuc Polsko. Próba analizy empirycznej. Instytut Rozwoju Wsi i Rolnictwa Polskiej Akademii Nauk, Warszawa (2011)
42. Strauss, A.L., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedures and Techniques, 2nd edn. Sage, London (1998)
43. UNCSD: Indicators of Sustainable Development Framework and Methodologies. UNCSD, New York (1996)
44. UNDSD: Indicators of sustainable development—review and assessment. Paper prepared for the United Nations Division for Sustainable Development (UNDSD) expert meeting. UNDSD, New York (2005)

45. United Nations: Indicators of Sustainable Development: Guidelines and Methodologies. New York (2007)
46. United Nations: Transforming Our World: The 2030 Agenda for Sustainable Development. New York (2015)
47. Wackernagel, M.: How Big is Our Ecological Footprint? Using the Concept of Appropriated Carrying Capacity for Measuring Sustainability. The Write Stuff, Vancouver (1994)
48. Wackernagel, M., Rees, W.E.: Our Ecological Footprint: Reducing Human Impact on the Earth. New Society Publishers, Philadelphia (1996)
49. Wiśniewska, J.: Wywiad jako technika gromadzenia danych w badaniach jakościowych. In: Ławniczak, K. (ed.) Metody jakościowe i ilościowe w badaniu organizacji i działania Unii Europejskiej, pp. 53–67. Wydział Dziennikarstwa i Nauk Politycznych UW, Warszawa (2013)

Chapter 4
Designing Public Policy According to Achievements of Behavioural Economics

Danuta Miłaszewicz

Abstract The assumptions of behavioural economics, about cognitive biases, heuristics, and observations that sometimes individuals are not the best judges of their own well-being, are now being used in the paternalist libertarian approach. It is the basis for creating a behavioural public policy. Previous experience in its implementation has allowed the development of strategies for policy and regulatory design. Applying this approach should lead to increasing the efficiency of the public policy, often thanks to quite small, simple and not too expensive interventions. The aim of the study is to analyse the existing possibilities of using the achievements of behavioural economics in the public policy with reference to the strategy of creating and conducting behavioural public interventions. This aim was achieved thanks to the application of critical literature analysis and desk research.

Keywords Behavioural public policy · Nudging · Principles of good choice architecture

4.1 Introduction

In developed market economies, the public policy finds support primarily in theoretical views existing in the mainstream economics and in the long-term experience in conducting public policy. Deriving from the mainstream economics, models of the impact of economic policy on the economy are built on the basis of neoclassical synthesis, which is based on the Keynesian theory of the aggregate demand and classical microeconomics. These models are constructed under numerous assumptions about the behaviour of markets and individual decision-makers. A basic assumption adopted in this models is that in effective markets, which sooner or later are able to reach stability, rational entities, so-called *homo oeconomicus* operate, which having full information about available alternatives, have unlimited possibilities of its

D. Miłaszewicz (✉)
Department of Economics, Faculty of Economics and Management, Institute of Economy, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: danuta.milaszewicz@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_4

processing, and when making decisions are guided only by their constant, selfish preferences, making choices that maximize their usability. The normative models created on these foundations are confronted with the real behaviours of people, and the results of this confrontation indicate that the more difficult and complex are the decisions actually made the less the individuals are similar to the model *homo economicus*, and the harder the effects of the policy are to predict.

The solution to this problem is findings of the behavioural economics, resulting from conducted experiments and improving the understanding of the human decision-making process. According to D. Kahneman, behavioural economics is economics that rejects the extreme model of rational market participants [1, p. 339]. In turn, R. Thaler perceives it as 'economics based on realistic assumptions and descriptions of human behaviour'. This is 'evidence based economics, which has more explanatory power because the models are better fit with the data' [2, p. 23]. It has three essential elements that R. Thaler refers to as 'three bounds'. These include: bounded rationality, bounded willpower and bounded self-interest [3, p. 258]. Behavioural economy is a combination of classical economic knowledge with psychological models, and such a relation of both sciences has resulted in the formation of theories in this field, defined as descriptive. They describe typical, average behaviour of decision-makers, which is very important for making public policies and interventions, which are supposed to solve problems of real people (*Humans*) and not ideally rational ones (*Econs*). Application of behavioural economics findings as innovative solutions in evidence-based public policy has been growing in popularity among scientists, policymakers and international organizations promoting it for decades.

The aim of the study is to analyse the existing possibilities of using the achievements of behavioural economics in the public policy with reference to the strategy of creating and conducting behavioural public interventions. The article has been divided into two main parts, which are summarized with a conclusion. The first part is a short presentation of the libertarian paternalism paradigm and the concept of choice architecture. The second part presents the strategies of behavioural changes currently used in the public policy and the principles of using behavioural insights in the public policy. The aim of the article was achieved thanks to application of critical literature analysis and desk research.

4.2 Libertarian Paternalism and Nudging

Using the findings of behavioural economics regarding cognitive biases, heuristics, and an observation that sometimes individuals are not the best judges of their own welfare C. Sunstein and R. Thaler's proposed the concept of libertarian paternalism[1] known as 'Third Way' [14, p. 252] or 'the middle way' [6, p. 914] between the

[1]This concept triggered a broad discussion both in the environment of scientists and public policy practitioners who presented both criticism and support. See e.g. [4–13]. The subject of the present study is not an additional voice in this discussion, and therefore, the discussion is not presented.

traditional paternalism and libertarianism. The authors of this approach to the public policy describe libertarian paternalism as a form of paternalism, libertarian in spirit that should be acceptable to those who are firmly committed to freedom of choice on grounds of either autonomy or welfare [15]. They also underline that libertarian paternalism 'attempts to influence the choices of affected parties in a way that will make choosers better off' [16, p. 1162].

Instead of mandatory rules, libertarian paternalism introduces default rules, which choosers may easily opt for a different rule, a 'so long as people can contract around the default rule, it is fair to say that the legal system is protecting freedom of choice, and in that sense complying with libertarian goals' [16, p. 1199].

This ensures not only increase in freedom of individuals but also 'to allow people to go their own way' [14, p. 249]. According to them, such a way of conducting public policy 'does not coerce anyone to do anything' [16, p. 1165]. It is also not onerous, because it 'is a relatively weak and nonintrusive type of paternalism, because choices are not blocked or fenced off. In its most cautious forms, libertarian paternalism imposes trivial costs on those who seek to depart from the planner's preferred option' [16, p. 1162].

The concept of libertarian paternalism can be reduced to the fact that if people systematically apply heuristics, which result in cognitive and motivational errors, they are hopelessly incompetent to make choices that actually further their own best interests, and the government should at least direct them towards the 'right' choice [13].

According Tverskiego and Kahneman, who created the basis for the research program heuristics and cognitive biases, the context and the framework of the decision are important because people make different decisions depending on the information presented to them, despite the fact that the information presented has not been substantively changed by the way they are presented [17]. This means that decisions and behaviour of individuals depend to a large extent on factors related to the context or situation (its framework) in which these units are located.

Individuals do not make decisions in the vacuum. Their decisions are influenced by perceived and unnoticed features of the environment in which these decisions are made, and for real people—*Humans* by Thaler and Sunstein [14]—details forming the conditions of the choice made can be important [18, p. 428]. According to Thaler and Sunstein, the decision-making environment can be appropriately created and changed, influencing the conditions of the decisions made, and the formulation of the interpretive framework for individual decisions has a huge impact on the result. They described such frames as choice architecture and manipulating the choice as nudging [14].

The choice architecture (or nudging) generate more deliberate, reflective and reasoned forms of decision-making. It is an approach to designing intentional intervention aimed at influencing behaviour change. It uses the evidence provided by behavioural economics to overcome mistakes we usually make in making decisions [9]. The choice architecture describes the fact that there is no neutral way of presenting the choice. People choose what is available and not what they absolutely want;

they do not make much effort when they make a choice and apply rules called heuristics; they do not have the main strategy to make decisions. People behave differently in different circumstances, and contextual influences are much more significant than we would like to believe. Behavioural economics makes it clear that every time a decision has to be made, it is reconstructed and depends on the context—what?, who?, how?, where?, when? [19].

The choice architecture is the background to the choices made, and many real people turn out to be choice architects, although most do not realize it. The architect of choice is responsible for organizing the context in which people make decisions. Wherever institutions, private or public, can create sets of options, from which people choose, we are dealing with choice architects [18, pp. 428, 430]. The choice architects create strategies for influencing the environment of individuals and construct the condition of the environment in which the decision is made.

For real people, the details that create the conditions of the selection can be important; therefore, the choice architects, that is, the people who create the environment of choice, can have a significant impact on decisions, e.g. by providing nudges. By providing default options and rules, nudges can help people make good decisions that are in their long-term interest without limiting their freedom [1]. According to Thaler and Sunstein, nudge 'is any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid' [14].

The broader definition of nudge was presented by Hansen [20] who explained that nudge is a function of choice architecture that changes people's behaviour in a predictable way regardless of: (1) prohibiting or adding any rationally important options, (2) changing their motivation, no matter for time, trouble, social or economic sanctions, etc. and (3) to provide factual information or rational argumentation. Nudges are necessary because of bounds and cognitive biases, routines, habits in individual and social decision-making and work through the use of these bounds, mistakes, routines and habits as integral parts of the choice architecture [20].

The concept of nudges is a non-regulatory approach that enables people to become behavioural change through slight changes in the environment of choice that people face [21]. According to Thaler and Sunstein [14, p. 252], nudges are everywhere, even if we do not see them. It can be 'any factor that significantly alters the behaviour of *Humans*, even though it would be ignored by *Econs*' [14, p. 8]. Nudges are interventions that push (direct) people in specific directions, but also allow them to follow their own path. To encourage specific behaviour (or change existing), people should be offered choices that make the decision easy and automatic—taken 'without fatigue' [14].

In order for the intervention to be treated as a nudge, it cannot impose significant material incentives and must fully keep freedom of choice [22, p. 57, 23]. Both the reminder and the warning are nudges. However, they are not subsidies, taxes, fines or imprisonment. If the intervention causes significant material costs for selected people, which of course may be justified, it is not considered as a nudge. Some

nudges work because they inform people, others act because they make choices easier, yet another work because of inertia and postponement (procrastination) [23, p. 21].

4.3 Principles of Creating Behavioural Public Policy

The principles of creating public policy based on the findings of behavioural economics are determined as behavioural change principles, basic principles of effective choice architecture or good choice architecture [18, pp. 432, 436] or strategies for policy and regulatory design [24, p. 77]. These rules usually have the form of catchy mnemonic instructions, indicating the main methods of influencing the decisions of individuals. At the same time, they are a set of guidelines for the architects of choice involved in the design of public interventions. The best-known acronym for this meaning is the word nudge described above as the behavioural change instrument.

At the same time, it is an acronym in which each letter is taken from words explaining a simple set of advice for choice architects [18, p. 437]:

- i**N**centives—create an incentive system;
- **U**nderstand choice mapping—understand how the selection process proceeds;
- **D**efaults matter—remember that settings and default options are crucial;
- **G**ive feedback—provide feedback that will help people to understand, and therefore, inform, warn against failure and praise for successes;
- **E**xpect error—remember that people make errors—a good project of behavioural change that takes into account and minimizes penalties for these errors;
- **S**tructure complex choices—try to simplify the possibilities, do not multiply the alternatives—the more choices people will have, the more complicated and problematic they will become.

NUDGES, as behavioural change principles, is a set of simple recommendations, what and how to do to be successful and focuses on counteracting cognitive errors, which are the source of an intuitive, fast thinking system (so-called System 1).

The second of strategies for policy and regulatory design called MINDSPACE refers both to System 1 of thinking and to so-called System 2—a system of reflective thinking. This set of rules to use when creating behaviour change policy includes [25, pp. 18–28, 26, 27, p. 197]:

- **M**essenger—we are heavily influenced by who communicates information;
- **I**ncentives—our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses;
- **N**orms—we are strongly influenced by what others do;
- **D**efaults—we ‘go with the flow’ of preset options;
- **S**alience—our attention is drawn to what is novel and seems relevant to us;
- **P**riming—our acts are often influenced by subconscious cues;
- **A**ffect—our emotional associations can powerfully shape our actions;

- **C**ommitments—we seek to be consistent with our public promises and reciprocate acts;
- **E**GO—we act in ways that make us feel better about ourselves.

MINDSPACE is a descriptive strategy, because by showing the mechanisms and principles of operation of the choice architects, it also creates choices in relation to the design of public activities [24, p. 78]. This strategy can help in situations if government is attempting to shape behaviour and it should do so as effectively as possible. The policymakers can use MINDSPACE to improve the effectiveness of existing and new behaviour change policies [27]. Some of the elements of MINDSPACE have been developed to explain largely automatic effects on behaviour (e.g. N, D, S, P, A) while other effects relate to elements that draw more on reflective processing (e.g. M, I, C, E) [25, p. 18].

Next strategies for behavioural changes take a different method of impact on individuals because they aim at activating reflective or free thinking systems and focus on creating the right institutional framework that enables conscious, critical and active participation of citizens in decision-making [28, pp. 18–20]. These approaches are not based on acronyms, because free thinking cannot be reduced to simple, automatic rules [24, p. 78].

One of these approaches is STEER, based on workshops that are designed to make participants aware of how their brain works and to make them aware of the mental shortcuts used in everyday decisions. The STEER strategy therefore stimulates in citizens the awareness of their limitations in making decisions and teaches them how to make important decisions in a more conscious way. In short—it teaches reflexivity in making decisions [29, pp. 177–182]. This is therefore one of the possibilities, the so-called debiasing.[2]

Second of these approaches is THINK, which can be seen as an alternative to nudge [28]. This strategy is giving citizens the space to think through and debate solutions and focuses on the process of developing public opinion and deliberate citizenships. The key ways of achieving behavioural change in this approach are debate self-ownership and collective decision-making and this strategy can be expensive to implement. Therefore, they seem to be better when concerns addressing thorny public problems than day-to-day issues of behavioural change [31, p. 18]. Its postulates can be described as follows [24, p. 79]:

- let people know a problem, let them gain knowledge;
- create a platform for discussion, expressing thoughts and beliefs—in discussion people learn, change their views and attitudes;
- give people the opportunity to participate in a process for which they will feel co-responsible, which they will be co-authors.

In this strategy, the basis for overcoming cognitive errors and activating reflective thinking is deliberation. It assumes the intellectual involvement of recipients in

[2]This is the process of 'applying techniques that intervene in and alter the situation that produces the boundedly rational behaviour, without operating on the degree of motivation or effort an actor brings to the task' [30, p. 16].

discussion, appealing to their beliefs, values and views. Its foundation is conscious and partnership-based commitment of individuals in deliberation over directions of social changes [24, p. 79].

According to Halpern, in the practical application should be used four simple ways to apply behavioural insights in public policy. The basic idea should be EAST, which means, according to him, 'that to encourage any desired behaviour, the choices should be made simple and understandable (easy), salient in citizens' lives (attractive), socially encouraged (social), and present at the key time in the decision-making process (timely)' [32 p. 60]. The EAST principles apply equally to policymakers themselves. But, this framework cannot be applied in isolation from a good understanding of the nature and context of the problem. Therefore, in practice, it was elaborated a fuller method for developing projects, which has four main stages [33, p. 7].

- Define the outcome—Identify exactly what behaviour is to be influenced. Consider how this can be measured reliably and efficiently. Establish how large a change would make the project worthwhile, and over what time period.
- Understand the context—Visit the situations and people involved in the behaviour and understand the context from their perspective. Use this opportunity to develop new insights and design a sensitive and feasible intervention.
- Build your intervention—Use the EAST framework to generate your behavioural insights. It is a mnemonic to encapsulate of the key success factors for developing an effective behavioural insights function, as a prompt to help others government to find the best way mainstreamed behaviour change as a core public policy mechanism for ensuring the public goods.
- Test, learn, adapt—Put your intervention into practice so its effects can be reliably measured. Wherever possible, use randomized controlled trials to evaluate its interventions. These introduce a control group so you can understand what would have happened if you had done nothing.

This framework is explicitly designed for applying insights in practice, and should be used alongside the existing strategies for policy and regulatory design such as MINDSPACE and Test, Learn, Adapt [32, p. 149].

The experiences in the conduct of behavioural public policies indicate that government or other public bodies are advised to pay heed to the following necessary components that can be summarized in the simple mnemonic of APPLES [34, p. 12, 35]:

- **A** for administration and the importance of gaining traction within the public administration;
- **P** for politics, because it is critical to get political buy-in;
- **P** for people who need to do the job and need to have the necessary skills and knowledge;
- **L** for the location of the unit or units who will work on behavioural insights;
- **E** for experimentation, as it helps to start with a few quick wins to show some results and develop a number of longer-term projects;

- **S** for scholarship and academia which can be an essential resource.

It is a mnemonic to encapsulate of the key success factors for developing an effective behavioural insights function, as a prompt to help others government to find the best way mainstream behaviour change as a core public policy mechanism for ensuring the public goods [36, p. 73].

4.4 Conclusion

The objective of the public policy is to solve a specific problem or even reduce its negative effects. It may be conducted using traditional tools, like restrictions, incentives and increased information or using the findings of behavioural economics to help design the environments in which humans make decisions to benefit society. The latter method of making the public policy in many countries is used more and more often, and as Goldin [37] emphasizes the literature on the subject shows that nudging is inevitable and a government cannot help but nudge: policymakers end up shaping choice architecture whether they intend to or not.

The behavioural approach to the public policy-making described as innovative is promoted by institutions such as the World Bank [38], the European Commission [39] or the Organization for Economic Cooperation and Development [40]. It is currently used in many countries as part of one of the three institutional models that mutually exclusive and can coexist [40]:

- specialized units within the Centre of Government;
- networks of teams across ministries and agencies, even involving academic institutions and the private sector or not-for-profit organisations;
- ad hoc approaches, where behavioural insights are used for specific projects and initiatives.

Experiences of conducting behavioural policy in these institutional frameworks have become the source of elaboration of strategies for policy and regulatory design presented in this article, aimed at facilitating and encouraging the more frequent use of nudges —elements of the choice architecture. Applying this approach should lead to increasing the efficiency of the public policy, often thanks to quite small, simple and not too expensive interventions. The main task of the choice architecture, i.e. a task of organizing the context in which the decision is made, is replacing one biasing context by another. As presented by Thaler and Sunstein, it is justified if there is evidence that one option is actually better than the other, and the new context leads to that option being chosen more often [14, p. 236].

However nowadays, in the literature on the subject, it is emphasized that the impact of choice architecture on decisions is not always easy to anticipate, because choice architects should also be aware of the implicit interaction taking place between the targets of the choice architecture and themselves. It is therefore proposed to supplement the current approach to behavioural public policy (defined as choice architecture

1.0) by analyzing the hidden social interactions inherent in choice architecture. This approach, defined as choice architecture 2.0 conceives of targeted individuals as 'social sensemakers' [41].

The literature on the subject also emphasizes that governments faced with frequent and unavoidable decisions regarding the choice architecture (selection framework, selection context), aimed at influencing the voluntary choices of people in a way that leads them to increase their welfare, are forced to specify the best direction to nudge. According to Goldin [37], it is quite simple in two cases, when:

- policymakers have strong (possibly paternalistic) views of their own about which of the available options would make decision-makers better off;
- the decision under consideration involves substantial externalities to other members of society.

In other cases the public policymakers face the problem of identification of behavioural preference, which they have to cope and must have some way of identifying which nudge will make decision-makers better off.

Another direction of current analyses related to the behavioural approach to making public policy is research about what people think about the choice architecture and nudge, conducted in various countries [42] and if individuals want to make the choices that they are being nudged towards [43]. The conclusions of these studies should contribute to creation of more effective public policy based on evidence.

References

1. Kahneman, D.: Thinking, Fast and Slow. Macmillan, New York (2011)
2. Q&A with Richard Thaler.: In: Samson, A. (ed.) The Behavioral Economics Guide 2016 (with an introduction by Gerd Gigerenzer), pp. 23–24. Available via http://www.behavioraleconomics.com (2016)
3. Thaler, R.H.: Misbehaving: the Making of Behavioral Economics. W.W. Norton, New York (2015)
4. Mitchell, G.: Libertarian paternalism is an oxymoron. Northwest. Univ. Law Rev. **99**(3), 1245–1277 (2004)
5. Smith, M.A., McPherson, M.S.: Nudging for equality: values in libertarian paternalism. Admin. Law Rev. **61**(2), 323–342 (2009)
6. Schlag, P.: Nudge, choice architecture and libertarian paternalism. Michigan Law Rev. **108**(6), 913–924. Available via https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1278&context=mlr (2010)
7. Vallgårda, S.: Nudge a new and better way to improve health? Health Policy **104**(2), 200–203 (2012). https://doi.org/10.1016/j.healthpol.2011.10.013
8. White, M.D.: The richness of personal interests: a neglected aspect of the nudge debate. LSE British Politics & Policy, 23 Oct. Available via https://blogs.lse.ac.uk/politicsandpolicy/the-richness-of-personal-interests-a-neglected-aspect-of-the-nudge-debate/ (2013a)
9. Mills, Ch.: Why Nudges matter: a reply to goodwin. Politics **33**(1), 28–36 (2013). https://doi.org/10.1111/j.1467-9256.2012.01450.x
10. White, M.D.: The Manipulation of Choice: Ethics and Libertarian Paternalism. Palgrave Macmillan, New York (2013)

11. Lunn, P.: Regulatory policy and behavioural economics. OECD Publishing. https://doi.org/10.1787/9789264207851-en
12. Persad, G.: Libertarian Patriarchalism: Nudges, Procedural Roadblocks, and Reproductive Choice. Georgetown Law Faculty Publications and Other Works 1508. Women's Rights Law Reporter **35**(3/4), 273–298. Available via https://scholarship.law.georgetown.edu/facpub/1508 (2014)
13. Schweizer, M.: Nudging and the principle of proportionality—obliged to Nudge? Available via https://www.academia.edu/37264133/Nudging_and_the_principle_of_proportionality_-_obliged_to_nudge (2015)
14. Thaler, R.H., Sunstein, C.: Nudge: Improving Decisions About Health, Wealth and Happiness. Yales University Press, New Haven, CT (2008)
15. Thaler, R.H., Sunstein, C.R.: Libertarian Paternalism. Am. Econ. Rev. **93**(2), 175–179. https://doi.org/10.1257/000282803321947001. Available via http://www.aeaweb.org/articles.php?doi=10.1257/000282803321947001 (2003)
16. Sunstein, C.R., Thaler, R.H.: Libertarian paternalism is not an oxymoron. Univ. Chic. Law Rev **70**(4), 1159–1202. Available via https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=5228&context=uclrev (2003)
17. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. Science **30**, 453–458 (1981)
18. Thaler, R.H., Sunstein, C.R., Balz, J.: Choice architecture. In: Shafir, E. (ed.) The Behavioral Foundations Of Public Policy, pp. 428–439. Princeton University Press, Princeton (2012)
19. Gordon, W.: Behavioural economics and qualitative research—a marriage made in heaven? Int. J. Market Res. **53**(2), 2–12 (2011)
20. Hansen, P.G.: The definition of nudge and libertarian paternalism: does the hand fit the glove? Eur. J. Risk Regul. **7**(1), 155–174 (2016)
21. Oliver, A.: From Nudging to budging: using behavioural economics to inform public sector policy. J. Soc. Policy **42**(4), 685–700 (2013). https://doi.org/10.1017/S0047279413000299
22. Sunstein, C.R.: Why Nudge? the Politics of Libertarian Paternalism. Yale University Press, New Haven (2014)
23. Sunstein, C.R.: The Ethics of Influence. Government in the Age of Behavioral Science. Cambridge University Press, Cambridge (2016)
24. Olejniczak, K., Śliwowski, P.: Towards behaviorally informed public interventions. Manag. Bus. Admin. Cent. Eur. **23**(2), 61–91 (2015). https://doi.org/10.7206/mba.ce.2084-3356.142
25. Dolan, P., Hallsworth, M., Halpern, D., King, D., Vlaev, I.: MINDSPACE: influencing behaviour through public policy. Institute for Government, London, UK. Available via https://www.instituteforgovernment.org.uk/sites/default/files/publications/MINDSPACE-Practical-guide-final-Web_1.pdf (2010)
26. Dolan, P.: Influencing the financial behaviour of individuals: the mindspace way. In: Oliver, A. (ed.) Behavioural Public Policy, pp. 191–208. Cambridge University Press, Cambridge (2013)
27. Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., Vlaev, I.: Influencing behaviour: the mindspace way. J. Econ. Psychol. **33**, 264–277 (2012). https://doi.org/10.1016/j.joep.2011.10.009
28. John, P., Cotterill, S., Richardson, L., Moseley, A., Smith, G., Stoker, G., Wales, C.: Nudge, Nudge, Think, Think: Using Experiments to Change Civic Behavior. Bloomsbury Academic, London, New York (2011)
29. Jones, R., Pykett, J., Whitehead, M.: Changing behaviours. On the rise of the psychological state. Edward Elgar, Cheltenham (2013)
30. Jolls, C., Sunstein, C.R.: Debiasing through law. Harvard Law School John M. Olin Center for Law, economics and business discussion paper series 495. Available via http://lsr.nellco.org/harvard_olin/495 (2004)
31. John, P.: How far to Nudge? assessing behavioral public policy. Series: new horizons in public policy. Edward Elgar Publishing, Northampton, MA, US. https://doi.org/10.4337/9781786430557 (2018)
32. Halpern, D.: Inside the Nudge unit. WH Allen, London (2015)

33. Service, O., Hallsworth, M., Halpern, D., Algate, F., Gallagher, R., Nguyen, S., Ruda, S., Sanders, M., Pelenur, M., Gyani, A., Harper, H., Reinhard, J., Kirkma, E.: EAST. Four simple ways to apply behavioural insights. The Behavioral Insight Team UK, Behavioural Insights Ltd. Available via http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/ (2014)
34. OECD: Behavioural insights and new approaches to policy design. The views from the field Summary of an international seminar. OECD Publishing, Paris. Available via https://www.oecd.org/gov/regulatory-policy/behavioural-insights-summary-report-2015.pdf (2015)
35. Halpern, D., Sanders, M.: Nudging by government: progress, impact, and lessons learned. Behav. Sci. Policy **2**(2), 53–65 (2016)
36. Pykett, J., Howell, R., Lilley, R., Jones, R., Whitehead, M.: Governing mindfully: shaping policy makers' emotional engagements with behaviour change. In: Jupp, E., Pykett, J., Smith, F.M. (eds.) Emotional States: Sites and Spaces of Affective Governance, pp. 69–88. Routledge, Taylor & Francis Group, Abingdon, Oxon, New York, NY (2017)
37. Goldin, J.: Which way to Nudge? uncovering preferences in the behavioral age. Yale Law J. **125**, 226–270 (2015)
38. World Bank: Mind, Society, and Behavior. World Bank Group, Washington DC (2015)
39. Lourenço, J.S., Ciriolo, E., Almeida, S.R., Troussard, X.: Behavioural insights applied to policy: European report 2016. EUR 27726 EN. https://doi.org/10.2760/903938 (2016)
40. OECD: Behavioural insights and public policy: lessons from around the world. OECD Publishing, Paris. https://doi.org/10.1787/9789264270480-en (2017)
41. Krijnen, J.M.T., Tannenbaum, D., Fox, C.R.: Choice architecture 2.0: behavioral policy as an implicit social interaction. Behav. Sci. Policy **3**(2), 1–18 (2017)
42. Sunstein, C.R., Reisch, L.A., Rauber, J.: A worldwide consensus on Nudging? not quite, but almost. Regul. Gov. **12**, 3–22 (2018). https://doi.org/10.1111/rego.12161
43. Sudgen, R.: Do people really want to be nudged towards healthy lifestyles? Int. Rev. Econ. **64**, 113–123 (2017). https://doi.org/10.1007/s12232-016-0264-1

Chapter 5
Application of Behavioral Economics Insights to Increase Effectiveness of Public Awareness Campaigns

Anna Borawska

Abstract The traditional (neoclassical) economics assumes that every individual is egoistically oriented toward achieving their main goal (their own interest), which is to maximize utility. However, in many studies referring to various aspects of human behavior, behavioral economics proves that human choices vary depending on the circumstances, place, time, norms and social influences, emotional judgments, cognitive distortions and biases, simplifying reasoning principles applied (heuristics), and at the same time on how and in what circumstances the choice is made (the choice architecture). The objective of the article is to define the concepts of behavioral economics which are the most interesting from the point of view of increasing the effectiveness of public awareness campaigns. In order to determine which concepts of behavioral economics are the most interesting from the point of view of increasing the effectiveness of social campaigns, an analysis of the results of the systematic publication search in the Google Scholar database has been carried out. Literature overview has shown that the effectiveness of social campaigns can be increased by using the knowledge provided by behavioral economics on the subject of reflexive, unreflective, unwise, and fast cognitive processes carried out by individuals.

Keywords Public awareness campaigns · Effectiveness · Behavioral economics

5.1 Introduction

The traditional (neoclassical) economics assumes that every individual is egoistically oriented toward achieving their main goal (their own interest), which is to maximize utility. Using all available information, a fully rational individual carefully calculates and evaluates each of the available options to choose the one that ensures goal achieving. The rational choice theory, according to which people act rationally, are driven by self-interest and strive to maximize their benefits while minimizing costs,

A. Borawska (✉)
Faculty of Economics and Management, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: anna.borawska@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_5

leads the homo oeconomicus model, which is an idealistic model of individual choice [2]. This model stipulates that people provided with information, incentives, orders, or prohibitions will also react in a rational and thought-out manner, calculating their benefits and losses and considering all "pros and cons" [18, pp. 1–2].

However, in many studies referring to various aspects of human behavior, behavioral economics proves that human choices vary depending on the circumstances, place, time, norms and social influences, emotional judgments, cognitive distortions and biases, simplifying reasoning principles applied (heuristics), and at the same time on how and in what circumstances the choice is made (the choice architecture). The findings of behavioral economics, resulting from numerous experiments, indicate that both the consequences of unintentional (unconscious) and organized (ordered) decision-making do not follow the traditional model of homo oeconomicus, but rather lead to a specific choice depending on the conditions in which it is made.

Knowledge derived from behavioral economics explorations is more and more applied in practice, and its findings may be of fundamental importance for institutions (public and private) which create social campaigns. This knowledge helps in better understanding, explanation, and prediction of the real reactions of people to messages contained in social campaigns, contributing to increase their effectiveness. This is of a special social importance, because their creators try to solve specific social problems, and the main goal of such campaigns is to pursue interests for the common good—educating and persuading recipients to a certain value, assimilating social ideas, promoting pro-social attitudes, changing behavior or getting them to certain activities desirable from a social point of view, and overcoming social pathologies [27, p. 8].

The objective of the article is to define the concepts of behavioral economics which are the most interesting from the point of view of increasing the effectiveness of social campaigns. The article has been divided into 3 parts. The first part describes the nature of behavioral economics and its foundations, the second one presents the idea of social campaigns and measurement of their effectiveness, and the third one includes the outcomes from the literature review on the use of behavioral economy findings in social campaigns. The article concludes with a summary.

5.2 Behavioral Economics and Its Foundations

In its traditional form, economics assumes that individuals always make decisions in order to maximize utility. However, many observations and studies conducted to confirm or reject economic theories show that humans do not always behave in such rational way. These remarks led to the occurrence of a new trend within the field of economics. It is commonly known as behavioral economics (BE), and it adds to economics more realistic psychological foundations [5]. Its role is to extend the traditional framework, but it is not intended to replace it [11]. The behavioral economics enquiries include two elements: recognizing the ways in which behavior

differs from the standard model and showing how this behavior is important in economic contexts [21].

The origins of BE can be traced back to the first half of the twentieth century. Already at that time, there were economists that used some psychological factors in their work (Irving Fisher, Vilfredo Pareto, and John Maynard Keynes) (see [34]). But, it is Herbert Simon and his "bounded rationality" concept that became a real foundation of behavioral economics [24]. Further development of the BE domain happened during the 1970s, and it is strongly connected with the research of two psychologists: Tversky and Kahneman. Their contribution consists of mostly "heuristics and biases" [30] program and prospect theory [12] explaining the different departures from rational choice. Since then, field of behavioral economics has developed greatly as both economists and psychologists have expanded the work of above-mentioned pioneers [10].

Behavioral economics concepts do not focus only on the explanation of human behavior and deviations from rationality. Some researchers (i.e., [26]) have turned BE into a doctrine that gives the tools to influence human behavior change. This doctrine is known as "nudging" or "architecture of choice". Its concepts can be effectively used by institutions (public and private) which create social campaigns.

5.3 Public Awareness Campaigns and Measurement of its Effectiveness

The public awareness campaign is one of the elements of the social marketing [3, 14, 15] and, according to the frequently quoted definition, it is "a set of various activities planned at a specific time, addressed to a specific target group, whose aim is to increase knowledge, change in thinking, and behavior toward a specific social problem or to solve a social problem which blocks the achievement of the common good defined as a given marketing goal" [8]. As part of the campaign, information, intervention or information and intervention measures may be taken [13]. The selection of the right strategy is most often dependent on the problem that the campaign is to face.

The quality of a social campaign is usually determined on the basis of its effectiveness, i.e., the degree of achieving the intended goal [35]. As a rule, due to the complexity of the issue, the assessment is carried out on five different levels [7]:

- awareness of the problem,
- involvement,
- desired change in behavior,
- establishing a social norm (widespread and permanent change in behavior),
- general improvement of the social and environmental situation.

At each of these levels, there are other indicators that evidence the success achieved by the campaign's authors. In order to collect relevant data, different methods are also

applied at each stage. Table 5.1 presents a summary of commonly used indicators together with the methods of data collection specific for such indicators.

The selection of methods to measure effectiveness is conditioned by both the nature of the actions taken in the campaign and the stage at which the campaign is currently located. The effectiveness of information activities and interventions is checked differently. There are other ways to test the effectiveness of the campaign before it is published and some others after its completion. However, not all methods of obtaining data to assess the effectiveness of the campaign are equally reliable. The

Table 5.1 Indicators and methods to measure the effectiveness of social campaigns at the five levels

Level	Indicator	Method of measurement
Awareness	• The percentage of recipients aware of a problem	• Surveys among recipients
Involvement	• The percentage of recipients involved in deliberations and discussions about the problem • The percentage of recipients taking actions to obtain additional knowledge about the problem	• Surveys among recipients • Behavioral data (e.g., number of visits on the campaign's webpage)
Change in behavior	• The percentage of recipients who declare to change behavior • The percentage of recipients in whom changes were observed	• Surveys among recipients • Behavioral data (e.g., increase in number of people undergoing control tests)
Establishing a social norm	• The percentage of recipients who present positive attitude to the problem • The percentage of newspaper articles and opinions favorable to the campaign • legal regulations regarding the problem	• Surveys among recipients observations • Anecdotal feedback • Monitoring of media and politics
General improvement of a situation	• The percentage increase in social indicators • The percentage increase in environmental indicators	• Epidemiological data • Environmental data • Data regarding safety on roads • …

Source Varcoe [31]

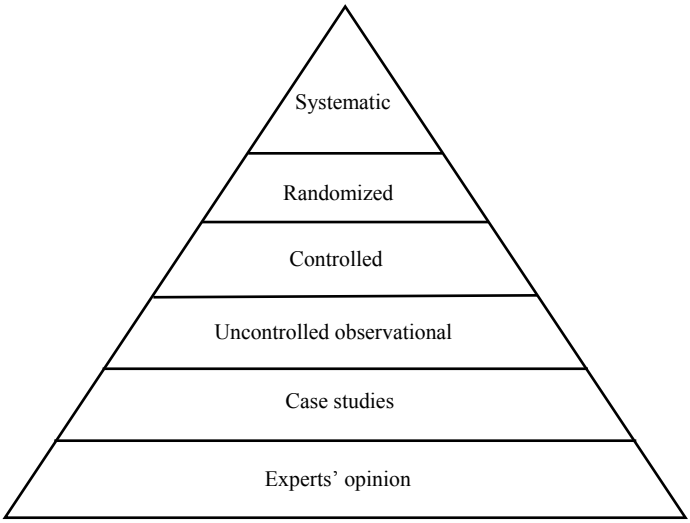


Fig. 5.1 The hierarchy of scientific evidence. *Source* Own elaboration under [25]

hierarchy of scientific evidence, presented in Fig. 5.1, ranks possible tests from the least (the base of the pyramid) to the most reliable (the top of the pyramid).

The presented hierarchy served as the basis for selecting examples of the application of the concept of behavioral economics to increase the effectiveness of social campaigns. First, publications presenting systematic reviews and randomized controlled trials were included. Then, controlled and uncontrolled observational studies were also taken into account in order to broaden the thematic scope of the analyzed campaigns.

5.4 Use of Methods of Behavioral Economics in Public Awareness Campaigns

In order to determine which concepts of behavioral economics are the most interesting from the point of view of increasing the effectiveness of social campaigns, an analysis of the results of the publication search in the Google Scholar database has been carried out. The usefulness of this base in conducting the literature review has been confirmed by researches (See [9]). The search was carried out initially with the following set of key words: "behavioral economics," "social marketing," "social campaign," "public awareness campaign," "public service announcement," and "social advertisement". This stage was used to separate publications that combined information about behavioral economics with those about social campaigns. Table 5.2 presents the results showing the number of publications meeting the preliminary search conditions.

Table 5.2 Number of search results of publications which contain specific key words

Key words		Number of search results
"Behavioral economics"+	"Social marketing"	1240
	"Public service announcement"	66
	"Public awareness campaign"	41
	"Social campaign"	11
	"Social advertisement"	10

Source Own elaboration

Analysis of the found sources allowed distinguishing those concepts of behavioral economics, which are most often considered and applied in relation to social campaigns. Among them, three main strategies can be distinguished:

- reference to social norms,
- financial incentives,
- message framing.

Next, key words were identified that allowed to search for sources reporting the effectiveness of using the above-mentioned concepts of behavioral economics in social campaigns, especially those containing quantitative results. Tables 5.3, 5.4, and 5.5 present a sequence of the selected publications on the possibility of referring to social norms, the use of financial incentives, and message framing. First of all, sources containing studies conducted with the use of methods that are high in the hierarchy of scientific evidence were selected for the compilation. At the same time, the broadest possible thematic spectrum of campaigns that can benefit from the use of the concept of behavioral economics was presented.

5.5 Conclusion

The effectiveness of social campaigns can be increased by using the knowledge provided by behavioral economics on the subject of reflexive, unreflective, unwise, and fast cognitive processes carried out by individuals. Behavioral economics proves that people's choices and behavior are always influenced by the decision-making context. In order to guide people to more sensible decision-making and making better choices (in their own judgment), without forcing anyone to achieve specific results, the creators of social campaigns can use heuristics and cognitive biases as well as nudges as elements of choice architecture.

The analysis of world literature allows us to state that the most commonly used in this area are strategies based on referring to social norms, appropriate message

Table 5.3 Studies on social campaigns referring to social norms

Area of application	Publication	Research method	Summary of results
Campaign to counteract driving a car after drinking alcohol (USA)	Perkins et al. [22]	Controlled observational studies	The campaign, based on references to social norms, allowed to reduce the number of cases of driving a car after alcohol
Campaign to limit speed (Great Britain)	Toy et al. [29]	Uncontrolled observational studies	Individual initiatives to convince people to reduce speed are ineffective. In order for changes to be made, it is necessary to create an appropriate social norm
Campaign for the reuse of towels by hotel guests	Bohner and Schlüter [4]	Controlled observational studies	Studies have proved that the message based on the social norm made the rate of reuse of towels increased
Promoting healthy eating	Robinson et al. [23]	Randomized controlled trials	Messages communicating information about social norms were more effective in encouraging respondents to healthy eating than messages with basic health information.

Source Own elaboration

framing, and the use of financial incentives. However, these approaches should not exhaust the list of possibilities to influence people's behavior in order to focus them on making good decisions. Taking into account the aforementioned multiplicity of heuristics and cognitive biases as well as the interventions in the form of nudges used in architecture, it can be concluded that research should be carried out to increase the use of behavioral economics in designing effective social campaigns.

The new role of the choice architects, which should be played by the creators of social campaigns, means choosing and creating decision-making contexts in a way that promotes behavior that is in our own, as well as the general public interest. By creatively shaping the choice of recipients of social campaigns, mainly through the use of soft persuasion and targeting desirable behaviors, their creators can easily and cheaply influence our decisions, giving us free choice.

Table 5.4 Studies on social campaigns using financial incentives

Area of application	Publication	Research method	Summary of results
Weight reduction	Volpp et al. [33]	Randomized controlled trials	Participants in the study, who were supported by financial incentives, lost on average more kilograms than those in the control group
Quitting smoking	Volpp et al. [32]	Randomized controlled trials	In the group with a financial incentive, the percentage of people who quit smoking was significantly higher than in the group that only received information about smoking cessation programs
Promoting use of public transport	Thøgersen [28]	Randomized controlled trials	Handing the free monthly card to the group of respondents resulted in a significant increase in the frequency of using public transport
Promoting walks among older adults	Kullgren et al. [16]	Randomized controlled trials	The financial incentive turned out to be the most effective method of promotion (in comparison to feedback or creating social networks)
Encouraging children and young people to eat fruits and vegetables	Morrill et al. [20]	Randomized controlled trials	Students participating in the program under study increased consumption of fruits and vegetables, especially in schools, where they received material prizes

Source Own elaboration

Table 5.5 Research on social campaigns based on message framing

Area of application	Publication	Research method	Summary of results
Social advertising against drug use	Cho and Boster [6]	Uncontrolled observational studies (post-test)	Messages which emphasize the loss resulting from the use of drugs are more effective than those that emphasize their profits from quitting
Promoting physical activity	Latimer et al. [17]	Controlled observational studies	Messages highlighting the benefits of physical activity can increase the effectiveness of promotional materials
Promoting vaccination	Abhyankar et al. [1]	Randomized controlled trials	Messages which emphasize losses resulting from not applying vaccines were found to be more effective than those indicating gains from undergoing a vaccination
Motivating youth to quit smoking	Mays et al. [19]	Randomized controlled trials	People with high self-esteem responded better to messages highlighting the losses resulting from smoking, and people with low self-esteem presented a reverse trend

Source Own elaboration

Acknowledgements The project was financed with the National Science Centre funds allocated according to the decision DEC-2016/21/B/HS4/03036.

References

1. Abhyankar, P., O'connor, D.B., Lawton, R.: The role of message framing in promoting MMR vaccination: evidence of a loss-frame advantage. Psychol. Health Med. **13**, 1–16 (2008)
2. Amadae, S.: Rational choice theory. In: Bevir, M. (ed.) Encyclopedia of Governance, pp. 785–791. Sage Publications (2007)
3. Andreasen, A.R.: A social marketing approach to changing mental health practices directed at youth and adolescents. Health Mark. Q. **21**(4), 51–75 (2004)

4. Bohner, G., Schlüter, L.E.: A room with a viewpoint revisited: descriptive norms and hotel guests' towel reuse behavior. PLoS ONE **9**(8), (2014)
5. Camerer, C., Loewenstein, G.: Behavioral economics: past, present, future. In: Camerer, C., Loewenstein, G., Rabin, M. (eds.) Advances in Behavioral Economics, pp. 3–51. Princeton University Press
6. Cho, H., Boster, F.J.: Effects of gain versus loss frame antidrug ads on adolescents. J. Commun. **58**, 428–446 (2008)
7. Donovan, R., Henely, N.: Social Marketing: Principles and Practice. IP Communications (2003)
8. Fundacja Komunikacji Społecznej.: Definition of public awareness campaign. http://kampaniespoleczne.pl/kampania-spoleczna-definicja-fundacji-komunikacji-spolecznej/. Accessed 23 Mar 2019 (2010)
9. Gehanno, J.F., Rollin, L., Darmoni, S.: Is the coverage of Google Scholar enough to be used alone for systematic reviews. BMC Med. Inform. Decis. Mak. **13**(1), 7 (2013)
10. Heukelom, F.: Behavioral Economics: A History. Cambridge University Press (2014)
11. Ho, T.H., Lim, N., Camerer, C.F.: Modeling the psychology of consumer and firm behavior with behavioral economics. J. Mark. Res. **43**, 307–331 (2006)
12. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica **XLVII**, 263–291 (1979)
13. Kotler, P., Lee, N.: Social Marketing: Influencing Behaviors for Good. SAGE Publications (2011)
14. Kotler, P., Roberto, N., Lee, N.: Social Marketing: Improving the Quality of Life. SAGE (2002)
15. Kotler, P., Zaltman, G.: Social marketing: an approach to planned social change. J. Market. **35**, 3–12 (1971)
16. Kullgren, J.T., Harkins, K.A., Bellamy, S.L., et al.: A mixed-methods randomized controlled trial of financial incentives and peer networks to promote walking among older adults. Health Edu. Behav. **41**(1 Suppl), 43S–50S (2014)
17. Latimer, A.E., Rench, T.A., Rivers, S.E., et al.: Promoting participation in physical activity using framed messages: an application of prospect theory. Br. J. Health. Psychol. **13**(Pt 4), 659–681 (2008)
18. Low, D.: Cognition, choice and policy design. In: Low, D. (ed.) Behavioural Economics and Policy Design: Examples from Singapore, pp. 1–13. World Scientific Publishing Company (2011)
19. Mays, D., Turner, M.M., Zhao, X., et al.: Framing pictorial cigarette warning labels to motivate young smokers to quit. Nicotine Tob. Res. **17**(7), 769–775 (2014)
20. Morrill, B.A., Madden, G.J., Wengreen, H.J., et al.: A randomized controlled trial of the Food Dudes Program: tangible rewards are more effective than social rewards for increasing short-and long-term fruit and vegetable consumption. J. Acad. Nutr. Dietetics **116**(4), 618–629 (2016)
21. Mullainathan, S., Thaler, R.H.: Behavioral economics No. w7948. National Bureau of Economic Research (2000)
22. Perkins, H.W., Linkenbach, J., Lewis, M., et al.: Effectiveness of social norms media marketing in reducing drinking and driving: a statewide campaign. Addict. Behav. **35**, 866–874 (2010)
23. Robinson, E., Fleming, A., Higgs, S.: Prompting healthier eating: testing the use of health and social norm based messages. Health Psychol. **33**(9), 1057 (2014)
24. Simon, H.A.: A behavioral model of rational choice. Q. J. Econ. **69**(1), 99–118 (1955)
25. Steier, G., Steier, L.: Textbox: the hierarchy of scientific evidence. In: International Food Law and Policy, pp. 319–320. Springer (2016)
26. Thaler, R.H., Sunstein, C.: Nudge: Improving Decisions About Health, Wealth and Happiness. Yales University Press, New Haven, CT (2008)
27. Thesis, M., Andrews, M.: Social Campaigns, Art of Visual Persuasion Its Psychology, Its Semiotics, Its Rhetoric. http://www.mahku.nl/download/m_andrews_socialcampaigns.pdf. Accessed 23 Mar 2019 (2008)
28. Thøgersen, J.: Promoting public transport as a subscription service: effects of a free month travel card. Transp. Policy **16**, 335–343 (2009)

29. Toy, S., Tapp, A., Musselwhite, C., et al.: Can social marketing make 20 mph the new norm? J. Trans. Health **1**, 165–173 (2014)
30. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science **185**(4157), 1124–1131 (1974)
31. Varcoe, J.: Assessing the effectiveness of social marketing. https://www.dors.it/documentazione/testo/200906/Ass_SM_VARCOE.pdf. Accessed 23 Mar 2019 (2004)
32. Volpp, K.G., Troxel, A.B., Pauly, M.V., et al.: A randomized, controlled trial of financial incentives for smoking cessation. N. Engl. J. Med. **360**(7), 699–709 (2009)
33. Volpp, K.G., John, L.K., Troxel, A.B., et al.: Financial incentive-based approaches for weight loss. A randomized trial. JAMA **300**(22), 2631–2637 (2008)
34. Wilkinson, N., Klaes, M.: An Introduction to Behavioral Economics. Macmillan International Higher Education (2017)
35. Wymer, W.: Developing more effective social marketing strategies. J. Soc. Market. **1**(1), 17–31 (2011)

Chapter 6
Experimental Methods in Behavioural Economic Sciences

Dominika Korzeniowska and Łukasz Sułkowski

Abstract The aim of the chapter is to present, basing on neo-evolutionary paradigm, the types of experimental methods used in behavioural economic sciences supported by research examples. Neo-evolutionary paradigm legitimates incorporation into behavioural economic sciences not only behavioural economics, behavioural finance and behavioural accounting but also neuroeconomics, neurofinance and neuroaccounting. The adoption of the aforementioned approach enables extending the methodological perspective by methods used in natural sciences, including experiment. The research methodology is based on a comprehensive review of the world's literature on behavioural economics, neuroscience and neo-evolutionism.

Keywords Experiment · Behavioural economics · Neuroeconomics · Neo-evolutionism

6.1 Introduction

Behavioural stream is one of the most rapidly developing areas within contemporary economics. The existing reflection is mostly founded on certain assumptions drawn from cognitive psychology, which deal with research into cognitive structures and processes, and into the organisation and functioning of the human mind. Cognitive psychology, however, treats these structures, processes and the ways the human mind works as existing elements. But it does not ask questions about why they developed in a way (that is a domain of evolutionary psychology), or how learning processes and constant interactions between organism and its environment shapes behaviour (as behavioural psychology does) [3]. Moreover, cognitive psychology postulates, by means of observation and behaviour analysis, the existence of cognitive systems, which in fact take the form of models further used to explain behaviours. So behaviour

D. Korzeniowska (✉) · Ł. Sułkowski
Społeczna Akademia Nauk w Łodzi, ul. Sienkiewicza 9, 90-113 Łódź, Poland
e-mail: dkorzeniowska@spoleczna.pl

Ł. Sułkowski
e-mail: lsulkowski@spoleczna.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_6

is a base to build a model that explains behaviour. It looks like vicious circle of inference. The possibility to reveal universal laws underlying human behaviour that would allow its prediction regardless of environmental conditions is also puzzling. Assuming we use brain that evolved hundreds of thousands of years ago and whose structure and "operation rules" remained relatively unchanged thus far, it seems possible to some extent.

Extending cognitive perspective by adopting another paradigm, i.e. the neo-evolutionary one, would provide the inspiration to develop new research and to make valuable discoveries. Studying the response of brain in economic situations by means of experimental methods could put some light on what happens between the input and output data, and consequently lead to the creation of models better predicting behaviour. Furthermore, neo-evolutionary paradigm, if applied into behavioural stream in economic sciences, legitimizes inclusion of neuroeconomics, neurofinance and neuroaccounting into that scope of reflection [35]. "Behavioural economic sciences" then embrace behavioural economics, finance and accounting together with their three "neuro-versions".

The aim of the chapter is thus to present experimental methods used in behavioural economic sciences that, thanks to employment of neo-evolutionary paradigm, expand their research methodology perspective by utilization of methods applicable for natural sciences. It bases on comprehensive literature review on behavioural economics, neuroscience and neo-evolutionism.

6.2 New Research Methods in Social Sciences

The evolutionary paradigm, in the sphere of methodology, is, as mentioned before, directed towards the application of the scientific methods of the natural sciences, especially experiment and controlled observation. Methodological constraints comprise representativeness and randomisation of trials, and control of variables, all of which facilitate the presentation of the results in a mathematical form. Experimentation and observation should lead to an incomplete verification or falsification of hypotheses resulting from theory. It, therefore, seems that the application of the experimental method in the social sciences requires its adaptation to problems of the highest level of complexity. At the same time, the methodological achievements obtained so far by the social sciences should not be rejected. The methodology of the neo-evolutionary trend in the social sciences is then of eclectic nature. Its core is shaped by methods such as social experiment and controlled observation, but it is also possible to reach for the quantitative and qualitative methodology of traditionally understood social sciences [34].

The largest body of research studies that can be seen as rooted in the neo-evolutionary trend has been conducted within evolutionary psychology [12]. The methods of experiment and quasi-experiment, observation and descriptive methods

dominate in this area. Psychology has developed methods of controlled social experiment that, when accompanied by elements of experiment from the sciences, may also be applied in experimental psychology [15].

Primatological research, related to the analysis of the behaviour of primates, is also included in the neo-evolutionary trend. It basically contains controlled observations in both artificial and natural conditions, experiments and quasi-experiments, including use of physiological examination, e.g. of the brain [29]. Such research studies allow to find answers to questions on the stage of development of cognitive and emotive skills, communication skills and social structures among apes. Comparison of the results allows us to discover the specifics of human behaviours.

There are three programmes of known studies in the social sciences within the neo-evolutionary trend: verification of the hypothesis of cultural universalism with evolutionary roots; indication of the perceptive and cognitive illusions related to the structure of the human cognitive system, and examination of the ability to detect evolutionary cheaters, the human moral sense and decision-making processes. Research programmes conducted by Tversky and Kahneman [39] over a number of years have allowed to describe a series of cognitive limitations on the human mind related to the assessment of probability, the search for algorithms in random sequences and other illusions falsifying the picture of reality and others. Tooby and Cosmides [37] carried out social experiments on the adaptation of the human mind to detecting evolutionary cheaters, i.e. individuals that only simulate group cooperation, while being untrustworthy. There are also famous observations and quasi-experiments conducted by Goodall, and later by de Waal [13], on chimpanzees. They show that man's simian relatives demonstrate many traits which used to be considered exclusively human, such as cooperation and formation of coalitions, wars, intrigues, communication, creation and use of tools [17].

6.3 The Essence of Behavioural Economics, Finance and Accounting

The "innate" experimental nature of the three behavioural disciplines was determined by their research subject or the other way round. Behavioural economics and its "sisters" base on the need for studying human imagination or emotional and cognitive structures that impact economic decisions [1, p. 658]. "Behavioural economics increase the explanatory power of economics by providing it with more realistic psychological foundations" [11, p. 3]. Brzezicka and Wiśniewski's quite comprehensive conceptualization [8, p. 354] accentuates experimental nature of this discipline stating that "behavioural economics is an experimental science that uses the scientific approach to test and better understand economic theories". Behavioural finance is, in turn, the "study of how human interprets and acts on information to make informed investment decisions" [23] or it "attempts to explain and increase understanding of the reasoning patterns of investors, including the emotional processes involved

and the degree to which they influence the decision-making process". Essentially, behavioural finance aims to explain the "what, why and how of finance and investing, from a human perspective" [28, p. 2]. While the fundamental assumption of behavioural accounting is to predict human behaviour in all possible accounting aspects. According to Belkaoui [5, preface] "behavioural accounting emphasizes the relevance of accounting information to decision-making as well as the individual and group behaviour caused by the communication of this information", what is more, "its purpose is to influence action or behaviour directly through the information content of the message conveyed and indirectly through the behaviour of accountants".

The behavioural trend thus opens new paths for understanding people's economic behaviour, departing from the rather hermetic assumption of *homo oeconomicus*. However, the cognitive paradigm dominating herein does not constitute sufficient basis for discovering certain mechanisms that decision-maker is not aware of as they are uncontrollable or unconscious. Adoption of the neo-evolutionary paradigm in methodology of research on economic behaviours favours taking into account a number of biological mechanisms in human body, causing that the choices made by individuals significantly deviate from the patterns proclaimed by neoclassical economy. These studies show the huge complexity of the decision-making processes that stems not only from imperfect human cognitive system but also his physiology. Such research would be impossible without medical techniques allowing for in vivo examination of the nervous system and the human brain. The dissemination of those non-invasive methods thus created new possibilities to explore what is hidden in human body and became a milestone in the emergence of neuroscience.

6.4 The Effect of Marriage Between Neuroscience and Behavioural Economics

By definition "neuroscience is the multidisciplinary science that analyses the nervous system to understand the biological basis for behaviour" [7, p. 3]. The term "neuroscience" appeared for the first time in the mid-60s of the last century, as a common name for many disciplines cooperating with each other, sharing a common language, common concepts and common goal, which is understanding the structure and function of a normal and abnormal brain [32].

Neuroeconomics is claimed to be nothing more than a natural development of behavioural economics [27] or behavioural economics is "the first arm" that neuroeconomics relies upon [31]. Some emphasise that neuroeconomics is, on the one hand, a subfield of behavioural economics as it uses empirical evidence to prove human limitations, and, on the other one—a subfield of experimental economics which requires mastering difficult experimental tools being a novelty for economists [9]. Neuroeconomics is then understood as "the study of the biological microfoundations of economic cognition and economic behaviour. Biological microfoundations are neurochemical mechanisms and pathways, like brain regions, neurons, genes

and neurotransmitters. Economic cognition includes memory, preferences, emotions, mental representations, expectations, anticipation, learning, perception, information processing, inference, simulation, valuation and the subjective experience of reward. In general, neuroeconomic research seeks to identify and test biologically micro-founded models that link cognitive building blocks to economic behaviour" [10, p. 153].

"Neurofinance proposes to identify and analyse the black box of the human brain to understand the physiological developments (including hormonal activities) that take place when people make financial decisions" [14]. Neurofinance seeks to explain the biological mechanisms (neuronal and physiological) of financial market participant's behaviours [38] and also provide a physiologically conditioned, alternative explanation for the obvious failure of standard finance theories. Apart from seeking explanations for individual and market behaviours as a function of classical financial variables, it aims to discover how neuronal and physiological signals influence individual differences in financial decision-making. To this end, it uses non-invasive methods to measure neuronal and physiological activity [25].

When it comes to neuroaccounting, the simplest way to explain its core is that there are those of the behavioural accounting research studies that use neuroscience tools. Their task is to discover what is happening in the brain of those involved in any accounting processes and what are the specific relations between brain activity and external behaviour (including data interpretation) [2].

Inclusion of neuroscience into research on economic behaviours can bring many benefits. First, as in psychology, it can facilitate the resolution of which of the competing cognitive theories has corresponding neuronal correlates and is therefore more reliable. Second, knowledge about the functioning of the brain and the nervous system may help to better target the hypotheses. Third, neurobiological methods may enable the measurement of variables which, despite their importance for the decision-making process, have been difficult to grasp so far, such as emotions [16].

6.5 Types of Experimental Methods Used in Behavioural Economic Sciences

When it comes to behavioural economics, finance and accounting, the experimental methods include:

A. laboratory experiments in which participants perform specific tasks in a controlled environment,
B. simulation experiments (also known as "experimental economics" or game theory) where participants take on the role of buyers and sellers in an artificially created market and
C. field experiments where participants function in their natural environment and their actions are observed by the researcher [18].

The specifics of those research methods together with research examples emphasizing the advantages of adopting the neo-evolutionary paradigm into experimental design will be presented in the following sections.

It should be also noted that all the three categories of experiments can be carried out using such types of neuroscientific measurement methods as:

(a) neuroimaging methods: EEG (electroencephalography), CAT (computer-assisted tomography), MRI (magnetic resonance imaging), PET (positron emission tomography), fMRI (functional magnetic resonance imaging), SPECT (single photon emission computed tomography), TMS (transcranial magnetic stimulation), NIRS (near-infrared spectroscopy), DOT (diffuse optical tomography), EROS (event-related optical signal), SNI (single neuron imaging);
(b) psychophysiological methods: GSR (galvanic skin response), blood pressure, eye movement;
(c) neuropsychological methods [42].

The characteristics of these methods will be briefly described below. More comprehensive descriptions can be found in the literature (e.g. [9, 20]).

If to combine experimental methods (A, B and C) with "neuromeasurement" instruments (a, b and c), it can be assumed that three types of "neuroexperiments" will emerge, i.e. neurolaboratory experiments, neurosimulation experiments and neurofield experiments. Their specifics will not be separately presented as from methodological point of view they do not significantly differ from their "mother" methods. The presentation will then focus on neuromeasurement methods description and the research example of their use in experimental conditions.

6.6 Laboratory Experiments

This is the most commonly exploited method in behavioural economic sciences. Its main advantage is, as stated by Shadish et al. [30, p. 18], "its ability to illuminate causal inference". Among other things, it is a particularly useful research method when testing or building a theory. Its aptness stems from several reasons. It facilitates maintaining the wide span of control over the experimental environment and the possibility to eliminate factors disturbing the view of the situation. In consequence, the measurement of results is quite precise. What is more, the researcher has a possibility of manipulating the experimental environment that is checking various factors or the intensity of a given factor. It is also underlined that due to randomization of participants, laboratory experiments are characterised by a high level of internal validity. It means that the likelihood of confounding is minimised. Another advantage of this method is an opportunity to study individual behaviour [18].

Nonetheless, there are also some challenges that experimenter needs to face. In order for the experiment to be credible (to ensure high internal validity), it is necessary to create highly artificial conditions far from those existing in reality. This, in turn, translates into low external validity. Therefore, inference and generalization

on the basis of the experiment must be made carefully. Experimental results entitle to conclude whether the phenomenon occurred or not, rather than about its magnitude. Another issue is the fact that conducting an experiment requires designing a special task. This is not easy. It is necessary to take into account a number of factors that can distort the results as they act as decision traps that may appear while resolving the experimental task. For example, the sequence of data presentation may influence subjects' perception and impact the final outcome.

Besides the experiment itself, finding the appropriate participants is often a serious obstacle. Ideally, they should be practitioners in the fields being tested, but usually experienced professionals are reluctant to take part in such projects. That is why the common practice is to replace them by kind of surrogates, e.g. students. Another challenge is to achieve the right level of motivation to complete the experimental task. Otherwise, the results are useless. It should also be remembered that an experiment is a kind of social situation that has its own specificity and rights. So, there is also danger that the experimenter himself will affect the results of the experiment to some extent.

In recent years, exciting research has been carried out on a mechanism of self-control. This mechanism undoubtedly is a result of evolutionary development of brain and social skills. Animals show rather impulsive behaviours, which is manifested in the fact that they choose smaller but immediate gratifications. Children have also problems with postponing the gratification, however, some of them exhibited various behaviours helping them to resist the temptation to eat sweets during the experimenter's absence (the so-called Marshmallow experiment) [26]. Adults should be then more self-possessed, however, the researchers proved that there are situations where the resources of self-control may be exhausted. Vohs and Faber [41] tested how the depletion of self-control resources influences consumer decisions. Thirty-five students (16 men, 19 women) at Case Western Reserve University were subjected to their laboratory experiment. It consisted of two parts: psychological one—the aim of which was to exhaust the resources of self-control and marketing one—where the students were to propose prices they would pay for given products. First participants were asked to watch a six-minute film without a sound with a woman being interviewed and then to share judgments about this woman's personality. During the film a series of mono-syllable words were displayed for 30 s at the bottom of the screen. The manipulation was that the control group did not receive any guidance as to the words at the bottom of the screen, and the test group was told to "not read or look at any words that may appear on the screen". After watching the film, the participants were told that they would participate in the marketing research. They were shown 18 colour images of products and were asked to give the price they would be willing to pay for each of them. It appeared that the test group proposed higher prices than the control group. It means that when the person drains their self-control resources not only will they price the products higher but the point where the price becomes too expensive to buy, will also move up.

6.7 Simulation Experiments

Vernon Smith is an originator of that method. It can be treated as a specific form of the experiment which consists in creating artificial market conditions to study behaviour of its participants. Economic situations such as: bargaining, buying through auctions, deciding on the distribution of resources, buying and selling securities, etc., are measured. If the experiment is to be a market simulation, it is necessary to make real payments, i.e. the participants are playing for real money. Economic experiments are a good source of information on financial behaviours occurring in a social context.

The strength of this method is that the experimental conditions are a little bit closer to reality than in laboratory experiments. It is particularly useful while trying to predict behaviour in a new, complex environment, e.g. after introducing innovative legal regulation or a new market institution. However, certain dose of artificiality is still unavoidable to maintain control and ensure internal validity. An important advantage of this method is that it enables controlled changing of individual variables or situational features and observation of how these changes affect the participants. Such variables can be, for example, the number of participants, the amount of the stake or the ability to communicate with one another before making a decision. Fortunately, in this case, it is not a problem to use students because they act as sellers and buyers in real life, so their behaviours should not vary from "non-student" respondents. It seems that the main disadvantage of this method is its cost. Researcher needs to mimic real market conditions, which is quite expensive. Besides, participants are very often remunerated for the results they receive [18].

Experimental economics study bases mainly on various economic games like ultimate game, dictator game, public goods game or trust game (a classic example of a trust game is that of Berg et al. [6]). Five researchers, in their ground-breaking study that combined trust game and risk game, were to provide convincing proof for the independence of trust and risk-taking [22]. The subjects were healthy male students from different universities in Zurich. In the first game (trust game), the amount of the sum sent back was decided by the trustee, while in the second (risk game), it was determined randomly. However, the researchers introduced a bold innovation into the experiment. Some of the subjects received a single intranasal dose of oxytocin, a hormone responsible for the process of mutual trust among people. Therefore, people who received oxytocin should have sent more money to the partner in the game than those who received the placebo. However, if trust and risk are two separate categories then administering a hormone should not increase the propensity to financial risk. Research results confirmed this assumption.

6.8 Field Experiments

Fieldwork is a research method where the researcher "steps out from behind the desk or a laboratory" and carries out his project in the analysed community (in the field). Thus, field experiment means a study where scientists conduct their research

within a certain company or market to investigate interrelations between variables. Usually, experimenters use real conditions and manipulate an independent variable to the extent permitted by the organization. It resembles laboratory experiment, but it takes place in more real environment. The subjects do not have the awareness of participating in the experiment, so they take decisions in conditions as close as possible to natural ones [21]. Therefore, there is a risk of ethical nature related to the fact that decisions made by subjects can be associated with real economic consequences, such as the purchase of a specific item, financial loss or failure to perform professional duties. On the other hand, the researcher has no influence on the course of the experiment as he has to remain anonymous and unseen to the participants. This method has two main disadvantages. First, it is difficult to observe the results. Second, assuring conditions close to reality results in the existence of many other factors that can "blur the picture". Field experiments can be classified under three categories: artefactual, framed and natural [19]. The first is the closest to a typical laboratory experiment. The main difference is that it uses "non-standard" subjects, that are participants of a specific market. The next step towards naturally occurring data is a framed experiment. It differs from the previous one in the fact that it contains important elements of the context of the natural environment such as goods, tasks, rates or stakes. However, it should be emphasized that in this type of experiment, the subjects are still aware that they are participating in the experiment, and their behaviours are recorded and analysed. In natural experiments, subjects do not know that they participate in the experiment as the tasks they perform are made by them on a daily basis. They do not know whether they were added to the sample or if their behaviour was being analysed. This approach combines the most attractive elements of the laboratory experiment and naturally occurring data: randomization and realism. Thus, many of the limitations cited above are not a problem when deducing from data generated by natural field experiments.

List, in his field experiment designed as a kind of public game but conducted in real-life conditions, was trying to test whether there is a link between age and behaviour. The idea of public game is that all participants can collect goods from the common pool until it is used up. In his research, he used basketball cards in the role of public good. He recruited subjects during sports card show organized in a large southern US city. When a potential subject approached the experimenter's table and inquired about the sale of the 1989 Michael Jordan Hoops basketball cards displayed there, he was asked to take part in the five-minute experiment. If the person agreed, he was presented with the game's rules and experimental task which was to divide ten cards between individual's own and group account (common pool). In typical simple public game, the amount paid into the common pool is multiplied by coefficient greater than 1. When everyone makes payments, the total value of the common good is equally divided between participants, regardless of the size of the initial contribution. Even if someone pays nothing, he may receive money from the common pool. In List's experiment, the marginal return to the group account was 0.40$ and any card fractions were paid in cash. Participants' answers were recorded on the sheet. The experimenter informed the respondents that they should return after some time to find out the results of the game and pick up the gains. If someone did

not come for the payment, they received it by post. In fact, List conducted his study in two ways. In the first, the respondents received ten cards at the beginning of the game, in the second, they had to buy them first for their own money. It appeared that more mature subjects (age > 49) invested a higher proportion of their endowment into the public good [24].

6.9 Neuromeasurement Methods

As mentioned above, there are three types of "neuromethods" that may be utilized by behavioural economic researchers. One of those types include neuroimaging measures whose popularity is constantly growing [40, 42]. These measures are briefly described below.

EEG bases on the fact that stimulated nerve cells produce an electric field. Electrodes placed on the head enable to register the electrical activity of the brain. Postsynaptic stimulation sent by the cell is recorded. Thanks to that it is possible to examine brain reactions to a specific stimulus and changes in its activity over time. The main advantage of this method is relatively low cost of research.

CAT is actually the equivalent of thousands of X-rays tests. X-rays scan the brain layer by layer to obtain a three-dimensional image of this organ.

MRI uses the magnetic properties of various atoms that resonate under the influence of different magnetic fields. The obtained image of brain structures has a much higher resolution than CAT and indicates the presence or absence of a substance being tested. As for brain, as the white matter and grey matter have a different water content, the density of water is often assessed.

Testing the brain activity using blood flow is possible with PET. It is known that the active nervous tissue needs more glucose that reaches it along with the blood. To determine where exactly the concentration of glucose is the highest, a radioactive substance, i.e. a marker is added to blood at regular intervals. The marker makes the glucose molecules send radiation that can be registered by the electronic PET scanner's system. Another method but less expensive than PET and also using markers is SPECT. SPECT makes possible to visualize the blood flow and metabolism of a given body area and by this means to create a 3D model of a brain. The main difference between PET and SPECT are the markers used and hence the types of cameras that can read certain sorts of radiation.

fMRI bases on the fact that oxyhaemoglobin has different magnetic properties from haemoglobin. It checks the level of oxygenation of certain brain areas. Those areas that are involved in resolving specific tasks need more oxygen. The interpretation of images obtained by fMRI is that when some part of brain activates it means involvement in the task.

TMS uses electromagnetic induction to produce a weak electrical voltage using a rapidly changing magnetic field. The brain is therefore "cannonaded" with a magnetic field, thanks to which it is possible to draw a map of its activity for a particular task.

The next three methods, NIRS, DOT and EROS, also based on the differences in the properties of oxidized and non-oxidized blood. But this time, the laser light at near-infrared wavelengths is harnessed, which makes the skull transparent. The amount of light at different frequencies reflecting from the brain is recorded to create brain activation maps.

SNI due to its invasiveness is used to study the animal brains (e.g. monkey). A super-tiny microelectrode is placed in the nerve cell and the rate of stimulation of this cell is thereby monitored.

Apart from the above-mentioned "medical" methods, there are also those examining the correlation of various mental functions with physiological sensations. Emotional stimulation is usually accompanied by an accelerated heartbeat, higher blood pressure or increased sweating. The last property is used in GSR, the method that studies electrodermal activity. It analyses electrical phenomena on the surface of the skin caused by sweating alterations. Another technique is eye tracking. Thanks to it, the information on the change of the eye position, time and directions of viewing the given area is registered. It helps not only in recognizing where and how the eyes are directed but also in revealing how visual and text code is processed by human brain [33].

Neuropsychological methods serve to determine the relations between human behaviour and the state of the brain. They allow to explain the essence and mechanism of mental processes dysfunctions and allow to conclude whether they are related to brain pathology [36]. It is known that certain brain areas correspond with particular mental activities. Hence, if those areas are damaged, taking such activities should be visibly impaired in comparison with the same activities performed by healthy people. Scientists discovered that people with frontal lobe lesion exhibit defects in decision-making consisting especially in inability to choose options beneficial in the long term. It was clearly visible in the experiment based on Iowa Gambling Task (IGT) and using GSR [4]. The aim of the research was to test the "somatic marker hypothesis" which was a kind of "risk as emotion" hypothesis that belongs to the group of those approaches that interpret decision-making as a process developed as a result of the biological evolution of the human mind. Briefly speaking, human choices are highly adaptive, so decision variants that trigger fear will be rejected, while those generating positive emotions will be accepted. It means that our "evolutionary brain" will decide "for us" and our consciousness will just try to find satisfactory justification for our choice.

In IGT, participants were presented with four decks of cards on the computer screen. On each deck, there were cards that bring a prize or a fine. The goal of the game was to win as much money as possible. Decks differed from one another by balancing reward cards and penalty cards. Two decks were "bad decks", and the other two were "good decks". Then, consistent selection of bad decks turned out to be unfavourable in the long term. People with frontal lobe lesion chose more cards from "bad decks" regardless of whether they did it for the first time or after an hour, a month or a half year. It was because they were missing with warning physiological response (measured by GSR), which was activated by a healthy brain. For healthy subjects, a warning signal (stronger GSR) appeared before the card was pulled out

of the "bad deck" even before they realized the real nature of the task. Hence, in the course of evolution, human brain was shaped so as to even unconsciously help us to choose more favourable strategies.

6.10 Conclusions

Experimental methods can be considered more and more important for the development of contemporary economics with its influential trend of behavioural economic sciences. They are quite attractive research methods because they make it possible to recognize the relationship between the isolated number of variables, which is "normal economic reality", where a whole bunch of factors impact human reaction, would be impossible. However, these methods, although it may seem they are not numerous considering above-mentioned types, are quite sophisticated and require expert knowledge and skills. Due to the fact that they are not used in "traditional" economics, they present quite a challenge for economists in terms of their theoretical foundation, methodology and interpretation of results.

Development of brain imaging techniques significantly expanded the possibilities of studying economic behaviour. Discovering brain areas that respond to specific economic problems can help in finding answers to theoretical dilemmas that have not yet been resolved by other methods. Thanks to "neuroexperiments", it is also possible to analyse some automatic response patterns or emotions, which are often not available to the consciousness, and therefore "not suitable" for research by means of, e.g. "pure" laboratory experiments. Despite the fact that neurobiological methods seem to be constantly gaining in importance in the nearest future, they should not be overestimated, as it is not always justified to use them. In some cases, linking the activity of specific brain areas with the decision problem can be very difficult or even impossible because the brain can be active simultaneously in many areas. How to isolate "the proper one" that is in line with the original assumptions then? Researchers may find it hard to resist the temptation to interpret the results in a manner consistent with their initial hypothesis. Sometimes, it may also happen that these methods are employed only because it is currently in fashion [42]. That is why laboratory and field experiments as well as experimental economics still can be fruitfully utilized in discovering rules underlying human economic behaviours.

Valuable experimental research in evolutionary psychology and neuroscience proves the significance and relevance of neo-evolutionist methodology. A number of research has verified the use of the general theory of evolution and its many specifications on various cognitive levels. The amount of empirical material gathered by research programs is rich and continues to grow. It allows to describe the human world as the subject of the social and humanistic sciences. Research results will help to create new image of the human mind, society and culture, which is largely inconsistent with traditional assumptions on human subjectivity accepted by the social sciences. Accounting for representation of the social sciences in scientific research and discussions is of great significance. This could be in the form of

integration of neo-evolutionary methods with other traditional methods of the social sciences. Such integration could be carried out within research programmes realized by representatives of various scientific disciplines. Complex problems of sociological, psychological, economic and anthropological natures could thus be studied from different perspectives.

References

1. Angner, E., Loewenstein, G.: Behavioral economics. In: Thagard, P., Woods, J. (eds.) Handbook of the Philosophy of Science. Elsevier, Amsterdam (2006)
2. Artienwicz, N.: Neurorachunkowość jako potencjalny, ale trudny kierunek rozwoju rachunkowości. Zesz. Teor. Rachun. **82**(138), 9–17 (2015)
3. Bąbel, P., Ostaszewski, P.: Współczesna psychologia behawioralna. Wyd UJ, Kraków (2008)
4. Bechara, A., Tranel, D., Damasio, A.R.: The somatic marker hypothesis and decision-making. In: Boller, F., Grafman, J. (eds.) Handbook of Neuropsychology, pp. 117–143. Elsevier, Amsterdam (2002)
5. Belkaoui, A.: Behavioral Accounting: The Research and Practical Issues. Quorum Books (1989)
6. Berg, J., Dickhaut, J., McCabe, K.: Trust, reciprocity and social history. Games Econ. Behav. **10**(1), 122–142 (1995)
7. Bloom, F.: Fundamentals of neuroscience. In: Squire, L., et al. (eds.) Fundamental Neuroscience. Elsevier, San Diego (2008)
8. Brzezicka, J., Wiśniewski, R.: Homo oeconomicus and behavioral economics. Contemp. Econ. **8**(4), 353–364 (2014)
9. Camerer, C.F.: Neuroeconomics: using neuroscience to make economic predictions. Econ. J. **117**, C26–C42 (2007)
10. Camerer, C.F., Cohen, J., Fehr, E., Glimcher, P., Laibson, D.: Neuroeconomics. In: Kagel, J., Roth, R. (eds.) Handbook of Experimental Economics, vol. 2. Princeton University Press, Princeton and Oxford (2015)
11. Camerer, C.F., Loewenstein, G.: Behavioral economics: past, present, future. In: Camerer, C.F., Loewenstein, G., Rabin, M. (eds.) Advances in Behavioral Economics. Princeton University Press, Princeton and Oxford (2004)
12. Crawford, C., Krebs, D.: Foundations of Evolutionary Psychology. Lawrence Erlbaum Associates, Mahwah, NJ (2008)
13. de Waal, F.: Chimpanzee Politics: Sex and Power Among Apes. The Johns Hopkins University Press, Baltimore (1982)
14. Gayathri, T., Ganesan, S.: An analysis of investment decision through neurofinance approach. Int. J. Bus. Insights Transform. **11**(2), 80–84 (2018)
15. Gerrig, R.J., Zimbardo, P.G.: Psychology and Life. Allyn & Bacon (2009)
16. Gęsiarz, F.: Po co ekonomom mózgi, czyli o pożytkach z neuroekonomii. Rocz. Kognitywistyczny **V**, 49–55 (2011)
17. Goodall, J.: Through a Window: My Thirty Years with the Chimpanzees of Gombe. Houghton Mifflin, Boston (1990)
18. Hageman, A.M.: A review of the strengths and weaknesses of archival, behavioral and qualitative research methods: recognizing the potential benefits of triangulation. Adv. Account. Behav. Res. **11**, 1–30 (2008)
19. Harrison, G.W., List, J.A.: Field experiments. J. Econ. Lit. **42**, 1009–1055 (2004)
20. Kable, J.W.: The cognitive neuroscience toolkit for the neuroeconomist: a functional overview. J. Neurosci. Psychol. Econ. **4**(2), 63–84 (2011)
21. Kerlinger, F.N., Lee, H.B.: Foundations of Behavioral Research. Thomson, New York (2000)

22. Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E.: Oxytocin increases trust in humans. Nature **435**(7042), 673–676 (2005)
23. Lintner, G.: Behavioral finance why investors make bad decisions. Planner **13**(1), 7–8 (1998)
24. List, J.A.: Young, selfish, and male: field evidence of social preferences. Econ. J. **114**(492), 121–149 (2004)
25. Miendlarzewska, E.A., Kometer, M., Preuschoff, K.: Neurofinance. Organ. Res. Methods 1–27 (2017)
26. Mischel, W., Ayduk, O.: Willpower in a cognitive-affective processing system: the dynamics of delay of gratification. In: Baumeister, R.F., Vohs, K.D. (eds.) Handbook of Self-regulation. Research, Theory, and Applications, pp. 99–129. New York (2004)
27. Mohr, P.N.C., Li, S.-C., Heekeren, H.R.: Neuroeconomics and aging: neuromodulation of economic decision making in old age. Neurosci. Biobehav. Rev. **34**, 678–688 (2010)
28. Ricciardi, V., Simon, H.K.: What is behavioral finance? Bus. Educ. Technol. J. **2**(2), 1–9 (2000)
29. Setchell, J.M., Curtis, D.J.: Field and Laboratory Methods in Primatology: A Practical Guide. Cambridge University Press, Cambridge (2003)
30. Shadish, W.R., Cook, T.D., Campbell, D.T.: Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Miffin, Boston (2002)
31. Sharp, C., Monterosso, J., Montague, P.R.: Neuroeconomics: a bridge for translational research. Biol. Psychiat. **72**, 87–92 (2012)
32. Squire, L., Berg, D., Bloom, F.E., du Lac, S., Ghosh, A., Spitzer, N.C. (eds.): Fundamental Neuroscience. Academic Press, Elsevier (2008)
33. Stolecka-Makowska, A., Wolny, R.: Możliwości zastosowania techniki okulograficznej w ilościowych badaniach marketingowych. Stud. Ekon. **195**, 197–205 (2014)
34. Sulkowski, L.: Neodarwinism in Organization and Management. PETER LANG, Internationaler Verlag der Wissenschaften (2012)
35. Sulkowski, L., Chmielecki, M.: Application of neuroscience in management. In: Neuroeconomic and Behavioral Aspects of Decision Making, pp. 49–62. Springer, Cham (2017)
36. Szepietowska, E.M.: Diagnostyka neuropsychologiczna i jej zastosowanie kliniczne. Pol. Przegl. Neurol. **8**(3), 107–113 (2012)
37. Tooby, J., Cosmides, L.: Toward mapping the evolved functional organization of mind and brain. In: Gazzaniga, M.S. (ed.) The New Cognitive Neuroscience. MIT Press, Cambridge, MA (2000)
38. Tseng, K.C.: Behavioral finance, bounded rationality, neuro-finance, and traditional finance. Invest. Manag. Financ. Innov. **3**(4), 7–18 (2006)
39. Tversky, A., Kahneman, D.: Availability: a heuristics for judging frequency and probability. In: Kahneman, D., Slovic, P., Tversky, A. (eds.) Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press, New York (1982)
40. Urbanik-Papp, G.: Neuroekonomia jako interdyscyplinarne podejście do badania decyzji ekonomicznych. Kult. Polityka **15**, 158–172 (2014)
41. Vohs, K., Faber, R.: Spent resources: self-regulatory resource availability affects impulse buying. J. Consumer Res. **33**, 537–547 (2007)
42. Zaleśkiewicz, T.: Neuroekonomia. Decyzje **9**, 29–56 (2008)

Chapter 7
An Attempt of Knowledge Handling for Experimental Economics Domain

Agnieszka Konys

Abstract As experimental economics proved to be promising and dynamically growing research field, a lot of methodological and practical issues are raised in scientific literature. However, practical possibility of knowledge handling in whole experimental economics domain is limited. On the other hand, a development of knowledge management mechanisms, especially focusing on knowledge engineering and ontologies, provides auspicious means to handle knowledge. As the result of the mentioned above, in this chapter, an attempt of ontology as a form of knowledge engineering for handling knowledge in experimental economics is proposed. To confirm the applicability of the proposed approach, several case studies referring to selected problems of experimental economics are provided.

Keywords Knowledge engineering · Ontology · Experimental economics · Knowledge management · Economic ontology · Financial ontology

7.1 Introduction

Global technology spending has increased in the last years and there are some expectations in effectivity of applied means. In a world of growing business dynamics, high rates of technological advances and structural changes and challenges, enterprises need to be effectively and continuously redesigned and reengineered in order to achieve strategic and operational success. It is quite difficult to expect the outcomes of these practices. To overcome these problems, effective means offered by experimental economics are increasingly significant sources of competitive advantage and a key to the success of contemporary organizations, bolstering the collective expertise of its employees and partners. Thus, to advance the power of experimental economics, the applicability of knowledge engineering seems to be promising approach. Infusing more knowledge into the domain and also successful handling

A. Konys (✉)
Faculty of Computer Science and Information Technology, West Pomeranian University of Technology in Szczecin, Żołnierska 49, 71-210 Szczecin, Poland
e-mail: akonys@zut.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_7

the knowledge generally helps to improve data interpretation, its management and prediction, and in the aftermath of this, making better decisions.

When analyzing experimental economics domain, it should be pointed that an adaptation of new tools, methods, mechanisms and also discovering new areas of practical usage of these means is well promising. For example, the refinement of these means may refer to an ontology-based approach and its applicability in experimental economics. The need of finding ways of consolidating the knowledge in one main aggregate solution might provide complete and comprehensive solution dedicated to experimental economics, although it requires putting the effort to successfully collect, handle and reuse existing knowledge derived from selected case studies. However, reviewing the literature and existing approaches, there were narrow attempts to adapt knowledge engineering mechanisms in experimental economics. Therefore, the chapter aims to provide an attempt for handling domain knowledge for experimental economics. The aim is not only an attempt to adapt knowledge engineering mechanisms in experimental economics, but to ensure formal, practical and methodological guidelines of how to find the comprehensive knowledge of applied decision situations.

The rest of the chapter is organized as follows: Sect. 7.2 presents literature review, providing essential information of knowledge management mechanisms and applicability of ontologies in experimental economics. In Sect. 7.3, formal foundations for ontology for experimental economics domain are featured. Section 7.4 yields an attempt for ontological representation for experimental economics domain supported by practical use cases.

7.2 Literature Review

7.2.1 Knowledge Management Mechanisms

According to the literature, knowledge management (KM) is a discipline concerned with the analysis and technical support of practices used in an organization to identify, create, represent, distribute, and enable the adoption and leveraging of good practices in organizational processes [1]. Effective knowledge management is an increasingly important source to build a competitive advantage as well as a crucial key of contemporary organizations, bolstering the collective expertise of its participants and collaborators [2]. On the other hand, an open problem in knowledge management concerns effective collecting, systemizing, handling and reusing of increasingly growing amount of knowledge [3]. This problem encompasses both tacit and explicit knowledge. Thus, an important challenge is weighing the relevance of knowledge engineering mechanisms to improve overall means of the progress towards knowledge handling [1, 4, 5].

Accordingly, ontologies show a wide spectrum of applicability in this area [4, 6]. Ontologies are one of the crucial components of the Semantic Web, providing

a formal and explicit description of the elements belonging to a domain, such as concepts, properties, relations, functions and axioms [7, 8]. A well-defined ontology construction requires to adapt a proper methodology, defining the main steps (e.g. Noy and McGuiness, Methontology and On-To-Knowledge [9–11]). Apart from methodologies, there are some tools to support and develop ontologies and knowledge engineering processes (e.g. OntoStudio, RapidMiner, Protégé), and recommended standards to build an ontology (e.g. OWL/XML and RDF/XML) [12].

7.2.2 *Ontologies in Experimental Economics (Economy and Finance)*

Nowadays, great effort has been put into building reusable semantic structures that can be informal vocabularies, catalogues, glossaries as well as more complex finite formal structures representing the entities within a domain and the relationships between those entities [13, 14]. A wide range of applications is emerging to handle with the problem of communication and exchange of information between different entities [15]. The pertinence of semantic structures, especially ontologies, is widely discussed in different applications covering economic and financial problems [16].

There are many domains of interests that can utilize ontologies to use them as the knowledge repository [17]. The field of economy and finance is a conceptually rich domain where information is complex, huge in volume and a highly valuable business product by itself [18–20]. A massive amount of valuable information is produced worldwide every day, but its collecting, handling and processing is a hard and time-consuming task [21]. Efficient handling knowledge is needed in the financial and economic fields, providing the access the contents that are most relevant in time, offering an effective way to exploit it [22, 23].

Based on the literature, previous works exploited the practical applicability of ontology by providing some basic ideas of existing relations between instances [17], providing an approach to creating domain ontologies for higher education in economics [24], using an ontology for the sharing economy [22], and elaborating a financial fraud ontology [16]. Moreover, the role of ontology in the field of finance and economy was considered by Lara et al. [21]. They emphasized the important role of the ontology as a management technique to enable an efficient generation, management and consumption of complex and big information resources.

Consecutive approaches consider also financial aspects in knowledge engineering, offering specification of financial knowledge [25], business intelligence modelling to predict bankruptcy using financial domain ontology with association rule mining algorithm [18] and using an ontology-based business intelligence in a financial knowledge management system [5]. Other works consider also development of a method for ontology-based empirical knowledge representation and reasoning [4] and building an ontology for financial investment [16].

The in-depth analysis of the literature highlights some missing aspects of exploitation of knowledge-based engineering mechanisms in economy and finance fields. For example, a representation of empirical knowledge in a structural way for experimental economics in order to help knowledge requesters clearly understand empirical knowledge seems to be promising approach [1–5, 24]. To do this, an ontology may be adopted to deduce empirical knowledge in experimental economics in order to share and reuse relevant empirical knowledge effectively. Unfortunately, the surveyed literature does not demonstrate an existence of a comprehensive and complex ontology-based approach supporting this process. The motivation to provide an attempt for handling domain knowledge for experimental economics is to fill this research gap and to ensure formal, practical, and methodological guidelines of how to find the comprehensive knowledge of applied decision situations in experimental economics domain.

7.3 Formal Foundations for Ontology for Experimental Economics Domain

This research aim is to offer an attempt to knowledge handling in experimental economics domain. Thus, this process requires a carefully elaboration of selected approaches dedicated to experimental economics domain. The proposed ontology contains a set of various case studies, however, each ontology may have supplement super classes, classes and subclasses that can be added or can be borrowed from another ontology concept.

The process of ontology construction requires collecting and further elaboration and classification of gathered information. In the aftermath of these activities, the knowledge is transformed from the unstructured form into the semi-structured form. To present the well-formed knowledge base, the set theory is used. The application of the set theory conveys the semi-formal mathematical description. Previously identified specifications of the selected approaches dedicated to experimental economics domain were a basis for the mathematical elaboration, shown below.

To define a range of a relation, it is assumed that R is a relation from set Ep and A, then the set of all taxons (all of the first components of the ordered pairs) belonging to R is called the domain of R. Thus, Dom is defined as follows:

$$(\mathrm{R}) = \{\mathrm{ep} \in \mathrm{Ep}{:}(\mathrm{ep}, \mathrm{a}) \in \mathrm{R} \text{ for some } \mathrm{a} \in \mathrm{A}\}$$

The set of all second components of the ordered pairs (the set of all taxons) belonging to R is called the range of R. Thus, the range of R is defined as follows:

$$(\mathrm{R}) = \{\mathrm{a} \in \mathrm{A}{:}(\mathrm{ep}, \mathrm{a}) \in \mathrm{R} \text{ for some } \mathrm{ep} \in \mathrm{E}\}$$

If experimental economics problem Ep and attributes A are two non-empty sets, then the Cartesian product E of Ep and A, denoted Ep × A, is the set of all ordered pairs (ep, a) such that ep ∈ Ep and a ∈ A:

$$\mathrm{Ep} \times \mathrm{A} = \{(\mathrm{ep}, \mathrm{a}) : \mathrm{ep} \in \mathrm{Ep}, \mathrm{a} \in \mathrm{A}\}$$

Attributes A contain the finite set of taxons, defined as follows:

$$\mathrm{A} = \{\mathrm{At}, \mathrm{Tt}, \mathrm{Kw}, \mathrm{Do}\}$$

where At is the author of the paper, Tt is the title of the paper, Kw refers to the keyword and Do is the digital object identifier.

Experimental economics problem Ep contains the finite set of taxons, which are subsets:

$$\mathrm{Ep}: \{\mathrm{Ep1}, \mathrm{Ep2}, \mathrm{Ep3}, \mathrm{Ep4}, \mathrm{Ep5}, \mathrm{Ep6}, \mathrm{Ep7}, \mathrm{Ep8}, \mathrm{Ep9}, \mathrm{Ep10}, \mathrm{Ep11}, \mathrm{Ep12}, \mathrm{Ep13}, \mathrm{Ep14}, \mathrm{Ep15}, \mathrm{Ep16}, \mathrm{Ep17}\}$$

where Ep1 is the first experimental economics problem, Ep2 is the second experimental economics problem, Ep3 is the third experimental economics problem, and Ep4 is the fourth experimental economics problem, Ep5 is the fifth experimental economics problem, Ep6 is the sixth experimental economics problem, Ep7 is the seventh experimental economics problem, Ep8 is the eighth experimental economics problem, Ep9 is the ninth experimental economics problem, Ep10 is the tenth experimental economics problem, Ep11 is the eleventh experimental economics problem, Ep12 is the twelfth experimental economics problem, Ep13 is the thirteenth experimental economics problem, Ep14 is the fourteenth experimental economics problem, Ep15 is the fifteenth experimental economics problem, Ep16 is the sixteenth experimental economics problem and Ep17 is the seventeenth experimental economics problem.

The formalization allows performing these formulations into model of concepts and relationships between them. Forerunning step of an ontology construction needs to a taxonomy elaboration. Taxonomical view offers a class hierarchy with assigned relations. It constitutes a multi-dimensional view and creates a meta-model for further ontological elaboration. On base of this, the set of concepts, properties and relationships was established. In the aftermath of this, the set of criteria and subcriteria with existing relations between them was implemented in the Protégé software. Overall, the final set of attributes was defined including 4 main criteria and 119 sub-criteria. The set of considered experimental economics problems contains 17 examples (Fig. 7.1) [1, 4, 5, 16–18, 21, 22, 24–34].

To preserve ontological rules, the set of object properties was constituted covering domain and range. It is assumed that each criterion has an attribute and also the inversion property is kept. Based on the formed semi-formal description using set theory, the descriptive notation is also provided.

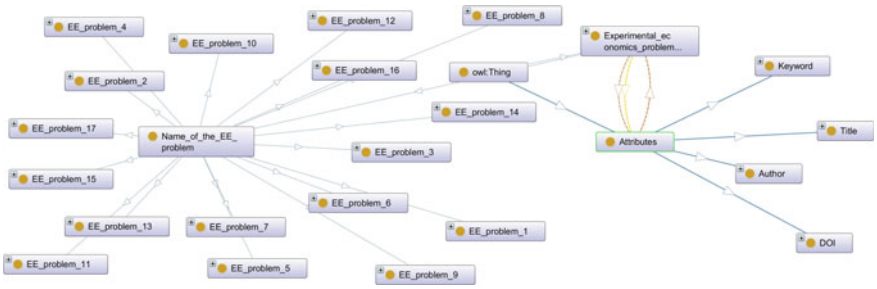


Fig. 7.1 Visualization of main schema using OntoGraf tool

- Author attribute contains the set of identified authors of the papers, especially including all of the authors assigned to a given paper. If it is possible, the full name is provided. For this ontology, 30 authors were selected: Waralak V. Siricharoen, Thitima Puttitanun, Pablo Castells, Borja Foncillas, Rubén Lara, Mariano Rico, Juan Luis Alonso, Dejan Lavbič, Olegas Vasilecas, Rok Rupnik, Josip Mesaric, Branimir Dukic, Moritz von Hoffen, Jerzy Korczak, Helena Dudycz, Mirosław Dyczkowski, A. Martin, M. Manjula, V. Prasanna Venkatesan, Hilary Cheng, Yi-Chuan Lu, Calvin Sheu, Zili Zhang, Chengqi Zhang, Swee San Ong, Yuh-Jen Chen, Małgorzata Łatuszyńska, Waldemar Tarczyński, Mateusz Piwowarski and Kesra Nermend (Fig. 7.2) [1, 4, 5, 16–18, 21, 22, 24–34].
- Title attribute covers various titles of the papers referring to experimental economics problems. For this ontology, 16 examples were added, but it is only representative, exemplary set: A Business Intelligence Model to Predict Bankruptcy Using Financial Domain Ontology with Association Rule Mining Algorithm, Semantic Web Technologies for Economic and Financial Information Management, An ontology-based business intelligence application in a financial knowledge management system, Cognitive neuroscience techniques in supporting decision making and the analysis of social campaign, An Approach to Creating Domain Ontologies for Higher Education in Economics, Creating Ontology Chart Using

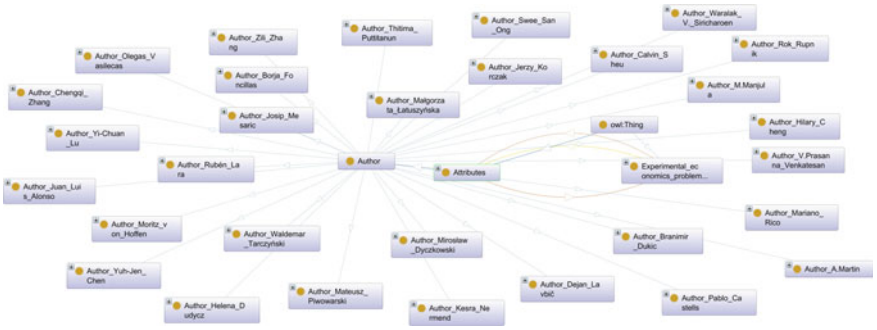


Fig. 7.2 Visualization of author attributes using OntoGraf tool

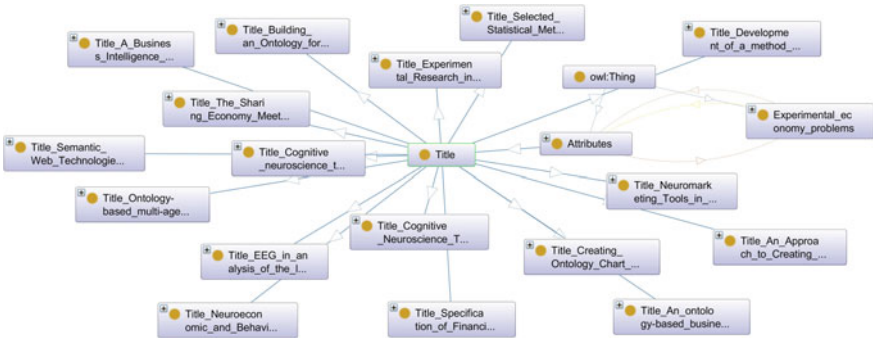


Fig. 7.3 Visualization of title attributes using OntoGraf tool

Economy Domain Ontologies, Neuromarketing Tools in Studies on Models of Social Issue Advertising Impact on Recipients, Experimental Research in Economics and Computer Simulation, The Sharing Economy Meets the Semantic Web: An Ontology for the Matchmaking of Peers, Development of a method for ontology-based empirical knowledge representation and reasoning, Selected Statistical Methods in Experimental Studies, Neuroeconomic and Behavioral Aspects of Decision Making, Cognitive Neuroscience Techniques in Examining the Effectiveness of Social Advertisements, Specification of Financial Knowledge—The Case of an Intelligent Dashboard for Managers, EEG in analysis of the level of interest in social issue advertising, Ontology-based multi-agent system to support business users and management, and Building an Ontology for Financial Investment (Fig. 7.3) [1, 4, 5, 16–18, 21, 22, 24–34].

- Keyword attribute comprises the listed keywords from the set of the analyzed papers. The identified set of keywords contains 55 examples and it is shown as follows: Ontology, Agent, Test methods of statistical regularities, Social advertisement, Topic maps, Decision making support, AIDA, Dashboard for managers, Neuroeconomic, Semantic Web, Semantic Web Technologies, Description Language, Economic and Financial Information Management, Ontology for the economic and financial information domain, Information retrieval, Ontology Chart, Knowledge base, Cognitive neuroscience, Economic ontology, Data warehouse, EEG, Behavioral aspects, Correlations, Financial Ontology, Knowledge reasoning, Multi-agent system, Sharing Economy, Business process management, Data Mining, Computer simulation, Empirical knowledge, Z-Score Model, Advertisement, Bankruptcy, Higher education, Social campaign, Domain Ontologies, Structure, Model of advertising impact, Financial domain Ontology, Dynamics and volatility, Financial Investment, Knowledge management, Decision support, Business intelligence, Business rules, Financial knowledge, Decision making, Collaborative Consumption, Economy, Bond ratings, Experimental economics, GSR, Knowledge representation, OWL and HR (Fig. 7.4) [1, 4, 5, 16–18, 21, 22, 24–34].
- DOI attribute provides the information of digital object identifier, used commonly to add unique identifier to the object, standardized by ISO. The aim of this

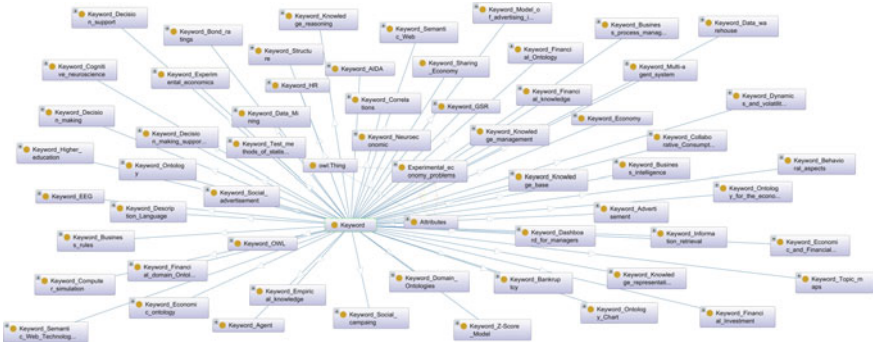


Fig. 7.4 Visualization of keyword attributes using OntoGraf tool

attribute is to help in finding the necessary information of the paper. In case of the problems with DOI, ISSN and ISBN are provided: https://doi.org/10.4156/jdcta.vol3.issue3.8, https://doi.org/10.1007/978-0-387-48531-73, https://doi.org/10.3846/tede.2010.21, https://doi.org/10.1109/iti.2007.4283747, https://doi.org/10.1109/icsc.2017.103, ISSN 1507-3858, ISSN 1694-0814, https://doi.org/10.1016/j.eswa.2008.02.047, https://doi.org/10.1007/s10506-005-4163-0, https://doi.org/10.1016/j.dss.2010.02.010, https://doi.org/10.1007/978-3-319-28419-4, ISBN 978-3-319-62938-4, https://doi.org/10.1007/978-3-319-99187-08, https://doi.org/10.1007/978-3-319-62938-422, https://doi.org/10.1016/j.procs.2018.08.056, https://doi.org/10.23918/icabep2018p1 (Fig. 7.5) [1, 4, 5, 16–18, 21, 22, 24–34].

However, it is possible to enrich the ontology adding additional attributes and relations. The presented set is only an empirical attempt for ontological knowledge representation in experimental economics domain. To conclude, the elaborated set of attributes and relations was implemented into the Protégé application using OWL (Ontology Web Language) standard for knowledge representation.

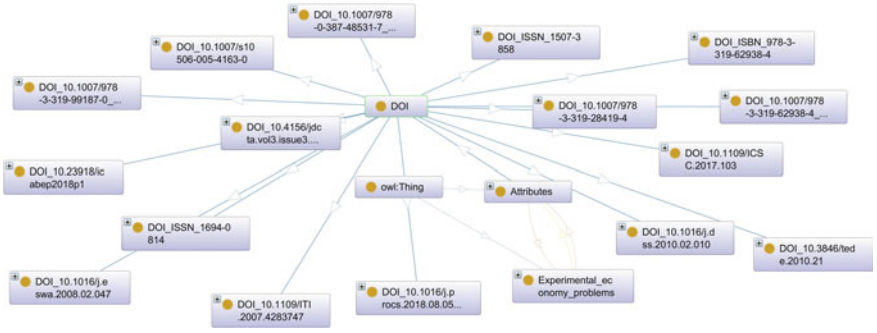


Fig. 7.5 Visualization of DOI attributes using OntoGraf tool

7.4 Ontological Representation for Experimental Economics Domain

To structure of the elaborated knowledge, the extracted concepts (attributes) were implemented. The formal description of presented knowledge representation allows for machine-readable processing, sharing, reuse and, finally, populating new knowledge for experimental economics. This formal representation captures domain knowledge about selected experimental economics problems and offers a wide spectrum to capitalize on the effects. The main structure of the ontology is shown in Fig. 7.6.

To show practical application of the presented ontology, the set of case studies was mapped out. The case study focuses on collecting the information of experimental economics problems containing the following keywords: domain ontologies and economic ontologies, and determining by https://doi.org/10.1109/iti.2007.4283747. Each of the criteria and sub-criteria should be fulfilled to belong to the final ranking. This specification was used to formulate the defined class reflecting necessary and sufficient conditions. Based on this definition, the reasoning and also validating process of the ontology was started. Validating the ontology allows to check both the consistency and correctness of the proposed ontology. To investigate whether the ontology is capable to complete information retrieval and interoperability needs, the results have been compared to the knowledge included in the comparison analysis of experimental economics problems.

To obtain the set of results, the query was computed by a Description Logic Query mechanism (Fig. 7.8), implemented in Protégé software. Visualizing the final ranking of the solutions is presented in Fig. 7.7. Based on the received results, only one of experimental economics problems matches these requirements (Fig. 7.8).

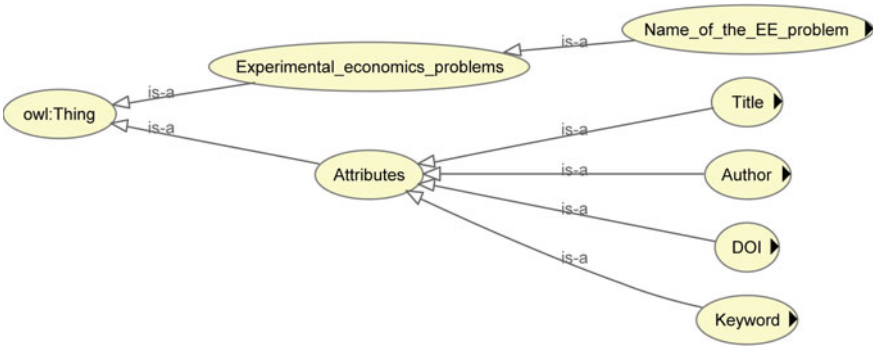


Fig. 7.6 Visualisation of the main structure of the ontology using OWLViz tool

Fig. 7.7 Visualisation of the main structure of the ontology using OWLViz tool

Additional experimental query has been posed to find the experimental economics problems including the set of features: keywords—domain ontologies, or economic and financial information management or economic ontology, or financial ontology, or financial domain ontology, or ontology for the economic and financial information domain. Assuming that this definition can be fulfilled partially, four experimental economics problems match the results. Thus, the application of the reasoning mechanism provides a set of results with regard to these pre-defined requirements, shown in Fig. 7.9. The definition implemented in the Description Logic Query mechanism is presented in Fig. 7.10.

Based on the experimental case studies, correctness of the obtained results confirms the validation process and, consequently, provides the set of results. Both of the presented queries were added manually. Enriching ontology by adding new experimental economics problems offers new opportunities for handling existing knowledge and obtaining new one.

Fig. 7.8 Visualisation of the main structure of the ontology using DL query mechanism

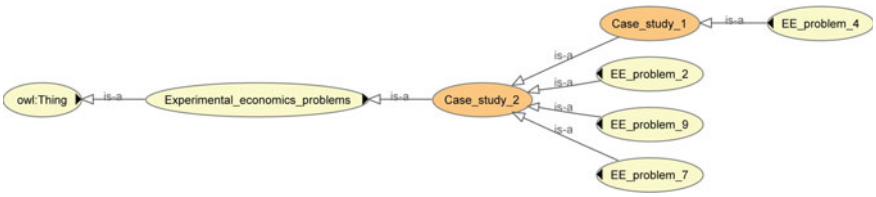


Fig. 7.9 Visualisation of the main structure of the ontology using OWLViz tool

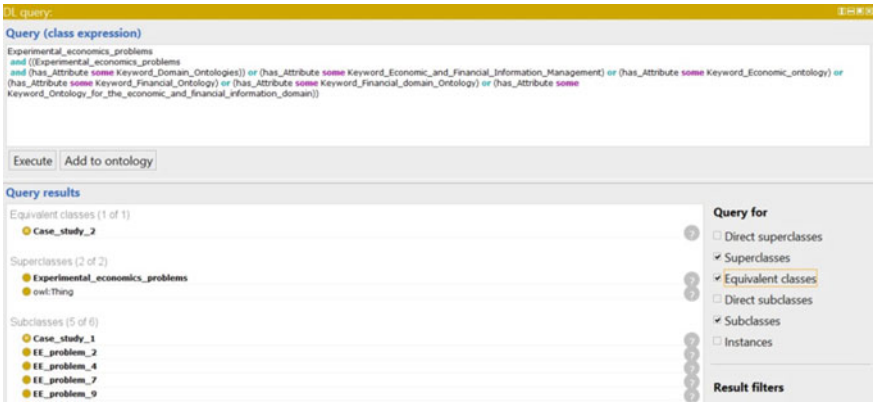


Fig. 7.10 Visualisation of the main structure of the ontology using DL query mechanism

7.5 Conclusions

This chapter concerned the compound problem of knowledge handling in experimental economics domain. Diversity of considered experimental economics problems and, on the other hand, the process of gathering knowledge in one single place, poses promising challenges on how to do it effectively and practically. A clear demand to adapt knowledge engineering mechanisms to modern economic problems highlights the validity to an attempt of applicability of ontology for handling knowledge in experimental economics. A revised literature confirmed the lack of knowledge systematization in this field. In response to bridge this gap, an author developed an approach based on knowledge management, tailored for the experimental economics domain. The main intention of ontologies' creation is to capture and formalize a domain of knowledge. In practical terms, the author proposes a reusable domain model of selected experimental economics problems to multi-usage, reusing, and knowledge sharing and dissemination.

The presented ontology contained several experimental economics problems, supported by the set of competency questions. This form of knowledge representation can be useful in case of finding quickly and intuitively any public available information (e.g. author, title, keyword and DOI) of the experimental economics problems. Technically, the formalized structure of the ontology allows the users to have a machine-readable access and handling data across the knowledge sources developed for experimental economics domain.

Through the introduced experiments yield satisfactory results in general proposed attempt of an ontology, there were some limitations in presented approach. For example, the presented ontology needs to be extended in the future by adding additional experimental economics problems to ensure complex and complete source of knowledge. Incorporating great amount of up-to-date information enhances the

applicability and comprehensiveness of proposed approach. Moreover, each ontology may have supplement super classes, classes and subclasses that can be added or can be borrowed from another ontology concept. It seems to be interesting to ensure this interoperability in the offered model.

References

1. Chong, A.Y.-L., Chan, F.T., Goh, M., Tiwari, M.K.: Do interorganisational relationships and knowledge-management practices enhance collaborative commerce adoption? Int. J. Prod. Res. **51**, 2006–2018 (2013)
2. Gunasekaran, A., Ngai, E.W.T.: Knowledge management in 21st century manufacturing. Int. J. Prod. Res. **45**, 2391–2418 (2007)
3. Desouza, K.C., Chattaraj, A., Kraft, G.: Supply chain perspectives to knowledge management: research propositions. J. Knowl. Manag. **7**, 129–138 (2003)
4. Chen, Y.-J.: Development of a method for ontology-based empirical knowledge representation and reasoning. Decis. Support Syst. **50**, 1–20 (2010)
5. Cheng, H., Lu, Y.-C., Sheu, C.: An ontology-based business intelligence application in a financial knowledge management system (2008)
6. Konys, A.: A framework for analysis of ontology-based data access. In: Nguyen, N.-T., Iliadis, L., Manolopoulos, Y., Trawiński, B. (eds.) 8th International Conference on Computational Collective Intelligence, ICCCI 2016, Part II. Lecure Notes in Computer Science, pp. 397–408. Springer International Publishing (2016)
7. Konys, A.: Knowledge-based approach to question answering system selection. In: Nunes, M., et al. (eds.) ICCCI 2015, Computational Collective Intelligence, Part I. LNAI, vol. 9329, pp. 361–370. Springer-Verlag (2015)
8. Konys, A.: Towards knowledge handling in ontology-based information extraction systems. Procedia Comput. Sci. **126**, 2208–2218 (2018)
9. Fernández-López, M., Gómez-Pérez, A., Juristo, N.: METHONTOLOGY: from ontological art towards ontological engineering. In: Spring Symposium on Ontological Engineering of AAAI, pp. 33–40. Stanford University, California (1997)
10. Konys, A.: An ontology-based knowledge modelling for a sustainability assessment domain. Sustainability **10**, 300 (2018)
11. Noy, N., McGuiness, D.: Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880 (2001)
12. Konys, A., Wątróbski, J., Różewski, P.: Approach to practical ontology design for supporting COTS component selection processes, ACIIDS 2013. In: Selamat, A., et al. (eds.) ACIIDS 2013, Part II. LNAI, vol. 7803, pp. 245–255. Springer, Heidelberg (2013)
13. Jankowski, J., Hamari, J., Wątróbski, J.: A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. Internet Res. **29**(1), 194–217 (2018)
14. Wątróbski, J., Jankowski, J., Piotrowski, Z.: The selection of multicriteria method based on unstructured decision problem description. In: International Conference on Computational Collective Intelligence, pp. 454–465. Springer, Cham (2014)
15. Wątróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., Zioło, M.: Generalised framework for multi-criteria method selection. Omega **86**, 107–124 (2019)
16. Zhang, Z., Zhang, C., Ong, S.S.: Building an ontology for financial investment. In: Leung, K.S., Chan, L.W., Meng, H. (eds.) Intelligent Data Engineering and Automated Learning—IDEAL 2000. Data Mining, Financial Engineering, and Intelligent Agents. IDEAL 2000. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg (2000)

17. Siricharoen, W.V., Puttitanun, T.: Creating ontology chart using economy domain ontologies. Int. J. Digit. Content Technol. Appl. **3**(2) (2009)
18. Martin, A., Manjula, M., Prasanna Venkatesan, V.: A business intelligence model to predict bankruptcy using financial domain ontology with association rule mining algorithm. IJCSI Int. J. Comput. Sci. **3**(2) (2011)
19. Karczmarczyk, A., Jankowski, J., Wątróbski, J.: Multi-criteria decision support for planning and evaluation of performance of viral marketing campaigns in social networks. PLoS ONE **13**(12), e0209372 (2018)
20. Jankowski, J., Kolomvatsos, K., Kazienko, P., Watróbski, J.: Fuzzy modeling of user behaviors and virtual goods purchases in social networking platforms. J. UCS **22**(3), 416–437 (2016)
21. Lara, R., Cantador, I., Castells, P.: Semantic web technologies for the financial domain. In: Cardoso, J., Hepp, M., Lytras, M.D. (eds.) The Semantic Web. Semantic Web and Beyond Computing for Human Experience, vol. 6. Springer, Boston, MA (2007)
22. Von Hoffen, M.: The sharing economy meets the semantic web: an ontology for the matchmaking of peers. In: IEEE 11th International Conference on Semantic Computing (ICSC) (2017)
23. Wątróbski, J., Ziemba, E., Karczmarczyk, A., Jankowski, J.: An index to measure the sustainable information society: the Polish households case. Sustainability **10**(9), 3223 (2018)
24. Mesaric, J., Dukic, B.: An approach to creating domain ontologies for higher education in economics. In: 29th International Conference on Information Technology Interfaces (2007)
25. Korczak, J., Dudycz, H., Dyczkowski, M.: Specification of financial knowledge—the case of an intelligent dashboard for managers. Bus. Inform. **2**(28) (2013)
26. Łatuszyńska, M.: Experimental research in economics and computer simulation. In: Nermend, K., Łatuszynska, M. (eds.) Selected Issues in Experimental Economics. Proceedings of the 2015 Computational Methods in Experimental Economics (CMEE) Conference. Springer Proceedings in Business and Economics, pp. 151–170 (2015)
27. Lavbič, L., Vasilecas, O., Rupnik, R.: Ontology-based multi-agent system to support business users and management. Technol. Econ. Dev. Econ. (TEDE) **16**(2), 327–347 (2010)
28. Nermend, K., Piwowarski, M.: Cognitive neuroscience techniques in supporting decision making and the analysis of social campaign. In: Proceedings Book International Conference on Accounting, Business, Economics and Politics (ICABEP-2018), Erbil, Iraq, pp. 1–12 (2018)
29. Piwowarski, M.: Cognitive neuroscience techniques in examining the effectiveness of social advertisements. In: Neuroeconomic and Behavioral Aspects of Decision Making, pp. 341–352. Springer, Cham (2017)
30. Piwowarski, M.: Neuromarketing tools in studies on models of social issue advertising impact on recipients. In: International Conference on Computational Methods in Experimental Economics, pp. 99–111. Springer, Cham (2017)
31. Piwowarski, M.: EEG in analysis of the level of interest in social issue advertising. Procedia Comput. Sci. **126**, 1945–1953 (2018)
32. Staab, S., Schnurr, H.P., Studer, R., Sure, Y.: Knowledge processes and ontologies. IEEE Intell. Syst. **16**, 26–34 (2001)
33. Tarczyński, W.: Selected statistical methods in experimental studies. In: Nermend, K., Łatuszynska, M. (eds.) Selected Issues in Experimental Economics. Proceedings of the 2015 Computational Methods in Experimental Economics (CMEE) Conference. Springer Proceedings in Business and Economics, pp. 91–106 (2015)
34. Vecchiato, G., et al.: Neurophysiological tools to investigate consumer's gender differences during the observation of TV commercials. Comput. Math. Methods Med. (2014). https://doi.org/10.1155/2014/912981

Part II
Application of Quantitative and Experimental Methods and Tools in Contemporary Economics

Chapter 8
Quantitative Methods in Real Estate Mass Appraisal

Mariusz Doszyń and Krzysztof Dmytrów

Abstract Proposed introduction of the ad valorem tax on the real estate market makes it necessary to perform mass appraisal of the real estates. The goal of the chapter was verification of usefulness of econometric and statistical methods in the mass appraisal process. Exponential econometric model and partial τ_B Kendall correlation coefficients were applied to identify the impact of attributes and location on unit real estate price. The so-called Szczecin algorithm of real estate mass appraisal was the basis in both econometric and statistical approach. Accuracy of both approaches was checked by means of percentage error (PE) and mean absolute percentage error (MAPE) distributions. Real database containing information about 113 transactions with undeveloped land for housing purposes in Szczecin was used. The research results suggest that both methods can be used for real estate mass appraisal; however, the econometric approach gave slightly better results.

Keywords Real estate mass appraisal · Szczecin algorithm of real estate mass appraisal · Qualitative variables · Econometric modelling · Partial τ_B Kendall correlation coefficient

JEL Codes C10 · C51 · R30

The work is financed by the National Centre of Science within the scope of project No. 2017/25/B/HS4/01813.

M. Doszyń · K. Dmytrów (✉)
Faculty of Economics and Management, University of Szczecin, ul. Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: krzysztof.dmytrow@usz.edu.pl

M. Doszyń
e-mail: mariusz.doszyn@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_8

8.1 Introduction

Nowadays, there is a need to use scientific and modern tools in mass evaluations. The number of real estate for which the value has to be settled is usually high. Mass evaluations have to be also often repeated. Therefore, in contemporary mass appraisals, quantitative methods should be applied, if possible.

Mass appraisal process is substantially different than individual real estate valuations. In mass appraisal, many real estate have to be evaluated at the same time, by means of the same methodology. Econometric and statistical methods seem to be useful in this context.

Therefore, the main hypothesis of the chapter is that quantitative (econometric and statistical) methods are helpful in supporting the process of real estate mass appraisal. These methods give the possibility to objectively estimate the impact of real estate attributes on value (or price). It does not mean that experts (appraisers) should be excluded from the mass appraisal process. There are situations in which it is not possible to apply quantitative methods. For example, databases would contain a small number of observations or real estate would not have certain states of attributes. Generally, all quantitative results should be judged by experts, before they are applied.

There are also many methodological issues connected with the application of econometric and statistical methods in the real estate mass appraisal. These methods are envisaged for modelling random variables, for which we have many realizations, for different states of control variables. That kind of conditions is mainly concerned in experimental sciences, where controlled (explanatory) variables could be freely determined.

While applying econometric and statistical methods, instead of different states of control variables, we have various real estate, with different states of their attributes. In the presented context, controlling the attributes' states could be understood as a choice of real estate with different attribute states.

To sum up, in the chapter, econometric and statistical methods will be applied to estimate the attributes' impact on the price of real estate. On the basis of price, real estate values will be calculated. Instead of control variables, real estate with different attributes states will be analysed.

8.2 Literature Review

There are many methodological proposals in the literature in the context of real estate evaluation but most of them regard individual appraisals. The extensive review of quantitative methods useful in mass valuation could be found in Kauko and D'amato [9].

Generally, they might be classified into four groups:

1. model-driven methods,
2. data-driven methods,

3. methods based on machine learning,
4. expert methods.

That kind of classification could be found in Kauko and D'amato [9]. Model-driven methods constitute the standard approach. They include standard econometric models, hedonic regression models and spatial econometric models. Data-driven methods involve nonparametric models, such as geographically weighted regression (GWR). Nowadays, the fastest growing groups are methods based on machine learning. They include varieties of quantitative tools, such as artificial neural networks (ANN), genetic algorithms, fuzzy logic and rough set theory. Also, expert methods, such as analytic hierarchy process (AHP), conjoint analysis, and contingent valuation (CV), constitute the fast-growing group. They might be helpful if the quality of databases is low and also could efficiently support mass appraisal process.

Despite the rapid development of other methods, multiple regression analysis is usually the standard approach in mass real estate valuations. Problems connected with application of these methods are discussed, e.g., in Benjamin et al. [1], Dell [3], Isakson [8], and Pagourtzi et al. [15]. Some researchers suggest using a priori information in the form of restrictions to avoid negative effects of multicollinearity [14].

Usefulness of econometric models with spatial effects in mass appraisal context was considered by Bourassa et al. [2], Kauko and D'amato [9]. In Zurada et al. [17], regression models are compared with artificial intelligence methods. An example of expert system application is provided in Kilpatrick [10].

In this chapter, in the last step, the Szczecin algorithm of real estate mass appraisal will be used. It is described, e.g., in Hozer et al. [7]. Problems connected with the econometric specification of that algorithm are presented in the chapter of Doszyń and Hozer [5]. The problems of introducing attributes into econometric models with respect to measurement scales are discussed in Doszyń [4]. Prediction accuracy measures useful in mass appraisal are discussed in McCluskey et al. [13].

8.3 Methods

In the chapter, two kinds of methods are used to estimate real estate attributes' impact on value: econometric and statistical. After this, the Szczecin algorithm of real estate mass appraisal will be applied to evaluate real estate. In the last step, usefulness of econometric model will be compared with statistical methods.

The Szczecin algorithm of real estate mass appraisal could be presented as follows [7]:

$$w_{ji} = \mathrm{wwr}_j \cdot \mathrm{pow}_i \cdot c_{\mathrm{baz}} \prod_{k=1}^{K} \prod_{p=1}^{k_p} (1 + a_{kp}) \tag{8.1}$$

where w_{ji}—market (or cadastral) value of ith real estate in jth location attractiveness zone,
wwr_j—market value coefficient in jth location attractiveness zone ($j = 1, 2, \ldots, J$),
J—number of location attractiveness zones,
pow_i—area of ith real estate,
c_{baz}—price of 1 m^2 of the cheapest (unimproved) land in the appraised area,
a_{kp}—the impact of pth state of kth attribute ($k = 1, 2, \ldots, K$; $p = 1, 2, \ldots, k_p$),
K—number of attributes,
k_p—number of states of kth attribute.

The basic price is the price of a real estate with the worst attribute states. The algorithm works such that the impact of appraised real estate attributes is added (by multiplication) to the basic price. In the chapter, attributes' impact (a_{kp}) will be estimated by means of econometric and statistical methods. Market value coefficients (wwr_j) are calculated for each location attractiveness zone. They reflect the impact of a broadly understood location.

Algorithm (8.1) served as a basis for the specification of econometric model. The surface area (pow_i) was moved to the left side of Eq. (8.1), and the logarithms were applied and error term was added:

$$\ln\left(\frac{w_{ji}}{\text{pow}_i}\right) = \alpha_0 + \sum_{k=1}^{K}\sum_{p=2}^{k_p}\alpha_{kp}x_{kpi} + \sum_{j=2}^{J}\alpha_j \text{te}_j + u_i \tag{8.2}$$

where α_0—intercept parameter,
K—number of real estate attributes,
k_p—number of states of kth attribute,
a_{kp}—impact of pth state of attribute k,
x_{kpi}—zero-one variable for pth state of attribute k for ith real estate ($i = 1, 2, \ldots, N$),
N—number of real estates,
α_j—market value coefficient for jth location attractiveness zone,
te_j—dummy variable equal to one for jth location attractiveness zone (and zero for others),
u_i—error term.

Model (8.2) is an econometric version of the algorithm (8.1). The dependent variable is a natural logarithm of a real estate unit value. Real estate attributes are qualitative variables measured on ordinal scale, so they were introduced into the model (8.2) as dummy variables, for each state of an attribute.

The model (8.2) includes a constant term. To avoid strict collinearity of the independent variables, dummy variables for the worst attributes' states were omitted, hence the summation of $p = 2, \ldots, k_p$ in (8.2). The ignored state of an attribute is a reference point for the remaining states. For the same reason, also the dummy variable for the first location attractiveness zone was ignored. Having regard to all of this, we could assume that constant term is an estimate of the basic price (c_{baz}). It is correct if, in the omitted location attractiveness zone, real estate was the cheapest

with respect to location. Also, database has to contain real estate with the worst states of each attributes. In other cases, better than worse attribute states will be omitted and constant term could not be treated as a basic price. For example, let us assume that we have surroundings as an attribute with three states: adverse, average and favourable. But in the database, we have real estate with only two states: average and favourable. Average (not adverse) surroundings would be a reference point in this case and the constant term would not be a basic price (c_{baz}).

Market value coefficients (α_j) are estimated by introducing dummy variables for each location attractiveness zone, but 'first' location attractiveness zone is omitted (because of strict collinearity of independent variables). The skipped location attractiveness zone creates a point of reference in interpreting market value coefficients for the remaining zones.

In the next step of the analysis, the statistical approach was used. It was based on the partial correlation coefficients. As the attributes and the location attractiveness zones were not numerical, the Kendall correlation coefficients were used. It can be calculated on the basis of the data measured on the ordinal scale and is based on ranks. Kendal correlation coefficient exists in three forms—one of them assumes that all ranks are distinct (τ_A). The other two (τ_B and τ_C) are based on the assumption that there are tied ranks. In the chapter, the τ_B coefficient was used. It is calculated by means of the following equation [16]:

- Observations are linked in all possible pairs: (x_i, y_i) and (x_j, y_j), $i \neq j$.
- If both $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$, then such a pair is called *concordant*. The number of such pairs is equal to n_{c}.
- If both $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$, then such a pair is called *discordant*. The number of such pairs is equal to n_{d}.
- If $x_i = x_j$ or $y_i = y_j$, then such a pair is neither *concordant* nor *discordant*. It is a *tied* pair.

Having determined the above values, τ_B Kendall coefficient is computed in accordance with the following formula:

$$\tau_B = \frac{n_{\text{c}} - n_{\text{d}}}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}, \tag{8.3}$$

where

$$n_1 = \frac{\sum_i t_i(t_i - 1)}{2},$$

$$n_2 = \frac{\sum_j u_j\left(u_j - 1\right)}{2},$$

t_i—number of observations in a sample featuring the same ith rank value of variable x,
u_j—number of observations in a sample featuring the same jth rank value of variable y.

When analysing the relationship between attributes and the value of 1 m^2, it may occur that there is a strong correlation between attributes. It may affect the strength and direction of the relation between attributes and the value of 1 m^2 of a real property. In order to eliminate the impact of the remaining variables during the correlation test between a given attribute and the value of 1 m^2, partial correlation coefficients are calculated [6]:

$$v_{yx\cdot z} = -\frac{R_{yx}}{\sqrt{R_{yy} \cdot R_{xx}}}, \tag{8.4}$$

where y—vector of the explained variable,
x—vector of the explanatory variable,
z—vector (or matrix) of the remaining variables,
R_{yx}—determinant of a matrix cofactor obtained by removing the row corresponding to y variable and the column corresponding to x variable,
R_{yy}—determinant of a matrix cofactor obtained by removing the row and column corresponding to y variable,
R_{xx}—determinant of a matrix cofactor obtained by removing the row and column corresponding to x variable.

$$w_k = \frac{\left|v_{yx_k\cdot z}\right|}{\sum_{k=1}^{K}\left|v_{yx_k\cdot z}\right|} \cdot 100\% , \tag{8.5}$$

where k—number of an analysed attribute,
K—number of attributes.

The greater the share of the absolute value of a partial correlation coefficient of a given attribute is in the sum of the absolute values of all attribute coefficients, the greater the weight ought to be assigned to a given attribute. Formula (8.5) is used for the assessment of weights of all attributes [11].

The estimation of the impact of pth state of attribute k $(1 + a_{kp})$ on the value of 1 m^2 of a real property can be calculated by means of the following formula [12]:

$$1 + a_{kp} = \left(1 - \frac{1}{2}w_k\right) + w_k \frac{l_{kp}}{k_p - 1}, \tag{8.6}$$

where w_k—weight of the kth attribute,
l_{kp}—pth state of attribute k,
k_p—number of states of kth attribute.

Having estimated all values of $1 + a_{kp}$, we can obtain the hypothetical value of 1 m^2 of real estate by means of the following formula:

$$\hat{W}_{hi} = \text{pow}_i \cdot c_{\text{baz}} \cdot \prod_{p=1}^{k_p} \prod_{k=1}^{K} (1 + a_{kp}), \tag{8.7}$$

where $\hat{W}_{hi}$—hypothetical value of the ith real estate.

The value of the wwr_i (for every real estate) is calculated by means of the following formula:

$$\text{wwr}_i = \frac{\text{WR}_{ri}}{\hat{W}_{hi}}, \tag{8.8}$$

where WR_{ri}—value if the ith real estate set by the appraiser.

Next, the wwr_j value for every attractiveness zone is given by means of the following formula:

$$\text{wwr}_j = \sqrt[l]{\prod_{i=1}^{l} \text{wwr}_i}, \tag{8.9}$$

where l is the number of real estates in the jth attractiveness zone.

Having calculated the values wwr_j, they are substituted in Eq. (8.1), to obtain the market (or cadastral) value of ith real estate in jth location attractiveness zone.

8.4 Empirical Results

The database contains information about 113 transactions from years 2014 to 2018. The analysis period was expanded to four years to increase the number of observations. These are real estate transactions with undeveloped land for housing purposes in Szczecin. According to market analysis, prices did not exhibit trend, so corrections (with respect to time) were not necessary.

Five attributes were taken as explanatory variables:

1. plots area: 1—large (over 1200 m^2), 2—medium (500–1200 m^2) and 3—small (less than 500 m^2),
2. plot utilities: 1—missing, 2—incomplete and 3—full,
3. surroundings: 1—burdensome, 2—unfavourable, 3—average and 4—favourable,
4. transport accessibility: 1—unfavourable, 2—average and 3—good,
5. plots shape: 1—unfavourable, 2—average and 3—favourable.

Attribute states are important determinants of real estate price (and value), but probably the most important is the location. In this research, location impact is estimated by introducing dummy (zero-one) variables for each precinct. In Szczecin algorithm of mass valuation, dummies are introduced for location attractiveness

zones. In this research, precinct are taken, because it is not possible to precisely identify in which location attractiveness zone is given real estate. Only precincts are precisely known.

Estimation results are presented in Table 8.1.

All explanatory variables are dummy variables. In case of all attributes, the worse state is omitted, because of strict collinearity. The 'first' precinct is omitted for the same reason. All omitted states and precinct create reference point. Their impact is present in a constant term.

All attribute states are statistically significant only in case of area.[1] Impact of plot utilities and transport accessibility is not statistically different from zero. This applies to all states of these attributes. For surroundings and plots' shape, only the best states are statistically significant. Therefore, the impact of area and surroundings and plots' shape (but only the best states) is statistically different from zero. All statistically significant estimates are positive, which is consistent with the theory.

The question should be asked, what is the reason that sometimes impact of attributes is statistically insignificant. At first, attributes are purely qualitative variables and their states are an effect of expert's decisions. Two experts might judge the same attribute's state differently and this is probably often true. Real estate is frequently also very similar. This makes variability of attributes low. It is known from econometric theory that the lower variability, the higher standard errors (and lower t-statistics). Also, collinearity of explanatory variables that increases estimators' inefficiency might be a reason. Collinearity makes standard errors of parameters high, so estimates might be statistically insignificant or could have inappropriate signs.

According to the location, impact of dummy variables is statistically insignificant only in case of four precincts. Most estimates are negative, which means that unit prices are lower than in the omitted precinct. Impact of the location is the most important.

The estimated model is well fitted, and determination ratio is $\bar{R}^2 = 0.806$. According to Jarque–Bera test, residuals distribution is not normal. In this test, $p_{\text{emp}} = 0.004$, which is lower than significance level equal to 0.05. Residuals are homoscedastic. In Breusch–Pagan–Godfrey test, empirical significance level is 0.129.

To sum up, in mass appraisals, econometrics results should be verified by experts (e.g., appraisers). It is especially true when estimates are statistically insignificant and have signs not consistent with the theory. Experts are also necessary when in the database some attributes' states are missing and their impact is necessary to evaluate different real estate.

Next, the statistical approach was applied. First, the partial τ_B Kendall coefficients were calculated, their significance was analysed and appropriate weights were calculated. These weights were further used for estimation of the $(1 + a_{kp})$ in Formula (8.1). The partial τ_B Kendall correlation coefficients and obtained on their bases weights are presented in Table 8.2.

As seen in Table 8.2, the influence of area and surroundings on the real estate unit values was statistically insignificant. Also, application of the partial coefficients

[1] In the chapter, significance level is equal to 0.05.

Table 8.1 Econometric model of real estate unit price (8.2) (own calculations made in EViews)

Variable	Coefficient	Std. error	*t*-statistic	*p*-value
Const	**5.963**	**0.146**	**40.852**	**0.000**
Area = 2	**0.151**	**0.047**	**3.222**	**0.002**
Area = 3	**0.561**	**0.168**	**3.337**	**0.001**
Utilities = 2	0.134	0.114	1.179	0.242
Utilities = 3	0.123	0.072	1.701	0.093
Surroundings = 3	0.091	0.068	1.340	0.184
Surroundings = 4	**0.269**	**0.125**	**2.150**	**0.035**
Accessibility = 2	0.053	0.051	1.030	0.306
Accessibility = 3	−0.043	0.075	−0.571	0.570
Shape = 2	0.146	0.074	1.977	0.052
Shape = 3	**0.301**	**0.083**	**3.628**	**0.001**
Precinct = 1042	−0.067	0.209	−0.323	0.748
Precinct = 1049	−0.174	0.143	−1.211	0.230
Precinct = 1050	−0.422	0.266	−1.583	0.117
Precinct = 2041	−0.184	0.209	−0.882	0.381
Precinct = 2129	**−0.629**	**0.290**	**−2.168**	**0.033**
Precinct = 3001	**−1.642**	**0.242**	**−6.785**	**0.000**
Precinct = 3006	**−0.975**	**0.204**	**−4.783**	**0.000**
Precinct = 3011	**−0.921**	**0.132**	**−6.957**	**0.000**
Precinct = 3021	**−1.118**	**0.274**	**−4.089**	**0.000**
Precinct = 3064	**−0.903**	**0.159**	**−5.677**	**0.000**
Precinct = 3065	**−0.724**	**0.133**	**−5.437**	**0.000**
Precinct = 3066	**−1.082**	**0.236**	**−4.588**	**0.000**
Precinct = 3068	**−1.097**	**0.209**	**−5.255**	**0.000**
Precinct = 3075	**−0.735**	**0.261**	**−2.811**	**0.006**
Precinct = 3076	**−0.702**	**0.237**	**−2.962**	**0.004**
Precinct = 3077	**−1.248**	**0.179**	**−6.959**	**0.000**
Precinct = 3079	**−0.806**	**0.209**	**−3.864**	**0.000**
Precinct = 3080	**−0.779**	**0.184**	**−4.236**	**0.000**
Precinct = 3081	**−0.891**	**0.133**	**−6.717**	**0.000**
Precinct = 3082	**−0.834**	**0.155**	**−5.384**	**0.000**
Precinct = 3084	**−0.848**	**0.144**	**−5.872**	**0.000**
Precinct = 3085	**−0.364**	**0.132**	**−2.766**	**0.007**

Statistically significant results are bolded (significance level 0.05)

Table 8.2 Partial τ_B Kendall correlation coefficients and weights (own calculations made in R and Excel)

Attribute	Partial τ_B Kendall coefficient	p-value	Weight
Area	0.103	0.112	0.142
Utilities	**0.183**	**0.005**	**0.251**
Surroundings	0.053	0.412	0.073
Accessibility	**0.202**	**0.002**	**0.278**
Shape	**0.187**	**0.004**	**0.256**

Statistically significant results are bolded (significance level 0.05)

cleared the influence between analysed attribute and the unit price of real estate from the influence of other attributes. By this manner, the signs of coefficients were positive and it is concordant with the assumption that the better the state of a given attribute is, the higher the unit price of the real estate is.

In a contrary to the econometric approach, the influence of location of the real estate is measured by the market value coefficients (wwr_j), calculated for each precinct by means of Formula (8.9). Their influence on the real estate unit value is slightly different than this obtained by means of the econometric model. The econometric model indicated that the theoretical unit real estate values (for the same values of all attributes) in the precinct 1018 were higher than in other precincts. The statistical approach, on the contrary, showed that the higher values were in the precinct 1042 (here, at the same values of all attributes, the theoretical unit real estate value was 2.275 times higher than the hypothetical value, obtained only by the values of attributes), followed by the precinct 1049. The theoretical unit real estate values in the precinct 1018 were third highest. Precinct 3001 was the least attractive (here, at the same values of all attributes, the theoretical unit real estate value was by 29.4% lower than the hypothetical value, obtained only by the values of attributes) (Table 8.3).

In the next step, estimated model is used to obtain the real estate unit values. Experts' modifications are not made, and their meaning will be considered in the future research. Removing insignificant variables might make remaining estimates biased. Therefore, real estate unit values are calculated on the basis of model presented in Table 8.1 (the econometric approach) and on the basis of weights presented in Table 8.2 and the market value coefficients presented in Table 8.3 (statistical approach). It should be added that the experts' corrections should be made in that kind of situations.

The obtained results show that the empirical distribution of unit prices is very similar to the distribution of prices calculated on the basis of the econometric model. However, application of the statistical approach gave worse results (Fig. 8.1). It may result from the fact that in the econometric approach, every state of each attribute was described by a distinct parameter. On the contrary, in the statistical approach, assigned weight for each attribute was shared for its every state. Therefore, the transitions between states, obtained by means of Formula (8.6), were linear, which

Table 8.3 Market value coefficients (wwr_j) (own work)

Precinct	wwr_j
Precinct = 1018	2.152
Precinct = 1042	2.275
Precinct = 1049	2.161
Precinct = 1050	2.150
Precinct = 2041	2.024
Precinct = 2129	2.139
Precinct = 3001	0.706
Precinct = 3006	0.875
Precinct = 3011	1.063
Precinct = 3021	1.002
Precinct = 3064	1.103
Precinct = 3065	1.378
Precinct = 3066	1.179
Precinct = 3068	0.813
Precinct = 3075	1.345
Precinct = 3076	1.663
Precinct = 3077	0.757
Precinct = 3079	1.087
Precinct = 3080	1.256
Precinct = 3081	1.211
Precinct = 3082	1.351
Precinct = 3084	1.352
Precinct = 3085	1.627

was not in the case of the econometric approach. Also, distribution of percentage errors (PE) seems to confirm this conclusion (Fig. 8.2).

Percentage error is calculated as $PE = (p_i - v_i)/p_i$, where p_i—a unit price of an ith real estate and v_i—unit real estate value obtained from the estimated model. Average percentage error, for all 113 transactions, was equal to −1.06% for the econometric approach. The statistical approach gave this error on the level of −1.43%. These results mean that predictions are almost unbiased. Average absolute percentage error, calculated as a mean for absolute PE (MAPE), is equal to 10.59% in case of the econometric approach and 12.62% for the statistical one. It means that the unit real estate values are different than real unit prices on the average by ±10.59% (in the econometric approach) and ±12.62% in the statistical approach. To conclude, estimated model predicts unit prices (which are treated as a unit real estate values) quite accurately.

Both econometric and statistical approach provide to a slightly negatively skewed distribution of percentage error (Fig. 8.2).

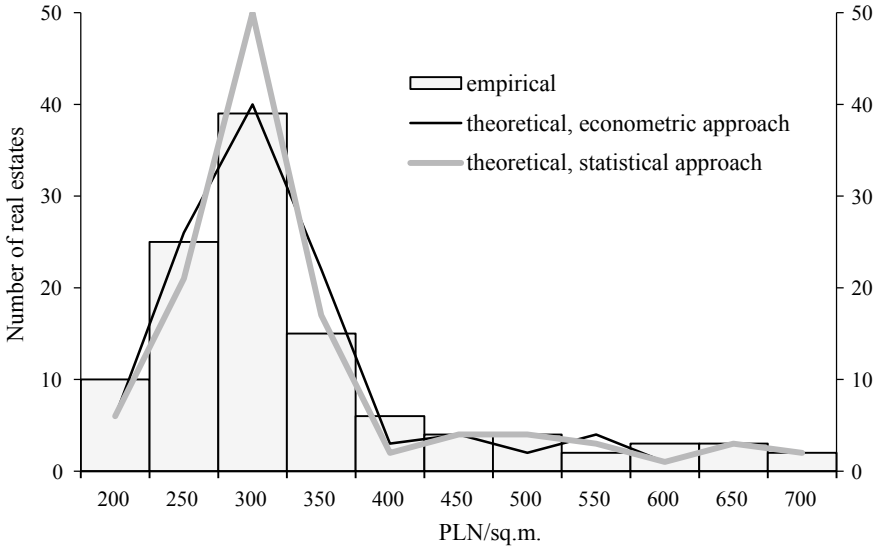


Fig. 8.1 Distribution of empirical and theoretical unit prices (own work)

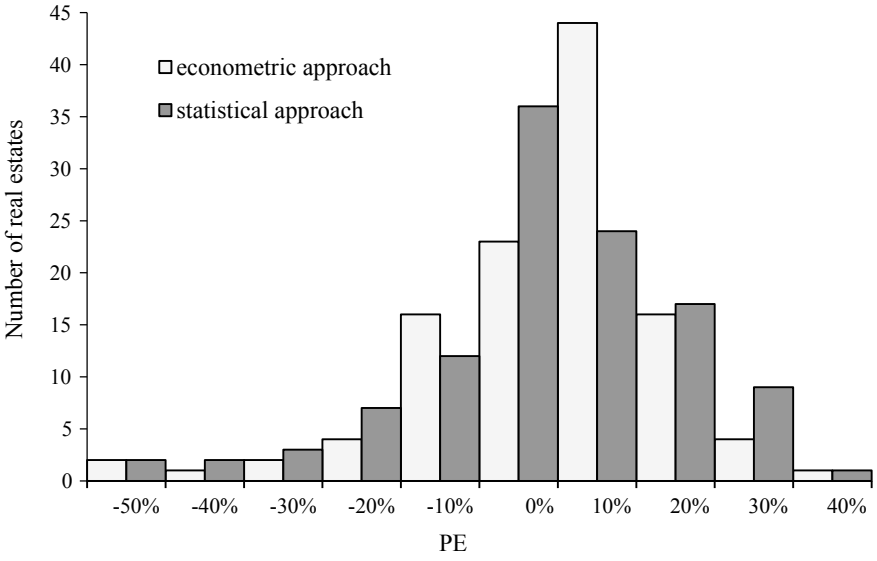


Fig. 8.2 Distribution of percentage error (own work)

8.5 Conclusions

In the chapter, the econometric and statistical approaches were applied for the real estate mass appraisal. The modified Szczecin algorithm of real estate mass appraisal was used. The results show that both approaches can be used in the process of real estate mass appraisal. The econometric approach yielded better results. Both methods generated quite low mean percentage errors, and however, they were smaller in the case of econometric approach. The reason for this was that in the econometric approach, each state of each attribute was described by distinct parameter, while in the statistical approach it was not—the transitions between states were assumed to be linear; however, in reality, it was not necessarily true. In order to improve the statistical approach, it would be necessary to propose the nonlinear transitions between the states of attributes of the real estates. On the contrary, this research applied the partial τ_B Kendall coefficients in order to estimate the impact of attributes and impact of variables depicting the location in a certain precinct. The analysis also showed that application of the partial coefficients allowed to clear the influence of given variable from the influence of other explanatory variables on the unit real estate price. Also, application of the τ_B Kendall correlation coefficients is appropriate with respect to the ordinal scales, by which the attributes are described.

References

1. Benjamin, J.D., Randall, S., Guttery, R.S., Sirmans, C.F.: Mass appraisal: an introduction to multiple regression analysis for real estate valuation. J. Real Estate Pract. Educ. **7**(1), 65–77 (2004)
2. Bourassa, S.C., Cantoni, E., Hoesli, M.: Predicting house prices with spatial dependence: a comparison of alternative methods. Paper presented at the 15th Conference of the Pacific Rim Real Estate Society (PRRES), Sydney, Australia, 18–21 Jan 2009
3. Dell, G.: Regression, critical thinking, and the valuation problem today. Apprais. J. **85**(3), 217–230 (2017)
4. Doszyń, M.: Statistical determination of impact of property attributes for weak measurement scales. Real Estate Manag. Valuat. **25**(4), 75–84 (2017). https://doi.org/10.1515/remav-2017-0031
5. Doszyń, M., Hozer, J.: Szczeciński algorytm masowej wyceny nieruchomości—podejście ekonometryczne (Szczecin's algorithm of real estate mass valuation—econometric approach). Stud. Pr. WNEiZ US **50**(1), 19–30 (2017). https://doi.org/10.18276/sip.2017.50/1-02
6. Han, L., Zhu, J.: Using matrix of thresholding partial correlation coefficients to infer regulatory network. BioSystems **91**, 158–165 (2008). https://doi.org/10.1016/j.biosystems.2007.08.008
7. Hozer, J., Foryś, I., Zwolankowska, M., Kokot, S., Kuźmiński, W.: Ekonometryczny algorytm masowej wyceny nieruchomości gruntowych (Econometric Algorithm of Land Properties Mass Valuation). Katedra Ekonometrii i Statystyki Uniwersytetu Szczecińskiego, Stowarzyszenie Pomoc i Rozwój, Szczecin (1999)
8. Isakson, H.R.: The review of real estate appraisals using multiple regression analysis. J. Real Estate Res. **15**(2), 177–190 (1998)
9. Kauko, T., D'amato, M. (eds.): Mass Appraisal Methods: An International Perspective for Property Valuers. Blackwell Publishing Ltd (2008)

10. Kilpatrick, J.: Expert systems and mass appraisal. J. Prop. Invest. Financ. **29**(4/5), 529–550 (2011). https://doi.org/10.1108/14635781111150385
11. Kolenda, M.: Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza obiektów wielocechowych (Numerical Taxonomy: Classification, Ranking and Analysis of Multivariable Objects). Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław (2006)
12. Lis, C.: Wykorzystanie metod ekonometryczno-statystycznych w procesie masowej wyceny nieruchomości lokalowych (Application of econometric and statistical methods in the process of mass valuation of residential properties). Typescript, doctor's thesis, Uniwersytet Szczeciński (2003)
13. McCluskey, W.J., McCord, M., Davis, P.T., Haran, M., McIllhatton, D.: Prediction accuracy in mass appraisal: a comparison of modern approaches. J. Prop. Res. **30**(4), 239–265 (2013). https://doi.org/10.1080/09599916.2013.781204
14. Pace, R.K., Gilley, O.W.: Estimation employing a priori information within mass appraisal and hedonic pricing models. J. Real Estate Financ. Econ. **3**, 55–72 (1990)
15. Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., French, N.: Real estate appraisal: a review of valuation methods. J. Prop. Invest. Financ. **21**(4), 383–401 (2003). https://doi.org/10.1108/14635780310483656
16. Parker, R.I., Vannest, K.J., Davis, J.L., Sauber, S.B.: Combining nonoverlap and trend for single-case research: Tau-U. Behav. Ther. **42**(2), 284–299 (2011). https://doi.org/10.1016/j.beth.2010.08.006
17. Zurada, J., Levitan, A.S., Guan, J.: A comparison of regression and artificial intelligence methods in a mass appraisal context. J. Real Estate Res. **33**(3), 349–388 (2011)

Chapter 9
Application of the Survival Analysis Methods in Contemporary Economics on the Example of Unemployment

Beata Bieszk-Stolorz

Abstract Contemporary economics uses newer and newer methods of data analysis. Survival analysis is one of such methods. It originated in the demography and reliability analysis. Duration of a phenomenon is observed until the moment of occurrence of specific event. It can be the duration of life, increase or decrease of share prices, debt payment, business life, unemployment. If the occurrence of the event does not take place in specific period, such observation is considered as censored. Survival function is the basic and primary one. It describes the probability of nonexistence of the event until the moment *t*. The second function is the hazard function that describes the intensity of occurrence of the event in the moment *t*. In the chapter, selected methods of the survival analysis will be applied for the analysis of duration of the registered unemployment with respect to the unemployed persons' features. Among the other things, models of risk of the competing events will be analyzed. They allow to estimate the probability and intensity of accepting the job, removal from the register and de-registration due to other causes.

Keywords Survival analysis · Cumulative incidence function · Lunn-McNeil model · Unemployment

9.1 Introduction

Survival analysis methods are commonly applied in demography and medical sciences for the analysis of human life duration. They are derived from the technical sciences, where they are used in the reliability theory. They can also be used for the analysis of the duration of social and economic phenomena. These methods require the knowledge of the individual data referring to each unit belonging to analyzed population. It is one of the most significant limitations of these methods because the access to such data can be very difficult. First, these difficulties can result from the

B. Bieszk-Stolorz (✉)
Faculty of Economics and Management, Institute of Econometrics and Statistics, University of Szczecin, ul. Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: beata.bieszk-stolorz@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_9

institutions' methods of data collection. Often only the aggregate data is available. Second, access to the individual data can be limited because of its legal protection. The survival analysis is based on the probability calculus. The unit's duration at given state, which is the random variable, denoted by T, until the moment of certain event ending its observation, is analyzed. It can be, for example, the company's duration, unemployment duration, or duration of credit repayment. The following events: cessation of the company's activity, taking up a job or credit repayment are the terminating ones in these cases. The methods of duration analysis can be used for the analysis of the companies' vitality [15], economic activity of the population [14], dynamics of poverty in urban and rural households [22], credit risk [17], or unemployment duration [3].

These methods are used in the labor market analyses. Except for the unemployment duration, the employment time [2, 6], influence of time of receiving allowance on time, and intensity of looking for employment are also analyzed [18]. The literature also considers the analysis of competing events (e.g., other forms of unemployment exit) [7, 21] and the analysis of recurrent periods of unemployment [4].

The goal of the chapter is the comparison of duration and intensity of unemployment exit due to various forms of de-registration: taking up a job, removal from the register, and other causes. The hypothesis was made that this differentiation was caused by the unemployment persons' features: gender, age, education, and also seniority and the number of subsequent registrations. In the chapter, the selected methods of the survival analysis and their application in the analysis of registered unemployment are analyzed. It was analyzed how the selected characteristics of the unemployed affected the probability and intensity of unemployment exit. The three forms of de-registration from the labor office were considered: taking up a job, removal from the register due to reasons attributable to the unemployed person, and other causes. It is worth mentioning that in the subject literature often instead of the term "de-registration", the term "outflow" is used [19].

9.2 Data Used in the Research

In the research, the individual data obtained from the Poviat Labor Office in Szczecin was used. Individual data referring to 22,274 unemployed persons registered in 2014 and observed until the end of 2016 was analyzed. The data contained information about the registration time, gender, age, education, seniority, number of registrations, and cause of de-registration of the unemployed person. The event that ended observation was the moment of de-registration from the office. The time from the moment of registration to the moment of de-registration for a specific cause was analyzed. Introduction of the Syriusz information system enabled gathering extensive data on the unemployed persons. Currently, registers give dozens of causes of de-registration. They are, inter alia, causes connected with taking a job, transition to a pension or

retirement, continuation of education, going abroad. These causes were grouped and three groups of competing events were considered: job, removal, and other causes. Each of the there main causes of de-registration contained the root causes (Table 9.2). Taking up job (abbreviated to "job") consisted of three root causes: taking up job or other employment, taking up subsidized job, business activity. Removal is the group of causes, in which the registered person did not show any willingness to cooperate with the office and was removed from the register through their own fault or at their own request. Other causes of de-registration (other) were less numerous, and as previously conducted analyses show, each of them had marginal impact on the probability of de-registration from the office; therefore, they were considered together. Part of the observations did not end with the event or de-registration in the analyzed period. These observations were right-censored. The sizes of each group are presented in Table 9.1.

Table 9.1 Structure of analyzed unemployed people (own study)

Group	Total	Job	Removal	Other	Censored
Total	22,274	9837	10,360	1752	325
Gender					
Women (K or 1)	9855	4943	3894	832	186
Men (M or 0)	12,419	4894	6466	920	139
Age					
18–24 (W_1)	3891	1296	2388	199	8
25–34 (W_2)	7367	3623	3309	375	60
35–44 (W_3)	4547	2090	2152	237	68
45–54 (W_4)	3236	1542	1373	249	72
55–59 (W_5)	2194	973	787	393	41
60–64 (W_6)	1039	313	351	299	76
Education					
At most lower secondary (S_1)	5439	1505	3498	354	82
Basic vocational (S_2)	4905	1947	2482	405	71
General secondary (S_3)	2896	1167	1447	250	32
Vocational secondary (S_4)	4040	1949	1624	392	75
Higher (S_5)	4994	3269	1309	351	65
Number of registrations					
Registered for the first time (Z_0)	5154	2421	2193	496	44
Registered subsequently (Z_0)	17,120	7416	8167	1256	281
Seniority					
Without professional experience (D_0)	8458	2769	5058	529	102
With professional experience (D_1)	13,816	7068	5302	1223	223

9.3 Research Methodology

The basic concept of the survival analysis is the survival function, defined as follows:

$$S(t) = P(T > t) = 1 - F(t) \tag{9.1}$$

where

T—survival time,
$F(t)$—the cumulative distribution function of the random variable T.

The survival function denotes the probability that the event will not occur until the moment t. In the case of demography and medical sciences, death is the analyzed event and the probability that analyzed unit will survive is estimated. In relation with the defined event, sometimes it is more convenient to analyze the cumulative distribution function $F(t)$ that denotes the probability that the event will occur at most at the moment t. The good example in this case is the analysis of the unemployment duration. If we define the event as taking up job by the registered unemployed person, the estimator of the survival function informs, what the probability of staying in the register is, and the estimator of the cumulative distribution function allows for the designation of the probability of taking up job.

If the distribution of the survival time of the analyzed occurrence is unknown, the survival function is mostly estimated by means of the Kaplan–Meier estimator [9]:

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \tag{9.2}$$

where

d_j—the number of events at the moment t_j,
n_j—the number of individuals at risk by the moment t_j.

Research connected with the survival analysis mostly consists in the observation of units belonging to defined cohort, or the set of units, selected from the population due to an event or process that takes place simultaneously for the whole set. The cohort should be selected on the basis of statistically significant features and should be homogeneous with respect to them. For each unit, the duration at given state or the duration of analyzed process is observed. If the period of observation of units is defined, for some of them the event may not occur before its end. In such cases, the duration for specific unit is known only partially. Such units constitute the right-censored observations. In research, such observations are also situations, in which the analyzed unit disappears from the view or there is an event excluding the occurrence of relevant event that ends the observation [20], or the competing event [1]. In medical applications, the analysis of causes of death is typical example. For example, in the analysis of cancer, death may occur because of recurrence or remission (caused by the treatment). Differentiation of these causes is very important

for the whole treatment process. In the engineering sciences, the competing risk is connected with the influence of the components on the whole system. Here, the failure of any component causes the failure of the whole system. The time, in which the system fails and which component to which degree has caused the failure, is observed. In practice, observation ended with the competing event can be treated as the right-censored one. However, the application of the competing risks can bring interesting results [10, 11].

Let T and C be continuous random variables describing the time to event and the time to censoring, respectively. When there are K types of competing risks, the observation encompasses the pairs (X, δ), where $X = \min(T, C)$ and $\delta = 0, 1, \ldots, K$. If a given observation is censored, then $\delta = 0$ and $\delta = 1, \ldots, K$ for the observations ending with an event (one of the K competing risk ones). In this context, one of the K events can be considered the event of elementary importance, while all the remaining ones—the competing risk events.

Estimator of the cumulative incidence was for the first time proposed by Kalbfleisch and Prentice [8]. It is the cumulative probability of occurrence of the event due to the cause k before time t with the assumption that analyzed unit is exposed to the existence of any competing risk k [5]. The cumulative incidence function is defined as follows [12, p. 52]:

$$\mathrm{CIF}_k(t) = P(T \leq t, \delta = k) = \int_0^t S(u)h_k(u)\mathrm{d}u = \int_0^t S(u)\mathrm{d}H_k \tag{9.3}$$

where

$H_k(t)$—specified (for a fixed $k = 1, 2, 3, \ldots, K$) function of cumulative hazard function,
$S(t)$—survival function.

Let $t_1 < t_2 < \ldots < t_i < \ldots < t_n$ be event times. Similar to the standard cumulative hazard function in the survival analysis, the cumulative hazard function $H_k(t)$ for the cause k can be expressed by the Nelson–Aalen estimator:

$$\hat{H}_k(t) = \sum_{j:t_{j\leq t}} \frac{d_{kj}}{n_j} \tag{9.4}$$

where

d_{kj}—number of events that have occurred due to the cause k,
n_j—number of individuals at risk at the time t_j.

Having combined the two above-mentioned estimators (9.3) and (9.4), we can estimate the cumulative incidence function due to the cause k [16] as:

$$\hat{\mathrm{CIF}}_k(t) = \sum_{j:t_j \leq t} \hat{S}\left(t_{j-1}\right)\frac{d_{kj}}{n_j} \tag{9.5}$$

The cumulative incidence function allows to determine the patterns of occurrence of an event due to the cause k and to assess, to which degree each cause contributes to the total failure. Since $\sum_{k=1}^{K} d_{kj} = d_j$, then the following relation is true:

$$\sum_{k=1}^{K} \hat{\mathrm{CIF}}_k(t) = 1 - \hat{S}(t) \tag{9.6}$$

In the absence of competing risk events, we have

$$\hat{\mathrm{CIF}}(t) = 1 - \hat{S}(t) \tag{9.7}$$

In order to assess the relative intensity of occurrence of analyzed event until the moment t, the Lunn-McNeil model can be applied. Data in this model must be grouped in particular way (Table 9.2). If there are K types of risk, initial data must be duplicated K times. The ith object with duration t_i exists K times. Each observation (also the duplicated one) is given the status e_i that informs about belonging of the ith object to the specific risk group: $e_i = 1$ if observation of the ith object ended with the occurrence of the event k ($k = 1, 2, \ldots, K$) and equals 0 otherwise. For the censored observation, $e_i = 0$ for every $k = 1, 2, \ldots, K$. Dummy dichotomic variables, $D_1, D_2, \ldots, D_K$, representing K types of risk are introduced in the model: $D_k = 1$ for the risk of the k—type and 0 otherwise.

If the risk $k = 1$ is assumed to be the reference, then the model considers $K - 1$ dummy variables $D_2, D_3, \ldots, D_K$. $X_1, X_2, \ldots, X_p$ are the explanatory variables.

If we denote the strata being the types of risk corresponding with K competing events, and the event $k = 1$ is considered as the reference by $g = 1, 2, \ldots, K$, then the Lunn-McNeil model can be defined as the stratified Cox model with interactions. It is described by means of the formula [13, p. 423]:

$$h_g^*(t, X) = h_{0g}^*(t) \exp\left(\sum_{j=1}^{p} \beta_j X_j + \sum_{k=2}^{K} \sum_{j=1}^{p} \delta_{kj} D_k X_j\right) \tag{9.8}$$

where

$X_1, X_2, \ldots, X_p$—explanatory variables,
$D_2, D_3, \ldots, D_K$—$K - 1$ dummy variables.

Also, the alternative form of the Lunn-McNeil model can be used [13, p. 423]:

$$h_g(t, X) = h_{0g}(t) \exp\left(\sum_{g=1}^{K} \sum_{j=1}^{p} \delta_{gj} D_g X_j\right) \tag{9.9}$$

where

$X_1, X_2, \ldots, X_p$—explanatory variables,

Table 9.2 Definition of variables in the Lunn-McNeil model (own elaboration)

Object	Time	Status	D_1	D_2	D_3	…	D_K	X_1	…	X_p
i	t_i	e_1	1	0	0	…	0	x_{i1}	…	x_{ip}
i	t_i	e_2	0	1	0	…	0	x_{i1}	…	x_{ip}
i	t_i	e_3	0	0	1	…	0	x_{i1}	…	x_{ip}
…	…	…	…	…	…	…	…	…	…	…
i	t_i	e_K	0	0	0	…	1	x_{i1}	…	x_{ip}

$D_1, D_2, \ldots, D_K$—K dummy variables.

Strata $g = 1, 2, \ldots, K$ correspond with K competing events. The second form of the model will be used in the presented analysis.

In the Lunn-McNeil model, the parameters δ_{gj} are not interpreted directly, but we interpret the form $\exp(\delta_{gj})$. If X_i is the explanatory dichotomic variable, then the expression

$$\mathrm{HR}_g\left(X_j = 1/X_j = 0\right) = \exp\left(\delta_{gj}\right) \tag{9.10}$$

is interpreted as relative hazard (relative intensity) of occurrence of the event of the k-type.

9.4 Analysis of Probability and Intensity of Exit from Unemployment

Five features of the unemployed persons were included in the research. In case of three of them, gender, seniority and the number of subsequent registrations, the two variants were selected and coded as dichotomic variables. Men, persons without professional experience, and persons registered for the first one were the reference groups (coded by 0). In case of education and age, five and six variants were distinguished, respectively. In such cases, variables were codes using the method −1, 0, 1 (quasi-experimental coding), what enabled to compare each variant with the mean value for all groups.

The first stage of the analysis consisted in the estimation of probability of de-registration with use of the CIF function. Analyzing all forms of exit from unemployment together, it could be noticed that median unemployment duration was about 3.5 months. It means that until this time half of the unemployed persons were de-registered (Fig. 9.1). Comparing three causes of exit from unemployment, it should be emphasized that the probability of exiting to job was similar to the probability of removal. However, after transition into the state of long-term unemployment (after 12 months since registration) the probability of removal increased.

The leap in value of probability of de-registration due to other causes after 7 months of unemployment is characteristic. It also visible in case of selected variants of analyzed features (Figs. 9.2, 9.3 and 9.4).

Direct analysis of individual data indicated that it was connected with increased number of de-registrations caused by granting the pre-retirement benefit. In case of three analyzed dichotomic features women, persons with professional experience and the unemployed persons registered for the first time had higher probability of taking up job than removal. For unemployed men, persons without professional experience and registered subsequently the situation was opposite (Fig. 9.2). In all variants of features, the probability of de-registration due to other causes was small

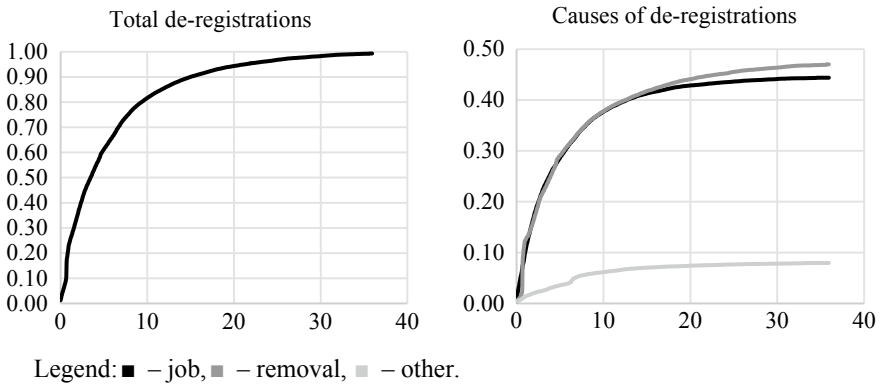


Fig. 9.1 Probability of de-registration in total and due to three causes (own study)

and did not exceed the value 0.1. Characteristic leap of its value in the seventh month of unemployment was particularly visible in case of persons registered for the first time.

Probability of de-registration due to three causes was differentiated with respect to education of the unemployed persons. Job was the soonest taken up by the persons with higher education, next with the vocational secondary education, and the latest with at most lower secondary (Fig. 9.3). Decision about resignation from mediation of the labor office was the soonest made by the persons with the lowest education and the latest with higher education.

In case of the age groups (Fig. 9.4) even 61.5% of persons at the age 18–24 year in three years since registration resign from the co-operation with the office. Probabilities of taking up job and resignation were similar for persons at the age 35–44 years. However, persons at the age groups 25–34 and 45–54 slightly more often take up job than are removed from the register. Persons older than 55 years start to enter the pre-retirement age. Therefore, they more rarely take up job are removed from the register. On the contrary, more often than in remaining groups they leave the register due to other causes. For persons at the age 55–59 year, it is slightly more than 18% and for persons older than 60 years—almost 30%. For persons at the age group W_5, the leap of the value of probability of de-registration due to other causes is high in seventh and eighth month and equals 200%. In the group of oldest persons, probabilities of de-registration due to all three causes are similar.

In the second stage of the research, by using the Lunn-McNeil model, the relative intensity of de-registration from the labor office was analyzed (Tables 9.3, 9.4 and 9.5).

Figures 9.5 and 9.6 present the hazard ratios. The lack of significance of parameter in the model means that the intensity for analyzed group was similar to the intensity of the reference group. The analysis showed that among persons registered in the labor office in Szczecin women (by 7%), persons with professional experience (by 38%), registered for the first time (by 14%), with higher education (by 58% more than the

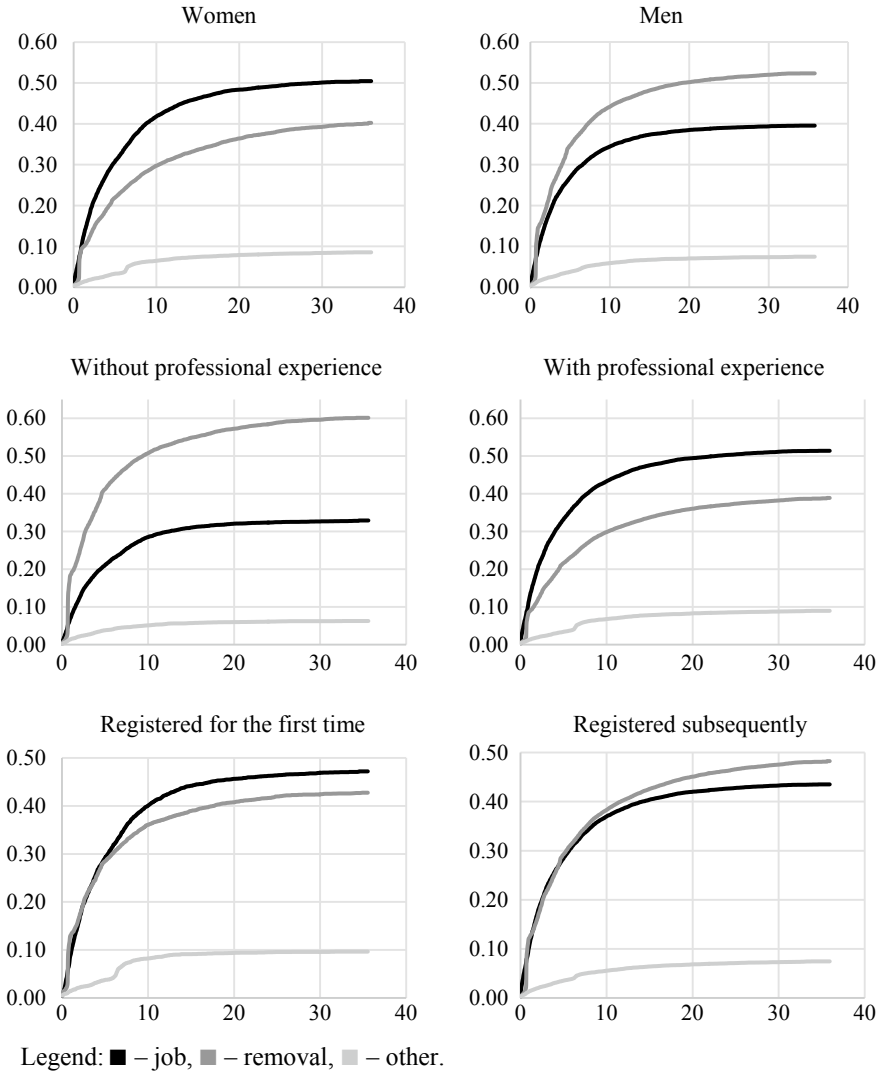


Fig. 9.2 Probability of de-registration with respect to gender, seniority, and number of registrations (own study)

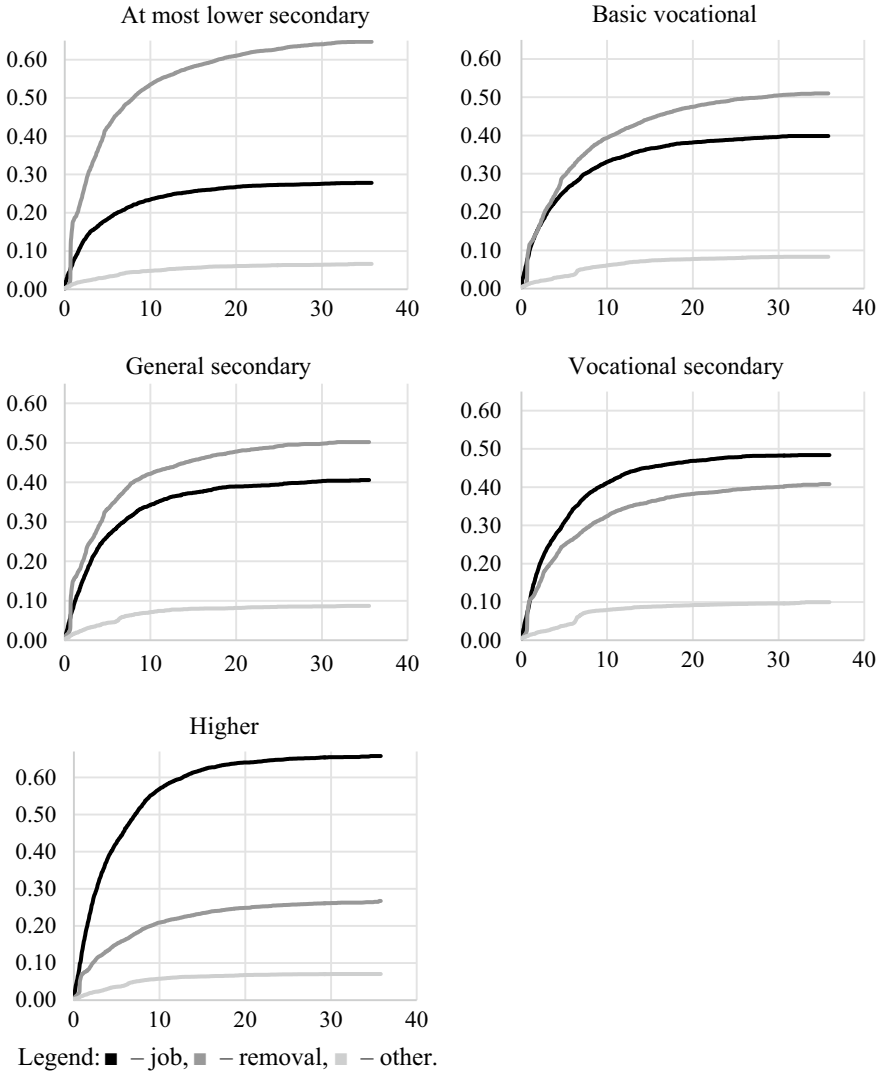


Fig. 9.3 Probability of de-registration with respect to education (own study)

group's mean) and at the age 25–34 years (by 39% more than the group's mean) took up job most intensively. On the contrary, men (by 61%), persons without professional experience (by 57%), with at most lower secondary education (by 46% more than the group's mean) and at the age 18–24 years (even by 228% more than the group's mean) were removed most intensively. The number of registrations did not differentiate the intensity of removal (lack of significance of the parameter). Intensities of de-registrations of men and women due to other causes were not significantly different. Persons with basic vocational education (by 20%), registered for the first time (by

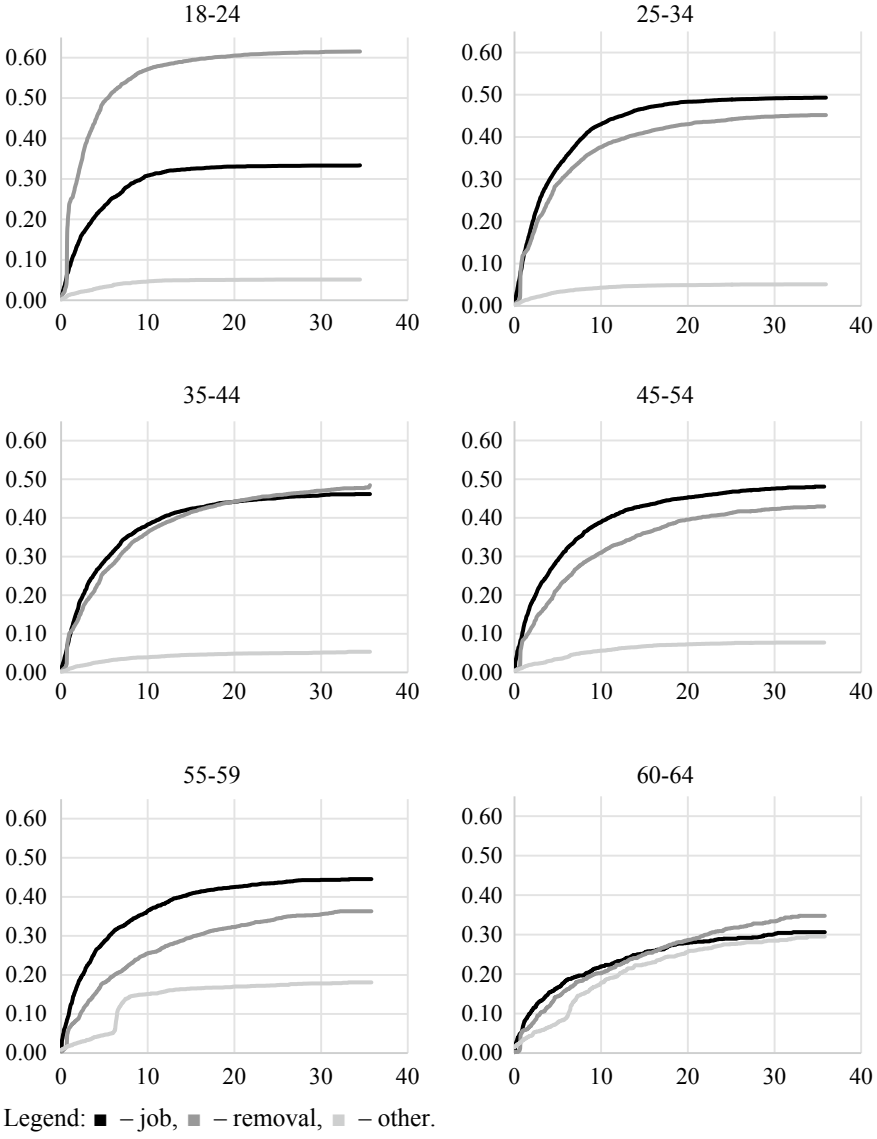


Fig. 9.4 Probability of de-registration with respect to age (own study)

Table 9.3 Estimators of the Lunn-McNeil model for variables: gender, seniority, and the number of de-registrations (own study)

Parameter	Parameter's estimator	Standard error	Wald's statistics	p	Relative hazard
Gender					
δ'_1	0.0635	0.0202	9.8638	0.0017	1.0656
δ'_2	−0.4732	0.0204	540.7017	0.0000	0.6230
δ'_3	−0.1027	0.0480	4.5763	0.0324	0.9024
$\chi^2 = 571.734, p = 0.0000$					
Seniority					
δ'_1	0.3197	0.0224	202.9225	0.0000	1.3767
δ'_2	−0.5791	0.0197	865.6541	0.0000	0.5604
δ'_3	0.1861	0.0521	12.7685	0.0004	1.2046
$\chi^2 = 1075.247, p = 0.0000$					
Number of registrations					
δ'_1	−0.1353	0.0234	33.3462	0.0000	0.8734
δ'_2	0.0461	0.0241	3.6689	0.0554	1.0472
δ'_3	−0.3555	0.0531	44.7774	0.0000	0.7008
$\chi^2 = 78.70215, p = 0.0000$					

Table 9.4 Estimators of the Lunn-McNeil model for education (own study)

Parameter	Parameter's estimator	Standard error	Wald's statistics	p	Relative hazard
Education S_1					
δ'_1	−0.4159	0.0227	337.1938	0.0000	0.6597
δ'_2	0.3797	0.0168	510.6982	0.0000	1.4618
δ'_3	−0.1885	0.0478	15.5521	0.0001	0.8282
Education S_2					
δ'_1	−0.1401	0.0206	46.4389	0.0000	0.8693
δ'_2	0.0588	0.0188	9.8226	0.0017	1.0606
δ'_3	−0.0514	0.0455	1.2769	0.2585	0.9499
Education S_3					
δ'_1	−0.0127	0.0251	0.2583	0.6113	0.9873
δ'_2	0.1640	0.0229	51.1901	0.0000	1.1782
δ'_3	0.1390	0.0547	6.4676	0.0110	1.1492
Education S_4					
δ'_1	0.1096	0.0205	28.4742	0.0000	1.1159
δ'_2	−0.1143	0.0219	27.2365	0.0000	0.8920

(continued)

Table 9.4 (continued)

Parameter	Parameter's estimator	Standard error	Wald's statistics	p	Relative hazard
δ'_3	0.1817	0.0460	15.5889	0.0001	1.1993
Education S_5					
δ'_1	0.4592	0.0173	707.2241	0.0000	1.5828
δ'_2	−0.4881	0.0239	418.6220	0.0000	0.6138
δ'_3	−0.0808	0.0480	2.8408	0.0919	0.9223
$\chi^2 = 1841.1077, p = 0.0000$					

Table 9.5 Estimators of the Lunn-McNeil model for age (own study)

Parameter	Parameter's estimator	Standard error	Wald's statistics	p	Relative hazard
Age W_1					
δ'_1	0.2120	0.0266	63.6224	0.0000	1.2362
δ'_2	0.8249	0.0216	1464.0910	0.0000	2.2816
δ'_3	−0.0484	0.0636	0.5779	0.4471	0.9528
Age W_2					
δ'_1	0.3306	0.0191	300.0242	0.0000	1.3918
δ'_2	0.2121	0.0193	120.8042	0.0000	1.2362
δ'_3	−0.3866	0.0490	62.3716	0.0000	0.6793
Age W_3					
δ'_1	0.0930	0.0223	17.4245	0.0000	1.0975
δ'_2	0.0656	0.0219	9.0154	0.0027	1.0678
δ'_3	−0.5740	0.0585	96.3522	0.0000	0.5633
Age W_4					
δ'_1	0.0717	0.0247	8.4292	0.0037	1.0744
δ'_2	−0.1109	0.0256	18.8181	0.0000	0.8950
δ'_3	−0.2574	0.0573	20.1511	0.0000	0.7731
Age W_5					
δ'_1	−0.0049	0.0294	0.0278	0.8676	0.9951
δ'_2	−0.2804	0.0319	77.4562	0.0000	0.7555
δ'_3	0.5837	0.0480	147.7115	0.0000	1.7926
Age W_6					
δ'_1	−0.7025	0.0482	212.4250	0.0000	0.4954
δ'_2	−0.7112	0.0457	242.4344	0.0000	0.4911
δ'_3	0.6827	0.0540	159.8288	0.0000	1.9793
$\chi^2 = 2315, p = 0.0000$					

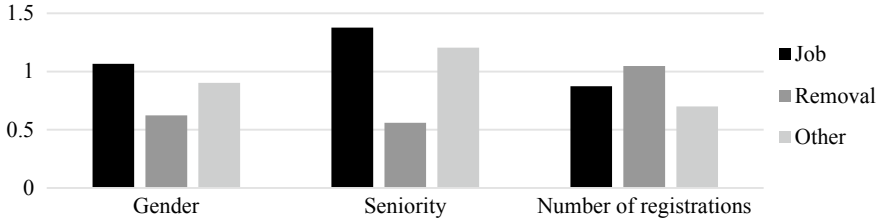


Fig. 9.5 Relative intensity of de-registration with respect to gender, seniority, and the number of de-registrations (own study)

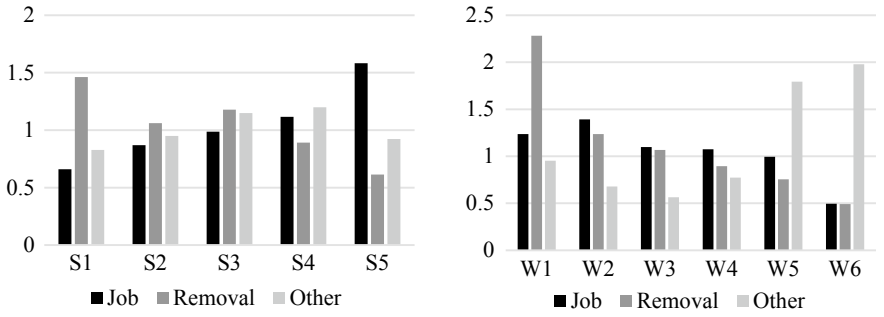


Fig. 9.6 Relative intensity of de-registration with respect to education and age (own study)

43%), with secondary vocational education (by 20%) and older than 60 years (by 98%), were de-registered due to other causes the most intensively. High relative intensity of de-registration due to other causes was also for the persons at the age of 55–59 years (by 79%).

9.5 Conclusions

Application of the survival analysis allowed to consider various types of competing risk. In the performed analysis, there were the three various forms of de-registration from the labor office: taking up job, removal due to the unemployed person's fault, and due to other causes. This last cause of de-registration had only marginal meaning, and for most analyzed features, its probability did not exceed 0.1. Only in the groups of the oldest persons, or at the age 55–59 and 60–64 years, this probability was almost 0.2 and 0.3, respectively. It was connected with transition of the unemployed persons to retirement or granting the pre-retirement benefit. Also, in case of this form of exit from unemployment, the significant leap in the value of the survival function in the seventh month since registration is visible. It is particularly high in case of women, persons with professional experience, registered for the first time, with basic and

secondary vocational education and for persons older than 55 years. Except for the two earlier mentioned causes, it was also connected with transition to a pension.

In order to lead efficient labor market policy, the identification of persons, who have problems with finding employment, is important. Among them, there can be persons discouraged by job-seeking, who eventually resign from the mediation of the labor office. Such persons were clearly predominant among men, persons without professional experience, registered subsequently, with at most secondary education, not older than 25 years. Performed analysis showed that for all unemployed persons, the probability of removal was higher than the probability of taking up job starting at 12 months since registration, or after transition into the state of long-term unemployment.

It seems that presented in the chapter, advanced methods of survival analysis are good supplement to typical statistical indicators used in the analysis of unemployment. Inclusion of the unemployment duration allows for the identification of groups of persons to whom the labor market policy should be directed.

References

1. Aalen, O., Borgan, O., Gjessing, H.: Survival and Event History Analysis. A Process Point of View. Springer, New York (2008). https://doi.org/10.1007/978-0-387-68560-1
2. Baumann, I.: Job search strategies and unemployment duration. In: The Plight of Older Workers. Life Course Research and Social Policies, vol. 5, pp. 91–107. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39754-2_5
3. Bieszk-Stolorz, B.: The impact of gender on routes for registered unemployment exit in Poland. Equilibrium. Q. J. Econ. Econ. Policy **12**(4), 733–749 (2017). https://doi.org/10.24136/eq.v12i4.38
4. Bieszk-Stolorz, B.: Stratified cox model with interactions in analysis of recurrent events. Acta Universitatis Lodziensis. Folia Oeconomica **3**(335), 207–2018 (2018). https://doi.org/10.18778/0208-6018.335.14
5. Bryant, J., Dignam, J.J.: Semiparametric models for cumulative incidence functions. Biometrics **60**(1), 182–190 (2004). https://doi.org/10.1111/j.0006-341X.2004.00149.x
6. Carling, K., Jacobson, T.: Modeling unemployment duration in a dependent competing risks framework: identification and estimation. Lifetime Data Anal. **1**, 111–122 (1995). https://doi.org/10.1007/BF00985262
7. Han, A., Hausman, J.A.: Flexible parametric estimation of duration and competing risk models. J. Appl. Econom. **5**, 1–28 (1990). https://doi.org/10.1002/jae.3950050102
8. Kalbfleisch, J.D., Prentice, R.L.: The Statistical Analysis of Failure Time Data, 2nd edn. Wiley, Hoboken, New Jersey (2002)
9. Kaplan, E.L., Meier, P.: Non-parametric estimation from incomplete observations. J. Am. Stat. Assoc. **53**, 457–481 (1958)
10. Klein, J.P., Bajorunaite, R.: Inference for competing risks. In: Balakrishnan, N., Rao, C.R. (eds.) Handbook of Statistics: Advances in Survival Analysis, vol. 23, pp. 291–312. Elsevier, Amsterdam (2004)
11. Klein, J.P., Moeschberger, M.L.: Asymptotic bias of the product limit estimator under dependent competing risks. Indian J. Prod. Reliab. Qual. Control **9**, 1–7 (1984)
12. Klein, J.P., Moeschberger, M.L.: Survival Analysis. Techniques for Censored and Truncated Data. Springer, New York (2003)
13. Kleinbaum, D., Klein, M.: Survival Analysis. A Self-Learning Text. Springer, New York (2005)

14. Landmesser, J.: Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce. Wydawnictwo SGGW, Warszawa (2013)
15. Markowicz, I.: Statystyczna analiza żywotności firm. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin (2012)
16. Marubini, E., Valsecchi, M.: Analysing Survival Data from Clinical Trials and Observational Studies. Wiley, Chichester (1995)
17. Matuszyk, A.: Zastosowanie analizy przetrwania w ocenie ryzyka kredytowego klientów indywidualnych. CeDeWu, Warszawa (2015)
18. Meyer, B.D.: Unemployment insurance and unemployment spells. Econometrica **58**(4), 757–782 (1990)
19. Osińska, M., Śliwicki, D.: Models and forecasts of outflows from unemployment in the Kujawsko-Pomorskie voivodship. Torun Bus. Rev. **16**(3), 5–19 (2017). https://doi.org/10.19197/tbr.v16i3.134
20. Pepe, M.S.: Inference for events with dependent risks in multiple endpoint studies. J. Am. Stat. Assoc. **86**(415), 770–778 (1991). https://doi.org/10.1080/01621459.1991.10475108
21. Røed, K., Zhang, T.: Unemployment duration and economic incentives—a quasi random-assignment approach. Eur. Econ. Rev. **49**, 1799–1825 (2005). https://doi.org/10.1016/j.euroecorev.2004.04.001
22. Sączewska-Piotrowska, A.: Dynamika ubóstwa w miejskich i wiejskich gospodarstwach domowych. Wiadomości Statystyczne **7**, 39–59 (2016)

Chapter 10
Could the Volatility Be a Measure of Investors' Optimism?

Sebastian Majewski

Abstract The optimism is rather the psychological term describing belief of someone in success or hopefulness and confidence in the decision-making process. After Kahneman and Tversky's "prospect theory" (Kahneman and Tversky in Econometrica 47:263–291, 1979 [9]) the perception of the role of psychology in economic activities of investors was totally changed. The works of these two authors revolutionized the academic attempt to human judgement. The using of terms "heuristic and biases" caused that errors of theoretical models became to be explained as the result of the influence of human nature (Gilovich and others in Heuristic and biases: the psychology of intuitive judgement. Cambridge University Press, UK, 2002 [5]). The trading volume on stock exchanges always confirms market trends and signals from technical analysis. It is possible to assume that during the growing trend, increase in trading volume describes a belief of investors in the bull market and during decreasing trend in bear market. Both situations could be treated as symptoms of optimism and pessimism of investors. The main goal of the article is to examine that the volatility could be an indicator of investors' optimism. So, the main hypothesis is that trading volume could be modelled by the market risk (standard deviation of rates of return) using econometric models. Such hypothesis is supported by additional—the sign of the rate of return is a significant binary variable for econometrical describing of trading volume. Different methods of estimation of the volatility parameter will be used in the research—from classical statistics to ARCH-type models. Different length of the time window of the standard deviation calculating was taken into account and the 180 trading days window was chosen for the analysis according to similar works from the past (Majewska in Przeg. Stat. 1–2, 161–170, 2000 [12]). The data used for verification of raised hypothesis was taken from the Warsaw Stock Exchange, and it concerns the group of chosen stock exchange indexes listed on the Warsaw Stock Exchange.

Keywords Behavioural finance · Financial econometrics

S. Majewski (✉)
University of Szczecin, ul. A. Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: sebastian.majewski@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_10

JEL Codes G40 · C58

10.1 Introduction

The trading volume is the one of the most important variables supporting investment decisions of individual investors on every capital market in the world. Joining trends of trading volume and prices stand the base rule signalling breakthrough in prices movements. The technical analysis charges to confirm signals flowing from moving averages, oscillators and other indicators by the changes of trading volume. It allows to assume that the changes of volume could express the state of investors' minds and emotions. It is not a new approach because some financial analysts have tried to use volume as a measure of market status [1, 8, 14]. The new in this approach is to treat indicators or results based on the trading volume as a part of behavioural finance—the explanation of the obtained results taking into account the behavioural point of view.

Generally, optimism is defined as hopefulness and confidence about the future or the success of something. It could be considered that firstly, it is an emotional state of human nature and secondly, it could drive to illusion of control. In 1979, Kahneman and Tversky, in their "Prospect Theory", indicated that investors ascribe different meanings to gains and loses—emotions after loses have two times more meaning than after gains. So, emotions should play a serious role in the process of making investment decisions. Additionally, the level of rates of return (negative or positive) will have influence on market feeling (optimism or pessimism of individual investors). It means that relations between trading volume and market volatility could have another character depending on the positive and negative rates of return [9].

The main goal of the article is to answer the question does the volatility be a measure of investors' optimism? It is necessary to raise some working hypothesis to verify usefulness of volatility in measuring of investors' mood. The hypothesis of the research assumes that the trading volume could be modelled by 180-working days volatility using ARCH-type models. The data set of the research is taken from the Warsaw Stock Exchange, and it involves quotations of five stock exchange indexes: WIG20, WIG-BUILDING, WIG-BANKS, WIG-INFO and WIG-TELECOM from their first quotation to the 31 October 2018. Indexes taken into account in the research representing both types: price and income. The chosen group of sectoral indexes is only a simple sample for verification of possibility of existing relationship between trading volume and market risk. The main reason for choosing such kind of indexes was the length of time series (number of observations describing how long is the index quoted on the market). The positive verification will be a starting point for the bigger and wider research for all indexes quoted on the Warsaw Stock Exchange.

10.2 Behavioural Background of the Research

Technical analysis assumes the possibility of the forecasting of market prices by means of analysis of data generated by the process of trading. It relies on the assumption that markets discount everything except information generated by market action. So, to forecast is needed only to have data generated by market action [16].

First of all, it should be underlined that technical analysis is one of the oldest techniques supporting investment decisions of individual investors. Its effectiveness depends on the level of the market efficiency. There is not the meaning of technical analysis if market is efficient even in weak form. But such kind of knowledge is insignificant for investors. Tversky and Kahneman also indicated weakness of technical analysis [17]. According to their work, it is possible to prove that technical analysis is an example of representativeness effect, which could be understood as the tendency of investors to predicting future uncertain events by taking a short history of data and asking what broader picture this history is representing. But the method is still used and it has a huge group of proponents. The question is why such big group of investors support technique despite scientific theories? There are some explanations why so a big number of investors still believe effectiveness of technical analysis. They are as follow:

1. Communal reinforcement—it is a social construction where an unusual belief is formed when a claim is repeatedly asserted by investors, rather than due to the existence of empirical evidence for the validity of the claim.
2. Selective thinking—is the process reliant on focusing on evidences confirming preconceived, i.e. investment strategy (to justify a belief), ignoring unfavourable evidences.
3. Confirmation bias—is a type of selective thinking, it is a tendency to notice and look for information that confirms one's existing beliefs, whilst ignoring anything that contradicts those beliefs. The literature gives two meanings of term confirmation bias: the first—to preference for confirming information and the second—a way of testing hypothesis on the basis of instances that are predicted by the hypothesis while ignoring instances beyond its scope (that is why it is called as confirmatory strategy of hypothesis testing) [11].
4. Self-deception—is the process of self-misleading tending to acceptance as true or valid fact believed to be false or invalid by ignoring evidence of the contrary position.

As it was presented above, the functioning of technical analysis drives to illusion of control, which is commonly called self-fulfilling prophecy [15]. The "illusion of control" phenomenon is defined as an expectancy of a personal success probability inappropriately higher than the objective probability [10]. The environment, chance situations (competition, choice, familiarity and involvement), skills and propensities could cause investor more overconfident and less careful in making investment decision process. Complexity of this problem underlines the fact that overconfidence as a behavioural phenomenon depends on many other factors.

The "overconfidence" is one of the most popular phenomena in the literature of behavioural finance. A term belongs to the wide range of effects called cognitive biases. There are among others [13]:

– overconfidence;
– confirmation;
– representativeness;
– anchoring effect;
– hindsight.

As the first, Barber and Odean [2] described overconfidence as a very important effect in decision-making process. They indicated that the gender, age and proficiency have a strong impact on the way of making investment decision. It could be considered that all mentioned phenomena have an influence on existing optimism on the capital market.

10.3 Methodology of the Research

The methodology of the research is based on econometric models. The first step is to determine trading volume dependent on the sign of rates of return. Dummy variables describing positive signs (DP) and negative signs (DN) were entered to econometric modelling. So firstly, the rates of return of chosen indexes were estimated using ARCH-type models and the Granger causality test, checking is the sign of rate of return that has an impact on the trading volume. After that rates of return were divided into two groups: above 0 (positive) and below 0 (negative). The second step relies on the verification of relationships between trading volume and volatility of rates of return of chosen indexes independently for positive and negative rates of return. The volatility was based on the moving averages and calculated using standard deviation in windows $n = 30$, 90 and 180 trading days. The length of time window refers to results of researches on the WSE index volatility from the article of Majewska [12]. This way time series of changes in trading volume (dVol—calculated as logarithms) and standard deviation (SD30, SD90, SD180) were obtained. Because of the White's test positive results, the ARCH-type models were used to verify the raised problem.

The rate of return of chosen indexes was explained by the following variables in all the models:

DP—positive signs of rates of return (dummy variable),
DN—negative signs of rates of return (dummy variable),
SD30—moving standard deviation for 30-day window,
SD90—moving standard deviation for 90-day window,
SD180—moving standard deviation for 180-day window.

The base equation in both steps is described by the general formula:

$$y_t = \gamma_0 + \sum_{k=1}^{n} \gamma_k X_{kt} + \varepsilon_t$$

where

y_t—logarithms of changes in trading volume of the index in period t;
X_{kt}—the value of k-th regressor in period t;
ε_t—the random component, *idd* and normally distributed $N(0, 1)$.

The ARCH(q) model is written [4]

$$h_t = \propto_0 + \sum_{i=1}^{q} \propto_i \varepsilon_{t-i}^2$$

where

h_t—conditional variance;
$\varepsilon_t | I_{t-1} \sim N(0, h_t)$
I_t—the information set available at time t;

$$\propto_0 > 0, 0 \leq \propto_i \leq 1, \varepsilon_t \sim N(0, h_t)$$

The general form of the GARCH model is [3]:

$$h_t = \propto_0 + \sum_{i=1}^{q} \propto_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $i = 1,\ldots, q, j = 1,\ldots, p$, to guarantee the non-negativity of the conditional variance.

The best fit of trading volume movement will attest to the possibility of successfully using non-economic factors in econometric modelling. The estimation procedure was maximum likelihood (L), the log of L being given by:

$$\ln L = -\frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{t=1}^{N} \ln h_t(\theta) - \frac{1}{2}\sum_{t=1}^{N} \frac{R_t^2}{h_t(\theta)},$$

where

N—the length of the series;
$h_t(\theta)$—the variance function expressed by the equation $h_t(\theta) = e^{\alpha_1 + \alpha_2 \cdot \ln \varepsilon_t}$
R_t—residuals from the regression.

The concept of Granger causality is based on prediction using econometric linear models [6, 7]. The idea assumes that the independent variable is the cause of dependent one if there is a possibility of better forecasting its value using historical data of independent variable that without them. The VAR models are used to testing that.

$$y_t = \propto_{10} + \sum_{j=1}^{k} \propto_{1j} y_{t-j} + \sum_{j=1}^{k} \beta_{1j} x_{t-j} + \varepsilon_{1t}$$

$$x_t = \propto_{20} + \sum_{j=1}^{k} \propto_{2j} x_{t-j} + \sum_{j=1}^{k} \beta_{2j} y_{t-j} + \varepsilon_{2t}$$

where

α, β—estimated parameters of VAR models;
j—time-lag.

The estimation was carried out using the GRETL programme.

10.4 Empirical Results

The research was conducted on the base of five stock exchange indexes listed on the Warsaw Stock Exchange. It was assumed that the minimum length of time series of indexes used in the research should be 5000 trading days. They are as follow: WIG20, WIG-Building, WIG-Banks, WIG-Info and WIG-Telecom. The period of time starts with the first quotation of indexes (accordingly it was: 21 February 1995 for WIG20, 21 September 1998 for WIG-Building, WIG-Banks and WIG-Telecom and 22nd September 1998 for WIG-Info) and ends at 30 October 2018. The hypothesis raised assumes that the relationship between volatility of prices and changes in trading volume exists. Therefore, the simplest idea was to build an econometric model based on these two variables. The results of estimation with using OLS classical models, FGLS model (feasible generalized least square with correction for heteroscedasticity) and ARCH-type models indicated there were no relationships between these two variables. Thus, the idea of using sub-samples time series was applied.

The first step was to verify the significance of signs (negative or positive) of rates of return in econometric modelling of trading volume. The results of estimation of ARCH-type models are presented in Table 10.1.

The results presented in Table 10.1 indicate that the sign of the rate of return has a strong impact on the changes in trading volume. In most cases (4 from 5), there is no causality between a positive sign of the rate of return and changes in trading volume in Grager's meaning. There is only one case of causality—for rates of return of index WIG20. Such situation could be a result of the existing of technical possibilities of building investment strategies based on this index. The WIG20 is an underlying asset for derivatives quoted on the Warsaw Stock Exchange, so it could be a subject of

Table 10.1 Results of the regression analysis of trading volume (dVol) of chosen stock exchange indexes in regard to rates of return's signs (DP or DN)

Index	Independent variable	Type of model	Parameter	*z*-stat	*p*-value	Granger causality[a]
WIG20	DP	ARCH(2)	0.0678222	7.9452	<0.00001	YES
WIG-Banks	DP	ARCH(1)	0.0654525	5.1278	<0.00001	NO[b]
WIG-Build	DP	ARCH(1)	0.0512677	3.3085	0.00094	NO[c]
WIG-Info	DP	ARCH(1)	0.0835452	5.8051	<0.00001	NO[d]
WIG-Tele	DP	ARCH(1)	0.0677225	4.6550	<0.00001	NO[e]

Source Own calculations
[a]VAR models are estimated for variables lagged from $t = 1$ to $t = 5$
[b]The positive sign is a cause of dVol only for $t = 1$ lag (p-value equal 0.1)
[c]The positive sign is a cause of dVol only for $t = 1$ lag (p-value equal 0.001)
[d]The positive sign is a cause of dVol only for $t = 1$ lag (p-value equal 0.001)
[e]The positive sign is a cause of dVol only for $t = 1$ lag (p-value equal 0.001)

building hedging and speculative strategies by individual and institutional investors. It seems to be legitimated to make an experiment, in which econometric models will be estimated separately for two groups of rates of return (positive and negative).

In the next step, models investigating dependency of changes in trading volume from moving standard deviations were estimated. In the pre-phase, the length of time series used for moving standard deviations was determined on the base of the econometric modelling. The chosen time window equals 180 trading days. The results of the estimation for the best models are presented in Table 10.2.

The results presented in Table 10.2 indicate that the relationships between changes in trading volume and volatility exist. Regression parameters have positive sign for the group of rates of return greater than 0 and negative for rates of return lesser than 0. The proper time window for estimation of standard deviation was determined on the base of likelihood logarithm—in both cases ("+" and "−") are the highest for 180 trading days.

Table 10.2 Results of econometric modelling for WIG20

Index/group	Independent variable	Type of model	Parameter	*z*-stat	p-value	Likelihood logarithm
WIG20/(+)	SD30	ARCH(1)	5.09088	3.1540	0.00161	−1234.838
	SD90	ARCH(1)	5.63147	3.2451	0.00117	−1234.563
	SD180	**ARCH(1)**	**6.68186**	**3.8041**	**0.00014**	**−1232.627**
WIG20/(−)	SD90	ARCH(2)	−2.88626	−1.9600	0.05000	−1077.213
	SD180	**ARCH(1)**	**−3.75376**	**−2.3531**	**0.01862**	**−1076.365**

Source Own calculations
Bolded values mean that the probability of making a type I error is small (p-value)

Table 10.3 Results of econometric modelling for chosen sectoral sub-indexes for 180-days standard deviation

Index/group	Independent variable	Type of model	Parameter	z-stat	p-value
There were no statistical significant results for the index WIG-Banks					
There were no statistical significant results for the index WIG-Building for the positive group					
WIG-Build/(−)	SD180	ARCH(2)	−9.07012	−2.4520	0.01421
WIG-Info/(+)	SD180	ARCH(1)	5.20718	2.2282	0.02587
WIG-Info/(−)	SD180	ARCH(1)	−6.10447	−2.2747	0.02292
There were no statistical significant results for the index WIG-Telecom					

Source Own calculations

Another interesting fact is that absolute values of regression parameters are almost two times higher in the case of the positive group than in the negative. Stronger fluctuations of volume in the group of rates of return greater than 0 could testify the volatility could be the measure of optimism of investors.

Next, the analogical analysis was repeated for sub-sectoral indexes. The results of estimation are presented in Table 10.3.

First, statistically significant results are obtained only for three of eight cases. In the most cases, moving standard deviation could not be treated as the regressor in econometric models. Therefore, the classical OLS models were also estimated during the research. Unfortunately, that models are also insignificant. Finally, the results for WIG-Info show that 18-days moving standard deviation could explain changes in trading volume, and in both cases ("+" and "−") the force of influence is similar. The highest absolute value of regression parameter in both Tables (10.2 and 10.3) was obtained for WIG-Building (−). Such situation could be caused by the speculative bubble, which takes place in building sector on the Warsaw Stock Exchange in 2007 and 2008.

10.5 Conclusions

The answer to the question contained in the title of the paper was the main goal of this research. The idea of the research assumes that emotions of investors are reflected in the trading volume and intensity of the trading could be observed in volatility of prices. Therefore, the existing of such relationship could bring an argument to positive verification of the hypothesis. The research gave three general conclusions.

First, that the including of dummy variables pretending positive and negative rates of return to econometric modelling of changes in trading volume causes improvement of results. It means that signs of rates of return have a statistical significant impact on the changes in trading volume. Obtained results indicate that positive rates of return generally caused increase in changes of trading volume. Such a situation could

testify that positive changes on the market build investors optimism. It is necessary to underline that changes in trading volume are results of decisions of investors observing changes in assets prices.

Second, the volatility of prices has a strong impact on the changes in trading volume of index WIG20. These changes are almost two times greater in case of positive rates of return than negative. Such a situation could testify the volatility could describe the investors' optimism. Therefore, for WIG20, the answer on the question raised is yes, the volatility could be a measure of investors' optimism.

Third, unfortunately, results for the other indexes are not unambiguous. The relationship between changes in trading volume and moving standard deviation for WIG-Info exists, but regression parameters do not indicate on greater force in case of positive rates of return than negative. Therefore, for WIG-Info, the answer is maybe or even no because of lack of relationship for the rest of analyzed indexes.

Fourth, the standard deviation calculated on the base of 180 trading days windows gives the best results in econometric modelling for trading volume of index WIG20.

The article is a starting point for the researches of relationships between rates of return, volatility and trading volumes. The further researches will take into account all kind of indexes from WSE and some important stocks from index WIG20 to confirm possibility of using trading such kind of optimism measure.

Concluding, it is possible that the volatility could indicate on optimism of investors, especially when the analyzed asset is a subject of huge number of transactions with high liquidity as the index WIG20.

Reference

1. Arms Jr., R.: Profits in Volume: Equivolume Charting. Marketplace Books, USA (1999)
2. Barber, B.M., Odean, T.: Boys will be boys: gender, overconfidence, and common stock investment. Quart. J. Econ. 261–292 (2001)
3. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. J. Econom. **31**, 307–327 (1986)
4. Engle, R.F.: Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica **50**, 987–1008 (1982)
5. Gilovich, T., Griffin, D., Kahneman, D. (eds.): Heuristic and Biases: The Psychology of Intuitive Judgement. Cambridge University Press, UK (2002)
6. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**(3), 424–438 (1969)
7. Granger, C.W.J.: Testing for causality: a personal viewpoint. J. Econ. Dynam. Control **2**, 329–352 (1980)
8. Gursoy, G., Yuksel, As., Yuksel, Ay.: Trading volume and stock market volatility: evidence from emerging stock markets. Investment Manage. Financ. Innovations **5**(4–1) (2008)
9. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica **47**(2), 263–291 (1979)
10. Langer, E.J.: The illusion of control. J. Pers. Soc. Psychol. **32**(2), 311–328 (1975)
11. Lewicka, M.: Confirmation bias. Cognitive error or adaptive strategy of action control? In: Kofta, M., Weary, G., Sedek, G. (eds.) Personal Control in Action, pp. 233–258. Plenum Press, New York (1998)

12. Majewska, A.: Szacowanie zmienności na rynku opcji. Przeg. Stat. 1–2, 161–170(2000) (in Polish)
13. Majewski, S.: The Influence of Behavioural Factors on the Market Stocks Valuation. Quantitative Approach. WNUS, Szczecin (2012). (in Polish)
14. Mestel, R., Gurgul, H., Majdosz, P.: The empirical relationship between stock returns, return volatility and trading volume on Austrian Stock Market. https://www.researchgate.net (2003). Accessed 12.02.2019
15. Murphy, J.J.: Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance, New York, USA (1999)
16. Sewell, M.: Technical Analysis. University College London, London, UK. http://www.technicalanalysis.org.uk/technical-analysis.pdf (2007). Accessed 13.02.2019
17. Tversky, A., Kahneman, D.: Judgment under uncertainty: heuristics and biases. Science **185**(4157), 1124–1131 (1974). New Series

Chapter 11
Factor Analysis of the Green Management in the Context of Poland Regions Sustainable Development

Adam Sulich and Malgorzata Rutkowska

Abstract The green management is one of many propositions of the actions required for adopting the new more sustainable and durable models of the economy which lead to the green economy. This kind of management can be observed on the different levels of economy, national and regional. Its assessment can be based on secondary data collected from the companies and presented by the national statistical offices. The chapter presents the idea of green management in the regional development context. This chapter aims to examine the role of green management in the context of sustainable development and use quantitative tools to solve contemporary problems of green management from the perspective of modern economics. In the first part of this chapter, green management as an element of the green economy is discussed. Next, to the quantitative, a qualitative—descriptive method was used. Therefore, this paper attempts to provide a basic overview of the concept of green management at the introductory level, which contributes to green job creation.

Keywords Green economy · Green management · Poland regions · Green labour market · Factor analysis of the development

11.1 Introduction

The green management is one of the innovative tools for sharing the vision of modern economics. In the literature, the idea of a sustainable economy has been discussed for decades. In the last few years, the concept of sustainable development has been sufficiently described, documented and developed, but this cannot be said about the idea of a green economy. The idea of "green economy" appeared relatively recently and immediately in documents of high rank of the European Union and the world [5].

A. Sulich (✉)
Wroclaw University of Economics, Wrocław, Poland
e-mail: Adam.Sulich@ue.wroc.pl

M. Rutkowska
Wroclaw University of Science and Technology, Wrocław, Poland
e-mail: Malgorzata.Rutkowska@pwr.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_11

What is more, the latest scientific studies and direct experiences of environmental devastation and climate change are making it clear, that the economic model needs to change. Therefore, the green economy was the central theme of the United Nations conference on sustainable development of Rio + 20 in June 2012 in Rio de Janeiro [20].

The use of modern econometric methods to measure green management meets the expectations of modern economics [6, 14]. On the one hand, it contributes to the broadening of the spectrum of analysed issues and ways of operationalisation of sustainable development and to increase the accuracy of calculations and formulation of conclusions from the other. In particular, it refers from the perspective of modern economics to such vital issues as sustainable regional development and green management. The research covers the chosen indicators data for all voivodeships of Poland related to the three main areas of sustainable development which is a subject of green management.

The research aimed to examine the role of green management in the context of sustainable development. To reach this goal, a factor analysis method was used.

11.2 Green Management as an Element of the Green Economy

In the literature, there are many publications and papers published in which proposed the definitions of the green economy, even though the term "green economy" is not defined consistently because it is still a nascent concept. This chapter presents the most important (according to the authors) of them. During the conference in Rio concluded that, the so-called green economy is just another facet of the current financial phase of capitalism, which also makes use of old and new mechanisms, such as the deepening of the public–private debt, the hyperstimulation of consumption, the concentration of ownership of new technologies, carbon and biodiversity markets, land grabbing, increased foreign ownership of land, and public–private partnerships, among others" [19].

The green economy is an economy that results in improved human well-being and social equity, while significantly reducing environmental risks and ecological scarcities [16]. The most widely used and authoritative green economy definition comes from UNEP [17]. According to this definition, "green economy is one that results in improved human well-being and social equity, while significantly reducing environmental risks and ecological scarcities [18].

Based on the principles proposed, a green economy can be seen to have the following characteristics: environmental, social, economic and others (Fig. 11.1). Moreover, grounded in the Rio + 20 outcome, the United Nations system views the green economy as a means for achieving sustainable development and eradicating poverty [19, 20].

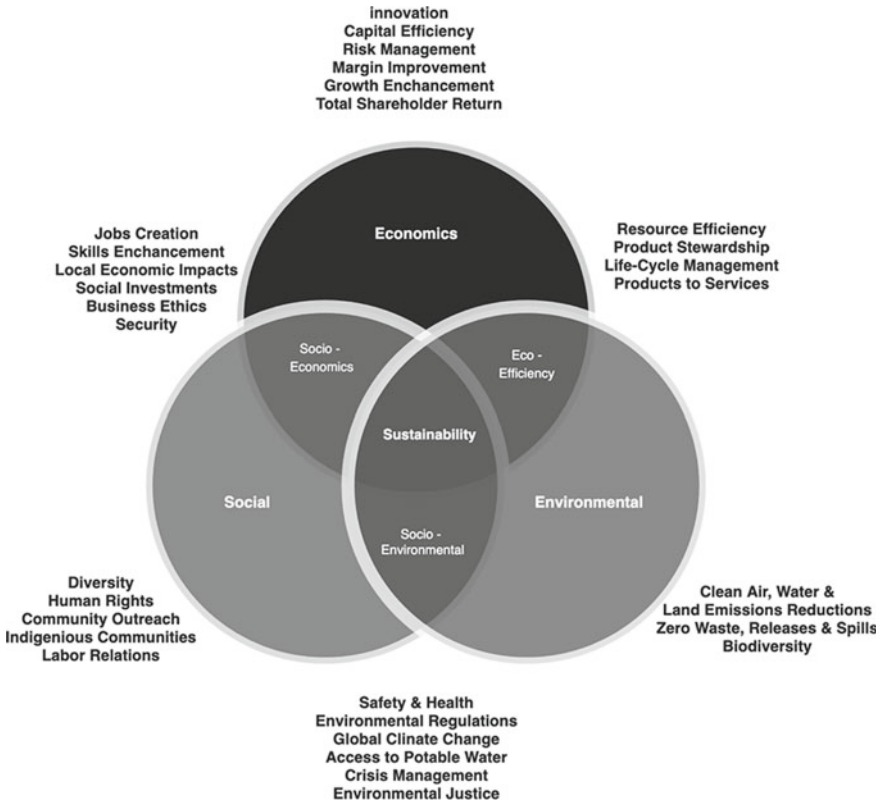


Fig. 11.1 Sustainability dimensions and related issues. *Source* Authors' own elaboration and Khambayat and Majumdar [4]

It is essential that management systems include the active policy to support the useful change of business models, regulators and control mechanisms that will be of increasing importance in the years to come. The use of necessary or voluntary tools will be fundamental for real mainstreaming of environmental governance as a cross-cutting action in all sectors of interest. One of the ways to effectively implement pro-environmental solutions is to implement the so-called green practices, which will contribute to limiting the negative environmental impact of the organisation itself as well as its processes and products. Therefore, it was recognised that green management is an element of the green economy. Green management is a relatively new term. It is assumed that green management is a situation in which a company tries to minimise processes that harm the environment. Environmental protection issues are a company's priority [9]. One of the most recent studies [8] on green management defined the term green management as a practice that produces environmentally friendly products and minimises the impact on the environment through green production, green research and development, and green marketing [10]. Green

management aim is to support ecological, pro-ecological, environmental management provided by companies, and to support environmental management by the government. This kind of management is visible mainly in green jobs and green processes and products.

Green management is an initiative aiming at continuously improving the foundation of environmental management, such as the development of personnel responsible for environmental activities, environmental management systems, and environmental communication as well as conservation of biodiversity. Loknath and Abdul Azeem [8] point that "green management is the organisation-wide process of applying innovation to achieve sustainability, waste reduction, social responsibility, and competitive advantage via continuous learning and development and by embracing environmental goals and strategies that are fully integrated with the goals and strategies of the organisation". Green management is an element of the green economy concept because of its contribution to the labour market, where this type of management creates green jobs.

11.3 Research Methods

Next, to the descriptive analysis of the idea of the green management and the green economy presented in previous subchapters, the quantitative method of the factor analysis was used to construct a model that concerns green management in the sustainable regional development. Factor analysis is a set of algorithms, utilising which a large number of studied variables describing a given object (a multidimensional random variable) is replaced with a new set of independent variables, including factors. The new set is usually less numerous, which means that the purpose of factor analysis is to reduce the number of variables. In the factor analysis conducted in this research, the subject of analyses was indicators of green management referring to three spheres of sustainable development: economy, society and the environment [1, 3, 12, 13]. These indicators refer mostly to the expenditures related to environment protection or represent the economic and social welfare. Factor analysis is widely used to support multidimensional analysis and taxonometric methods [12, 14, 15].

Factor analysis was carried out for 2016 for all voivodeships of Poland, based on secondary data obtained from the Central Statistical Office in Poland [2]. The studied statistical features and their symbols are presented in Table 11.1.

The indicators presented in Table 11.1 are mostly stimulants, except P11—registered unemployment rate. In the next step, in order to perform the factor analysis, the high indicator values (except for P3, P6 and P13) were recalculated per one inhabitant of the voivodeship to make them comparable. Only such data were subjected to factor analysis in the Statistica program. Then data were normalised. The main components method was selected for the isolation of factors, and the correlation of the analysed indicators was calculated (Table 11.2). In the first stage of the analysis, correlations between variables that are primary variables were calculated. Two primary variables strongly correlated with each other are reduced to one factor.

Table 11.1 Set of variables

Sphere	Symbol	Indicator characteristics
Environmental (E)	P1	Underground water exploitation resources ($hm^3/100\ km^2$)
	P2	Expenditure on the protection of atmospheric air and climate (thousands of PLN)
	P3	Fees and receipts for the environmental protection and water management fund—total revenues per capita (PLN)
	P4	Legally protected areas (ha)
	P5	Expenditure on fixed assets for water management (thousands PLN)
Economic (G)	P6	Gross domestic product per capita (PLN)
	P7	The gross value of fixed assets per capita (PLN)
	P8	Investment outlays in enterprises per capita of working age (PLN)
	P9	Capital expenditures on total fixed assets (PLN)
	P10	Gross operating surplus (billions of PLN)
Social (S)	P11	Registered unemployment rate (%)
	P12	Students of postgraduate studies and participants of doctoral studies (person/100 thousand citizens)
	P13	Average monthly health expenditure (PLN/person)
	P14	Green jobs (sectoral approach) (number of employees in thousand)

Source The author's study based on Główny Urząd Statystyczny [2]

In the general case, new variables are created for several primary variables as their linear combinations. Such linear combinations of correlated variables are the main components (main factors). For new factors, a new coordinate system is created, which is created by rotating the old coordinate system so that the highest possible variability of results along the new axis. When a new coordinate system is applied to the data, a different line can be defined that maximises the remaining variation. In this way, new main components are separated in such a way that further factors explain an ever-smaller percentage of variance.

11.4 Results

According to the adopted method of factor analysis, the indicator P6—gross domestic product per capita (PLN)—is highly correlated with the most significant number of other indicators: P4, P8, P9, P10, P11, P13 (Table 11.2) and represents the impact of variables on development of each region of Poland. The variances distinguished by factors explaining several variations of primary variables are called own values. Besides, calculations were carried out by introducing the maximum number of factors

Table 11.2 Correlation matrix for 14 variables used in the factor analysis

Factor	Correlation (Poland_2016.sta) No data (ND) deleted by cases N=16													
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14
P1	1	−0.21	0.25	−0.13	−0.39	0.21	−0.36	0.11	0.19	0.25	−0.16	−0.3	0.16	−0.07
P2	−0.21	1	0.42	−0.48	0.23	0.19	0.03	0.12	0.12	0.14	−0.3	0.51	0.38	−0.36
P3	0.25	0.42	1	−0.29	−0.08	−0.02	−0.04	−0.26	−0.25	0.02	−0.25	0.19	0.4	−0.03
P4	−0.13	−0.48	−0.29	1	−0.46	−0.55	0.55	−0.4	−0.38	−0.51	0.73	0.1	−0.48	0.11
P5	−0.39	0.23	−0.08	−0.46	1	0.34	−0.11	0.04	0.02	0.28	−0.25	−0.06	0.32	−0.28
P6	0.21	0.19	−0.02	−0.55	0.34	1	−0.43	0.74	0.8	0.99	−0.6	−0.01	0.68	−0.14
P7	−0.36	0.03	−0.04	0.55	−0.11	−0.43	1	−0.54	−0.48	−0.4	0.83	0.45	−0.16	−0.23
P8	0.11	0.12	−0.26	−0.4	0.04	0.74	−0.54	1	0.97	0.71	−0.56	0	0.44	0.27
P9	0.19	0.12	−0.25	−0.38	0.02	0.8	−0.48	0.97	1	0.78	−0.53	0.04	0.54	0.2
P10	0.25	0.14	0.02	−0.51	0.28	0.99	−0.4	0.71	0.78	1	−0.57	0.01	0.68	−0.12
P11	−0.16	−0.3	−0.25	0.73	−0.25	−0.6	0.83	−0.56	−0.53	−0.57	1	0.21	−0.38	0.01
P12	−0.3	0.51	0.19	0.1	−0.06	−0.01	0.45	0	0.04	0.01	0.21	1	0.34	0
P13	0.16	0.38	0.4	−0.48	0.32	0.68	−0.16	0.44	0.54	0.68	−0.38	0.34	1	0.08
P14	−0.07	−0.36	−0.03	0.11	−0.28	−0.14	−0.23	0.27	0.2	−0.12	0.01	0	0.08	1

Source The author's study

Table 11.3 Own values for all variables

Indicators	Own values (Poland_2016.sta) Isolated: main component			
	Own values	% whole	Cumulated	Cumulated
1	5.396976	38.54983	5.39698	38.5498
2	2.368947	16.92105	7.76592	55.4709
3	1.730185	12.35847	9.49611	67.8293
4	1.529098	10.92213	11.02521	78.7515
5	1.159405	8.28147	12.18461	87.0329
6	0.809870	5.78478	12.99448	92.8177
7	0.340684	2.43346	13.33517	95.2512
8	0.270400	1.93143	13.60557	97.1826
9	0.196317	1.40227	13.80188	98.5849
10	0.106252	0.75894	13.90813	99.3438
11	0.069922	0.49944	13.97806	99.8433
12	0.021061	0.15044	13.99912	99.9937
13	0.000865	0.00618	13.99998	99.9999
14	0.000018	0.00013	14.00000	100.0000

Source The author's study

and at the same time introducing an own minimum value (Table 11.3). Having a measure of the amount of variance that is distinguished by each subsequent factor, it is determined how many factors to leave. The Kaiser criterion is most commonly used when making decisions. It shows that it should be left only those factors that have eigenvalues greater than 1. Then these values were presented as a scree test plot (Fig. 11.2).

The scree test can also help to make a decision. It is a linear graph that presents the eigenvalues of separate factors. The chart looks for a place from which a gentle fall in own values occur to the right (Fig. 11.2).

The analysis of Fig. 11.3 and Table 11.3 was performed in the Statistica program, which revealed the three main factors—groups of examined indicators (presented previously in Table 11.1).

Taking the above into account, the method of the most significant likelihood factors was chosen in the Statistica program, because this method gives more excellent correlations of variables with factors. The obtained calculation results are presented in Table 11.4. There are three main factors which explain the described phenomena of regional development.

The values in Table 11.4 are different from those for 14 factors (Table 11.5). The reason for this is to leave the default method of extracting factors. The Statistica program displays three eigenvalues >1. Table 11.5 presents the results of calculations of factor loads greater than 0.55 in red. In Table 11.5, coloured part of the columns

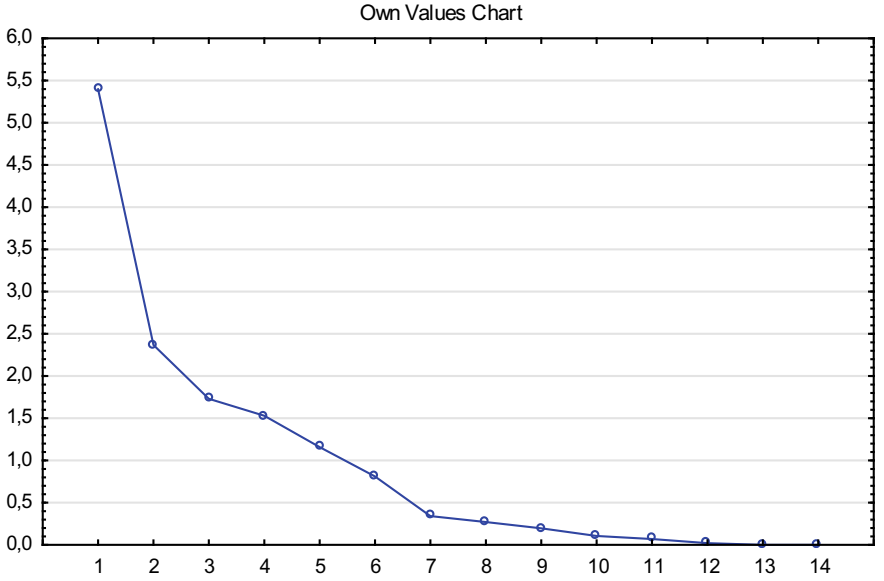


Fig. 11.2 Obtained own values in a scree test. *Source* The author's study

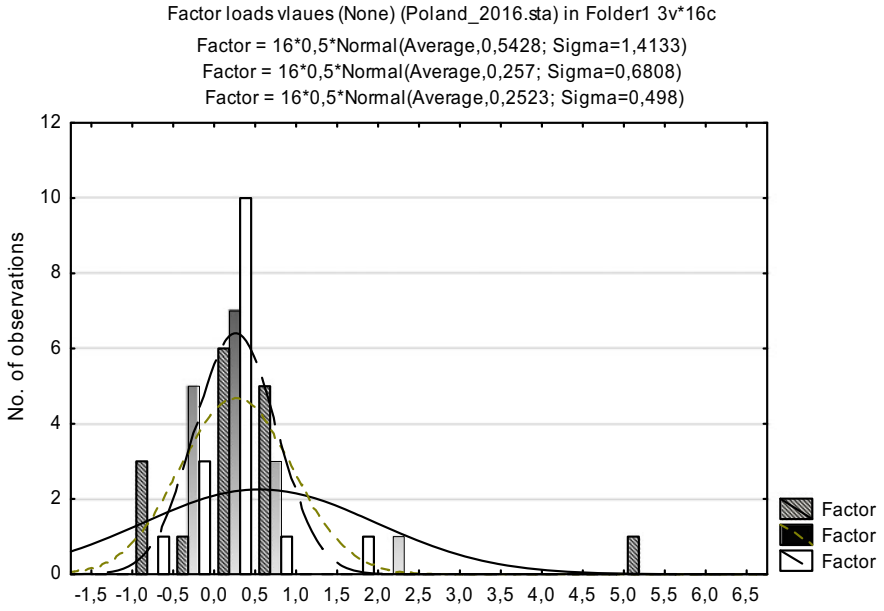


Fig. 11.3 Histogram multiple indicators. *Source* The author's study

Table 11.4 Own values obtained in the method of the highest likelihood factors

No.	Results (Poland_2016.sta); extracted: highest value factors			
	Value	% whole	Accumulated value	Cumulated
1	4.712870	33.66335	4.712870	33.66335
2	1.450720	10.36228	6.163589	44.02564
3	1.662077	11.87198	7.825666	55.89761

Source The author's study

Table 11.5 Factor loads

Variable	Factors loads (Poland_2016.sta) Isolated: factors of highest reliability (italic values are >0.550000)		
	Factor 1	Factor 2	Factor 3
P1	−0.208099	0.019499	0.098214
P2	−0.187337	0.041672	0.210609
P3	0.021330	0.364173	0.385046
P4	0.545922	−0.024944	−0.548772
P5	−0.333649	0.364298	0.183917
P6	*−0.989744*	0.022606	−0.000704
P7	0.439956	0.331333	*−0.685911*
P8	*−0.754489*	*−0.642673*	−0.009125
P9	*−0.813589*	*−0.557839*	−0.081122
P10	*−0.992195*	0.061291	−0.032757
P11	*0.602416*	0.172778	*−0.730152*
P12	0.013315	−0.008753	−0.335516
P13	*−0.676203*	0.085645	−0.027375
P14	0.124674	*−0.554183*	−0.013549
Explained variance	4.712870	1.450720	1.662077
Share	0.336634	0.103623	0.118720

Source The author's own study

is correlations between variables and factor loads and is called correlations [11] between primary variables and factors.

Table 11.5 shows that the first factor is strongly negatively correlated positively with variables P6, P8, P9, P11, P13. The second factor is correlated negatively with the variables P8, P9, P14, and the third factor is negatively correlated with the variables P7 and P11. Large loads distinguish the first factor with economic indicators (calculated per capita, unless the Central Statistical Office in Poland measurement assumes this):

P6—gross domestic product per capita (PLN),
P8—capital expenditures in enterprises per capita (PLN),
P9—capital expenditures for total fixed assets (PLN),

P11—registered unemployment rate (%),
P13—average monthly health expenditure (PLN/person).

A positive value of coefficient at P11 means that the value of factor 1 increases if the unemployment rate decreases, but decreases when it grows: gross domestic product, investment outlays and health expenditure.

High negative charges with indicators distinguish the second factor:

P8—capital expenditures in enterprises per capita (PLN),
P9—capital expenditures for total fixed assets (PLN),
P14—number of green jobs (sectoral approach) (number of employees in thousands).

The value of factor 2 decreases if investment outlays and the number of green jobs increase.

The third factor is correlated negatively with variables:

P7—the gross value of fixed assets per capita (PLN),
P11—registered unemployment rate (%).

This means that factor 3 decreases when the unemployment rate and the value of fixed assets per capita increase. In connection with the above:

- the first factor can be defined as green management,
- the second factor is anthropopressure,
- the third factor is social development.

In the above correlation analysis, it was shown that social development is connected with the labour market by the P11 index (negatively correlated) because in both cases these factors increase with the decrease of the registered unemployment rate. The next step in the factor analysis was to calculate the correlation obtained without rotation continuing the method of the highest credibility factors.

Voivodeship rankings were obtained due to: green management (factor 1); anthropopressure (factor 2) and social development (factor 3). Table 11.6 shows that:

(a) Taking into account the first factor concerning the labour market, the Mazowieckie voivodeship is the most developed one, and there is the highest impact of green management on the regional development.
(b) Taking into account the following provinces took the second factor, anthropopressure, the top positions: Wielkopolskie, Śląskie and Dolnośląskie.
(c) Taking into account the third factor, social development, the Śląskie Voivodeship is the most developed.

The next step in the analysis was to assign rank values by Statistica and add up these ranks for each province.

Table 11.7 shows that voivodeships: Kujawsko-Pomorskie, Śląskie and Dolnośląskie occupy the highest positions among voivodeships in terms of the rank of the analysed factors.

These are voivodeships for socio-economic reasons, developed most with simultaneous high pressure on the natural environment. Factor 1 was recognised as green

Table 11.6 Factors for individual provinces are sorted in ascending order

Voivodeship	Factor 1	Voivodeship	Factor 2	Voivodeship	Factor 3
Mazowieckie	−3.35059	Wielkopolskie	1.09075	Śląskie	1.52473
Podlaskie	−2.38666	Śląskie	1.02408	Kujawsko-Pomorskie	0.90928
Zachodniopomorskie	−1.89653	Dolnośląskie	0.99530	Zachodniopomorskie	0.78831
Śląskie	−0.98242	Lubuskie	0.88027	Łódzkie	0.71505
Łódzkie	−0.96086	Zachodniopomorskie	0.69472	Małopolskie	0.58935
Wielkopolskie	−0.50148	Opolskie	0.53339	Podkarpackie	0.54664
Opolskie	−0.48758	Łódzkie	0.40789	Dolnośląskie	0.54660
Podkarpackie	−0.44685	Kujawsko-Pomorskie	0.39953	Lubelskie	0.27542
Dolnośląskie	−0.34762	Świętokrzyskie	0.11717	Wielkopolskie	0.19930
Małopolskie	0.28580	Podlaskie	0.07774	Świętokrzyskie	0.07659
Pomorskie	0.40821	Warmińsko-Mazurskie	0.01239	Pomorskie	−0.22125
Warmińsko-Mazurskie	0.50920	Pomorskie	−0.81304	Warmińsko-Mazurskie	−0.71878
Świętokrzyskie	2.19526	Lubelskie	−0.90615	Opolskie	−0.75315
Kujawsko-Pomorskie	2.41998	Mazowieckie	−0.93062	Mazowieckie	−1.00043
Lubelskie	2.75844	Podkarpackie	−1.17336	Podlaskie	−1.41781
Lubuskie	2.78369	Małopolskie	−2.41005	Lubuskie	−2.05984

Source The author's own study

Table 11.7 Rank according to the value of factors for individual voivodeships

Voivodeship	Factor 1	Rank	Factor 2	Rank	Factor 3	Rank	Sum of ranks
Kujawsko-Pomorskie	2.419980	14	0.399530	9	0.909279	15	38.0
Śląskie	−0.982420	4	1.024077	15	1.524730	16	35.0
Dolnośląskie	−0.347618	9	0.995296	14	0.546598	10	33.0
Lubuskie	2.783687	16	0.880266	13	−2.059845	1	30.0
Wielkopolskie	−0.501482	6	1.090755	16	0.199301	8	30.0
Zachodniopomorskie	−1.896528	3	0.694718	12	0.788306	14	29.0
Lubelskie	2.758438	15	−0.906151	4	0.275420	9	28.0
Łódzkie	−0.960857	5	0.407887	10	0.715055	13	28.0
Świętokrzyskie	2.195264	13	0.117169	8	0.076594	7	28.0
Małopolskie	0.285804	10	−2.410053	1	0.589345	12	23.0
Warmińsko-Mazurskie	0.509195	12	0.012390	6	−0.718782	5	23.0
Opolskie	−0.487575	7	0.533391	11	−0.753153	4	22.0
Pomorskie	0.408206	11	−0.813037	5	−0.221252	6	22.0
Podkarpackie	−0.446848	8	−1.173364	2	0.546638	11	21.0
Podlaskie	−2.386657	2	0.077745	7	−1.417807	2	11.0
Mazowieckie	−3.350588	1	−0.930618	3	−1.000426	3	7.0

Source The author's own study

Table 11.8 Final rank according to the value of factors for individual voivodeships

No.	Voivodeship	Module F1	No.	Voivodeship	Module F1
1	Mazowieckie	3.350588	9	Łódzkie	0.960857
2	Lubuskie	2.783687	10	Warmińsko-Mazurskie	0.509195
3	Lubelskie	2.758438	11	Wielkopolskie	0.501482
4	Kujawsko-Pomorskie	2.419980	12	Opolskie	0.487575
5	Podlaskie	2.386657	13	Podkarpackie	0.446848
6	Świętokrzyskie	2.195264	14	Pomorskie	0.408206
7	Zachodniopomorskie	1.896528	15	Dolnośląskie	0.347618
8	Śląskie	0.982420	16	Małopolskie	0.285804

Source The author's own study

management. In Table 11.8, the module value of first factor (F1) was compared in order to present a list of voivodeships (regions) in Poland which are affected by green management both negatively and positively.

The green management has a direct impact on the development of voivodeships in Poland. Table 11.8 shows that the labour market in some provinces contains a significant number of green sectors in which green jobs are created. This information was compared with the spatial distribution of eco-development developed based on calculations.

The green management is visible sometimes in the specific areas of the economy such as the labour market, and therefore also some green jobs can be an indicator of the upcoming changes [13].

11.5 Conclusions

The green economy is a way of managing aimed at reducing the consumption of natural resources, producing less waste, reducing greenhouse gas emissions and expenditure on the protection of atmospheric air and climate.

The assumptions of the green management presented in work indicate that it is the right trend of economic theory and points to the proper relations between the economy and the natural environment. This concept is the realisation in the economic practice of the theoretical assumption of the idea of sustainable socio-economic development. It provides solutions for sharing the vision of modern economics.

In economics there is a significant re-evaluation associated with the formulation of new business objectives, keeping in mind human welfare and satisfying its needs that also include environmental qualities [7]. The growing popularity of eco-friendly products and services is a manifestation of the growing impact of the green economy on economic practice.

This preliminary study in future is going to be developed to be more related to the practical aspect of the green management in business sector.

Acknowledgements The project is financed by the Ministry of Science and Higher Education in Poland under the programme "Regional Initiative of Excellence" 2019–2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN.

References

1. Demków, K., Sulich, A.: Wybrane wyzwania związane ze społeczną odpowiedzialnością dużych i średnich podmiotów gospodarczych. Mark i Rynek **11**, 42–52 (2017)
2. Główny Urząd Statystyczny (2019) Bank Danych lokalnych. https://bdl.stat.gov.pl/BDL/start. Accessed 3 Apr 2019
3. Kasztelan, A.: Zielony wzrost jako nowy kierunek rozwoju gospodarki w warunkach zagrożeń ekologicznych. Stud Ekon Pol Akad Nauk Inst Nauk Ekon **2**, 185–208 (2016)
4. Khambayat, R., Majumdar, S.: Preparing teachers of today for the learners of tomorrow. J. Eng. Sci. Manage. Edu. **2**, 9–16 (2010)
5. Kryk, B.: Zielona gospodarka—teoretyczne podstawy koncepcji i pomiar jej wdrażania w Unii Europejskiej. Bożena Ryszawska, Wrocław 2013 : [recenzja]. Stud i Pr Wydz Nauk Ekon i Zarządzania **35**, 317–321 (2014)
6. Kryk, B., Skubiak, B.: The necessity of formulating sustainable regional policy in problem areas on the example of west Pomeranian province in Poland. Transform. Bus. Econ. (2015)
7. Laloux, F.: Reinventing Organizations: A Guide to Creating Organizations Inspired by the Next Stage of Human Consciousness. Diateino (2014)
8. Loknath, Y., Abdul Azeem, B.: Green management-concept and strategies. Natl. Conf. Mark. Sustain. Dev. **10**, 688–702 (2017)
9. Olszak, C.M., Mach-Król, M.: A conceptual framework for assessing an organization's readiness to adopt big data. Sustain **10**, 3734 (2018). https://doi.org/10.3390/su10103734
10. Peng, Y.S., Lin, S.S.: Local responsiveness pressure, subsidiary resources, green management adoption and subsidiary's performance: evidence from taiwanese manufactures. J. Bus. Ethics (2008). https://doi.org/10.1007/s10551-007-9382-8
11. Rabiej, M.: Analizy statystyczne z programami Statistica i Excel. Helion, Gliwice (2018)
12. Skubiak, B., Kryk, B.: Creation of the development potential of problem areas on the example of West Pomeranian voivodeship. Pr Nauk Uniw Ekon we Wrocławiu **1**, 318–328 (2016). https://doi.org/10.15611/pn.2016.418.32
13. Sulich, A., Zema, T.: Green jobs, a new measure of public management and sustainable development. Eur. J. Environ. Sci. **8**, 69–75 (2018). https://doi.org/10.14712/23361964.2018.10
14. Tarczynska-Luniewska, M., Tarczyński, W., Tarczyński, P.: Scoringowa metoda wyznaczania sektorowego wskaźnika siły fundamentalnej na przykładzie spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie. Zesz Nauk Uniw Szczecińskiego Finans Rynk Finans Ubezpieczenia **86**, 21–31 (2017). https://doi.org/10.18276/frfu.2017.86-02
15. Tarczyńska-Łuniewska, M., Tarczyński, W., Tarczyński, P.: Scoringowa metoda wyznaczania sektorowego wskaźnika siły fundamentalnej na przykładzie spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie. Zesz Nauk Uniw Szczecińskiego Finans Rynk Finans Ubezpieczenia **86**, 21–31 (2017). https://doi.org/10.18276/frfu.2017.86-02
16. Tarczyński, W., Łuniewska, M.: Stability of selected linear ranking methods—an attempt of evaluation for the Polish stock market. In: Baier, D., Wernecke, K.D. (eds.) Innovations in Classification, Data Science, and Information Systems. Studies in Classification, Data Analysis, and Knowledge Organization, pp. 523–532. Springer, Berlin, Heidelberg (2005)

17. UNEP: Towards a green economy: pathways to sustainable development and poverty eradication. Sustain. Dev. https://www.unenvironment.org/explore-topics/green-economy (2011)
18. UNEP: Green jobs: towards decent work in a sustainable, low-carbon world (2008)
19. United Nations: The future we want I. Our common vision (2012)
20. United Nations: Indicators for sustainable development goals (2014)

Chapter 12
ESG Risk Perception in Sustainable Financial Decisions. Quantitative Methods Perspective

Magdalena Ziolo, Iwona Bak, Ria Sinha and Manipadma Datta

Abstract The role of ESG risk in both decisions of financial institutions and enterprises is systematically growing. The financial sector is particularly predisposed to the exposure of ESG risk, which is an increasingly important element taken into account in the credit risk management process. Therefore, sustainable financial decisions are those that take into account the ESG risk in the decision-making process. The paper discusses the quantitative methods used in the ESG risk analysis. The critical literature review, induction, and deduction methods were implemented to diagnose the significance of qualitative methods in ESG assessment process. Within the methods enabling the analysis on the financial market, the mathematical, statistical, and econometric methods are of particular use. The results of the study confirmed that usage of quantitative tools in the study of ESG factors is beneficial for the analysis of economic and financial conditions of entities.

Keywords Sustainability · ESG · Finance · Financial decisions · Qualitative methods

JEL G21 · G23 · G24 · C15 · C18

M. Ziolo (✉)
Faculty of Economics and Management, University of Szczecin, Szczecin, Poland
e-mail: magdalena.ziolo@usz.edu.pl

I. Bak
Faculty of Economics, West Pomeranian University of Technology Szczecin, Szczecin, Poland
e-mail: iwona.bak@zut.edu.pl

R. Sinha · M. Datta
Department of Business and Sustainability, TERI School of Advanced Studies, New Delhi, India
e-mail: rias.echo@gmail.com

M. Datta
e-mail: manipadma.datta@terisas.ac.in

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_12

12.1 Introduction

Financial institutions, when determining the criteria for assessing the risk of transactions, select entities by excluding entities that do not meet the requirements specified by them from the availability of financial services. In turn, entities fulfilling the criteria, depending on the assessment of the level of transaction risk, conditioned by the degree of implementation of the criterion, are differentiated by financial institutions in terms of service (including price, range of services, level of monitoring, legal repayment security, etc.). Criteria for assessing the risk of transactions change under the influence of economic changes. This is particularly evident in the conditions of "greening" the economy and social inclusion. These two phenomena referring to the environmental and social pillar of sustainable development strongly weigh on the necessity of extending the risk assessment criteria by financial institutions for the ESG risk (environmental, social, and corporate governance).

Since 1992, the UNEP FI has been pointing out the need for financial institutions to integrate environmental, social, and corporate governance factors (ESG factors) into the decision-making process [1]. The banks' internal regulations also show that environmental risk must be taken into account with respect to their own ethical guidelines, prestige, and reputation risk [2]. Among financial market actors, the so-called environmental, social, and governance rating agencies (ESG rating agencies or sustainability rating agencies) emerge as key actors connecting how companies are managing their contribution to sustainable development with the decisions taken by the financial markets [3].

The global financial crisis of 2007 that resulted from the credit crunch has forced financial markets and companies to rethink systemic risk exposure. The ESG risk has particular relevance to financial institutions, especially banks in relation to their role as financial intermediaries and as capital raising agents. Financial institutions are significant catalysts in promoting economic development. This role needs to include and integrate the promotion of sustainable business practices failing which, banks and other financial institutions will end up facilitating practices which have significant negative environmental and social impacts and will miss opportunities to create new products and services that capitalize on ESG issues [4]. Finansinspektionen (FI) notes that several large banks have made an effort to adapt their lending procedures to take more into account environmental and sustainability issues, for example, by following various international principles or joining in international initiatives and standards. FI highlights the importance of banks being more transparent so that their customers, investors, counterparts, and other stakeholders can form a clearer opinion of how the banks take account of environmental and sustainability issues in their lending procedures [5]. Sustainability rating agencies provide environmental, social, and corporate governance information for the elaboration of sustainability indices. These indices help to track financial performance of the leading sustainability-driven firms [6] in a capital market context. The aim of the study is to indicate the importance and role of quantitative methods in sustainable financial decisions, with particular emphasis on ESG factors impacting on these decisions. The paper is organized as

follows: Sect. 12.2 includes a review of literature related to ESG factors incorporated in risk assessment, Sect. 12.3 presents the quantitative methods used in ESG risk assessment, and Sect. 12.4 concludes.

12.2 ESG Factors Incorporated in Risk Assessment Toward Sustainable Financial Decisions

Incorporating ESG factors in financial risk assessments is gaining ground across developed economies. ESG investing market is expanding at a progressive rate in USA, Canada, and Europe where ESG investing strategies have increased from $589 billion to $945 billion in the time period 2012–2014.[1] Institutional investors are seeking higher alpha in investments.

Academic literature is indicative of the drivers and motives of such integration. Przychodzen et al. [7] investigate the motives, drivers, and characteristics of mutual fund managers to integrate ESG factors in investment decisions. The study finds that the propensity to incorporate ESG factors is positively correlated with risk aversion. The other motives for ESG integration include stakeholder pressure, mandatory compliance to regulatory framework, investor's reputation, and fiduciary duty of asset managers. Evidence is also indicative of lower volatility in stocks that incorporate ESG risk factors [8]. The different forms of ESG risks include reputational, litigation, competitive, physical, and legal.

HSBC has indicated that risk management forms one of the key drivers of ESG investing among pension and sovereign funds. Other pertinent drivers include company strategy, stakeholder pressure, regulatory requirements, and financial planning. In fact, globally 48% of issuers and 62% of investors have an ESG strategy. This includes investors from Asia, Gulf States, and the USA [9].

The most commonly used approaches by AMC's include positive screening, negative screening, best-in-class, thematic investment, index-based approaches, and shareholder engagement. Some AMC's and analytics organizations have also developed ESG risk assessment frameworks to integrate sustainable factors in portfolios. CFA in association with UNPRI has developed a comprehensive analytical framework consisting of ESG research, security valuation, and portfolio construction. This identification of key factors leads to developing an ESG matrix which can have an impact on firm. The other processes include SWOT analysis, voting, collaborative engagement, checking the red flags, developing company questionnaires, security valuation, and portfolio construction [10].

A similar approach has been adopted by MSCI and Morningstar. In MSCI, ESG ratings are developed by identifying 37 ESG issues such as carbon intensity, water intensity, pollution, and waste. ESG ratings seek to understand the significant ESG risks pertaining to core industries, companies' exposure to such risks, and its ESG management system as compared to its peers within the same industry. Morningstar

[1] https://pdfs.semanticscholar.org/3597/efbeea7542e0c649cc366fea523678ca06a5.pdf.

uses company ESG ratings from Sustainalytics and compares it to its peers by constructing a portfolio sustainability score which is calculated as: Portfolio Sustainability score = Portfolio ESG score − Portfolio Controversy score. The score reflects how well the company is managing its ESG issues based on preparedness, performance, and disclosure [11, 12].

RiskLab has built upon a framework for modeling ESG risk factors. This model incorporates different types of sustainability risks such as global warming, human rights, bribery, and corruption and performs sensitivity analysis of ESG factors on equity returns. Each of the sustainability factors is modeled as a stochastic process and the analysis is made sector specific [13].

The above sections are indicative of the different approaches undertaken by AMC's to capture, integrate, and leverage upon the ESG risks. However, ESG risks can be accounted for when the firm or business entity makes a disclosure. There exists no standard disclosure framework; however, some of the pertinent ones include Global Reporting Initiative, The European Federation of Financial Analysts Societies, and ISO 26000. Development of a comprehensive ESG risk management system is necessary to account unsystematic risks as appropriate quantification and accounting lead to increased value of investments.

12.3 Quantitative Methods Used in ESG Risk Assessment

ESG risk management systems should be based on best information, including historical data, experiences, feedback from stakeholders, forecasts, observations, expert assessments, etc. In the area of sustainable business, many organizations try to define the performance indicators, to measure the level of motivation and involvement of employees, to evaluate customer satisfaction indicators, and to assess the degree of fulfillment of social expectations. Their decisions are influenced by the quantitative and qualitative factors. Qualitative factors are not quantifiable and are tinted with subjectivity. In contrast, the quantitative factors enable an analysis based on numeric values which often connects with each other and forms mutual relationships. They can therefore be treated as objective factors.

Quantitative methods in mathematics, including econometrics, statistical methods, and financial econometrics, are used to analyze and assess economic problems and enable forecasting economic developments in the future. Skillfully used synthetic figures support the decision-making processes and allow reducing the risk associated with them. Application of financial econometrics can be divided into two parts: Classic financial econometrics, relating to the analysis of financial series (exchange rates, rates of return, and their variability and stability); financial econometrics relating to the analysis of the relationship between the prices of financial instruments and financial-economic figures [14].

The models of multiple regression widely used in socio-economic studies are also implemented in the ESG risk studies in the following form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_k x_k + \varepsilon, \tag{12.1}$$

where y—the dependent variable (the endogenous one),

$x_1, x_2, \ldots, x_k$—independent variables (the exogenous ones),

β_j—model parameters (regression coefficients) describing the effect of jth variable,

ε—a random component.

In the paper [8], we can find examples of the use of multiple regression in studies of ESG propensity depending on subjective and objective variables, obtained on the basis of surveys in five selected countries of the world.

The probit and logit models, also referred to as binomial, may constitute a useful tool in determining factors that have a significant impact on the financial and economic situation of the examined entity (e.g., company, bank, etc.) [15–19]. They are one of the most commonly used models for the description of the quality phenomena. In their basic versions, dichotomous variables are considered, to the variants of which values 1 or 0 are assigned. If Y_i means the value of 0 or 1 for the relevant variants of a quality variable, and where i changes following cases then:

$$P(Y_i = 1) = p_i \text{ and } P(Y_i = 0) = 1 - p_i \tag{12.2}$$

In probit and logit model, the dependent variables are the probabilities defined herein above. In the probit model, the probabilities correspond to the values of cumulative distribution function of the standard normal distribution, and in the logit model, the values of the cumulative distribution function correspond to the logistics distribution. Therefore, these models have the following form:

probit model:

$$p_i = \int_{-\infty}^{x_i^T \beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t \tag{12.3}$$

logit model:

$$p_i = \frac{1}{1 + \exp\left(-x_i^T \beta\right)} \tag{12.4}$$

where x_i—explanatory variable values for individual cases,

β—structural parameters.

After estimating the parameters of the models, one can determine the theoretical values of the variable Y according to the standard rule of the forecast [19]:

$$\hat{y}_i = \begin{cases} 1 \text{ when } 0.5 < \hat{p}_i \leq 1, \\ 0 \text{ when } 0 \leq \hat{p}_i \leq 0.5. \end{cases} \tag{12.5}$$

The correctness of the estimated model can be verified by using the likelihood ratio test, used to verify the null hypothesis stating that all parameters of the model, except the constant, are equal to zero. The test statistic is expressed by the formula [19]:

$$\chi^2 = 2(\ln L_{\mathrm{UR}} - \ln L_{\mathrm{R}}) \tag{12.6}$$

where L_{UR}—the value of likelihood function for the full model,

L_{R}—the value of the likelihood function for the model reduced to the constant.

The extension of the binomial models is polynomial models for qualitative variables. Polynomial variables can be ordered or unordered variables (modeled by a beam of models, several models describe the one variable). In the case of a category that can be ordered, we deal with one model and its parameters are clearly interpretable [20]. The examples of ordered variables y are, for example:

- corporate bond rating: *junk* ($y_i = 1$), *low grade* ($y_i = 2$), and *investment grade* ($y_i = 3$),
- customer credit risk: *very low* ($y_i = 1$), *low* ($y_i = 2$), *medium* ($y_i = 3$), *high* ($y_i = 4$), and *very high* ($y_i = 5$),
- the financial risk of the company: *small* ($y_i = 1$), *moderate* ($y_i = 2$), and *high* ($y_i = 3$).

In the model for ordered categories, it is assumed that there is an unobservable continuous variable y^*, such that, for example (for five categories):

1. $y_i = 1$ if $-\infty - \tau_0 \le y_i^* < \tau_1$
2. $y_i = 2$ if $\tau_1 \le y_i^* < \tau_2$
3. $y_i = 3$ if $\tau_2 \le y_i^* < \tau_3$
4. $y_i = 4$ if $\tau_3 \le y_i^* < \tau_4$
5. $y_i = 5$ if $\tau_4 \le y_i^* < \tau_5 = \infty$.

A linear model for y^* is:

$$y_i^* = \mathbf{x}_i^T \beta + \varepsilon_i, \tag{12.7}$$

where $\mathbf{x}_i^T$ is a vector (line) of explanatory variable values for the ith observation. For the ith observation (of the ith individual, the ith company, etc., $i = 1, 2, \ldots, n$) the probability of p_i of the adoption by Y category m ($m = 1, \ldots, J$) is equal to:

$$p_i = P\left(y_{i=m}/\mathbf{x}_i^T, \beta, \tau\right) = F\left(\tau_m - \mathbf{x}_i^T \beta\right) - F\left(\tau_{m-1} - \mathbf{x}_i^T \beta\right), \tag{12.8}$$

where F is the cumulative distribution function. For logistic distribution, we get an ordered logit model in which:

$$F\left(\tau_m - \mathbf{x}_i^T \beta\right) = \exp\left(\tau_m - \mathbf{x}_i^T \beta\right)/[1 + \exp\left(\tau_m - \mathbf{x}_i^T \beta\right)], \tag{12.9}$$

and $F(\tau_0) = 0$ and $F(\tau_1) = 1$. The parameters β and τ of the ordered logit model are estimated using the method of maximum likelihood. Gruszczyński [21] gives examples of the application of ordered logit models in finance for modeling: rating awarded by US banking supervision for financial holdings based on inspections (three categories) [22], credit risk rating: from 1 (minimum) to 5 (maximum) [23], the level of financial supervision using a variable (FAC Index) with the values arranged from 1 to supervision completely divided to 7 for supervision in one agency [24, 25], the financial threat level of households in the UK (three categories) [26], and the monetary decisions of the central bank (three categories: growth, no change, and decrease) [27].

There is no doubt that in practice almost every decision has a multi-dimensional character, which results from mutual interactions of many phenomena. Therefore, good and reliable information needs to be processed so that the decisions made on its basis are effective. The most important method in this regard includes quantitative methods and, in particular, multi-dimensional comparative analysis (MCA) [28].

Within MCA, the taxonomic methods are most widely applied, which consist of comparing objects by organizing sets of objects and grouping them into subsets of units similar to each other due to their characteristic properties, and then selecting representatives of the received groups of objects.

The taxonomic research consists of following stages [29–37]: specification of the subject matter and the purpose of the research; construction of a database of diagnostic features collected in the form of intensity indicators describing the subject of the study; selection of diagnostic features (the substantive method and the formal and statistical one); determining the nature of features (stimulants, destimulants); standardization of diagnostic features; determination of the synthetic measure; ordering and classification of objects due to the achieved level of the synthetic measure; and checking the quality of the classification.

The basis for the analysis is the list of m potential diagnostic measure: $X_1, X_2, \ldots, X_m$, which in the next stages of research can be used to organize and classify n objects $\Gamma = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_n\}$, for example, the member countries of the OECD, in terms of their level of sustainable development or banks in terms of their financial condition. The organization of objects consists of their ranking according to the value of the synthetic measure designated by a selected taxonomic method. In contrast, the classification is understood as the division of the analyzed set of objects into disjoint and relatively homogeneous subsets (clusters), composed of objects characterized by similarity of development, described by the set of m features $X_1, X_2, \ldots, X_m$, and $X_j = (x_{1j}, x_{2j}, \ldots, x_{nj})$, where x_{ij} $(i = 1, \ldots, n; j = 1, \ldots, m)$ is the observed value of the jth feature for the ith multi-dimensional object Γ [32, 35]. The object Γ_i $(i = 1, 2, \ldots, n)$ is treated as m-dimensional vector (or alternatively a point) from the space $\mathbb{R}^m$ whose coordinates are the corresponding values of the features: $\Gamma_i = (x_{i1}, x_{i2}, \ldots, x_{nj})$, $i = 1, 2, \ldots, n; j = 1, 2, \ldots, m$. The object described in this way can be presented in the form of a matrix with the following dimensions: $n \times m$, which is called *the observation matrix* as:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{nm} \end{bmatrix} \tag{12.10}$$

There are researched objects in the rows of this matrix, while in the columns, there are values of further diagnostic features.

The ways to organize a set of objects can be divided into linear and nonlinear ordering methods. The first of these enable to determine the hierarchy of objects due to the specified criterion. However, the use of nonlinear ordering methods does not allow to establish the hierarchy of objects, but only to indicate objects similar to the values of the features describing them [36].

Among the methods of linear ordering, one can distinguish, for example, the procedures based on a synthetic variable. Pattern and non-pattern measures are used for that purpose [30, 31]. In the non-pattern methods, the synthetic variable is the function of normalized values of input variables. However, in the pattern methods, the concept of a reference object, i.e., a model object with the desired values of input variables is used. The synthetic measure is constructed on the basis of measuring the distance between the observed object and the reference object [28, 36].

Linear ordering methods applicable in the classification of, for example, enterprises or financial institutions include: Taxonomic measure of investment attractiveness TMAI_i, generalized distance measure GDM, relative measurement of development indicator BZW, and positional taxonomic measure based on Weber median. These measures are used to describe (evaluate) objects using one aggregated quantity.

1. The synthetic development measure is determined by the following formula:

$$\text{TMAI}_i = 1 - \frac{d_i}{d_0}, \tag{12.11}$$

d_i—the distance of the ith object from the reference object determined according to the formula:

$$d_i = \sqrt{\sum_{j=1}^{m} w_i \cdot \left(z_{ij} - z_{0j}\right)^2}, \quad (i = 1, 2, \dots n), \tag{12.12}$$

d_0—the standard assuring that TMAI_i adopts the values from 0 to 1:

$$d_0 = \bar{d} + a \cdot S_d. \tag{12.13}$$

The limit value for constant a is determined by the formula:

$$a \geq \frac{d_i \max - \bar{q}}{S_d}, \tag{12.14}$$

where d_i max—this is the maximum value of d_i.

The formula using the variation coefficient of a feature is adopted as the weights w_j in the distance formula:

$$w_j = \frac{V_j}{\sum_{j=1}^{m} V_j}. \tag{12.15}$$

2. In the case of a generalized GDM distance measure, the synthetic measure is determined using the following formulas:

$$d_{ik} = \frac{1}{2} - \frac{\sum_{j=1}^{m} w_j \cdot a_{ikj} \cdot b_{kij} + \sum_{j=1}^{m} \sum_{l=1}^{n} w_j \cdot a_{ilj} \cdot b_{klj}}{2 \cdot \left[\sum_{j=1}^{m} \sum_{l=1}^{n} w_j \cdot a_{ilj}^2 \cdot \sum_{j=1}^{m} \sum_{l=1}^{n} w_j \cdot b_{klj}^2 \right]^{\frac{1}{2}}} \tag{12.16}$$

where d_{ik}—distance measure (a synthetic meter), w_j—the weight of the jth characteristic fulfilling the conditions:

$$w_j \in (0, m), \qquad \sum_{j=1}^{m} w_j = m. \tag{12.17}$$

For variables measured on the ratio scale, the substitution is used:

$$a_{ipj} = x_{ij} - x_{pj} \tag{12.18}$$

$$b_{krj} = x_{kj} - x_{rj} \tag{12.19}$$

where x_{ij} (x_{kj}, x_{lj})—the ith (kth, lth) observation of the jth variable.

3. The indicator of the relative level of development is based on the idea of the method of sums of standardized values:

$$W_i = \frac{\sum_{j=1}^{k} z_{ij}}{\sum_{j=1}^{k} \max_i (z_{ij})}, \tag{12.20}$$

$$z_{ij} = x_{ij}^{*} + \left| \min \underbrace{(x_{ij}^{*})}_{i} \right|, \tag{12.21}$$

$$x_{ij}^{*} = \frac{x_{ij} - \bar{x}_j}{S_j}, \tag{12.22}$$

where W_i—the indicator of the relative level of development,

x_{ij}—the value of jth feature for ith object,

$\bar{x}_j$, S_j—respectively, the arithmetic mean and the standard deviation of the feature number j.

4. The positional variant adopts a different standardization formula in comparison to the classical approach (TMAI_i), based on the ratio of the deviation of the feature value to the corresponding Weber median coordinate and the weighted median absolute deviation using the Weber median [37]:

$$z_{ij} = \frac{x_{ij} - \theta_{0j}}{1.4826 \cdot \widetilde{\mathrm{mad}}\left(X_j\right)}, \tag{12.23}$$

where $\theta_0 = (\theta_{01}, \theta_{02}, \ldots, \theta_{0m})$ is the Weber median, $\widetilde{\mathrm{mad}}\left(X_j\right)$ is the absolute median deviation, in which the distance from the features to the Weber vector is measured,[2] i.e., $\widetilde{\mathrm{mad}}\left(X_j\right) = \mathrm{med}_{i=1,2,\ldots,n} \left| x_{ij} - \theta_{0j} \right|$ $(j = 1, 2, \ldots m)$. The aggregate measure is calculated with the formula:

$$\mu_i = 1 - \frac{d_i}{d_-}, \tag{12.24}$$

where $d_- = \mathrm{med}(d) + 2.5\,\mathrm{mad}(d)$, where $d = (d_1, d_2, \ldots, d_n)$ is a distance vector calculated with the formula: $d_i = \mathrm{med}_{j=1,2,\ldots,m} \left| z_{ij} - \varphi_j \right|$ $i = 1, 2, \ldots, n$,

$\varphi_j = \max_{i=1,2,\ldots,n} z_{ij}$—the coordinated of the development pattern vector, which constitutes of the maximum values of the normalized features.

Another proposal to use MCA is the discriminant analysis, which on the one hand allows the classification and the division of objects (companies and enterprises) into attractive and unattractive from the point of view of the adopted criterion, and on the other hand enables to select companies to a portfolio of securities [28]. It is used primarily to assess the financial condition of enterprises, to evaluate credit applications, audit opinions, and the early warning system. The essence of the discriminant method consists in the estimation of one-dimensional linear observation functions (discriminant functions), on the basis of which the groups of objects are differentiated. The most important issues of discriminant analysis, which should be solved before estimating the parameters of the discriminant function, are: the selection of

[2]The Weber median was calculated in *R program: l1median* of *pcaPP* package.

a grouping variable and the selection of the optimal set of diagnostic features. The task of selecting the grouping variable is the initial division of objects into several groups. It can be done, for example, with the use of *k*-means method, which belongs to cluster analysis (classification). This method consists in the division of objects into a predetermined number of clusters, differing as little as possible within the framework of a given cluster and as much as possible between different clusters. The grouping variable and the original values of diagnostic variables are used to estimate the parameters of discriminant function. Discriminant functions have the following form:

$$D_{kj} = \beta_0 + \beta_1 x_{1kj} + \cdots + \beta_p x_{pkj}, \tag{12.25}$$

where p—the number of discriminatory variables,

n—the sample size,

D_{kj}—the value of the discriminant function for the kth case in the jth group,

$k = 1, \ldots, n$ and $j = 1, \ldots, g$,

x_{1kj}—the value of the ith discriminant variable for the kth case in the jth group ($i = 1, \ldots p$),

β_i—the discriminant function coefficients determined on the basis of the properties of this function.

The multi-criteria decision-making (MCDA) method is used to make balanced decisions in financial institutions and enterprises that take into account ESG risk. MCDA is a set of mathematical methods and tools that enable comparison of decision variants with the consideration of different, often conflicting criteria [38]. It is striven to achieve such an effect, which will maximize the multi-criteria function of the objective with the following form:

$$F(x) = \max\big(f_1(x), f_2(x), \ldots, f_j(x)\big), \text{ with the restriction } x \in A^{\text{dop}}, \tag{12.26}$$

where A^{dop}—a set of acceptable solutions,

$f_j(x)$—individual partial criterion functions for $j = 1, 2, \ldots J$.

The MCDA application process usually consists of several stages [39]; Choice of decision variants that will be analyzed during the decision-making procedure, selection of criteria (measures) serving as a basis for the assessment and ranking of criteria according to their importance by assigning them weights, and selection and application of the appropriate method.

There is a whole range of multi-criteria methods and categories of their division and rich literature describing their use.[3] First of all, they can be classified into two groups: the methods aimed at solving continuous problems (multi-criteria linear programming) and the discrete ones. The discrete MCDA methods, where a finite

[3] See, for example, Mendoza and Martins [40], Huang et al. [41], Buchholz et al. [42], Figueira et al. [43], and Nermend [44].

set of decision variants and a finite set of criteria on the basis of which a decision should be made exist, include[4]:

- the methods of multi-attribute usability theory (e.g., MAUT, AHP, DEMATEL, and UTA),
- the methods based on outranking relations (e.g., the ELECTRE method group, PROMETHEE methods, the TOPSIS method, ORESTE, and TACTIC). Table 12.1 presents examples of application of quantitative methods for ESG risk in the literature.

12.4 Conclusions

The present paper tries to decipher the impact of inclusion of ESG factors by financial institutions in the decision-making process and identify quantitative methods used in the ESG factors assessment. Several academicians have resorted to empirical studies both at the company and portfolio levels to assess the impacts. The impact of incorporating ESG factors on financial institutions risk assessment is analyzed using qualitative and quantitative methods. The paper provides insights on quantitative methods used in the ESG risk assessment process. Among others, MCDA methods are implemented to assess the role of ESG factors in the decision-making process of financial institutions (risk assessment) and companies (sustainable value creation). Finding relationships among: types of financial institutions; and the impact of ESG factors on financial institutions' risk assessment process on sustainable value, based on related studies are analyzed with panel data analysis, multivariate regression model. In sum, quantitative analysis, in the context of sustainable financial decisions, is used to search relationships between ESG factors incorporated by financial institutions and business models' performance (economic, social, and environmental level); selecting the factors with significant impact; finding the impact of type of financial institutions (bank/market oriented) on the companies; and finding the impact of the type of ESG risk on companies' performance (environmental, social, governance, and financial).

Acknowledgements Research results presented in this paper are an element of research project implemented by the National Science Center Poland under the grant OPUS13 no UMO-2017/25/B/HS4/02172.

[4] A detailed description of these methods can be found, for example, in the works: Figueira et al. [45], Triantaphyllou [46].

Table 12.1 Examples of application of quantitative methods for ESG risk in the literature

Author	Publication	Aim of study	Method
Frondel et al. [47]	End-of-pipe or cleaner production? An empirical comparison of environmental innovation decisions across OECD countries	Identification of factors that can increase the company's tendency to implement environmentally friendly production technologies	Polynomial logit model
Moussiopoulos et al. [48]	Environmental, social, and economic information management for the evaluation of sustainability in urban areas: a system of indicators for Thessaloniki, Greece	Developing a system of indicators to assess sustainable development of urban areas	Multi-dimensional comparative analysis
Inglehart [49]	Public support for environmental protection: objective problems and subjective values in 43 societies	Identification of variables affecting environmental protection support	A multiple regression model
Bassen et al. [50]	The influence of corporate responsibility on the cost of capital	Analysis of the relationship between corporate responsibility (CR) and finances in the aspect of ESG risk	Surveys, multivariate comparative analysis, and multiple regression models
Husted and Allen [51]	Strategic corporate social responsibility and value creation: a study of multinational enterprises in Mexico	The study of the relationship between financial results (value creation of enterprises) and variables related to corporate social responsibility	Surveys, correlation analysis, and multiple regression models

References

1. Changhong, Z., Yu, G., Jiahai, Y., Mengya, W., Daiyu, L., Yiou, Z., Jiangang, K.: ESG and corporate financial performance: empirical evidence from China's listed power generation companies. Sustainability **10**(8), 1–18 (2018)
2. Finansinspektionen: How Can the Financial Sector Contribute to Sustainable Development. https://www.fi.se/contentassets/123efb8f00f34f4cab1b0b1e17cb0bf4/finansiella_foretags_hallbarhetsarbete_eng.pdf (2016). Accessed 19 Nov 2018

3. Muñoz-Torres, M.J., Fernández-Izquierdo, M.A., Rivera-Lirio, J.M., Escrig-Olmedo, E.: Can environmental, social, and governance rating agencies favor business models that promote a more sustainable development? Corp. Soc. Responsib. Environ. Manag. (2018). https://doi.org/10.1002/csr.1695
4. Hachigian, H., McGill, S.M.: Reframing the governance challenge for sustainable investment. J. Sustain. Financ. Invest. **2**(3–4), 166–178 (2012)
5. World Wildlife Fund: Living Planet Report. https://www.worldwildlife.org/pages/living-planet-report-2014 (2014). Accessed 19 Nov 2018
6. Bravo, R., Matute, J., Pina, J.M.: CSR as a vehicle to reveal the corporate identity: a study focused on the websites of Spanish financial entities. J. Bus. Ethics **107**(2), 129–146 (2012)
7. Przychodzen, J., Gomez-Bezares, F., Przychodzen, W., Larreina, M.: ESG issues among fund managers: factors and motives. Sustainability **8**, 1078 (2016)
8. Kumar, N.C.A., Smith, C., Badis, L., Wang, N., Ambrosy, P., Tavares, R.: ESG factors and risk-adjusted performance: a new quantitative model. J. Sustain. Financ. Invest. **6**(4), 292–300 (2016). https://doi.org/10.1080/20430795.2016.1234909
9. HSBC: Sustainable Financing and ESG Investing Report. https://www.gbm.hsbc.com/insights/sustainable-financing/sustainable-financing-and-esg-investing-report (2018). Accessed 31 Dec 2018
10. FA Institute: Environmental, Social and Governance (ESG) Survey. http://www.cfainstitute.org/-/media/documents/survey/esg-survey-report-2017.ashx (2017). Accessed 31 Dec 2018
11. Morningstar: The Morningstar Sustainability Rating: Helping Investors Evaluate the Sustainability of Portfolios. https://www.morningstar.com/articles/745467/morningstar-sustainability-rating.html (2017). Accessed 31 Dec 2018
12. MSCI: MSCI ESG Ratings Methodology. https://www.msci.com/documents/10199/123a2b2b-1395-4aa2-a121-ea14de6d708a (2018). Accessed 31 Dec 2018
13. Risklab: ESG Risk Factors in a Portfolio Context. https://www.ipe.com/esg-risk-in-a-portfolio-context/34522.article (2010). Accessed 31 Dec 2018
14. Łuniewska, M.: Ekonometria finansowa. Analiza rynku kapitałowego. PWN, Warszawa (2008)
15. Wiśniewski, J.W.: Ekonometryczne badanie zjawisk jakościowych. Studium metodologiczne. Uniwersytet Mikołaja Kopernika, Toruń (1986)
16. Chow, G.C.: Ekonometria. PWN, Warszawa (1995)
17. Johnston, J.: Econometric Methods. McGraw-Hill Book Company (1991)
18. Zeliaś, A., Pawełek, B., Wanat, S.: Prognozowanie ekonomiczne. Teoria, przykłady, zadania. PWN, Warszawa (2003)
19. Gruszczyński, M. (ed.): Mikroekonometria. Modele i metody analizy danych indywidualnych, pp. 63, 75. Wolters Kluwer Polska Sp. z o.o., Warszawa (2010)
20. Gruszczyński, M.: Corporate governance and financial performance of companies in Poland. Int. Adv. Econ. Res. **12**(2), 251–259 (2006)
21. Gruszczyński, M.: Uporządkowany model logitowy: zastosowana biznesowe i finansowe. In: Tarczyński, W. (ed.) Rynek kapitałowy. Skuteczne inwestowanie. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin (2007)
22. Curry, T., Fissel, G., Hanweck, G.: Market information, bank holding company risk, and market discipline. FDIC Working Paper 2003-04 (2003). https://doi.org/10.2139/ssrn.886687
23. Berger, A.E.: Potential competitive effects of Basel II on banks in SME credit markets in the United States. J. Financ. Serv. Res. **26**(1), 5–36 (2006)
24. Masciandaro, D., Porta, A.: Single authority in financial markets supervision: lessons for UE enlargement. Paper presented at conference Financial Intermediation in the New Europe: Economics, Policies and Institutions, Mediolan (2003)
25. Freytag, A., Masciandaro, D.: Financial supervision fragmentation and central bank independence: the two sides of the same coin? University of Lecce Economics Working Paper 76(37) (2005). https://doi.org/10.2139/ssrn.837124
26. Del-Rio, A., Young, G.: The impact of unsecured debt on financial distress among British households. Bank of England Working Paper 262 (2005). https://doi.org/10.2139/ssrn.824147

27. Gascogine, J., Turner, P.: Asymmetries in Bank of England monetary policy. Appl. Econ. Lett. **11**(10), 615–618 (2004). https://doi.org/10.1080/1350485042000227296
28. Łuniewska, M., Tarczyński, W.: Metody wielowymiarowej analizy porównawczej na rynku kapitałowym, pp. 41–43, 57. PWN, Warszawa (2006)
29. Hellwig, Z.: Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. Przegl. Stat. **4**, 307–326 (1968)
30. Grabiński, T., Wydymus, S., Zeliaś, A.: Metody prognozowania rozwoju społeczno-gospodarczego. Wydawnictwo Akademii Ekonomicznej, Kraków (1982)
31. Pociecha, J., Podolec, B., Sokołowski, A., Zając, K.: Metody taksonomiczne w badaniach społeczno-ekonomicznych, p. 71. PWN, Warszawa (1998)
32. Nowak, E.: Metody Taksonomiczne w Klasyfikacji Obiektów Spo-łeczno-gospodarczych. PWE, Warszawa (1990)
33. Jajuga, K.: Statystyczna analiza wielowymiarowa. PWN, Warszawa (1993)
34. Malina, A.: Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw. Wydawnictwo AE w Krakowie, Kraków (2004)
35. Młodak, A.: Analiza taksonomiczna w statystyce regionalnej, pp. 136–137. Centrum Doradztwa i Informacji DIFIN, Warszawa (2006)
36. Panek, T.: Statystyczne metody wielowymiarowej analizy porównawczej, pp. 57–58. SGH w Warszawie, Warszawa (2009)
37. Walesiak, M.: Uogólniona miara odległości GDM w statystycznej analizie wielowymiarowej z wykorzystaniem programu R. Wydanie drugie poprawione i rozszerzone. Wydawnictwo Uniwersytetu Ekonomicznego, Wrocław (2016)
38. Vicke, P.: Multicriteria Decision-Aid. Wiley, Chichester (1992)
39. Broniewicz, B., Dziurdzikowska, E.: Metody wielokryterialne w równoważeniu procesów społeczno-gospodarczych. Pr. Nauk. Uniw. Ekon. Wrocław. **491**, 53–62 (2017). https://doi.org/10.15611/pn.2017.491.05
40. Mendoza, G.A., Martins, H.: Multi-criteria decision analysis in natural resource management: a critical review of methods and new modelling paradigms. For. Ecol. Manage. **230**(1–3), 1–22 (2006). https://doi.org/10.1016/j.foreco.2006.03.023
41. Huang, I.B., Keisler, J., Linkov, I.: Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. Sci. Total Environ. **409**(19), 3578–3594 (2011). https://doi.org/10.1016/j.scitotenv.2011.06.022
42. Buchholz, T., Rametsteiner, E., Volk, T.A., Luzadis, V.A.: Multi criteria analysis for bioenergy systems assessments. Energy Policy **37**(2), 484–495 (2009). https://doi.org/10.1016/j.enpol.2008.09.054
43. Figueira, J., Greco, S., Ehrgott, M. (eds.): Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, London (2005)
44. Nermend, K.: Metody analizy wielokryterialnej i wielowymiarowej we wspomaganiu decyzji. PWN, Warszawa (2017)
45. Figueira, J., Mousseau, V., Roy, B.: ELECTRE methods. In: Figueira, J., Greco, S., Ehrgott, M. (eds.) Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, New York (2005)
46. Triantaphyllou, E.: Multi-criteria Decision Making Methods: A Comparative Study. Kluwer Academic Publishers (2000)
47. Frondel, M., Horbach, J., Rennings, K.: End-of-pipe or cleaner production? An empirical comparison of environmental innovation decisions across OECD countries. Bus. Strategy Environ. **16**, 571–584 (2007). https://doi.org/10.1002/bse.496
48. Moussiopoulos, N., Achillas, C., Vlachokostas, C., Spyridi, D., Nikolaou, K.: Environmental, social and economic information management for the evaluation of sustainability in urban areas: a system of indicators for Thessaloniki, Greece. Cities **27**(5), 377–384 (2010). https://doi.org/10.1016/j.cities.2010.06.001
49. Inglehart, R.: Public support for environmental protection: objective problems and subjective values in 43 societies. Polit. Sci. Polit. **28**(1), 57–72 (1995). https://doi.org/10.2307/420583

50. Bassen, A., Meyer, K., Schlange, J.: The Influence of Corporate Responsibility on the Cost of Capital. http://ssrn.com/abstract=984406/ (2006). Accessed 30 Dec 2018
51. Husted, B.W., Allen, D.B.: Strategic Corporate Social Responsibility and Value Creation: A Study of Multinational Enterprises in Mexico. https://core.ac.uk/download/pdf/9426392.pdf (1996). Accessed 31 Dec 2018

Chapter 13
Purchase Decision Support with Internet of Things-Based Systems

Monika Kulisz, Jerzy Lipski and Agnieszka Bojanowska

Abstract While current studies in the field of Internet of Things (IoT) tend to focus on the technical aspects, such as programming, hardware and software, publications on the behavioural aspects of IoT remain few. Therefore, the primary objective of this chapter is to contribute to the discussion by providing a conceptual framework for an IoT-based system supporting customer purchase decisions. The system operates on several levels. The major level of operation is customer behaviour and purchase decision support. The second level concerns generating information regarding the shelf content status and reporting on the dynamics of stock-level changes, in order to shape an appropriate marketing strategy and develop effective stocking management processes. The third foundation of the system encompasses the technology and the essential elements ensuring efficient operation of the system. The capabilities of IoT are constantly progressing and expanding, thus leading to the development of software and hardware tools compatible with particular phases of service and customer support process. The presented study employs two research methods: self-observation and the literature study. The proposed solution requires that the system is equipped with necessary software and sensors connected in a wireless data transmission network of Internet of Things.

Keywords Internet of Things · Consumer · Purchase decision support

13.1 Introduction

The rise of the Internet of Things (IoT) paradigm and enabling technologies (such as augmented reality (AR), cyber-physical systems, artificial intelligence (AI), blockchain or edge computing) promise a new era for retail. Rather than on the technical aspects of IoT, this study analyses its behavioural implications. Therefore, bearing in mind the subject matter of this analysis, two research methods were employed: own observation and the literature study. Self-observation is a deliberate

M. Kulisz (✉) · J. Lipski · A. Bojanowska
Lublin University of Technology, Nabystrzycka 38, 20-618 Lublin, Poland
e-mail: m.kulisz@pollub.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_13

search for facts, the act of learning with the help of the senses; the scientific observation is a process of attentive and intentional perceiving. The results of scientific observation are scientific insights included in the further part of the chapter, whereas the section below presents a brief overview of the literature in the field of IoT.

IoT allows people and things to be connected anytime, anywhere, with anything and anyone, ideally using any path/network and any service. Internet of Things does not really focus on things, but on data that is collected through broadly understood *things*, i.e. physical objects in Internet-connected networks that collect process or transfer data to each other independently or in groups [2]. The scope of IoT applications is constantly expanding and in future is expected to encompass: smart homes, buildings, cities, home appliances and multimedia, healthcare systems, retail sales, smart industry, security systems, logistics and others [17].

The following section reviews some existing customer-related IoT solutions. Fashion retail is one of the areas where IoT systems are readily implemented, e.g. in the form of smart mirrors—as a means of sales support. Smart mirrors offer an array of features; e.g. Neiman Marcus's Memory Mirror shows a 360-degree view of a buyer in the outfit on and allows them to compare outfits side by side. Furthermore, the photographs taken in front of the mirror are shared with social media friends to ask for their opinion before concluding the purchase. Ralph Lauren's mirror is located in the fitting room, where shoppers can interact with it by, e.g., adjusting the lighting in their fitting room and selecting different sizes or colours of their outfit, immediately provided by a shop assistant [5]. Another example of the use of IoT in this industry is recording customer behaviour, as a reflection of their hidden preferences. The obtained behavioural intentions can be used by a salesperson to recommend products to customers and for supply chain planning [3]. IoT is also found in retail sales. The authors of the study [12] proposed a smart unstaffed retail shop scheme based on artificial intelligence (AI) and IoT, aimed at exploring the feasibility of implementing the unstaffed retail shopping. Based on 11,000 images in different scenarios containing 10 different types of stock keeping units (SKUs), an end-to-end classification model trained by the Mask-RCNN method was developed for SKU counting and recognition. The solution in question eliminates the process of character segmentation and therefore avoids the error brought by character segmentation. The scientific literature in the field reports on other solutions relying on artificial intelligence (AI) and Internet of Things (IoT) technologies to serve customers; e.g. 7-Eleven has already opened two unstaffed X-Stores [11], and Amazon Go—four stores in Seattle and Chicago; here, customers enter, shop and leave without checking out at a register. China's connected store models are based on an "enter, scan, pay" model and are equipped with personalised concierge services that give customers product information and discount deals via their smartphone apps or BingoBox. Similar solutions are being effectively implemented by other chains, such as French Auchan—exploring the mini- and micro-store format in China—and Wheelys, which is deploying mobile walk-in unmanned mini-stores [9].

Apart from in-store applications, the Internet of Things technology is capable of fulfilling the needs of home shoppers. An example of such an application is presented by Kaur and Kaur [8]. The work advances the concept of sensor-based measurement

containers (SBMCs) coupled with an Android application. SBMCs provide the cloud IoT-enabled grocery management system (CE-GMS) with pressure readings for the management of necessary household commodities—sensing their quantities on hand by means of an ultrasonic sensor. It allows an alert to be generated when the item quantities reach the reorder point. The automated information flow from a customer's SBMCs is handled by the nearest store [8]. Another solution, the Amazon's Dash-button trigger operates at the press of a button, ordering products immediately and without reminder, so that the customer would not waste time searching for products [1, 14].

Another reason for focusing the scope of this chapter on IoT in consumer and retail sales is the current demographics. According to "ambient commerce"—a recent report by GlobalData, it is predicted that Generation Z, which consists of 2.3 billion of "digital natives" born after 1995, will be increasingly choosing ambient commerce as the preferred purchasing method. Ambient commerce is expected to fuel the expenditure on IoT technologies, software and services in the retail sector. Over the next two years, the report said the retail sector is projected to emerge as the fifth-largest customer of IoT software and services, giving way only to the government, transportation, utilities and manufacturing sectors [10].

The ongoing trends for the application of IoT software and services in various industries imply that retail exhibits real potential, thus making it a potent area of scientific research, which is another justification for selecting this field for the following study.

13.2 Purchase Decision Support

The consumer decision process is a complex combination of various factors, which may be categorised according to the situation, consumer psychology and the environment; however, companies have a limited impact on customer purchase decisions, as certain factors are non-modifiable. The decision is frequently the resultant of several variables, such as customer's personality, knowledge, subconscious decisions, competitive pressure, selection of choices on the market, season and fashion. While customers' personality is beyond the influence of retailers, the latter may, however, stimulate the acquisition of knowledge about the product or attempt to evoke positive connotations by appealing to their sub-consciousness (targeted branding). Through their actions, companies may escape the pressure of competition and propose innovative products to customers. They may also adapt to current fashions or seasons and react dynamically to changes in the environment, simultaneously having a small impact on how these factors shape.

In the presence of numerous decision factors beyond their control, businesses must strive to skilfully manage, e.g. by means of IoT, the ones they have a real impact on. Internet of Things in sales is sure to boost the competitive potential of businesses, as innovativeness is regarded as one of the rudimentary prerequisites of increasing the competitiveness [4]. The customer of the modern age is eager to use

new technologies, as evidenced by the growing number of smartphone users, but is also extremely demanding and constantly searching for novelty.

It is critical that companies show high sensitivity to how customers respond to attempts to influence their decision. While their reaction may be highly individual, it is critical to establish the current stage of the buying process. Some may respond positively to assisted purchase, whereas others may take it as unnecessary or hostile solicitation, which could delay their purchase decision turning to competition. In both B2C and B2B scenarios, the effectiveness of these efforts largely depends on the overall customer–company relationship. Loyal customers in long-term partnerships are likely to trust the *spur*, which could in turn discourage new clients or those with former subpar customer experience.

The in-store customer decision-making may be successfully managed provided that the relevant factors affecting the customer are taken into consideration, e.g. through merchandising. The main purpose of merchandising is to activate sales. Attractive appearance applies to the entire area of the retail store, whereas a balanced layout of high-performing areas creates the impression of a wide assortment of products. Building an atmosphere conducive to customers' sense of comfort positively not only affects the amount of shopping but also develops in the customers' subconscious preferences regarding the place [13].

One of the key elements of the buying decision process, consumer emotions can be consciously and subconsciously regulated [6]. Emotions are crucial and are present at various stages of purchase decision-making. In the phase of information search, customers tend to be more rational, whereas immediately after or even during the purchase is the time when a feeling of euphoria emerges. It is important to influence the customer when they make most rational decisions.

To generate positive emotions in the customer, stores may stimulate different senses, by, e.g., spraying a familiar smell of positive connotation in crucial areas (the smell of baked bread at the entrance to the market or freshly brewed coffee at gas stations). However, the vision offers the highest perception potential and is the key to gaining customer support and loyalty. A visual image of the dish, the purchased products can be transformed into, shifts the focus from the sight to the taste and memory regions of the brain and is bound to spur the purchase decisions of the customer and conclude the sales.

13.3 Implementation of IoT in Purchase Decision Support

In retail sales scenarios, the role of Internet of Things is currently narrowed down to ensuring that sales are directly aimed at a targeted customer, based on the user preferences and the purchase history. Thus, obtained information package enables profiling the customer with regard to products that they may potentially find interesting. However, due to the constant development of IoT technologies, they may soon play a leading role in supporting purchase decisions when shopping. The data on the products shelved in the vicinity of the current position of the customer's cart can

be wirelessly sent to a cloud-based software application via IoT. The programme combines the potentially best use of the products for a given customer, and using the IoT sends recipe suggestions to the cart—in order to aid their purchasing [7].

The IoT-based system presented in the sections to follow is designed as a solution for supporting purchase decisions. In addition to its core function, this multi-level system's capabilities include updating shelf content status and reporting on the dynamics of stock-level changes. Thus, obtained information can be used to increase the company's competitiveness as well as to optimise the logistic process control. On the technological level, the system description includes the necessary elements ensuring its efficient operation. The proposed innovative implementation of IoT technology requires equipping the store sales system with appropriate software and sensors, as well as establishing wireless transmission of data between the networked elements of Internet of Things.

13.3.1 Purchase Decision Support System

The developed system focuses on the customer and aims to influence their shopping decisions. On entering the store, the customer collects a smart cart with an embedded display, where all kinds of information and messages will appear. Simultaneously, the customer becomes automatically assigned to the selected smart cart and once the necessary permissions are granted, they are recognised by the store identification system. The user stories presenting basic features of the system are listed below, whereas Figs. 13.1 and 13.2 illustrate the scheme of its operation.

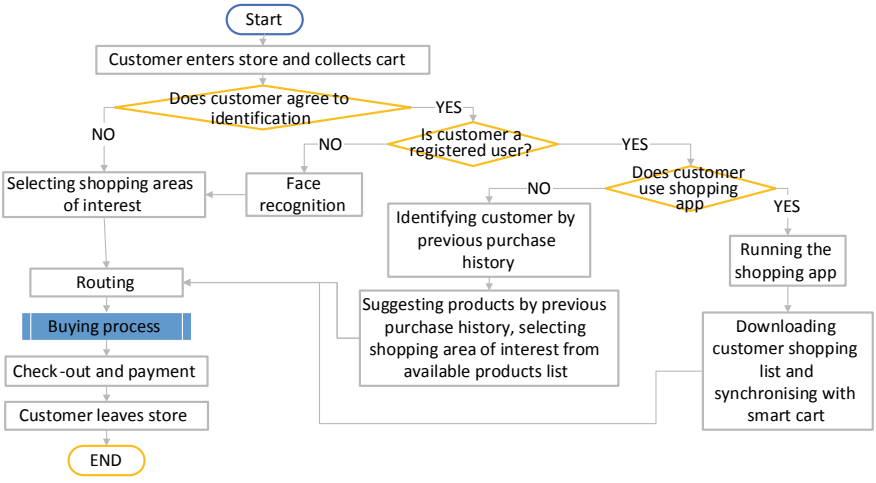


Fig. 13.1 Schematic algorithm of the system

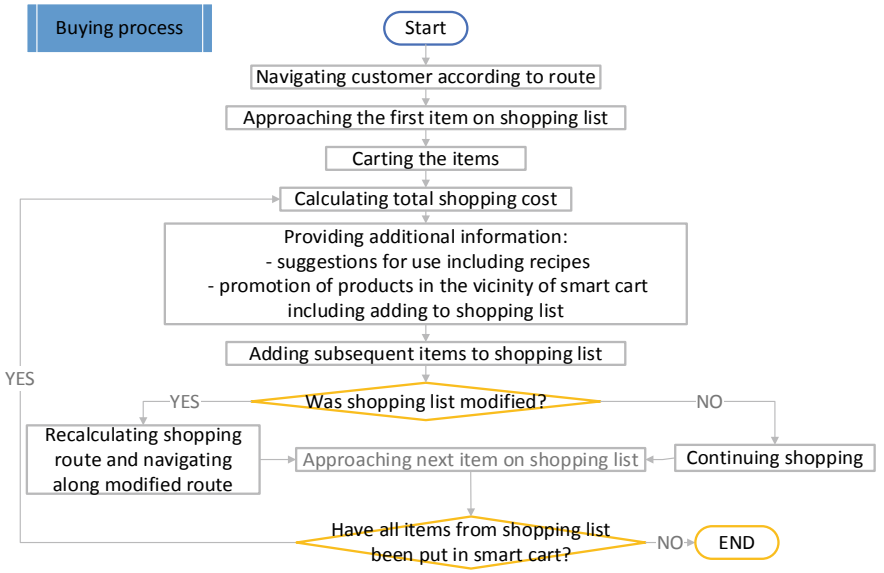


Fig. 13.2 Schematics of the proposed purchasing system process algorithm

1. The customer is not registered in the system. After picking up the smart cart, the shopping area of interest is established, upon which the system determines the shopping path. The smart shopping cart display views purchase suggestions to the user as they move through alleys. The suggestions shown to the user are based on data communication picked up by the cart from the shelves on its path in the sales hall. The IoT system controls the flow of information to the customer depending on the stock on shelves and the availability of other products from the established set of potentially desired products available in the store. A customer not registered in the system who, however, agrees to face recognition and their shopping activities being registered has their model image created and saved for future shopping. This consequently allows the store to register future purchases and to create a customer profile.
2. The customer is registered in the system and has a shopping list prepared in a dedicated application. After assigning a smart cart, the system recognises the customer via Bluetooth, downloads the shopping list, generates the shopping path and guides the customer through the alleys to reach the listed goods. In addition, after downloading each product, the screen displays recipe proposals, which can be downloaded, and the necessary ingredients are added to the shopping list. The customer is subsequently navigated through the alleys and receives notifications concerning the location of the products from the list.
3. The customer is registered in the system but does not use the application, so there is no pre-made shopping list. After positive identification (face recognition), they select elements (represented by icons) from different categories, which are then

added to the system. The smart cart displays products that the customer might be interested in. The information is generated on the basis of previous customer purchases. The purchase decision support system displays on the smart cart screen the route to the alley with the selected products. Once the item is removed from the shelf and placed in the smart cart, the information about how to prepare a given product, including a shopping list suggestion, appears on the screen. The customer may use one of the recipes, and as in the previous case, the list of ingredients necessary to prepare a given dish is uploaded to the shopping list in the smart cart. Then, a shopping route is drawn to navigate the customer through the alleys to pick up the necessary products. The route is recalculated whenever the customer leaves the shopping path, e.g. in search of another product. When a new product lands in the cart, the recipe suggestions, the shopping list and the route are automatically updated.

4. The offer is subjected to further personalisation, based on previous shopping history, for customers who have agreed to face recognition and identification.
5. The total value of the shopping cart is updated upon each new addition to the cart and displayed live on the smart cart screen. Similarly, in the case when a previously carted product is removed, e.g. because the customer has found a cheaper alternative, the value of the cart is recalculated. Once the shopping is complete—all products from the list are in the smart cart—the customer approaches the cash register. Due to the fact that the content of the cart has already been automatically identified, the checking out and payment processes are simplified and the latter executed via contactless payment.

In terms of customer decision support, this system may offer an additional feature. Given that the customer agrees to identification and has browsed for various offers in online stores and added them to favourites prior to entering the smart store, it is possible that when such a product is available, the customer will receive a notification reading, e.g., "*Products on Your Favourites list are available in alley X*". By providing the real-time access to customer activity data, as well as indicating various possibilities to involve customers in the purchasing process, the system supports their decision-making.

13.3.2 Employing Data from the IoT-Based System to Stimulate Desired Shopping Decisions

The second of the system's levels of operation, referred to as in the introduction, concerns generating information about the state of and the dynamics of changes in the shelves' stock levels. After the product is placed in the customer cart, the status of the shelves is updated. The analysis of this information enables exerting appropriate control over the restocking process. After reaching the minimum level, the order information is sent to the supplier (or the warehouse), which is possible due to automated ordering process. Ongoing stock-level control will ensure continuity

of sales and simultaneously minimise the risk of generating high storage costs. In addition, the constantly provided information on goods flow and the dynamics of stock-level changes provides data for analysis of, e.g., the reasons for poor sales of certain goods. The analyses may result in implementing changes in the layout of goods in the shopping hall or the form of promotional campaigns—with a view to increasing the turnover. IoT can track the efficiency of the store layout or the reaction to product promotions. These data are analysed in real time and can be used to determine customer preferences and their purchase paths in great detail.

It is important that the IoT-based system is compatible with the knowledge gained on customer behaviour in a commercial outlet and, furthermore, how the store itself stimulates the customer movement. Proper display of goods and prompting the customer to approach the product may counteract excessive stock levels of a particular product on sale. A customer wandering about the store needs about one-third of the store's area to make a shopping choice. However, it is in the best interest of the store that most of the store's surface is penetrated by the customer. The area at the entrance, where the customers tend to buy fewer products, is typically referred to as the decompression zone. It is approximately 6–7 m in length. To increase its attractiveness and visibility, promotional stands are usually located there [18]. One of the key factors that should be managed is the speed of customer flow. It is assumed that the average speed of customers in large stores is about one metre per second [18]. Retail stores can stimulate faster or slower customer flow. Messages addressed to customers (e.g. concerning current promotions) should be adjusted to this speed, as well as to the fact that customers usually move counterclockwise and that moving on the right side of the alley they show a tendency to look to the right. It is also assumed that women pay the most attention to goods situated at the height of their belt, while men at eye level. With IoT, one can check whether their store conforms to statistics and whether their customer behaviour is typical or atypical. If no adjustments are possible, the data can be used to display the goods to be sold to customers in the first place.

The IoT-based systems offer a breadth of data that may be harnessed to stimulate the consumer behaviour, and this includes the data regarding purchase history, previously selected products and their amount, the purchase decision-making speed, changes and trends.

13.3.3 The Functions of Customer Assistance Systems and Automatic Identification Technologies in Customer Support Solutions

IoT technologies enable the transmission of information between elements of the physical system, without human intervention. The system event that triggers the transmission, the source of information and the addresses and transmission contents are programmed by the system programmer. The obtained data can be used by the

receiver to control various objects or to generate events suitable for the purpose of particular system elements. These events can trigger calculation procedures that result in controlling physical objects. Alternatively, they can simply constitute system events for information transmission (wired/wireless) to another IoT node.

We can distinguish three groups of activities performed by a digital built-in module for a smart shopping cart or devices that perform an auxiliary role in the process of retail distribution (such as storage or control equipment):

- registration, initial processing, selection of addressees and sending information transmission to other functionally related modules built into the customer service process, as well as to the cloud computing system;
- acquiring information from other networked modules in order to generate correct control decisions and information for the customer (e.g. generating an optimal path for the client during the shopping process);
- processing of collected local and acquired data from other embedded modules and generation of control decisions and sending them to the customer's information system (display screen).

The wide range of capabilities offered by IoT solutions makes this technology (software/hardware) applicable in various phases of service and customer support. The basic problem to solve is the method and sources of obtaining information about both customers and their articulated needs, as well as about the stock resources and their features. Taking into account the sequence of events from the customer's arrival to the store to leaving it, the following functions of the assistance system can be distinguished:

- identifying or registering the customer (with his consent);
- assigning a smart shopping cart;
- customer–system interaction allowing to set preferences;
- generating a smart cart route ensuring optimal implementation of the customer's preferences;
- information communication about product features that may potentially interest the customer and are on the route or are offered on promotion;
- handling the contents of the smart cart and calculating the sopping cost;
- handling payment.

Each of the listed functions requires selecting a data acquisition method and an algorithm for its processing. Therefore, the subsequent paragraphs will present alternative technical solutions necessary for the performance of the required functions.

Considering the customer's right to save personal data, the cases described in point 3.1 should be considered and described by algorithms. To support the customer identification phase and to generate the path of travel and customer decision support, applications for the transmission of Bluetooth data and the face recognition system will be required. A further requirement concerns establishing the communication between the smart cart with the customer database and the database of products. The communication can employ Wi-Fi technology and Internet connection for communication with cloud computing, in which these databases could be stored. These

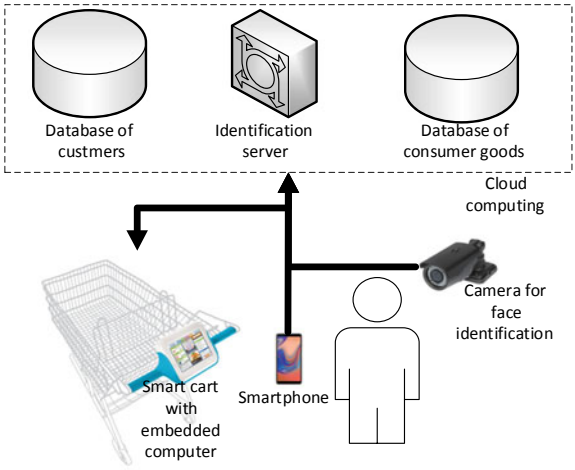


Fig. 13.3 Data flow in the customer identification and registration phase

databases may be accessed by commercial outlets of the same company located in different districts or cities, thus ensuring the customer receives the same level of service quality at different store locations of the same company. The communication scheme, in the purchasing phase, of the smart customer support system is shown in Fig. 13.3.

In the case of customers who registered their shopping list, the support system displays instructions on the screen—allowing them to find products from this list. At the same time, when analysing the previous purchase history and relevant promotions, it makes suggestions of consumer goods that are on the route of the smart cart. Accomplishing this objective requires harnessing communication technology with a beacon module. Beacon, based on Bluetooth smart technology, constantly sends a radio signal. The signal is read only by a dedicated mobile application active on the computer embedded in the user's smart cart with Bluetooth 4.0.

The application running on this computer receives a radio signal, checks the unique beacon ID and—in accordance with the built-in logic—provides the user with appropriate content displayed on the cart's screen. Miniature beacon transmitters can be placed on store shelves near which the customer moves. This allows individual contact with the user of the mobile application when it is near the beacon transmitters. When generating such a message, the loyalty system can use all previously accumulated knowledge (in the customer database) about the given user and his previous purchase decisions, enabling him to present not only the benefits of participating in the loyalty programme, but also offers for products and services that are likely to interest him. Unlike GPS, this technology works well inside buildings, enabling not only direct communication with the user in real time, but also navigation and geolocation.

The process of filling the smart cart with purchased consumer goods should be monitored in order to accelerate the final stage of purchase, i.e. the payment. The customer support system calculates the amount due and displays the information

live—when subsequent goods are placed in the cart. At this stage, it is also possible to cancel the purchase of items previously selected and carted.

Technologies that fulfil such formulated tasks are already and successfully implemented in warehouse management systems. Product identification during the transfer from shelf to cart and from cart to shelf can be made by the implementation of:

- Bar code reader placed on the packaging of products—the direction of the shift is determined from the changes in the weight of the cart.
- Radio-frequency identification (RFID), a method employing radio waves to transmit data and power the electronic system (RFID tag)—goods are labelled with RFID tags. The smart cart CPU reads the item code written on the label of the product placed in the cart and after finding the increase in the weight of the content of the smart cart adds the value to the receipt [15].
- An innovative Amazon Go technology currently tested by Amazon—it is advanced software devised to perform such tasks as product identification—regardless of whether they are on the shelf or not, even in the case when the item is partially covered by another object. The store does not use more conventional technologies, such as RFID tags to track items, but may obtain additional confirmation from the regular bar code on the product packaging. The information is obtained by analysing the image from multiple cameras [1].

Based on the current state of technology, it can be concluded that the most feasible solution in this phase of customer support is RFID technology [16]. Numerous companies have already introduced to the market certain elements of this technology infrastructure, such as labels with antenna and chip imprint and information readers. An example of using RFID technology for the identification of the process of manipulation of goods is illustrated in Fig. 13.4. Since the total cost of shopping is updated constantly throughout the entire shopping time, the customer uses convenient card

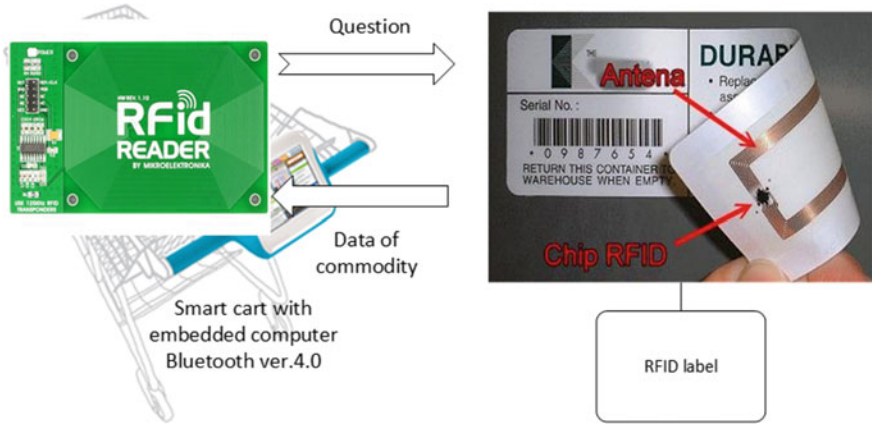


Fig. 13.4 Application of RFID technology in consumer goods handling

payment in traditional terminals requiring pin payment acceptance or launches their bank's smartphone app and confirm the payment.

13.4 Conclusions

This chapter proposes a conceptual framework for the IoT information technology in customer purchase decision support. Therefore, its primary objective, i.e. to contribute to the discussion by providing a conceptual framework for an IoT-based system supporting customer purchase decisions, appears to have been achieved. The system is described in terms of the algorithm for customer service using multiple data sources, as well as regarding the data acquisition and processing technology. The system is designed to facilitate the selection of products that best meet customers' needs and solidify their company loyalty. This multi-level system, by implementation of necessary hardware and software solutions, operates on customer behaviour and purchase decision support. In addition, it generates the shelf content status data and reports on the dynamics of stock-level changes, thus aiding an appropriate marketing strategy shaping and optimising stocking management. The constant progress of IoT capabilities leads to higher compatibility of designed software and hardware tools with particular phases of service and customer support process. The proposed solutions will also optimise customer service costs and facilitate continuous analysis of the structure and the dynamics of sales.

References

1. Amazon.com: Amazon Dash Button. https://www.amazon.com/b/%3fnode%3d10667898011%26sort%3ddate-desc-rank%26lo%3ddigital-text (2019). Accessed 3 Jan 2019
2. Bojanowska, A.: Customer data collection with Internet of Things. MATEC Web Conf. **252**, 03002 (2019). https://doi.org/10.1051/matecconf/201925203002
3. Chan, C.O., Lau, H.C.W., Fan, Y.: IoT data acquisition in fashion retail application: fuzzy logic approach. In: 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, pp. 52–56 (2018)
4. Charucka, O.: Kluczowe czynniki konkurencyjności MSP i ich wpływ na rozwój gospodarki. Zesz. Nauk. Uczel. Vistula **35**, 45–67 (2014)
5. Gaur, L., Singh, G., Ramakrishnan, R.: Understanding consumer preferences using IoT smart-mirrors. Pertanika J. Sci. Technol. **25**(3), 939–948 (2017)
6. Gross, J.J.: Emotion regulation: affective, cognitive, and social consequences. Psychology **39**, 281–291 (2002)
7. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of Things (IoT): a vision, architectural elements, and future. Future Gener. Comput. Syst. **29**(7), 1645–1660 (2013)
8. Kaur, J., Kaur, P.D.: CE-GMS: a cloud IoT-enabled grocery management system. Electron. Commer. Res. Appl. **28**, 63–72 (2018)
9. Kokoszka, P.: 11 trends in ambient commerce that will change retail over next two years. https://www.retail-insight-network.com/features/trends-in-ambient-commerce/ (2018a). Accessed 7 Jan 2019

10. Kokoszka, P.: Ambient commerce set to boost retail spending on IoT tech to $5.3bn, report. https://www.retail-insight-network.com/news/ambient-commerce-boost-retail-spending/ (2018b). Accessed 7 Jan 2019
11. Liao, C., Hwang, A.: 7-Eleven opens unstaffed store in Taipei. https://www.digitimes.com/news/a20180726PD201.html?chid=9 (2018). Accessed 7 Jan 2019
12. Liu, L., Zhou, B., Zou, Z., Yeh, S., Zheng, L.: A smart unstaffed retail shop based on artificial intelligence and IoT. In: 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Singapore, pp. 1–4 (2018)
13. Marszał, K.: Sterowanie ruchem klienta w obiektach handlowych. In: Grzegorczyk, A., Wiśniewska, A. (eds.) Merchandising, pp. 56–73. Wyższa Szkoła Promocji, Warszawa (2014)
14. Möhring, M., Keller, B., Schmidt, R., Pietzsch, L., Karich, L., Berhalter, C., Kilian, K.: Using smart edge devices to integrate consumers into digitized processes: the case of Amazon dash-button. In: Teniente, E., Weidlich, M. (eds.) Business Process Management Workshops. BPM 2017. Lecture Notes in Business Information Processing, vol. 308, pp. 374–383. Springer, Cham (2018)
15. Sai Ganesh, S., Sahithi, B., Akhila, S., Venumadhav, T.: RFID based shopping cart. Int. J. Innov. Res. Eng. Manag. (IJIREM) **2**(3), 24–30 (2015)
16. Sawant, R., Krishnan, K., Bhokre, S., Bhosale, P.: The RFID based smart shopping cart. Int. J. Eng. Res. Gen. Sci. **3**(2), 275–280 (2015)
17. Tao, B.: Vivante Internet of Things (IoT) Solutions. https://bensontao.wordpress.com/2013/10/06/vivante-internet-of-things/ (2019). Accessed 3 Jan 2019
18. Witek, L.: Merchandising w małych i dużych firmach handlowych. C.H. Beck, Warszawa (2007)

Chapter 14
Application of EEG Metrics in the Decision-Making Process

Mateusz Piwowarski, Uma Shankar Singh and Kesra Nermend

Abstract The decision-making process is a complex task uses the multi-criteria methods in the formalized decision support. Decisions are direct reflection of decision maker preferences. Multi-criteria methods use different methodological approaches (algorithms) to determine the final assessment of decision variants (e.g., ranking). Decision maker must do many actions (partial evaluations) in some of these methods. Issues of the decision maker's engagement in the assessment process arise which can be identified using measurements by EEG. It is possible to identify various internal processes occurring with the decision maker during individual stages of the calculation procedure. Various types of EEG metrics are used for this, such as the index of frontal asymmetry, engagement, distraction, etc.

Keywords EEG metrics · Decision making · Multi-criteria methods

14.1 Introduction

Making important decisions is usually a difficult and complicated task, and during the decision process a common question arises about possible alternatives (variants). What could we gain or what could we lose by making such or other decisions? This question is important because each decision has its specific effect, which is often possible to identify a rational decision considering the conditions of decisions. Each

M. Piwowarski (✉) · K. Nermend
Department of Computational Methods in Experimental Economics, Faculty of Economics and Management, Institute of IT in Management, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: mateusz.piwowarski@usz.edu.pl

K. Nermend
e-mail: kesra.nermend@usz.edu.pl

U. S. Singh
Faculty of Administrative Sciences and Economics, Tishk International University, Erbil-Kurdistan, Iraq
e-mail: umashankar.singh@ishik.edu.iq

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_14

action taken has some effects that can be assessed positively, negatively, or indifferently (neutrally). In the context of rationality, decisions should be understood as a choice of a variant of action that ensures the best ratio between the positive and negative effects of the assessed decision [1, 2]. Rationally, a human decision is guided by reason, and they consider the arguments and counter-arguments available to them and make the decision closest to their preferences before making a decision. Irrationally, human behavior is usually guided by emotions often with false imaginations which frequently results in a wrong view of the situation and unrealistic assessments of the effects of taken decisions. Emotions give a false impression of attractiveness. Decision making is closely related to the functioning of the brain. Issue of rules of the brain's operation and the mechanisms of human thinking including the role of intuition and experience in decisions making has been given by numerous studies [3, 4]. It shows that choices and decisions made by people depend on the mechanisms of thinking, the so-called quick thinking (unconscious, automatic) and slow thinking (controlled). Quick thinking does not require effort, because it is based on thought patterns (life experience) or heuristics (so-called shortcut thinking, automatic reactions to stimuli from the environment). Quick thinking allows for almost immediate decision making, but it is associated with the risk of not analyzing a given situation. In the case of slow thinking, a significant intellectual effort is required to analyze the information. It boils down to verifying fast (automatic) thinking and taking control of emerging choices [4].

The decision-making process can be enriched by various less formal factors, such as experience, intuition, common sense, or developed practical rules in many other decision-making situations. All of these approaches can be helpful, but they cannot replace reliable analysis. It allows for more detailed consideration of decisions variant and their assessment due to the decision maker's preferences [5]. This applies particularly to situations in which there is a need for detailed consideration of decision variants. Mathematics language and the recognition of decision situations in the form of models reflecting real process come in the picture here. Approaches in this field derive from econometrics and its derivative scientific disciplines, such as statistics (analysis and processing of data reflecting reality) or operational research (optimization of decisions). Participants in the decision-making process (decision makers) are placed in the spotlight in the case of operational research. They assess decision-making situations based on their own preferences (subjective point of view) taking into account the set of adopted assessment criteria. They face a decision problem in which different conflicting situations often have to be taken into account. Such a situation is referred to a multi-criteria decision problem and characterized by the subjectivity of defining the problem, objectives, defining assessment criteria as well as evaluating their effects. This approach is supported by multi-criteria decision support methods. Some decision-making situations require consideration of many decision-making variants assessed on the basis of many criteria [6]. The solution of this type of decision problems (using multi-criteria methods) usually requires a significant participation of the decision maker. For some methods, the decision maker must make many comparisons (assess) before the final ranking of variants is created [7]. The issue of the decision maker's involvement in such a formalized

decision-making process arises here with too little commitment (or even fatigue) and partial evaluations can be unreliable. Ultimately, this will translate into a final grade that will not fully reflect the decision maker's preference. An important issue seems to be the identification of the decision maker's involvement in the decision-making process. The purpose of this chapter is to present the possibility of using the various EEG metrics (e.g., related to engagement, distraction) in the decision-making processes. The problem of identifying engagement is important because it can distort the effectiveness of certain multi-criteria methods (e.g., AHP, ANP) to solve complex decision problems. The importance of the problem is also raised by the fact that these methods are commonly used to solve contemporary economic problems.

14.2 Limitations in the Applicability of Selected Multi-Criteria Methods

The developed methodology under operational research allows to determine a set of acceptable solutions, moreover, to search for the optimal solution (in the Pareto sense) based on methods used in the multi-criteria decision support. It takes place by evaluating solutions under the consideration of using a defined set of criteria. These methods are included in the group of multi-criteria optimization methods with a finite set of value [8]. Literature in the subject area discusses these issues are described as multiple criteria decision-making (MCDM) or multi-criteria decision analysis (MCDA) [9]. The main goal of multi-criteria decision support is to provide decision makers with tools that will facilitate in solving decision-making problems where there are different criteria for assessing decision-making variants. MCDM is related with the American mainstream, which assumes the existence of a multi-attribute utility function providing synthetic assessments of the analyzed decision variants. MCDA is combined with the European mainstream. This trend is based on the occurrence of an outranking method, in which the aggregation of partial grades is carried out on the basis of a set of conditions at a global over-ratio outranking [10]. MCDM/MCDA methods can be grouped according to different criteria. One of the possibilities of their divisions is shown in Fig. 14.1.

A detailed description of these methods can be found in the literature of this subject commonly available among others on the Internet [11–19]. The computational algorithms of the different methods are different, although they partly use similar solutions (e.g., pairwise comparison: AHP, ANP, MACBETH, REGIME, etc.). Analytic hierarchy process (AHP) method is one of the best known and widely used [20] among them. The vector of the scale is obtained here by comparing the pair of decision variants in the context of individual criteria and comparing the assessment criteria with each other. The components of the scale vector enable us to obtain the order of the variants, and thus, chose the best. The pairwise comparison process uses a nine-point rating scale called the Saaty scale or a corresponding verbal description.

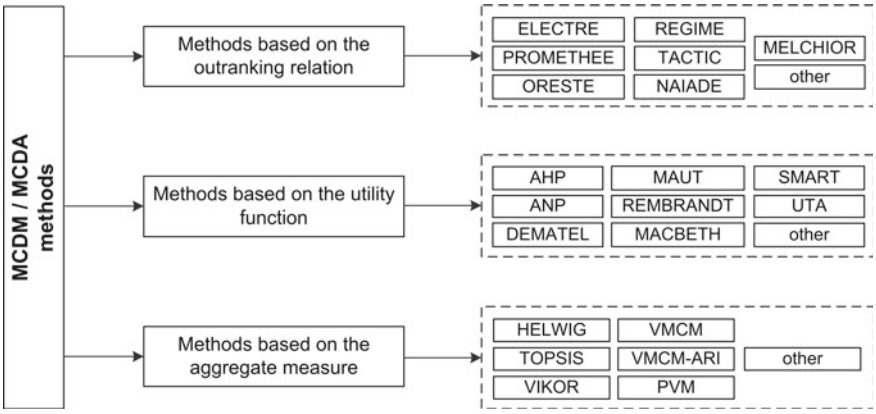


Fig. 14.1 Classification of MCDM/MCDA methods

On the top of the hierarchical structure describing the decision problem is a main goal of the decision problem, the next level includes assessment criteria (subcriteria) and at the bottom of this hierarchy are considered decision variants. The analytic network process (ANP) method is a development of the AHP method by considering links between the criteria and links between the variants and the criteria [21]. The hierarchical structure is used to describe the decision problem similar to the AHP method. The ANP method is being able to have horizontal relationships between the assessment criteria and the feedback links between the criteria and the decision variants [22, 23]. The criteria can be combined in components which may be elements of the type of criteria and decision variants. Both methods (AHP and ANP) are time-consuming for a large number of main and partial criteria. With a larger number of evaluation criteria and decision variants, a large number of comparisons should be made. This number is expressed by the formula for each criterion $\frac{n^2-n}{2}$, where n means the number of variants. For example, for six decision variants and eight evaluation criteria, the number of comparisons will be 120. With a relatively small number of variants and criteria (for this case), the decision maker would have to do the tedious work related to the comparative assessment. As our previous research [24] shows, the effectiveness of a long-term decision-making process decreases with the increasing number of necessary object comparisons. EEG measurements can be used in order to check the level of decision maker engagement in the process with the point of pairwise comparisons. Electroencephalography gives the opportunity to study the frontal asymmetry. This phenomenon gives the opportunity to study the brain's response to stimuli, e.g., resulting from the process of comparing objects. These neural measurements allow for the determination of various metrics, for example, related to engagement.

14.3 EEG as a Measuring Technique

Electroencephalography (EEG) is one of the brain neuroimaging methods. Other measuring techniques also have been used except EEG is positron emission tomography (PET), functional magnetic resonance (FMR), and functional magnetic resonance imaging (FMRI) [25]. These methods allow for observation of changes in brain activity and the brain waves analysis. EEG is highly practical and important in the context of economics research. The temporal resolution of EEG is very high (milliseconds), which makes it possible to track brain activity for rapidly changing stimulation (as advertising shots, consumer reactions to viewed goods on store shelves, etc.) [26]. The EEG reads (measures) the recording of bioelectrical brain activity. The source of this activity (signal) is pyramidal cells, which is arranged parallel to each other in such a way their arrangement causes adding up their electrical activity. It creates sufficiently electrical potential that can be registered as the effect. The produced electrical potential (in a given place) is read with a device called an electroencephalograph. The electrodes are used to read the bioelectric activity of the brain. They are placed on the head most often according to the international standard 10–20 (or its extensions) (Fig. 14.2).

The positions of each electrode are described by the letter denoting the lobe of cerebral cortex and the digit defining the part (side) of the brain (even-right side, odd-left side). The *F* letter means the frontal lobe, *T*—the temporal lobe, *P*—parietal lobe, *O*—occipital lobe, *C*—Rolando's fissure, *z*—line connecting the nasal cavity with occipital nodule [27, 28]. Recorded electrical potentials are also influenced by potentials generated by other parts of the body (e.g., eyes, heart, muscles) and external factors (measuring apparatus, power network). These signals are called

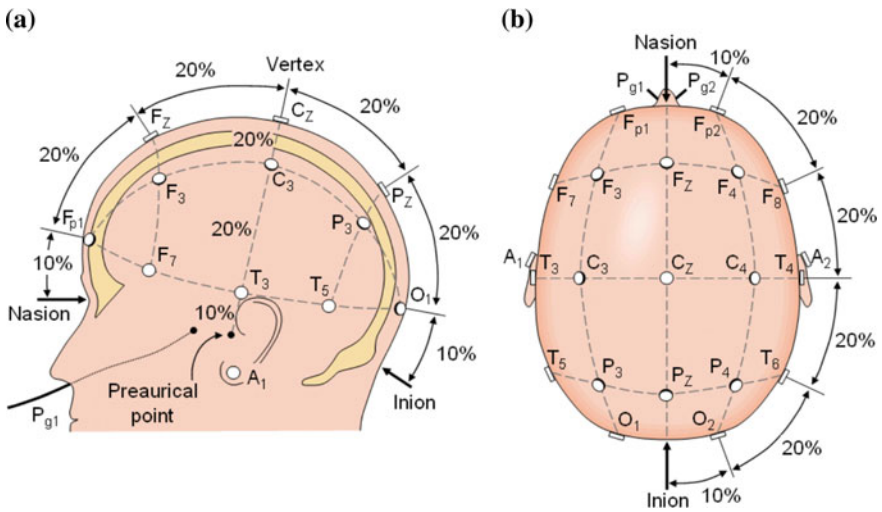


Fig. 14.2 System of 10–20 locations of electrodes, **a** left part of head, **b** top part of head [27]

artifacts, and signals which are interested should be cleaned. Signals generated by the human brain can be grouped into rhythms (EEG waves) with different amplitude and frequency. The most important types of waves (in the context of economics research, but not only) include alpha (frequency 8–13 Hz, amplitude approx. 30–100 μV), beta (12–30 Hz, up to 30 μV), theta (4–8 Hz), delta (up to 4 Hz), and gamma (26–100 Hz) [25]. The individual wave bands are associated with various activities, processes occurring in a human being. For example, alpha waves are associated with a low level of human cognitive activity and a state of relaxation.

14.4 The Metrics Used in Decisions Making

In economics study (e.g., consumer behavior, decision making), various types of techniques and scientific methods are used to get information. They are usually based on data that is statements and opinions of respondents. They reflect subjective reflections and own interpretations of reality. However, consumer declarations are not always consistent with their behavior. It can therefore lead to incorrect results and interpretations. Often people cannot explain their decisions and behaviors in a rational way. Behavioral economics and techniques based on cognitive neuroscience can help them. The study carried out with the EEG, GSR and other techniques allow to reach motives that consumers do not realize. Such methods allow to build knowledge about a human and his behavior at the level of internal psychophysiological and cognitive processes. They allow us to look into human minds and their basic physiological states. They allow access to unfiltered reactions and feelings. Many different metrics can be obtained from the EEG data acquired. Various phenomena can be identified based on such metrics such as emotions, motivations, engagement, distraction, etc. These metrics are created from purified EEG measurement data (filtering, deleted artifacts) and using specific frequency bands (waves). EEG signals are measured in the right places of the brain responsible for specific mental states and higher-level of cognitive processes. Examples of metrics those are used to identify human behavior (including when making decisions) include frontal asymmetry, engagement, distraction, workload, etc.

14.4.1 Frontal Asymmetry

One of the phenomena has been identified as neuroscience study is called the frontal asymmetry measured in the alpha band frontal alpha asymmetry (FAA). It is the difference between left and right alpha activity over the frontal lobes, recorded with EEG. This phenomenon is used among others to study emotion, motivation, cognitive control, or temperament and personality [29–32]. In practical, two approaches are used to study frontal asymmetry. First of these involves the study of the FAA while resting as a variable trait related to psychological constructs and predicting

future emotional behaviors [33]. Second one is related to the study of the state of changes in frontal asymmetry as a function of the current emotional state or behavior [34]. As it shows, most studies use frontal asymmetry as an indicator of psychological phenomena (motivation, well-being, etc.). The registration of FAA frontal asymmetry is similar to the measurement of other EEG signals. However, it should be noted that it may be the result of individual characteristics as well as change due to unintentional state manipulation. This applies both to the time of the test itself and also to the period before the signal is recorded. All kinds of emotional reactions that may occur in the subject (before or during the study) have an effect on the results of EEG asymmetry [35]. In order to reliably determine the frontal asymmetry, it is necessary to ensure an adequate amount of data and remembering not to temporarily overload the participants of the study. During longer tests, the emotional state of the subjects may change. Good FAA credibility can be obtained from 100 epochs. This corresponds to the time of data registration without artifacts in the range of 1–3 min [36]. In order to ensure the best conditions for FAA measurements and interpretation of results, practices in this area should be strictly observed [32]. The frontal asymmetry index (FAI) can be easily determined on the basis of raw EEG data obtained from the F_3/F_4 and F_7/F_8 electrodes of the 10–20 system:

$$\text{FAI} = \ln\left(\frac{\text{alpha power right } F4}{\text{alpha power left } F3}\right) \tag{1}$$

It should also be noted that both behavioral tests and hemodynamic measures have confirmed that the power of alpha waves is inversely proportional to the activity of brain fragments [37, 38].

14.4.2 Engagement

Engagement is defined in various ways. Some approaches focus on the result of engagement. In such cases, it can be referred to explicit or covert attention of allocation [39]. It is also described as a tendency to attract and maintain attention [40]. Some researchers define engagement as a state of involvement of an individual in an activity in which the other proceedings are blocked [41]. Other approaches to engagement focus on the characteristics of the task that would force more or less involvement. This applies to the degree that determines the difficulty of the task, interactivity, durability, novelty, or the amount of feedback [42]. Using the EEG measurement data, the engagement index can be obtained as a ratio of appropriate quantities (bands). There are several known ways to determine it [43]. We can, for example, count it according to the formula:

$$\text{EnI} = \frac{\text{Beta}}{(\text{Alpha} + \text{Theta})}, \tag{2}$$

averaged across all sensor locations [44–46]. The engagement index can also be obtained in the following way:

$$\text{EnI} = \frac{\text{Theta}}{\text{Alpha}}, \tag{3}$$

average frontal midline theta and average parietal alpha [47, 48].

There is another approach involves the engagement index with the theta value from the front area [49].

14.4.3 Distraction

Distraction is associated with the fact of involvement in activities other than cognitive tasks on which one should focus. It may cause due to boredom or frustration. Different external factors (coming from the environment), e.g., noise or any other phenomena occurring nearby may affect the distraction. Distraction can also be caused by factors directly related to the subject of some errors or mistakes. The index of cognitive distraction widely used in EEG measurements includes alpha activity (8–14 Hz), beta (14–35 Hz), and theta (4–8 Hz) [50]. Beta and theta activities (in the front part of the brain) are associated with cognitive functions such as decision making, problem solving, or working memory [51]. Additionally, gamma activities (31–40 Hz) can be taken into account. The distraction index using these frequency bands (alpha, beta, gamma, theta) can be determined according to the formula [52]:

$$\text{DI} = \frac{\text{theta}}{\text{alfa}} + \frac{\text{alfa}}{\text{beta}} + \frac{\text{beta}}{\text{gamma}}, \tag{4}$$

where *theta* is the average EEG power between 4 and 8 Hz, *alpha*—between 8 and 12 Hz, *beta*—between 12 and 30 Hz, and *gamma*—between 31 and 40 Hz.

The placement of these EEG frequency bands was associated with capturing their variable participation in the distraction.

14.4.4 Workload

Workload is associated with an increase in the work memory load. It appears during analytical reasoning, solving various types of problems (arithmetic, decision-making, etc.). We deal with the workload when we start a group of cognitive functions needed to focus on tasks those are important to us. The load levels recorded by the EEG show a correlation with objective efficiency as well as a subjective assessment of the load on tasks of various difficulties. For example, the load is the basic measure used in the

usability test [53]. The load is related to oscillations in the alpha band, alpha peaks, as well as theta band activities.

In the literature, it is clearly mentioned that many other metrics are determined on the basis of EEG measurements and measurements of other psychophysiological features (e.g., GSR, HR). These are such records as memorization index [54–56], approach–withdrawal index [57], emotional index [58, 59]. These types of records allow identification of various psychophysiological states. They enable building knowledge about human behavior reactions to various external stimuli. It is of great practical importance in the context of research into the sources of consumer decision making. Certainly, it can facilitate reaching consumers, improve the personalization of services, increase the effectiveness of marketing activities, etc. Using psychophysiological measurements to monitor the decision-making process using multi-criteria methods (MCDA/MCDM) can give a new look to the effectiveness of some methods. Although, for example in the AHP method, there is a cohesion factor (checking the assessment logic) where it is not known how decision makers react to the prolonging assessment process. The metrics obtained on the basis of EEG measurements in many situations help to identify phenomena occurring in a person (positive and negative).

14.5 Conclusions

Decisions making is related to the process of choosing the right course of action. This often requires choosing the best option from many possible solutions. There are many types of decision problem (less complicated and very complex) that requires a significant participation of the human factor. Problem-solving methods such as the MCDA/MCDM methods are often used to solve such problems. Despite the proven effectiveness of these methods, there are also risks associated with their selection for a given issue. In the case of complex decision problems (many variants and evaluation criteria), some methods require considerable effort and commitment from the decision maker, e.g., the decision maker must perform many comparison operations (at different levels). A long process of evaluation may cause the appearance of many unfavorable phenomena such as distraction, emotions, reduced engagement, too much workload, etc. Confirmation of these concerns is the result of research carried out by the authors on the use of the AHP method [24]. To be able to respond to such situations, it should get identified first. EEG measurements (and other techniques) allow to identify the phenomena related to such type of factors. Frontal asymmetry gives the opportunity to show how the brain reacts to specific stimuli (e.g., for a product or for multiple comparisons in the decision-making process—multi-criteria methods). The calculated frontal asymmetry index may be the determinant of the approach–withdrawal which reflects the motivation. Increased activity of the left part of the brain can serve as a determinant of motivation (index of approach motivation) and emotions like joy or anger. On the other hand, increased activity of the right part of the brain can serve as a determinant of avoiding motivation and

such emotions as fear or sadness. Greater involvement of the left, frontal side of the brain relative to the right is associated with greater engagement. The computational algorithms developed allow for the conversion of psychophysiological signals into derived cognitive metrics, such as engagement, motivation, distraction, workload, etc. Therefore, we are able to assess the effectiveness of external stimuli affecting decisions.

Acknowledgements The project financed within the Regional Excellence Initiative programme of the Minister of Science and Higher Education of Poland, years 2019-2022, project no. 001/RID/2018/19, financing 10 684 000,00 PLN.

References

1. Grünig, R., Kühn, R.: Successful Decision-Making. A Systematic Approach to Complex Problems, 3rd edn. Springer, Berlin (2013)
2. Pei, Z.: Rational decision making models with incomplete weight information for production line assessment. Inf. Sci. **222**, 696–716 (2013). https://doi.org/10.1016/j.ins.2012.07.060
3. Aronson, E., Wilson, T.D., Akert, R.M.: Social Psychology, 9th edn. Pearson, Boston (2016)
4. Kahneman, D.: Thinking, Fast and Slow. Penguin Books, London (2012)
5. Wang, Y.: The Theoretical Framework of Cognitive Informatics. Int. J. Cognitive Inform. Nat. Intell. **1**, 1–27 (2007). https://doi.org/10.4018/jcini.2007010101
6. von Winterfeldt, D., Edwards, W.: Decision Analysis and Behavioral Research. Cambridge University Press, Cambridge (1986)
7. Roy, B.: Multicriteria Methodology for Decision Aiding. Springer, US (1996)
8. Vincke, P.: Multicriteria Decision-Aid. J. Wiley, New York (1992)
9. International Society on MCDM.: Multiple criteria decision making. http://www.mcdmsociety.org (2015). Accessed 22 Apr 2016
10. Mota, P., Campos, A.R., Neves-Silva, R.: First look at MCDM: Choosing a decision method. Adv. Smart Syst. Res. **3**, 25–30 (2013)
11. Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis State of the Art Surveys. Springer, New York (2005)
12. Huang, I.B., Keisler, J., Linkov, I.: Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. Sci. Total Environ. **409**, 3578–3594 (2011). https://doi.org/10.1016/j.scitotenv.2011.06.022
13. Hwang, C.-L., Yoon, K.: Multiple Attribute Decision Making. Springer, Berlin, Heidelberg (1981)
14. Karczmarczyk, A., Jankowski, J., Wątróbski, J.: Multi-criteria decision support for planning and evaluation of performance of viral marketing campaigns in social networks. PLoS ONE **13**, e0209372 (2018). https://doi.org/10.1371/journal.pone.0209372
15. Nermend, K.: Metody analizy wielokryterialnej i wielowymiarowej we wspomaganiu decyzji (in Polish). Wydawnictwo Naukowe PWN, Warszawa (2017)
16. Piwowarski, M., Miłaszewicz, D., Łatuszyńska, M., et al.: Application of the vector measure construction method and technique for order preference by similarity ideal solution for the analysis of the dynamics of changes in the poverty levels in the European union countries. Sustainability **10**, 2858 (2018). https://doi.org/10.3390/su10082858
17. Roy, B.: Multicriteria Methodology for Decision Aiding. Springer, New York, NY (2013)
18. Wątróbski, J., Jankowski, J., Ziemba, P., et al.: Generalised framework for multi-criteria method selection. Omega **86**, 107–124 (2019). https://doi.org/10.1016/j.omega.2018.07.004

19. Zanakis, S.H., Solomon, A., Wishart, N., Dublish, S.: Multi-attribute decision making: A simulation comparison of select methods. Eur. J. Oper. Res. **107**, 507–529 (1998). https://doi.org/10.1016/S0377-2217(97)00147-1
20. Saaty, T.L.: Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process. RWS Publications, Pittsburg (2000)
21. Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburg (1996)
22. Cypryjański, J., Grzesiuk, A.: Expressing our preferences with the use of AHP: The game is not worth the candle? In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics, pp. 155–165. Springer, Cham (2018)
23. Ziemba, P., Wątróbski, J., Jankowski,J., Piwowarski, M.: Research on the properties of the AHP in the environment of inaccurate expert evaluations. In: Nermend, K., Łatuszyńska, M (eds) Selected Issues in Experimental Economics: Proceedings of the 2015 Computational Methods in Experimental Economics (CMEE) Conference. Springer International Publishing, Cham, pp 227–243 (2016)
24. Nermend, K., Piwowarski, M.: Cognitive neuroscience techniques in supporting decision making and the analysis of social campaign, pp. 1–13. Ishik Univeristy, Erbil, Iraq (2018)
25. Bear, M.F., Connors, B.W., Paradiso, M.A.: Neuroscience: Exploring the Brain, 3rd edn. Lippincott Williams & Wilkins, Baltimore, Md (2007)
26. Nunez, P, Srinavasan, R.: Electric fields of the brain (2006)
27. Malmivuo, J., Plonsey, R.: Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields. Oxford Univ. Press, New York, NY (1995)
28. Sharbrough, F., Chatrian, G.E., Lesser, R., et al.: American Electroencephalographic Society guidelines for standard electrode position nomenclature (1991)
29. Ahern, G.L., Schwartz, G.E.: Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. Neuropsychologia **23**, 745–755 (1985). https://doi.org/10.1016/0028-3932(85)90081-8
30. Coan, J.A., Allen, J.J.B.: Frontal EEG asymmetry as a moderator and mediator of emotion. Biol. Psychol. **67**, 7–50 (2004). https://doi.org/10.1016/j.biopsycho.2004.03.002
31. Harmon-Jones, E., Gable, P.A., Peterson, C.K.: The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. Biol. Psychol. **84**, 451–462 (2010). https://doi.org/10.1016/j.biopsycho.2009.08.010
32. Smith, E.E., Reznik, S.J., Stewart, J.L., Allen, J.J.B.: Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. Int. J. Psychophysiol. **111**, 98–114 (2017). https://doi.org/10.1016/j.ijpsycho.2016.11.005
33. Stewart, J.L., Bismark, A.W., Towers, D.N., et al.: Resting frontal EEG asymmetry as an endophenotype for depression risk: Sex-specific patterns of frontal brain asymmetry. J. Abnorm. Psychol. **119**, 502–512 (2010). https://doi.org/10.1037/a0019196
34. Nusslock, R., Shackman, A.J., Harmon-Jones, E., et al.: Cognitive vulnerability and frontal brain asymmetry: Common predictors of first prospective depressive episode. J. Abnorm. Psychol. **120**, 497–503 (2011). https://doi.org/10.1037/a0022940
35. Blackhart, G.C., Kline, J.P., Donohue, K.F., et al.: Affective responses to EEG preparation and their link to resting anterior EEG asymmetry. Pers. Individ. Differ. **32**, 167–174 (2002). https://doi.org/10.1016/S0191-8869(01)00015-0
36. Towers, D.N., Allen, J.J.B.: A better estimate of the internal consistency reliability of frontal EEG asymmetry scores. Psychophysiology **46**, 132–142 (2009). https://doi.org/10.1111/j.1469-8986.2008.00759.x
37. Cook, I.A., O'Hara, R., Uijtdehaage, S.H.J., et al.: Assessing the accuracy of topographic EEG mapping for determining local brain function. Electroencephalogr. Clin. Neurophysiol. **107**, 408–414 (1998). https://doi.org/10.1016/S0013-4694(98)00092-3
38. Davidson, R.J., Chapman, J.P., Chapman, L.J., Henriques, J.B.: Asymmetrical brain electrical activity discriminates between psychometrically-matched verbal and spatial cognitive tasks. Psychophysiology **27**, 528–543 (1990). https://doi.org/10.1111/j.1469-8986.1990.tb01970.x

39. Tops, M., Boksem, M.A.S.: Absorbed in the task: Personality measures predict engagement during task performance as tracked by error negativity and asymmetrical frontal activity. Cognitive Affect. Behav. Neurosci. **10**, 441–453 (2010). https://doi.org/10.3758/CABN.10.4.441
40. Pintrich, P.R., de Groot, E.V.: Motivational and self-regulated learning components of classroom academic performance. J. Educ. Psychol. **82**, 33–40 (1990). https://doi.org/10.1037/0022-0663.82.1.33
41. Csikszentmihalyi, M.: Flow: The psychology of optimal experience. Nachdr, Harper [and] Row, New York (2009)
42. O'Brien, H.L., Toms, E.G.: What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inform. Sci. Technol. **59**, 938–955 (2008). https://doi.org/10.1002/asi.20801
43. Mcmahan, T., Parberry, I., Parsons, T.: Evaluating Electroencephalography Engagement Indices during Video Game Play (2015)
44. Berka, C., Levendowski, D.J., Lumicao, M.N., et al.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviat. Space Environ. Med. **78**(Supplement 1), B231–B244(14) (2007a)
45. Freeman, F.G., Mikulka, P.J., Prinzel, L.J., Scerbo, M.W.: Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. Biol. Psychol. **50**, 61–76 (1999). https://doi.org/10.1016/S0301-0511(99)00002-2
46. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biol. Psychol. **40**, 187–195 (1995). https://doi.org/10.1016/0301-0511(95)05116-3
47. Gevins, A.: High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb. Cortex **7**, 374–385 (1997). https://doi.org/10.1093/cercor/7.4.374
48. Smith, M.E., Gevins, A.: Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. In: Caldwell, J.A., Wesensten, N.J. (eds.) Orlando, p. 116. Florida, USA (2005)
49. Yamada, F.: Frontal midline theta rhythm and eyeblinking activity during a VDT task and a video game: useful tools for psychophysiology in ergonomics. Ergonomics **41**, 678–688 (1998). https://doi.org/10.1080/001401398186847
50. Lin, C-T., Chen, S-A., Ko, L-W., Wang, Y-K.: EEG-based brain dynamics of driving distraction. In: The 2011 International Joint Conference on Neural Networks. IEEE, San Jose, CA, USA, pp. 1497–1500 (2011)
51. Almahasneh, H., Chooi, W.-T., Kamel, N., Malik, A.S.: Deep in thought while driving: An EEG study on drivers' cognitive distraction. Transp. Res. Part F Traffic Psychol. Behav. **26**, 218–226 (2014). https://doi.org/10.1016/j.trf.2014.08.001
52. Bajwa, G., Fazeen, M., Dantu, R.: Detecting driver distraction using stimuli-response EEG analysis. (2019). arXiv:190409100 [cs]
53. Berka, C., Levendowski, D.J., Lumicao, M.N., et al.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviat. Space Environ. Med. **78**(Supplement 1), B231–B244(14) (2007b)
54. Borawski, M.: Use of computer game as an element of social campaign focusing attention on reliability of information in the internet. In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics, pp. 127–139. Springer, Cham (2018)
55. Summerfield, C., Mangels, J.A.: Coherent theta-band EEG activity predicts item-context binding during encoding. NeuroImage **24**, 692–703 (2005). https://doi.org/10.1016/j.neuroimage.2004.09.012
56. Werkle-Bergner, M., Müller, V., Li, S.-C., Lindenberger, U.: Cortical EEG correlates of successful memory encoding: Implications for lifespan comparisons. Neurosci. Biobehav. Rev. **30**, 839–854 (2006). https://doi.org/10.1016/j.neubiorev.2006.06.009
57. Davidson, R.J.: What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. Biol. Psychol. **67**, 219–234 (2004). https://doi.org/10.1016/j.biopsycho.2004.03.008

58. Astolfi, L., De Vico, Fallani F., Cincotti, F., et al.: Neural basis for brain responses to TV commercials: A High-resolution EEG study. IEEE Trans. Neural Syst. Rehabil. Eng. **16**, 522–531 (2008). https://doi.org/10.1109/TNSRE.2008.2009784
59. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. Cogn. Emot. **23**, 209–237 (2009). https://doi.org/10.1080/02699930802204677

Chapter 15
Examining the Degree of Engagement of a Participant in Economic Games Using Cognitive Neuroscience Techniques

Konrad Biercewicz and Mariusz Borawski

Abstract The popularity of economic games is used mainly for learning. It causes a need for methods allowing to evaluate their content in terms of player's satisfaction already at the pre-production stage. There are already methods to determine a player's interest in a game, but they do not always allow for objective and unambiguous determination of a player's involvement. Cognitive neuroscience methods can give such an assessment. The aim of the research presented in this chapter is to develop a concept of a procedure for investigating a player's involvement in a game using cognitive neuroscience methods. The chapter presents the concept of the research procedure, the survey, the prototype of the game and the review of the engagement indexes. On the basis of the analysis of the results of the survey, it was stated, among others, that the respondents do not like to take a big risk related to money. Therefore, in the designed economy game, the player should be accustomed to taking risks in order not to be discouraged from the game. In addition, through the use of cognitive neuroscience, we are able to have knowledge of the level of engagement of the player in each part of the game. Then, game developers will be able to improve them in order to get the greatest satisfaction from the player. In the case of economic games, this will translate into a longer time spent by the player on the game, and thus his skills acquired during the game will be greater.

Keywords Economic game · Cognitive neuroscience · EEG · Engagement

15.1 Introduction

Business games support the integration of educational experiences at different levels of education and between different disciplines: economics, management and social

K. Biercewicz (✉) · M. Borawski
Department of Multimedia Systems, Faculty of Computer Science and Information Technology, West Pomeranian University of Technology Szczecin, Żołnierska 49, 71-210 Szczecin, Poland
e-mail: kbiercewicz@poczta.onet.pl

M. Borawski
e-mail: mborawski@wi.zut.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_15

sciences. Most games are played on modeling and simulation courses, microeconomics, logistics and management. Different applications of economic games allow adapting games of this type to different missions in schools, universities, corporate or training centers. It is, therefore, necessary to develop a methodology for the creation of economic games that will provide good entertainment for gamers while at the same time allowing them to understand certain economic processes and phenomena.

A simulation game can be used for this purpose. The task of this kind of computer games is to imitate the impression of different sensations that a person experiences in everyday reality. That kind of game is a combination of the following three elements [1]: game (created as a set of rules), role (assigned to each participant) and simulation. Simulation games are basically computer programs based on mathematical models, which are a simplified way to imitate specific economic processes or phenomena, such as running a business in a competitive environment. On the Internet, we can find many economic games.

A properly designed economy game should involve the player to a high degree. The most common methods used are the following commitment test methods: [2]:

a. Survey self-reported—the examined person determines the level of involvement in particular elements of the game;
b. Attention-based engagement (ABE)—the time during which the test person looked at the monitor is measured. Relating this to the time when the person did not look at the monitor, one can deduce how much attention of the tested person was focused on the game;
c. Electrodermal activity (EDA) otherwise called galvanic skin response (GSR)—allows to determine the emotions of the person tested on the basis of the measurement of skin conductivity;
d. Mouse clicks, mouse movement—measurements of the number and places of clicks and mouse movement allow you to determine the level of player activity during the game.

The above methods are very limited. In the survey self-reported, the researcher relies on the observations of the respondent. The test person may have difficulty in remembering his or her feelings during the whole game. It is difficult to determine the exact time frame within which the growth of interest in the game begins and ends on the basis of this study. In addition, the respondent may formulate the answers in such a way that they fit into specific social expectations. For example, there are elements of violence, so the respondent may be negative about it, even though they actually aroused his or her interest. Attention-based engagement (ABE) depends not only on the player's engagement but also on the type of game and the situation in the game. During the fight, the player's focus on the game will be very high, because he/she has to react quickly to the opponent's actions, while during the wandering around the city or conversations with NPCs can be smaller. If the game itself is interesting, but the fights are monotonous and boring, the player will still be very focused during the fight, so as not to die and be able to continue playing. ABE will not say anything about the attractiveness of fighting during the game.

Galvanic skin response (GSR) allows you to define emotions in the first place. However, not always a player's involvement can be emotional. Certain elements of the game may not generate emotions until some success or failure is achieved.

Mouse clicks, mouse movement are strongly dependent on the scenario of the game itself. They can be useful if you are able to refer to other players. You can tell from them which player is more involved and which one is less involved. For example, during a fight, the mouse movements will depend on the weapon chosen by the player and the way the opponent fights. They may, therefore, be incomparable between opponents.

It is necessary to look for such methods of engagement research that will allow determining the level of engagement at any time in the game, while not being dependent on other factors. An example of such methods is a method of cognitive neuroscience. They are becoming more and more useful because they allow us to get to know the current state of the brain. This task is facilitated by the indexes calculated on the basis of the recorded signals. In the literature on the subject, numerous indices of engagement can be found, which will be presented later in this chapter. They allow us to know the level of human involvement in a given activity in a given moment of time.

The aim of the research presented in this chapter is to develop a concept of a procedure for investigating a player's involvement in a game using cognitive neuroscience methods. The chapter will present a diagram of the research procedure, the preliminary results of the research, i.e., the results of a survey aimed at getting to know the preferences of players and, additionally, a review of commitment indices. A prototype of the game will also be presented, which will be used for further research.

15.2 Proposal for a Test Procedure

The test procedure proposed in the test is shown in Fig. 15.1. The first stage of the test procedure is to formulate the problem. It is presented in the introduction to this chapter. The second stage of the research procedure is to prepare a questionnaire. The survey aims to find out the preferences of players—their expectations of the game content. This will make it possible to enrich the economic game with elements that will arouse their interest. The results of the survey will be the basis for writing a scenario of the game and its design.

Checking how much the game can involve players will require creating a game. The game will be played using the Unity game engine in C# language. It will be executed in such a way that it is possible to record events in the game in order to synchronize them with the registered EEG signals. The EEG signals will be registered in a group of 30 people. First, a pilot study will be carried out on several people to verify the correctness of the registration, synchronization, etc. After the pilot study and possible correction of all errors, a proper study will be carried out.

The recorded signals will be used to calculate the EEG indices and specific game elements that cause low player engagement. If such elements are detected, the game

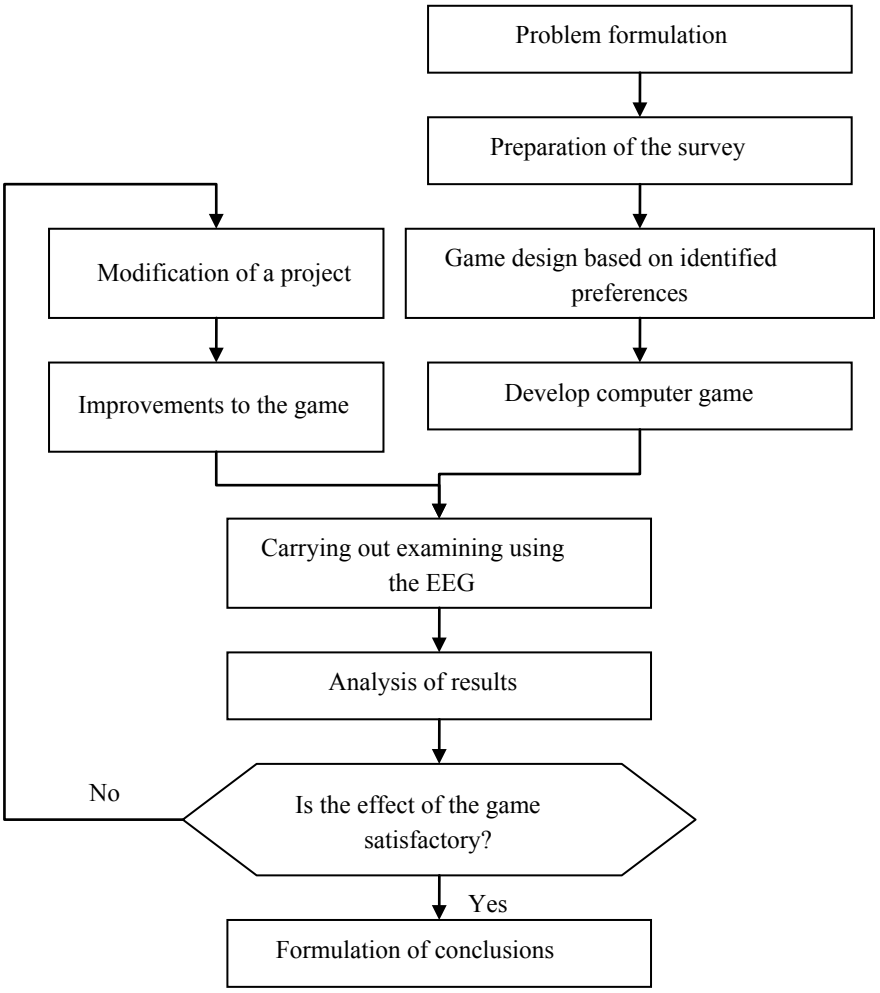


Fig. 15.1 Proposal for a test procedure

design and the game itself will be amended and the examining will be repeated. This part of the procedure will be repeated until satisfactory results are obtained. Finally, conclusions will be drawn on how to design the game with the greatest possible player involvement in it.

15.3 Questionnaire Survey

Among 52 people (35 men and 17 women), a survey was carried out to determine what type of games people like to play for recreation (frequency of playing = less than once a week) and what elements of these games should be implemented to create an economic game in order to stimulate greater engagement. The questionnaire was divided into four parts:

- the first part included questions evaluating previous experiences with video games and other personal characteristics;
- the second part included questions related to the evaluation of the activities performed in the computer game;
- the third part was the so-called Case Study, which consisted of imagining the situation and making the best decision;
- the fourth part of the questionnaire, the individual activities in the game had to be sorted from the most interesting to the least interesting.

Using the above division of questions, the answers to the questions asked for the purpose of the survey were obtained.

The analysis of the questionnaire showed that the average age of the respondents was 25 years. The obtained target group is an ideal group because they are people who have already played several games in their lives. So, they can clearly tell you what their expectations are in relation to virtual gameplay. Therefore, we can assume that the answers given have to be taken into account in order to achieve the result of greater engagement.

The participants were asked to answer the question of what factors influence the game that they get bored after some time. The most frequent answers were monotony, boring plot, too much difficulty, lack of competition or ugly graphics. In addition, it was found that those surveyed liked to play games related to Esport, such as Counter-Strike: Global Offensive, League of Legends, but also strategic games.

Analyzing the part related to the evaluation of the activities in the game, (Figs. 15.2, 15.3 and 15.4) it was noticed that the majority of the respondents in a game likes the management of the military. Similar interest is connected with controlling expenses related to the development of the city and with building and development of the city. On the contrary, you can see that most respondents do not like investing money on the stock market. Generally, they do not like taking risks in situations related to money. At this point, the question should be asked whether they do not like it because they are not properly educated, or it is due to the fact that they may lose too much cash.

In Part III—Case Study, it was learned that the most opinion responders (almost 70%), who would find themselves in a situation where the investment brings them losses. They would be cutting the losses and waiting for the situation to develop further. A similar number of respondents answered the question related to the case of what would you do when your village will be destroyed by the enemy. They said: "I will rebuild the village and set out to fight the enemy." On the other hand, ¾ of

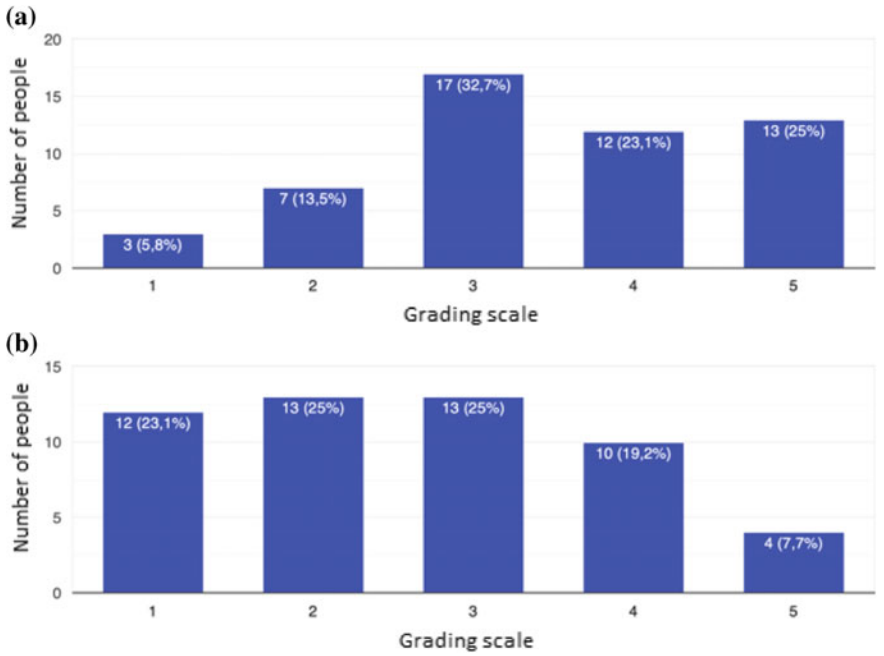


Fig. 15.2 Diagram presenting the opinion on **a** military/tank/people/etc., management, **b** investing money on the stock exchange

the respondents said that if they landed with a second person on the planet which is likely to be uninhabited, they would find a place where they can settle down. Also, it is worth looking at the following question: "On your way, you meet an NPC who has an item you are interested in. What action will you choose to get it (note that choosing one action rules out the possibility of choosing another one)": fight, exchange, theft, dice roll?

Most of the respondents responded rationally, i.e., they chose to exchange goods for goods (Fig. 15.5). This is probably due to the fact that most of the people questioned did not choose to roll dice, because there is a random factor associated with this game. Then, we do not have a 100% guarantee that we will win, and in addition, we may lose everything. Similarly, it is worth looking at the last question: "What would you invest your money in if you found an oil field and started cooperation with sheiks?" (Fig. 15.6) Most of the respondents chose the answer—the money they will receive, they will invest in real estate. This may be due to the fact that it is currently a form of investment, thanks to which we obtain a passive profit.

The last part of the analysis, i.e., the fourth part, will be devoted to how the survey participants sorted the activities in the game, where 1—means the most interesting and 6—the least.

Looking at Fig. 15.7, you can see that waiting for buildings to be built is the most interesting, but also the least interesting. What could it be caused by? Well, some

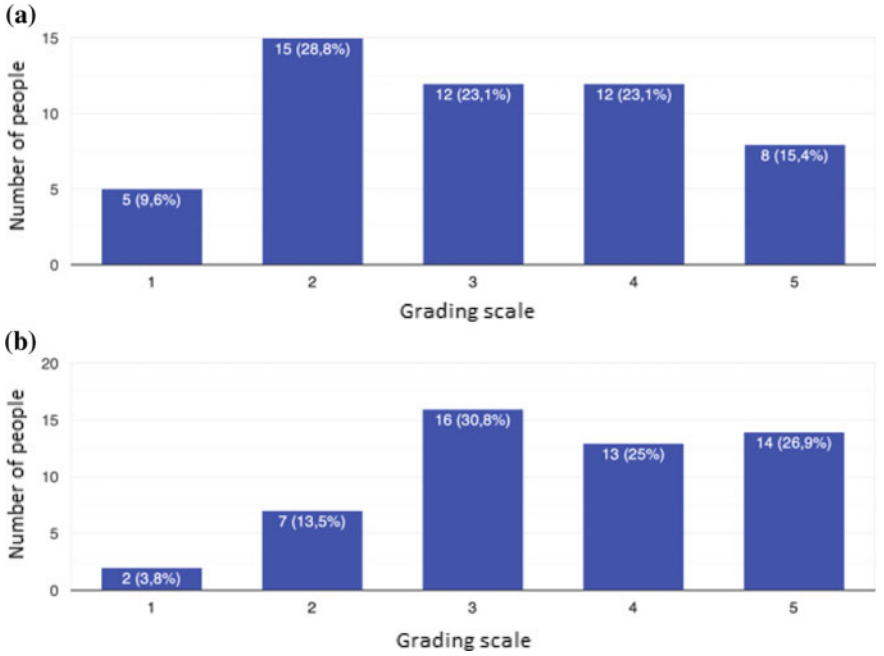


Fig. 15.3 Diagram presenting an opinion on taking risks in money situations, **b** controlling expenses related to the development of the city/planet

people who say that it is the most interesting cannot wait to build a building, thanks to which, for example, we will be able to produce much better armor and swords, which in the end can translate into defeating the enemy. Therefore, we are looking forward to the successful completion of the building. On the other hand, it does not cause emotions, because it is just an ordinary building. It is only an instrument for further action. An interesting activity is also buying new things.

15.4 The Conception of the Economic Game

On the basis of the survey, an economic game was designed, which will probably be more engaging than the games that exist on the market so far. In order to check this, we will invite people to the study to play a game already existing on the market and design. For examining the engagement, we will use EEG and indices of the engagement which is the more later part of the chapter will be introduced. We know that we cannot create a typical stock market game because it will be monotonous, and there will be no plot or if it is not very interesting.

The aim of the game is to colonize, explore the planet Mars and make alliances or wage wars with other planets. A mechanism will be put in place to build buildings,

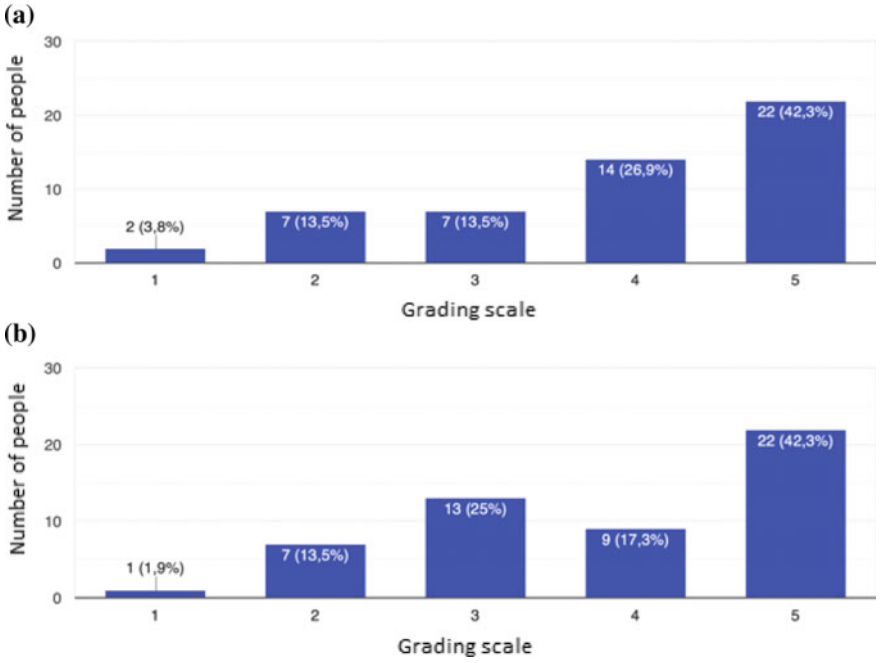


Fig. 15.4 Diagram presenting an opinion on **a** building a town/village from scratch, **b** developing a town/village/etc.

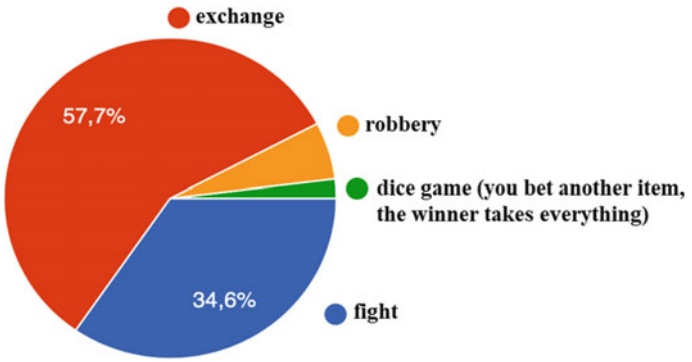


Fig. 15.5 Circular diagram showing your opinion of what you will do in a situation when you meet an NPC on your way with an object you are interested in

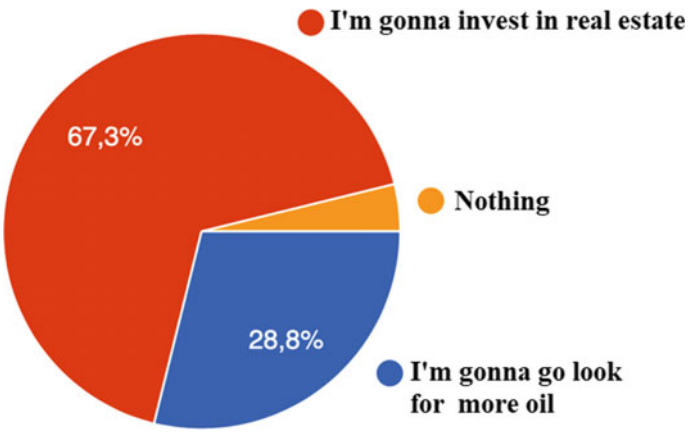


Fig. 15.6 Circular diagram showing your opinion of what you will do in a situation when it turned out that you came across a crude oil plant and started cooperation with sheiks. What will you do with the cash you receive?

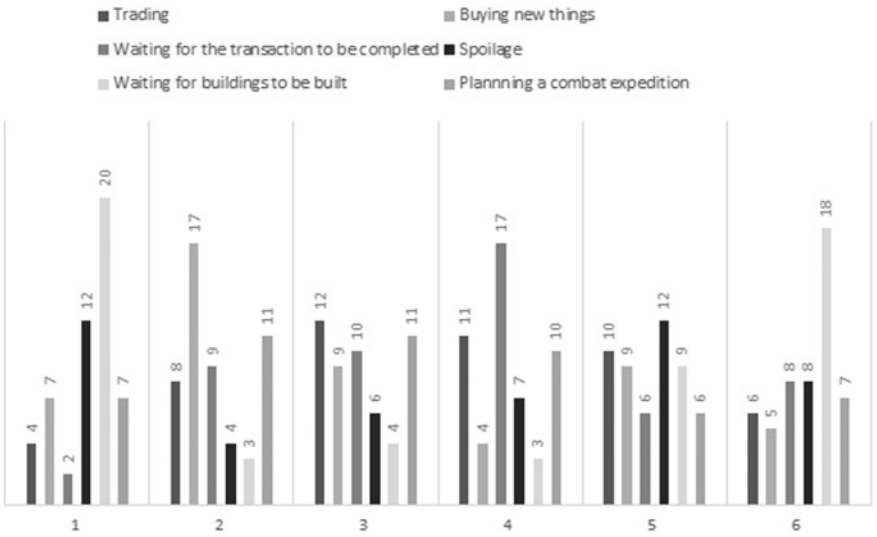


Fig. 15.7 Bar chart showing the sorting activities in the game

weapons, ships. In addition, there will be an opportunity to trade or you will be able to bet your items and roll dice against the opponent. Winner takes everything. There will also be implemented a stock exchange system. If a planet is at war or trading, for example, the exchange rate will be unstable. The news will be published as if it were a realistic stock exchange, as well as various stock market indicators available for us to use to make decisions. There will be a tutorial on how to use such indicators.

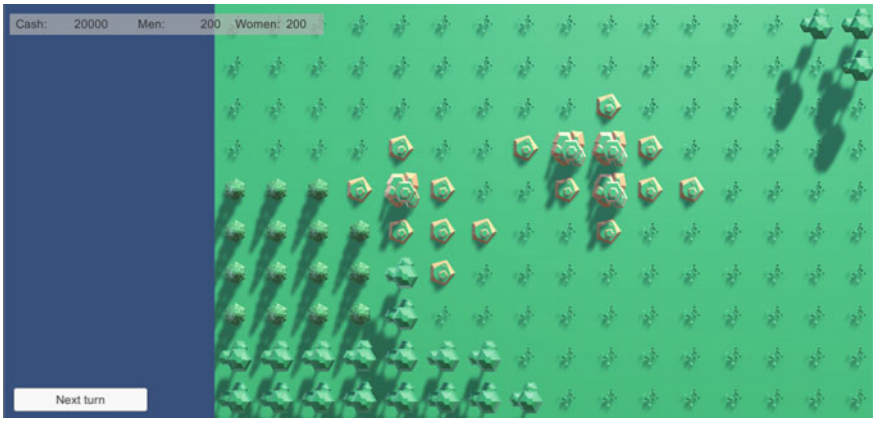


Fig. 15.8 Game board

In addition, the player will be accustomed to taking risky transactions, i.e., at the beginning the risk will be very low, but with each stage in the game it will increase. Apart from the stock exchange and trade exchange, it will be possible to invest virtual money, e.g., in real estate on a friendly planet, modern technology.

In the beginning, a part of the game containing the most important elements of the game will be realized and examined. The main element of the game is a board with a terrain (Fig. 15.8). It consists of square tiles symbolizing different types of terrain. Each type of land has a different cost of preparation for construction. The player must choose the area of the land in a thoughtful way so that he has enough money to build the buildings necessary for the functioning of the city.

The player can choose between different types of terrain. Four types were selected for preliminary testing: housing estate, cultivated fields and two types of factories (Fig. 15.9). At the beginning of the game, the player has a certain number of men and women. He/she has to create enough housing estates to accommodate all the residents. A certain amount of space is needed, as the population may increase if the conditions are favorable. Farmland is required to feed the population. In turn for agricultural production are required materials and machines which are produced by factories.

The game is a turn-based. One tour covers one year. Once the player has completed all the intended actions, he presses the "Next turn" button, which starts the simulation. Multi-agent simulation was used in the game. Figure 15.10 shows a diagram of classes inheriting from the TAgent class. These classes are used to create objects that are agents in this simulation.

The TAMan and TAWoman classes are represented by men and women, respectively. They can join together to form a TAHousehold property after the TAHousehold class representing the household. The creation of a household is carried out in three

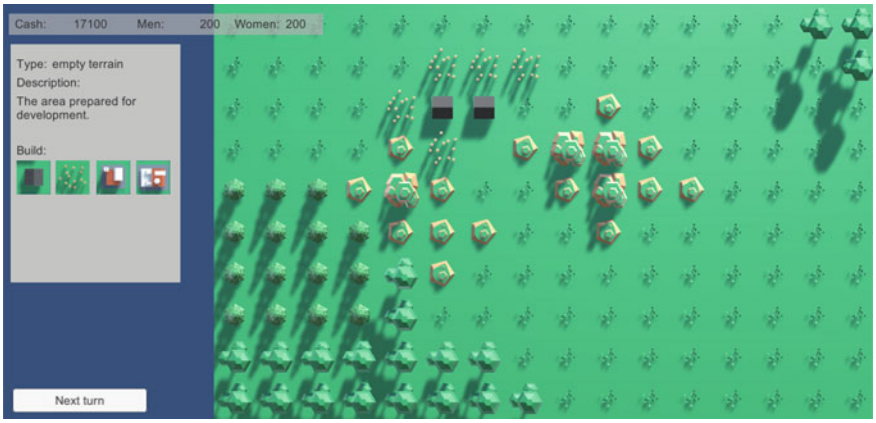


Fig. 15.9 Types of areas available to the player

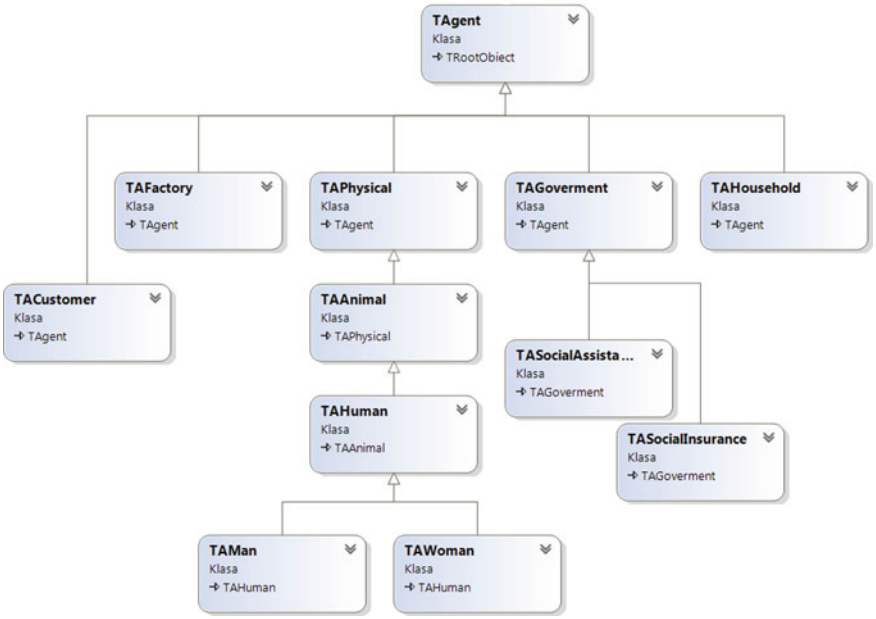


Fig. 15.10 Diagram of classes showing classes inherited from the TAgent class

steps. In the first step, agents created in TAMan and TAWoman classes send matrimonial offers. The offer can only be sent after the age of 18 and is randomly selected depending on the age of the agent.

After reaching the age of 18, they start sending each other matrimonial offers. The sending or acceptance of an offer depends on the probability, which is determined by the age of the agent. The offer includes expectations toward the partner. Currently, they cover the expected age range of the partner. In the next stage, the offers are delivered to the addressees. Addressees younger than 18 years of age ignore the offers. Addressees aged 18 or over shall check whether the broadcaster is within their expected age range, if so, they shall draw lots with a probability depending on age and whether they respond to the offer. At the last stage, the senders of offers receive responses. After selecting one of the offers (the first one that came is currently selected), the household agent is created. Female and male agents separate themselves from previous households and join the newly created one, receiving the status of husband and wife. From now on, if both partners are alive, they will not send or receive matrimonial offers.

For a woman who has the status of a wife in the household, the number of children to be born is drawn by lot. Furthermore, in every step of the simulation one draws whether she got pregnant or not. After nine months, a new agent is created for whom the sex is drawn. This agent is assigned to the household as a child. The minimum interval between successive pregnancies is assumed to be one year. For every agent representing, a man from his birth in every lottery is drawn whether or not he died. For every agent over 18 years of age, there is a draw for a profession. The level of earnings is related to the profession.

In addition, every human agent must consume certain products. The source of the products is the factories represented by the TAFactory class. Each factory can have any number of production lines on which it manufactures its products. For each production line, it is determined how many products can be produced per unit of time, as well as the demand for raw materials and employees. Any product manufactured in a different factory can be a raw material. There is also a possibility of production without the need to obtain raw materials. An example of such an object is arable land, which is a simulation that requires only human resources. Each production line requires the assignment of a certain number of employees in the right professions.

The prices of the products are determined in advance; however, the course of their changes over time can be taken into account. Before starting production, the factory checks that the production costs do not exceed the market price. If there is any possibility to earn money from production, the factory sends out job offers. When the number of employees is sufficient to create one product unit, the factory purchases raw materials and begins production. The production lasts as long as it is profitable or until the warehouse is overfilled. Cessation of production results in the dismissal of employees. Overfilling of the warehouse may be a result of a lack of orders for products.

Orders for goods can be placed by people, other factories or objects created after the TACustomer class. The object represents the recipients of the goods outside the simulation. This object buys a certain number of goods in a unit of time. This number

may vary over time. People have to buy certain goods for every certain unit of time. They will use the money from the household budget for this purpose. Every member of the household who earns money in factories contributes it to the common pot. At the end of his professional career at the age of 65, an agent retires. He will receive the money from a TASocialInsurance class facility that represents the Social Insurance Institution (ZUS). Money is paid out from the age of 65 until the death of the agent. If the income of household members is too low, the household will receive a benefit from a TASocialAssistance class object representing social welfare. The condition for receiving the benefit is that the household member's income falls below a certain threshold.

15.5 Study of EEG

After the creation of the game, a study will be conducted to investigate the involvement of players. Referring to the scientific literature, several studies have used questionnaires that are not entirely a good form of investigating a player's experience [3]. The problems arise from the formulation and context of these forms [4]. Therefore, the EEG should be used to determine the involvement of the user. Frequency bands are determined from the EEG signal using a spectral method, e.g., Fourier transform. The EEG measured on the scalp corresponds to a recording at frequencies of 0.5–30 Hz. Four basic bands are recognized in this range [5]:

- Delta (0.5–4 Hz).
  The brain waves of the delta are generated in the deepest meditation and sleep. Delta waves suspend external consciousness and are a source of empathy. In this state, treatment and regeneration are stimulated, which is why deep restorative sleep is so important for the healing process.
- Theta (4–8 Hz).
  Theta waves occur most frequently in sleep but are also dominant in deep meditation. Theta waves are noticed during learning or remembering.
- Alpha (8–12 Hz).
  Alpha activity is best seen in the back regions of the brain and is typical for relaxation. It occurs when closing the eyes.
- Beta (12–30 Hz).
  The beta activity can be divided into: low activity waves (12–15 Hz), medium waves (15–20 Hz) and high waves (18–30 Hz). The average range of beta activity is associated with increased energy, anxiety, performance and concentration. The most visible is in the leading regions.

Using the EEG device, we are able to determine which moment of the game is not very interesting, and we can improve it so that the player is fully active in the game. New EEG devices are increasingly being used outside of medicine and are finding more and more new applications.

The resulting data will be analyzed using MATLAB. Events such as blinking of the eyes, head movements or body movements may cause undesired EEG registration data. Most EEG analyses require the removal of such events in order to identify medical problems. However, this is not a problem to analyze the game. Such events are common in everyday play [6].

The EEG spectral signal will be analyzed using a Fast Fourier transform (FFT) and an overlapping three-second time frame with a one-second jump for Delta (1–4 Hz), Theta (4–7 Hz), Alpha (7–13 Hz), Beta (13–25 Hz), Gamma (25–43 Hz), Delta-1 (0,5–3,9 Hz), Theta-1 (4–7,9 Hz), Alpha-1 (8–10,9 Hz), Alpha-2 (11–13,9 Hz), Beta-1 (14–19,9 Hz), Beta-2 (20–29,9 Hz), Gamma-1 (30–35 Hz).

Measuring the level of engagement is one part of determining the player's experience when playing a computer game. Pope [7] has demonstrated that the exposure index can be calculated by assuming the ratio of the EEG bands (Beta/(Alpha + Theta)) [Index 1] (Table 5.1). Gevins and Smith [8] have introduced another task engagement index that analyzes Theta's frontal lobe against Alpha (Theta's frontal lobe/Alpha's parietal lobe) [Index 2]. The third index has been identified by Yamada [9], which analyzes the activity of the frontal lobe Theta [Index 3], which indicates increased attention. The fourth and fifth indexes use the same formula as Pope, but Hockey [10, 11] used a different distribution of electrodes.

Table 15.5.1 Description of the indices used in the examine

Index number	Formula	Method of calculation
Index 1	Beta/(Alpha + Theta)	Average registration value of all electrodes on the head
Index 2	Theta/Alpha	Average registration value from electrodes placed on the frontal lobe of Theta and parietal lobe of Alpha
Index 3	Theta	The average value of registration from electrodes placed on the frontal lobe of Theta
Index 4	Beta-1/(Alpha-1 + Theta-1)	Average registration value from electrodes: F3, F4, F7, F8, Cz, P3, Pz, P4
Index 5	Beta-2/(Alpha-2 + Theta-1)	Average registration value from electrodes: F3, F4, F7, F8, Cz, P3, Pz, P4

• Index 1 [12]: Beta/(Alpha + Theta) has been calculated for each participant using the following electrodes: AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2
• Index 2 [12]: Theta/Alpha has been calculated using the average registration value from electrodes placed on the frontal lobe of Theta: F3, F4, FC5, FC6 and divided by the average registration value from electrodes placed on the parietal lobe of Alpha: P7, P8
• Index 3 [12]: Theta has been calculated using the average registration value from the electrodes placed on the frontal lobe of Theta: AF3, AF4, F3, F4, F7, F8, FC5, FC6
• Index 4, 5 [13]: Beta/(Alpha + Theta) has been calculated using the average registration value of the following electrodes: F3, F4, F7, F8, Cz, P3, Pz, P4

15.6 Examining Procedure

Before starting the study, the subjects will be informed about the course of the study. They will then sign a consent form and sit in a comfortable chair with access to the keyboard and mouse. The next step is to put on a cap and connect the electrodes to the skin of the participant's head and connect them to a device that will record data from the participant's brain. After completing the above steps, the tested person will be informed what the game will be about and how to move around in it. Immediately after the end of the game, each participant will be interviewed about their computer and gaming experience and which elements of the game made them more (or less) involved in the game.

Before the game starts, a black screen will appear to silence the player, and after 60 s, a spider will appear on the screen. Its purpose is to cause a reaction in the tested person, which will be reflected in the measurement and will be used to shift the graph by the time of his response. Each participant's game will be intercepted at a resolution of 1360×768 by means of a programmed in-game registration. Each screenshot will generate a time stamp for EEG data to determine the position of the beginning and end of each section. In addition, screenshots will be taken, which will be saved for later reference during the data analysis phase. After the game is over, an Excel file will be saved, in which the screenshot number and the corresponding time will be stored. In addition, during the game, events (such as keyboard presses) will be recorded and saved to an Excel file. In addition to EEG, we will also use Eye Tracker in the study to track what the tested person pays particular attention to.

15.7 Conclusions

This chapter presents a proposal for a study of player involvement using the EEG and the preliminary results of the study. Investigating the player's involvement is a very important task because it allows you to identify the elements of the game that bore the player, and thus reduce the attractiveness of the game. Removing or improving these elements may translate into a commercial success of the game. The use of an EEG signal in the registration procedure makes it possible to identify unambiguously which parts of a game the player is involved in, and to decide whether to remove or correct them.

Further work will include the selection of engagement index(es) and more extensive research on the 30-person group. Each of the examined persons will take part in a 30-minute game during which the changes in their level of involvement will be examined. Once the least engaging parts of the game have been identified, they will be modified. The procedure is iterative, it is repeated until the satisfactory result is achieved, i.e., the appropriate degree of involvement of the player in the game.

Thanks to the knowledge of the level of involvement of the player in each part of the game, game developers will be able to improve them in order to achieve the

greatest possible satisfaction of the player. In the case of economic games, this will ensure a longer time spent by the player on the game, and thus his or her skills acquired during the game will be greater.

References

1. Latuszyńska, M.: Experimental research in economics and computer simulation. In: Nermend, K., Latuszyńska, M. (eds.) Selected Issues in Experimental Economics, pp. 151–169. Springer, Cham (2016)
2. Martey, R.M., Kenski, K., Folkestad, J., et al.: Measuring game engagement: Multiple methods and construct complexity. Simul. Gaming **45**(4–5), 528–547 (2014). https://doi.org/10.1177/1046878114553575
3. Chiang, Y., Cheng, C., Lin, S.S.J.: The effects of digital games on undergraduate players' flow experiences and affect. In: 2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning, pp. 157–159 (2008). https://doi.org/10.1109/digitel.2008.11
4. Witmer, A., Slater, M.: Measuring presence: A response to the Witmer and singer presence questionnaire. Presence Virtual Augmented Reality **8**(5), 560–565 (1999). https://doi.org/10.1162/105474699566477
5. Hosťovecký, M., Babusiak, B.: Brain activity: Beta wave analysis of 2D and 3D serious games using EEG. J. Appl. Math. Stat. Inform. **13**(6), 39–53 (2018). https://doi.org/10.1515/jamsi-2017-0008
6. Plass-Oude Bos, D., Reuderink, B., van de Laar, B. et al.: Brain-computer interfacing and games. In: Brain-Computer Interfaces: Applying Our Minds to Human-Computer Interaction, pp. 149–178. (2010). https://doi.org/10.1007/978-1-84996-272-8_10
7. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biol. Psychol. **40**(1–2), 187–195 (1995). https://doi.org/10.1016/0301-0511(95)05116-3
8. Smith, M., Gevins, A.: Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. Proc. SPIE Biomonitoring Physiol. Cogn. Perform. During Mil. Oper. **5797**, 116–126 (2005). https://doi.org/10.1117/12.602181
9. Yamada, F.: Frontal midline theta rhythm and eye blinking activity during a VDT task and a video game: Useful tools for psychophysiology in ergonomics. Ergonomics **41**(5), 678–688 (1998). https://doi.org/10.1080/001401398186847
10. Hockey, G.R.J., Nickel, P., Roberts, A.C., Roberts, M.H.: Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control. Appl. Ergonomics **40**(6), 1011–1018 (2009). https://doi.org/10.1016/j.apergo.2009.04.008
11. Nassef, A., Mahfouf, M., Linkens, D.A., et al.: The assessment of heart rate variability (HRV) and task load index (TLI) as physiological markers for physical stress. In: Dössel, O., Schlegel, W.C. (eds.) World Congress on Medical Physics and Biomedical Engineering, September 2009, Munich, Germany, pp. 146–149. Springer, Heidelberg (2010)
12. McMahan, T., Parberry, I., Parsons, T.D.: Evaluating player task engagement and arousal using electroencephalography. In: 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences vol. 3, pp. 2303–2310. (2015). https://doi.org/10.1016/j.promfg.2015.07.376
13. Kamzanova, A.T., Matthews, G., Kustubayeva, A.M., Jakupov, S.M.: EEG indices to time-on-task effects and to a workload manipulation (Cueing). World Acad. Sci. Eng. Technol. **80**, 19–22 (2011)

Chapter 16
Simulation Modeling to Verify the Efficiency of Information Transfer

Katarzyna Rostek and Radosław Zając

Abstract Research problem discussed in this chapter refers to simulation model for verification communication effectiveness. Communication and knowledge management processes usually suffer from susceptibilities or malfunctions. These result in delays, communicate distortion and reduced efficiency of knowledge-to-use transformation. Thereupon, organizational goals' achievement, staff engagement and overall management efficiency are lowered. Hence, there is a need for comprehensive, ready-to-use model for application of communication and knowledge management processes. Model that allows to determine assumptions. Boundaries and simulating behavior of real communication and knowledge management processes. In this chapter, only graph modeling part derived from general ATENA model will be discussed.

Keywords Organizational communication · Communication · Communication management

16.1 Introduction

Communication, especially meant as distribution of information, retention and conversion information into knowledge or action, becomes essence of organizational performance and existence. Drucker's concept of knowledge workers highlights role of information and knowledge management in organizations that base on this kind of labor. Rapid changes reflected in idea of turbulent business environment [1] in organizational and individual perspective means that organizations, teams and every one team member should learn how to learn new—acquire knowledge and forget old—get rid of knowledge no longer up-to-date. Compared to human ability to remember,

K. Rostek (✉) · R. Zając
Faculty of Management, Warsaw University of Technology, Warsaw, Poland
e-mail: katarzyna.rostek@pw.edu.pl

R. Zając
e-mail: radoslaw.zajac@pw.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_16

understand, learn and put into practice, this trend clearly leads to concept of organization as an information and knowledge ecosystem. Both as an organizational core value for all the stakeholders and as a solid base needed for existence of any given organization at the same time.

Regarding this point of view, key question for any organization becomes: How to model and assess communication management effectiveness within organization.

This will allow organization to manage, monitor and control processes lying beneath various management and organizational domains. In fact communication and knowledge management are not only immanent parts of those domains—but determine efficiency and stability of all organizational levels.

It that approaches organizations should define measures (KPIs) for communication effectiveness and efficiency. At the same time, organizations should learn how to model communication processes to ensure synchronization of all the connected areas, as communication and knowledge management can be seen as an inherent part of management, training, innovation and change management. Thus, main goals of the study will be:

G1: Identification of set of measures for communication effectiveness and efficiency.
G2: Building enhanced graph model for communication processes and using it in simulations.

Those goals might be transformed into the following hypothesis:

H1: Definite set of measures enables to assess communication model effectiveness and efficiency.
H2: Graph model might be used to model and forecast communication process performance and obstacles.

Second goal may be treated as a utility for the first one, but in fact graph model enables formulating set of additional metaphors that, when analyzed back in organizational terms, highlights some communication issues that cannot be easily seen either in business or in scientific approach focused on standard models for communication (i.e., Shannon).

From mathematical point of view, organizations, networks and graphs might be treated as one entity—based on nodes and edges—and this aspect is discussed in the first chapter referring to communication and networks. Next chapter, presenting ATENA model, introduces background for core analysis of this chapter. Measures for communication process chapter introduce proposal of definite set of exact parameters derived from ATENA model. Next two chapters discuss, respectively, assumptions for graph model of communication processes and simulation results. Last chapter is about conclusions regarding initial goals and hypothesis and further research steps.

16.2 Communication and Networks

Research on communication explores two main areas: one-on-one communication with all the personal and subjective attributes and more general network-like communication. First approach concentrates on synchronizing author and recipient, rate of similarity and psychological aspects. Second approach is based on network or graph definition and dependencies—nodes and edges. For some time, there started some research combining elements of those approaches—in fact, different points of view define organizational reality.

Nordin [2] discusses interrelations between conflict management, communication climate and communication management. Very important from efficiency point of view, only specific open-like culture based on trust and certain level of tolerance allows to share knowledge and doubts. Only then, network (graph) can work properly and efficiently—when information, knowledge and question feed are based on curiosity, engagement rather than facade actions or common social behavior. Koçoğlu et al [3] investigate organizational symbols as element of collective unconscious and scripts—enabling to embed communicates in the context. This is crucial to both—embed single communicates in continuity of organizational strategy and prevent from bastardization organizational language itself. When truth or success will be several times equated with tricks and corporate corpo-language covering failures, where corpo-language means jargon used to create corporate community identity—like challenges instead of problem and all the buzzwords used by white collars within corporations. Redza [4] investigates interconnections between communications: willingness, behavior, quality and commitment, presenting how those soft aspects affect communication process and its efficiency.

Modern organizations, products and services require continuous, efficient collaboration. King [5] discusses transactive memory systems (TMS) as measure for teamwork and team performance. Faraj [6, 7] refers to knowledge collaboration in online communities. Kleinsmann [8] discusses team knowledge integration needed for new product development. Patterson [9] studies team communication and highlights aspect of communication centralization or decentralization. Meltzer [10] connects social methods directly to quality improvement. All those researches reflect inevitable trend of expert team-centered, partially virtualized communities needed for sustainable organizational existence and continuous development resulting in all kinds of innovation (product, process, organizational).

Researchers extensively investigate criteria and measures for efficient communication from social network analysis, such as clusters/partitions (Meltzer)—that might be addressed by model P described in this chapter. Others [8] highlight interfaces and channel aspect. That structures communication model and helps to determine communicate type or purpose.

On the other hand, top-down, Faraj [7] stipulates integrated view of collaboration and communication in communities. Similar approach, anchoring communication in other domains, is presented by Borca [11], where successful internal organizational communication depends on: decision making by management, the way the

organization is structured for communication, how organization distributes information internally and allowing company and allowing employees the ability to adapt to technology. Madalina [12] discussed types of conflict: conflict messages/semantic problems, use of time/learning period necessary, individual transmitting/receiving messages, the managerial style, and employees use their time than it should do, and the time required to, difference of perception. Those aspects "translated" into communication and knowledge management directly identify issues affecting efficiency and effectiveness in this area.

Those papers represent incremental evolution of research and organizational self-awareness. Interdependence between levels—individual, team and organization, aspects—culture, processes and applications is more and more investigated and becomes tangible. As well as embedment of communication models in organizational reality and impact on other organizational and managerial domains. That cannot exist and work in stable manner, as well as be developed without efficient communication processes. And, at the same time those domains affect heavily communication efficiency. For example innovation or training should base on communication and knowledge management because innovation needs to grow out of previous knowledge and training is in fact a channel of communication and knowledge transfer.

As a result, both directions should be developed parallelly:

1. Framework for organizational communication (and conversion communicate into knowledge and application).
2. Measures for communication models.

This second aspect is the key goal of this chapter, but it must be embedded in broader framework—linking single measures and enabling interpretation of single parameters and its specific values.

In addition, framework for creating models for communication and knowledge management has been developed by Zajac from 2016. Figure 16.1 presents concept of this model.

16.3 ATENA Communication Model

ATENA model was created as a result of not working partial models or approaches for communication and knowledge management. CKM differs from other management domains because in fact it is a tool and essence of them. Without communication and knowledge, no organization can exist, neither corporation nor group of people.

This diffusion heavily affects both sides of each domain—especially communication and knowledge management. It results in many failures, and malfunctions which analyzed separately—as effect detached from actual reasons and sources—do not allow to solve those problems.

ATENA communication model (Fig. 16.1) consists of four components:

1. Principles—assumptions coming from good practices, being axioms enabling to distinct various aspects of organizational culture.

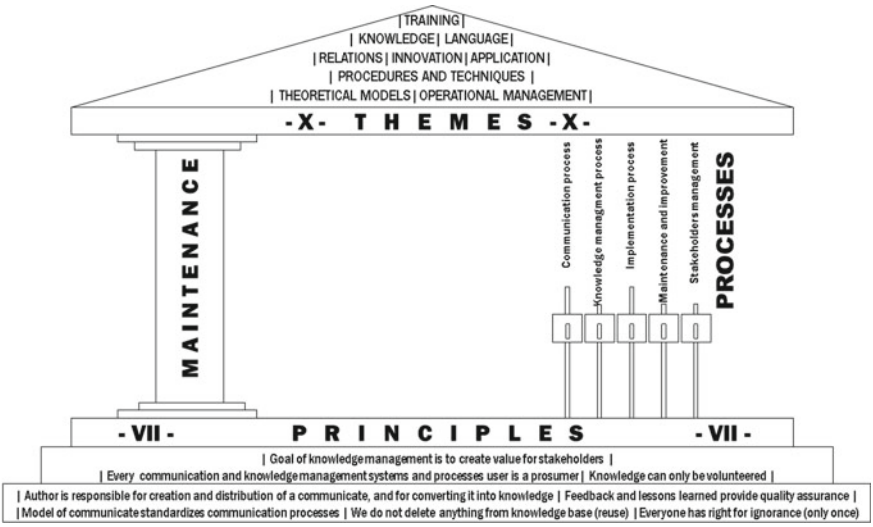


Fig. 16.1 ATENA communication model

2. Themes—aspects to be monitored, controlled and synchronized with communication and knowledge processes and—more general—management.
3. Processes—describing high-level processes for communication and knowledge management (CKM).
4. Maintenance—key process and approach making CKM working or not.

That is origin of model ATENA's structure similar to four components of PRINCE2 and general construction of any other model. That is also the reason for introducing finite set of comprehensive measures for communication model estimation in order to control CKM improvement or ineffectiveness. On the other hand, ATENA-based measures coming from organizational reality create superset for graph model being core of this model. Next chapter presents those measures and their operationalization.

Some of those parameters are included in graph model (network model) being analyzed in this chapter. In other words, ATENA model and superset of measures are necessary for embedding graph model in organizational context and assure complex view of CKM domain.

Proposed measures might be used by either scientists or communication and knowledge managers to estimate initial values justifying starting project for redesigning CKM domain—especially communication processes. During the project, those parameters might be monitored, as well as at the end of the project—to assess benefits coming from project product implementation. The same measures should be used during maintenance stage—in order to support assuring CKM quality level as well as support CKM (quality) continuous improvement, where CKM quality level satisfies process approach requirements (as in general quality management). And CKM continuous improvement refers rather to staff engagement—at both the individual

and the team levels. This is part of ATENA's maintenance component—how to make the CKM vivid and working for long time.

16.4 Measures for Communication Processes

Business observations and research lead to complementary set of measures that satisfy stipulations above: complexity and environment-creating approach based on ATENA communication model and specific parameters enabling to measure CKM efficiency:

M1: Number/ratio of doubled communicates for given group of recipients.
M2: Number/ratio of various communicates about the same topic (e.g., various answers for the same question).
M3: Number/ratio of questions generated by communicates.
M4: Overall training expenditures.
M5: Number/ratio analogous mistakes in various teams over time.
M6: Expenditures for disk space for communication (e.g., mailboxes).
M7: Mean time needed for getting to communicates older than 6 months.
M8: Net Promoter Score (NPS) referring to communication system users' satisfaction.
M9: Number of local knowledge bases.
M10: Time needed for successful transferring communicates and put them into practice (delay).
M11: Communicate loss of quality (loss of knowledge).

The above list proofs H1 by describing connections and dimensions of proposed set of measures in the context of real organization and ATENA model.

Some of those measures will be used in graph (network) model introduced in the following paragraphs. Whole set is useful for assessing communicate model. Before implementing change in communication and knowledge management model, by identifying numerically initial state in order to monitor and set baseline to control communication model throughout the transformation. And after change, to check benefits and communication model performance. In specific, some of the measures M1–M6 can be easily transformed info benefits from project implementing new communication model. Disk space freed and saving time in medium or big organizations justify expenditures on intranet portal implementation.

16.5 Graph Model—Theoretical Assumptions

Graph or network describes set of nodes and edges. Graph definition might be presented in such manner [13] (definition and first two parameters):

$$G = (P, S),$$

where P is a set of employees $(p_1, p_{2,} \ldots, p_n)$, and S set of edges representing communication paths $(s_{11}, s_{12,} \ldots, s_{1n}, s_{21}, \ldots s_{nn},)$, where s_{ij} refers to communication path between person i and person j. Path s_{ii} (loop) in this model will be meant as inherent ability for communicate perception and conversion communicates into knowledge and practice—specific for every single person. This loop allows to implement context of the recipient, one's experience and knowledge, according to Lungefors and Sundgren.

This model can be parametrized by:

1. D_{ijkt}—delay of communicate t with deterministic and stochastic components—D_{ijpi} (standard delay for the person) and $D_{ijtpilos}$ (stochastic delay of specific communicate t) [M10].
2. Q_{ijkt}—synthetic quality of communicate t with, respectively, deterministic and stochastic components of quality from interval [0,1]—Q_{ijpi} (typical quality connected with s_{ii}) and $Q_{ijtpilos}$ (stochastic quality component for specific communicate t) [M11].

Following models might reflect reality of any organizational or business unit. We might treat it as representation of company that uses at the same time:

1. Mail cascade for part of communicates replacing former official letters.
2. Internal portal, Web site or Facebook site, respectively, for more static communicates (like forms or general documents) and for social communication.

In this context—Use 1. (model K introduced beneath) refers to a mail cascade going downwards through organization cross one or more levels of managers. Responsible for execution of this task, motivation and support to actual doers. In this model, manager becomes at the same time coach and secretary—responsible for delivering communicate, being understood by doer and putting it into practice. Use 2., portal or site (model M introduced beneath), is either static Web site or any social media site. What is important, regarding ATENA's this static Web site should support comments to enable feedback for authors. In simple model, comments might be replaced by mail address for comments added for any site.

On one graph node set, there might be spread two models:

- Model K—cascade, typical for hierarchical structures and mail-based communication (Fig. 16.2).
- Model P—horizontal, describing portal-based communication environment (Fig. 16.3).

This model might be enhanced by introducing set of additional parameters:

1. RC_{jkt}—number/rate of redundant (doubled) communicates with, respectively, deterministic and stochastic components RC_{jpi} and RC_{jpilos} M1].
2. DA_{jkt}—number/rate of various communicates with, respectively, deterministic and stochastic components DA_{jpi} and DA_{jpilos} [M2].

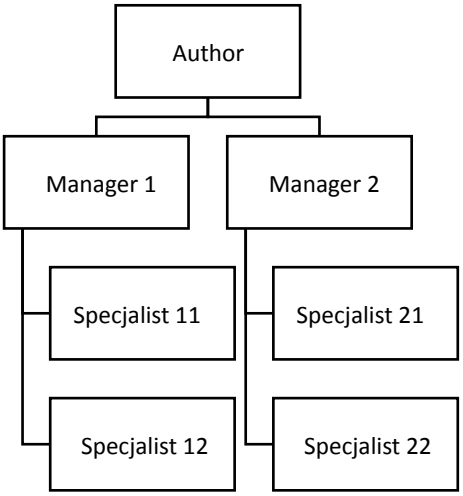


Fig. 16.2 Model K—cascade

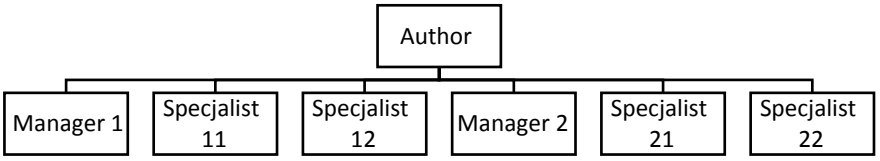


Fig. 16.3 Model P—horizontal

3. CQ_{jkt}—number/rate of questions generated by communicates with, respectively, deterministic and stochastic components CQ_{jpi} and CQ_{jpilos} [M3].

Those parameters describe communicate recipient reality, when on the same topic there are many communicates (RC_{jkt}—M1), some of them are varying (DA_{jkt}—M2), and implementation or conversion into knowledge requires additional questions (CQ_{jkt}—M3).

Numerical analysis based on model spread on organization based on the following nodes:

- A—author.
- $D_1, \ldots, D_4$—[5] directors (or deans, at university faculty).
- $M_{11}, \ldots M_{44}$—(16), managers—where M_{ij} is subordinate to D_i.
- $P_{111}, \ldots P_{444}$—(64), experts/specialists—where P_{ijk} is subordinate to M_{ij}.

This abstract model refers to three- or four-level organization and 85 staff members (author can be replaced by chairman or executive director). This organization is sufficient to model real-life organizations like faculty, department or even whole organization—model tailoring is reduced to changing the number of members of certain managerial level. In fact, most of the organizations have no more than just four levels communicating day by day for assuring business continuity.

For the simulation, the following values were set for above parameters:

1. D_{ijpi} (standard delay for the person): for author, directors, managers, experts, respectively: 1, 1, 2, 4 h.
2. $D_{ijpilos}$ (stochastic delay of specific communicate t) on average 0.5 multiplied by stochastic number from interval [0,1] of standard delay.
3. Q_{ijpi} (typical quality connected with s_{ii}) equal to 1 and $Q_{ijpilos}$ equal to 0,1 multiplied by stochastic number from interval [0,1] for all individuals.

16.6 Results

Enhanced communication graph-based model reaffirms results from Zając [13]—that flat communication model (model P) provides lower variance of delay (D_{ijkt}) and stability of communicate quality (Q_{ijkt}). Furthermore, additional communication process characteristics, respectively: doubled communicates (RC_{jkt}), alternative versions (DA_{jkt}) and number of questions (CQ_{jkt}), are significantly reduced "by design."

Author using portal instead of mail cascade has all the authorized users within direct reach. There is no need for anybody to be intermediate or explain "what author meant," and all the clarifications, questions and suggestions can be found (and should take place) in comments or communicates linked to initial one. It prevents from differing interpretation and multiplying questions (Table 16.1).

Those exemplary results present differences between mail cascade-based and portal-based communication processes. Referring to virtual organization consisting of 4 directors, 16 managers, 64 experts and one author of the single communicate, we might discuss above results.

Comparing model K to model P, from level of managers there can be observed reduction of average delay (D_{ijkt}), respectively, 2.56 h and 6.06, that makes on average half of total time for distribution of single communicate. In practice, it means that managers read e-mails next working day and experts at most next day or even two days after sending initial e-mail. What is important, it is a kind of "net" delay being induced by the communication process itself. Analyzing "doubled" communicates in model K, it can be seen that at the level of 64 experts there are 194 doubled communicates (RC_{jkt}). In practice, it means more than 388 communicates to be read and filtered, contrary to 10 (20) doubled communicates in model P that might be interpreted as intensification of communication about given topic. Those communicates multiplied by any time period (e.g., 5 min) result in 32 h wasted for reading redundant communicates. In other words, distribution of one communicate might result in wasting one week of expert work—that might be converted into money related to average expert salary. Multiplied by the number of such communication distribution processes per month—e.g., 5—it leads to one out of 83 individuals only

Table 16.1 Simulation results

Level	Model K				
	Avg. D_{ijkt}	Avg. Q_{ijkt}	Sum RC_{jkt}	Sum DA_{jkt}	Sum CQ_{jkt}
A [11]	1	1.5	0	0	0
D_{ij} [6]	2.31	0.9	0	0	9
M_{ij} (16)	6.39	0.83	16	10	35
P_{ij} (64)	12.31	0.80	194	24	120
Model P					
	Avg. D_{ijkt}	Avg. Q_{ijkt}	Sum RC_{jkt}	Sum DA_{jkt}	Sum CQ_{jkt}
A [11]	1	1.5	0	0	0
D_{ij} [6]	2.31	0.9	0	0	9
M_{ij} (16)	3.83	0.89	3	5	10
P_{ij} (64)	6.25	0.94	10	4	24
Comparison (model K vs. model P)					
	Avg. D_{ijkt}	Avg. Q_{ijkt}	Sum RC_{jkt}	Sum DA_{jkt}	Sum CQ_{jkt}
A [11]	0	0	0	0	0
D_{ij} [6]	0	0	0	0	0
M_{ij} (16)	2.56	0.06	13	5	25
P_{ij} (64)	6.06	0.14	184	20	96

Source Own elaboration

reading redundant e-mails. Redundant communicates result also in various communicates (DA_{jkt}) that induce confusion and another work needed for investigating right version.

Furthermore, comparing both models in the context of communicates induced by not clear or adequate communicate, model K generates ("by design") 120 questions for one communicate that are to be formulated, distributed and (hopefully) answered. In model P, this number is significantly mitigated by forums or comments that enable preventing from doubling questions. Using time period (5 min), those questions can easily be converted into waste of resources—50 h on average—and money. This shows how defined above set of measures and ATENA we might model and forecast communication process performance and obstacles. For example by comparing different communication models and finding sources of ineffectiveness in communication processes.

16.7 Conclusions

"Democratic", turquoise-like organization is materialization of idea of flat organization, but this modern approach will fail—if not supported by tailored processes

and tools. But this lean, agile, turquoise trend is very important for communication efficiency. Trendy myths and scripts create space for organizational culture evolution—readiness for flatting organizational structures as well as communication models. Democratization utilizes one more ATENA's principle—every communication model user is becoming prosumer. Empowerment of every single staff member is connected with one's right to speak as well as obligation to do so. This shift of responsibility creates space for self-development and self-awareness, as well as space for doubts and questions. This discourse in turn leads to improvement, sharing lessons learned (both positive and negative) and staff engagement as part of responsibility and becoming co-creators of business reality.

Referring to chapter goals and hypothesis, set of measures (M1–M11) enabling to assess communication effectiveness and efficiency was presented, briefly in ATENA model and organizational contexts (G1 and H1).

Two main applications of this set of measures are:

1. Assessing initial, several intermediate and target CKM states, in order to support creating business case for transformational project. Proposed measures enable assessing and controlling such project product benefits, like expenses for disk space or waste of staff time (and money for their salaries).
2. Comparing different CKM models as well as providing tool for CKM maintenance within given organization.

Initially, ATENA-based measures were intended for justifying business and organizational benefits coming from use of ATENA. But the same measures are also tools for implementation and maintenance of ATENA-based CKM models.

Based on that, graph model with some of parameters derived from above measure set was created and analyzed (G2 and H2).

Results prove that flat, portal-based model P ensures better quality of communication processes than mail model K. The same approach might be used for any organization, after tailoring number of management/communication levels and number of individuals within actual teams.

ATENA-based model's efficiency and effectiveness are mitigated by real fulfillment of ATENA's principles and organizational readiness—evolving from awareness of past communication malfunctions in given organization. Comparison of models K and P clearly shows that also size of organization, number of management or communication levels and team sizes affect ATENA's application. Last of main limitations of ATENA model application comes from organizational attitude for change—resulting in introducing ATENA in name only or real and profound change.

As a general consequence for any given organization, communication channel management and synchronization of various domains described in themes should be implemented in order to reduce values of proposed measures, what will result in reducing costs and CKM domain malfunctions. As shown in table presenting measures, only the whole ATENA model gives comprehensive and consistent view at communication and knowledge management in organization. Principles are essential for defining corporate and communication culture in the context of maturity and readiness for communication and knowledge management. Themes, representing

good practices and interrelated domains, enable ensuring communication quality. Processes and maintenance components, respectively, lead to introduction and continuous existence communication and knowledge management in given organization.

References

1. Kotler, P., Caslione, J.A.: Chaotics: The Business of Managing and Marketing in the Age of Turbulence. AMACOM Div American Mgmt Assn (2009)
2. Nordin, S.M., Sivapalan, S., Bhattacharyya, E., Ahmad, H.H.W.F.W., Abdullah, A.: Organizational communication climate and conflict management: Communications management in an oil and gas company. Procedia-Soc. Behav. Sci. **109**, 1046–1058 (2014)
3. Koçoğlu, İ., Akgün, A.E., Keskin, H.: The collective unconscious at the organizational level: The manifestation of organizational symbols. Procedia-Soc. Behav. Sci. **235**, 296–303 (2016)
4. Redza, A., Nordin, S.M., Saad, S., Wahab, H.: Inter-organization communication management between organizations in a subsidized fertilizer market in Malaysia. UMK Procedia **1**, 33–41 (2014)
5. King, K.G.: Measuring teamwork and team performance in collaborative work environments. In Evidence-Based HRM: A Global Forum for Empirical Scholarship, vol. 5, issue. 2, pp. 196–205. Emerald Publishing Limited (2017)
6. Faraj, S., Jarvenpaa, S.L., Majchrzak, A.: Knowledge collaboration in online communities. Organ. Sci. **22**(5), 1224–1239 (2011)
7. Faraj, S., Kudaravalli, S., Wasko, M.: Leading collaboration in online communities. MIS Q. **39**(2), 393–412 (2015)
8. Kleinsmann, M., Buijs, J., Valkenburg, R.: Understanding the complexity of knowledge integration in collaborative new product development teams: A case study. J. Eng. Tech. Manage. **27**(1–2), 20–32 (2010)
9. Patterson, P.D., Pfeiffer, A.J., Weaver, M.D., Krackhardt, D., Arnold, R.M., Yealy, D.M., Lave, J.R.: Network analysis of team communication in a busy emergency department. BMC Health Serv. Res. **13**(1), 109 (2013)
10. Meltzer, D., Chung, J., Khalili, P., Marlow, E., Arora, V., Schumock, G., Burt, R.: Exploring the use of social network methods in designing healthcare quality improvement teams. Soc. Sci. Med. **71**(6), 1119–1130 (2010)
11. Borca, C., Baesu, V.: A possible managerial approach for internal organizational communication characterization. Procedia-Soc. Behav. Sci. **124**, 496–503 (2014)
12. Madalina, O.: Conflict management, a new challenge. Procedia Econ. Finance **39**, 807–814 (2016)
13. Zając, R., Rostek, K., Wiśniewski, M.: Model grafowy komunikacji jako narzędzie tworzenia wartości w organizacji. Przegląd Organizacyjny **8**, 19–24 (2018)

Chapter 17
Hybrid System Dynamics—Agent-Based Simulation for Research in Economics and Business

Małgorzata Łatuszyńska

Abstract In a rapidly evolving business environment, the question of searching of effective methods and tools for researching and analyzing economic entities becomes more and more important. One of the ways of studying this world of growing dynamic complexity and supporting the process of decision-making is the use of computer simulation. According to the reference literature, there are three most commonly used methods of computer simulation for research in management: system dynamics, discrete event, and agent-based simulation. However, the complex, multifaceted nature of modern-day economic and business systems can pose considerable challenges for single-methodology simulation approach. In such cases, it may be that (should be removed) an alternative simulation approach, using either another modeling paradigm or a hybrid approach, could provide a simpler, more natural, or more efficient solution. Hybrid simulation, defined as a combination of two or more computer simulation methods, has become an increasingly common approach to modeling complex systems in the past two decades. The study concentrates on one particular hybrid—that involving agent-based simulation (ABS) and system dynamics (SD). It aims to discuss the issue of combining system dynamics and agent-based simulation approaches for research in economics and business. First, the two methods will be briefly characterized with the indication of their differences and similarities. Then, the possible ways of combining them in a single model will be described. Finally, examples of hybrid SD-ABS model applications in economics and business will be presented.

Keywords Computer simulation · System dynamics · Agent-based simulation · Hybrid simulation

M. Łatuszyńska (✉)
University of Szczecin, Szczecin, Poland
e-mail: malgorzata.latuszynska@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_17

17.1 Introduction

In a rapidly evolving business environment, the question of searching of effective methods and tools for investigating economic entities becomes more and more important. One of the ways of studying this world of growing dynamic complexity and supporting the process of decision-making is the use of computer simulation.

In the narrow sense, the term of computer simulation refers to the use of a computer to solve an equation that we cannot solve analytically or more generally to explore mathematical properties of equations where analytical methods fail (e.g., [1–4]). In the broad sense, computer simulation refers to the entire process of constructing, using, and justifying a model that involves analytically intractable mathematics (e.g., [1, 2, 4]). Such a model is called a computational or a simulation model.

The advantage of simulation models is their capability to embrace complex real-life systems characterized by dynamic nonlinear relationships. Another substantial benefit is that what-if scenarios can be tested, but intervention in reality is not required [5]. Consequently, simulation models enable increasing the understanding of the system under study and testing policies with the objective to support decision-making and at times policy implementation [6].

According to the reference literature [7, 8], there are three most commonly used methods of computer simulation for research in economics and business: system dynamics, discrete event, and agent-based simulation. However, the complex, multi-faceted nature of modern-day economic and business systems can pose considerable challenges for single-methodology simulation approach. In such cases, it may be that (should be removed) an alternative simulation approach, using either another modeling paradigm or a hybrid approach, could provide a simpler, more natural, or more efficient solution [9].

Hybrid simulation, defined as a combination of two or more computer simulation methods [9, 10], has become an increasingly common approach for modeling complex systems in the past two decades [11–13]. This chapter concentrates on one particular hybrid—which involves agent-based simulation (ABS) and system dynamics (SD)—and discusses the issue of linking both approaches for research in economics and business. First, the two methods will be briefly characterized with the indication of their differences and similarities. Then, the possible ways of combining them will be described. Finally, on the basis of the literature review, examples of hybrid SD-ABS model applications in economics and business will be presented.

17.2 System Dynamics Versus Agent-Based Simulation

DS and ABS approaches are iconic in the way that they are often presented as exclusive alternatives to analyze complex systems [13–15]. Many scholars argue that the choice of an appropriate approach to adopt in a particular case should depend on the nature of the system at hand and the purpose of the model [13, 14, 16, 17]. Swinerd

and McNaught [13] suggest that many systems can be modeled in equivalent ways by both paradigms. However, they conclude that sometimes one paradigm presents a more natural choice than the other [18].

System dynamics approach was found more than 50 years ago by Forrester [19] around two notions from systems theory [20]:

– Aggregated-level variables affect each other through feedback loops.
– System's structure drives system's behavior [5].

According to Coyle's [21] definition, SD deals with the behavior of systems over time in order to describe (1) the system through qualitative and quantitative models and (2) the reaction to feedback, and (3) to design appropriate feedbacks and steering methods. SD gives the chance and basis to understand and examine all the mutual interrelations within the systems as well as to determine the character of changes in these systems over time.

The building blocks in specifying an SD model are stocks, flows, and auxiliary variables [19, 22]. Stocks represent the accumulation of material and information, caused by the action of inflows and outflows. While stocks are mathematically described by integral equations, flows are described by differential equations [23, 24]. These sets of equations are solved through numerical integration at discrete time steps. The solution describes the system's aggregated state changing continuously over time and depending on the previous state of the system [19, 22, 25].

Summarizing, SD models are feedback-based, and they model systemic problems at an aggregate level over time. The model of causal relationships is the result of a cognitive abstraction process on behalf of the modeler [26].

In contrast to SD, ABS models the structure of a system as the result of decentralized decisions of individual entities or agents over time [23, 27]. Therefore, instead of assuming a given system structure, agents' decisions shape and change the state and structure of the system. In turn, agents react to the dynamic changes in the system, which can potentially alter their decision rules [5].

Various disciplines worked out their own ways of understanding the term "agent." It is commonly accepted that agents are placed in a certain environment and they are able to take autonomous actions [28–30]. From a practical point of view, it can be assumed that an agent has got the following features [27]: (1) It is an identifiable unit which possesses a certain set of features and rules governing its behavior and decision-making capabilities; (2) it is placed in environment in which it cooperates with other agents; (3) its actions can be directed at achieving a particular goal; (4) it is autonomous, and it can function independently in its environment and in contacts with other agents, at least in certain defined situations; (5) it is flexible, and it has got the ability to learn and adapt [31].

The main building blocks of ABS models are: autonomous agents, their decision rules and actions, and the environment in which they interact [32–35]. ABS approach generally focuses on micro-level interactions that may explain emergent patterns such as transient dynamics on a system level [14].

Summarizing, the aim of agent-based simulation is to look at global consequences of individual or local interactions in a given space (environment) and agents are seen as the generators of emergent behavior in that space ([36, 26]).

Both SD and ABS can explore complex problems, but even from the above-presented description, it appears that they differ significantly. One of the first researches on confronting both approaches was made by Scholl [26], and more recent works have enriched those comparisons [5, 23, 37–42].

SD is usually used to analyze problems from a macro-perspective and holistic-thinking perspective. It is a "top-down" modeling approach that can avoid the limitations of one-sided thinking (e.g., the micro-perspective) and help to understand the structure behind a complex phenomenon [43]. However, SD cannot give a profound explanation of the micro-behaviors in the system, because it ignores the relationship between the macro- and micro-behaviors [38]. In turn, ABS follows a "bottom-up" procedure that emphasizes the spatial or social interactions between individuals and their environment [44]. ABS is an effective cross-scale modeling method that combines time dimension with space dimension, and bears the characteristics of heterogeneity, space discretization, time discretization, and discrete states [44]. Through computer simulation, the microscopic mechanism of complex macro-phenomena can be revealed [38]. While ABS models are used to describe disaggregated parts of a system, SD models represent the aggregated system [14]. SD focuses on the "flow" relationships and feedbacks that can longitudinally simulate a system's dynamic behavior. It is appropriate to analyze the interactions between different elements and cumulative longitudinal effects. However, spatial factors are not covered in the SD modeling process. In contrast, ABS considers the spatial interactions. However, the feedback effect of various social and economic factors on agents is ignored [38, 45]. A more detailed comparison of both approaches is presented in Table 17.1.

17.3 Linking SD and ABS

SD and ABS have their own strengths and weaknesses depending on the type of component under consideration. Hence, many studies have underlined the respective suitability of the two approaches to handle diverse types of problems [46–50] or rectify each other's results [51]. However, there are also studies arguing that what have traditionally been seen as antagonistic properties turns out to be powerful synergistic features when SD and ABS are merged (e.g., [26, 15, 40]). The approaches then become highly complementary and essential to the dynamic modeling of some classes of systems. The fusion of SD and ABS makes the models' structure more natural and realistic. Moreover, results in an important gain in modeling capabilities, analytical power and simplicity [15] and in consequence enable achieving a more powerful research tool.

Ding et al. [38] indicate three reasons justifying the combination of both approaches:

Table 17.1 Comparison of SD and ABS approaches (based on Behdani [7]; Ding et al. [38])

	System dynamics	Agent-based simulation
Key concept	The simulation system changes continuously, in countless points in time: smooth and steady changes	The simulation system changes the action or interactions of agents mainly at discrete points in time. It can also occur continuously
Orientation	System-oriented: The focus is on modeling aggregates	Individual-oriented: The focus is on modeling the entities and interactions between them
Model	1. Open-loop model: feedback loop 2. Stock and flow model a. Stocks: basic stores of objects (= quantities) b. Flows: the movement of objects between stocks in the system (= time period) c. Delays: delays between the measuring and then acting on that measurement	1. Low-level model: discrete time-based agent interaction and decision-making 2. High-level model: multi-agents' network a. Autonomous agents: self-directed objects b. Rules: that agents follow to achieve their objectives
Modeling methodology	Strategic-level modeling Stocks: products, items, jobs Flows: decision trends or patterns Time delays: The delay parameter usually uses an exponential distribution, and deterministic delays are special constructs	Statechart inside agent modeling Agents are usually active and exhibit behavior Time delays: stochastic delay or deterministic delay
Building blocks	Equations, feedback loops, stock and flow diagrams	Individual agents and their decisions, and statechart diagrams
System structure	Fixed and stable	Flexible and not stable
Structure type	Homogenized entities, all entities are assumed to have similar features, averaging of values	Heterogeneous entities
Application type	Problem-solving	Exploring
Handling of time	Continuous or quasi-continuous functions	Mainly discrete functions, but can also be continuous
Mathematical formalization of the system	Stock and flow	Agent and environment

(continued)

Table 17.1 (continued)

	System dynamics	Agent-based simulation
Experimentation	By changing the system structure	By changing the agents' behavior rules (internal/interaction rules) and in system structure
Conceptual modeling technique	Stock and flow diagram, causal loop diagram, flowchart, hexagons, archetypal structure, influence diagram	Unified Modeling Language (UML), including class diagrams and instance diagrams, but especially statechart diagrams, agent–object relationship (AOR) diagrams, cognitive mapping, and business process modeling (BPM)

– SD models do not take into account different levels of aggregation, while ABS gives the possibility of mapping a fine level of detail. Thus, SD can be used to reflect the highest level of abstraction, while ABS can be used to reflect the lower levels of abstraction, depending on the nature and scale of the elements.
– SD does not allow the inclusion of heterogeneous elements. Each stock is composed of homogeneous elements. The diversity of elements must therefore be modeled by adding new stocks. However, the heterogeneous elements (agents) can be easily established using ABS.
– SD is equation-based and needs quantified relationships between variables; thus, it is not suitable for complex systems with unknown structures. However, ABS can reasonably represent complex systems based on a limited number of relatively simple rules to reveal emergent behavior.

Some pioneering work on the linking system dynamics and agent-based simulation was conducted by Akkermans [52], Scholl [26, 53], Pourdehnad et al. [54], Schieritz [55], and Größler et al [56]. Till date, many studies have been made on combining SD and ABS and the use of this hybrid to study problems in various disciplines, including studies on how to integrate the two approaches. For instance, Swinerd [57] presents a framework for design concepts of hybrid SD-AB simulation that develops further a categorization introduced by Shanthikumar and Sargent [58]. In this framework, three classes of hybrid simulation are suggested: interfaced, sequential, and integrated (Fig. 17.1).

In the interfaced simulation, different system parts are modeled with independent models of different paradigms, without an exchange of information during the simulation. These models may be run in parallel with their outputs combined as required to represent the desired output as a function of time. In the sequential class, the information flow is restricted to one point in time, when the other model takes over. In this design, one model has to be run first and its output then fed to the next [13, 57].

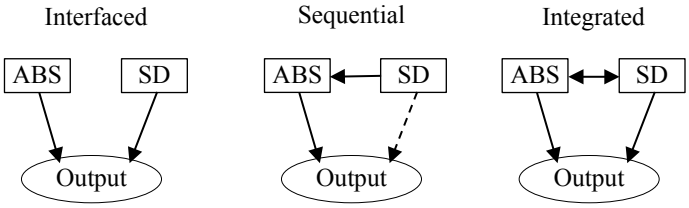


Fig. 17.1 Three classes of hybrid SD-ABS [57]

Integrated models, finally, are linked over the entire simulation. In this, class incorporates feedback between models representing a continuous fluid process. The feedback arrows between models in the integrated simulation and the arrows from models to output in the interfaced simulation do not constrain flows to a single point in time. Indeed, the expectation is that such flows will usually take place several times during simulation [13, 57].

There are at least three options available to implement the SD-ABS-integrated hybrid class (Table 17.2): First, the SD model is within an agent ("agents with rich internal structure"); second, a stock in a SD model bounds the behavior of agents ("stocked agents"); third, an emergent property of an ABS influences a parameter in the SD model ("parameters with emergent behavior") [18]. It is important to note that the process of feedback between the SD and AB modules within these three interpretations of the integrated design class is not constrained. Feedback is the key feature that provides for an integrated hybrid approach to simulation [13].

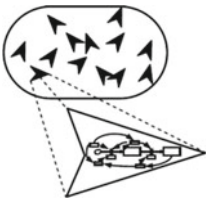
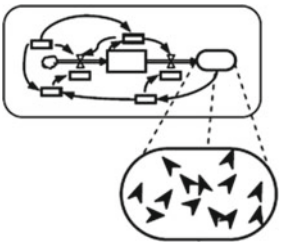
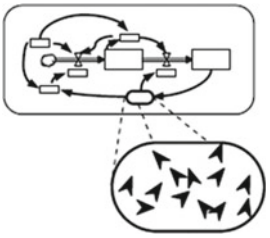
A complementary categorization to this functional view of hybrid-integrated simulation classes was offered by Vincenot et al. [15], who related to the spatial structure of SD-ABS models. They distinguish four cases (Table 17.3): (1) individuals interacting within a single SD model, (2) individuals interacting with spatially disaggregated instances of a SD model, (3) SD sub-models embedded in individuals, and (4) components swapping between the SD and the ABS approaches [18].

Although both proposals were independently developed, they correspond well with each other. Subtle differences arise from different points of view of the authors. Whereas Vincenot et al. [15] seem to take a stronger ABS perspective, Swinerd and McNaught [13] tend to conceptualize the system to be governed by the SD model. In the "individuals interacting with a single SD model" class for instance, Vincenot et al. [15] state that the SD usually represents a dynamic system environment in which the ABS operates or in special cases computes SD parameters dynamically. The latter view matches with Swinerd and McNaught [13] concept of "parameters with emergent behavior" [18].

A separate important issue is the implementation of the hybrid simulation model. The eventual ways to create DS-ABS hybrid models are summarized in Table 17.4.

It is possible to use a low-level programming language to create hybrid simulation model, but it is time-consuming and requires right skills as there would not be any specialized program in automating the work. Many hybrid models were constructed by only using a SD program (e.g., [52, 60]). This usually requires the use of arrays

Table 17.2 Classification of the integrated hybrid SD-ABS (based on Swinerd and McNaught [13]; Wallentin and Neuwirth [18]; Kim et al. [59])

Option	Schema	Description
Agents with rich internal structure		An SD model is built into the individual of the ABS model A flow of information between SD and ABS models can be bidirectional
Stocked agents		An ABS model is built as part of the SD model A level within an SD module is used to bind an aggregate measure of an ABS model An ABS model plays a role as a stock in the SD model that represents an entity within the system An information flow tends to be from the SD model to the ABS model only
Parameters with emergent behavior		Similar to "stocked agents," an ABS model is built as part of the SD model The output of the ABS model affects the parameters, not stock, in the SD model An aggregate measure or observation of an ABS model is used to influence a parameter within an SD model A flow of information between SD and ABS models can be bidirectional

to be able to cope with added complexity [40]. Some models (such as [61, 62]) used another tool in combination with a SD program. This can be done by either using some middleware or creating a separate code using a low-level programming language [40]. Currently, there is also existing commercial software, which is able to do both SD and ABS modeling, for instance AnyLogic™. And final way to create the hybrid simulation models is to build an own platform for the work. However, this is similar to the use of a low-level programming language as the whole software needs to be programmed first [40].

Table 17.3 Classification of the integrated hybrid SD-ABS proposed by Vincenot et al. (based on Kim et al. [59]; Wallentin and Neuwirth [18])

Class	Description
Individuals interacting with a single SD model	Among the components constituting the system, individual and environment (e.g., resource and environmental condition) are modeled by ABS and SD, respectively An ABS model and SD model interact with each other
Individuals interacting with a space made of SD models	Individuals and the space in which they can be distributed are modeled by ABS and SD, respectively By allowing the ABS model to interact with multiple SD models (e.g., various types of space), a more realistic individual–environment interaction can be described
SD sub-models embedded in individuals	Individual properties are computed dynamically by the SD model It can be applied not only to the state of the individual but also to the modeling of movement and behavior of the individual The same what "agents with rich internal structure"
SD-ABS model swapping	SD and ABS approaches are selectively applied during the simulation, considering changes in the importance of the model components and the computational cost

Table 17.4 Different ways to implement hybrid simulation models [40]

Method	Advantages	Disadvantages
Low-level programming language	Totally flexible	Time-consuming Requires good expertise in programming
System dynamics program	Relatively easy to use Includes all the necessary pieces for a SD model	Structure of the model usually fixed Might not be able to incorporate complex events
System dynamics program with middleware	Includes all the necessary pieces for a SD model Flexibility regarding agent model	Might require some own programming Synchronization might be an issue
Hybrid simulation model toolset	Includes the basic elements of both SD and ABS Easy to integrate	Few exist
Construct simulation software	Flexible	Time-consuming Requires good expertise in programming

17.4 Hybrid Simulation for Research in Economics and Business—A Literature Review

This review reports publications targeting DS-ABS hybrid in economics and business research. For the purpose of this chapter, two scientific databases were used: Scopus and ScienceDirect located among the largest and main multidisciplinary databases. The search was performed on March 2019. The literature review followed a structured approach, based on the studies by Jahangirian et al. [63], Barbosa and Azevedo [64], and Scheidegger et al. [65], and adopted the steps in Fig. 17.2.

An initial search was performed to define the most appropriate keywords to address the research objective. Next, searches based on the defined searching criteria (1–4) were performed. Some general filters were applied in an attempt to improve the search results for the aim of this paper. These filters can be seen in Fig. 17.2 and include: (1) search fields: title, abstract, and keywords; (2) years: all; (3) research areas related to business, management, accounting, economics, econometrics, and finance; (4) articles published in peer-reviewed journals, conferences, and books; and (5) papers written in English. The results are shown in Table 17.5. Searching the databases allowed for the selection of 58 in Scopus and 31 in ScienceDirect publications matching the assumed criteria.

Then, a screening process of the abstracts was performed. A total of 89 abstracts were read. Forty-five out of them were identified as the most relevant according to the

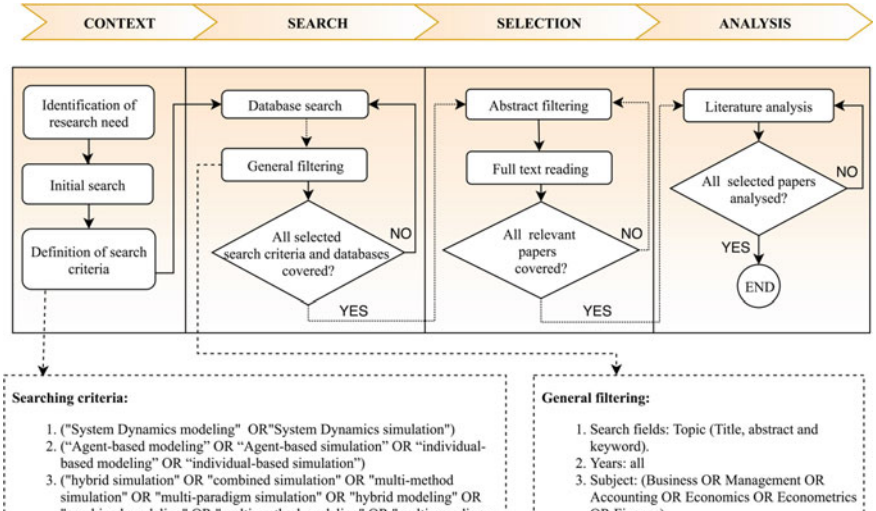


Fig. 17.2 Steps of the structured literature review (based on: Jahangirian et al. [63], Barbosa and Azevedo [64], and Scheidegger et al. [65])

Table 17.5 Number of papers published and the first year of publication for the chosen searching criteria and filters

	Searching criteria	Scopus		ScienceDirect	
		Total number of papers	Years of publication	Total number of papers	Years of publication
1	("System dynamic* modeling" or "system dynamic* simulation")	1929	1976–2019	1015	1995–2019
2	("Agent-based modeling" or "agent-based simulation" or "individual-based modeling" or "individual-based simulation")	1128	1998–2019	1085	1995–2019
3	("Hybrid simulation" or "combined simulation" or "multi-method simulation" or "multi-paradigm simulation" or "hybrid modeling" or "combined modeling" or "multi-method modeling" or "multi-paradigm modeling")	1813	1969–2019	1570	1995–2019
4	Within results of the search no. 3: ("agent-based model*" or "individual-based model*") and ("system dynamics model*")	58	2008–2018	31	2011–2019

objective of this study. This means that papers who discussed only the integrated DS-ABS hybrid but not any case study were not selected. Then, the full-text screening process was started in order to capture the papers presented only the application of the mentioned hybrid in the field of business and economics. As a result, a total list of 25 papers was created as shown in Table 17.6. Six papers were found in ScienceDirect and 20 in Scopus—one paper was duplicated in both databases [66].

A deeper analysis of the found publications shows that they concern various issues in the field of economics and business (Table 17.7). Most of them refer to innovation management and then to project management and product management.

17.5 Conclusion

Hybrid simulation comes in many shapes and forms. It has been argued by many researchers that hybrid simulation provides more and better insights into the real-life system as it allows modelers to assess its inherent problems from different dimensions. As a result, hybrid simulation is becoming an important field within the modeling and simulation arena.

It is significant that while the hybrid simulation was started relatively early (the first studies appeared already in the last century), the first publications on the possible combination of agent-based simulation and system dynamics appeared only in 2001, and reports on the use of this hybrid in economics and business in 2008—according to the results of the presented review. Of course, it should be emphasized that despite the considerable number of papers included in this study, it is possible that some papers may have been missed—because for the purpose of this study only two databases were selected—based on its size and relevance to the field. Due to the large number of available databases and publications in the field of economics and business, performing the search in all databases would be impractical in a timely manner. Nevertheless, the results of the review clearly indicate that the number of studies on the use of DS-ABS hybrid in various areas of economics and business is increasing year by year (Fig. 17.3). But there are still very few of them.

Undoubtedly, integrating SD and AB makes it possible to avoid their individual constraints and allows for a more complete use of the potential of their complementary features; hence, more accurate representations of complex dynamic systems can be provided. The further widespread adoption of hybrid SD-AB models depends on the availability of easy-to-use tools that are able to effectively integrate these modeling paradigms. Therefore, an important area of research is the creation and improvement of free and open-source tools for hybrid simulation to support decision-making.

Table 17.6 List of papers discussing application of hybrid SD-ABS in economics and business (own elaboration)

	Authors	Year	Purpose of hybrid model
1	Alvi et al.	2018	Assisting authorities in formulating policies toward achieving effective management of available water resources and efficient water supply delivery coupled with responsible demand-side management
2	Barbosa and Azevedo	2018a	Supporting performance assessment in manufacturing environments combining different production strategies
3	Barbosa and Azevedo	2018b	Understanding the impact that different development projects and production variables have on the manufacturing system performance
4	Elia et al.	2018	Assessing quantitative impacts of adopting dynamic scheduling in waste from electric and electronic equipment collection and comparing different collection schemes
5	Nasirzadeh et al.	2018	Modeling construction projects in the context of construction workers' safety
6	Nassehi and Colledani	2018	Investigation of the influence of the interaction between customer behavior and strategic decisions of the manufacturer on the economic viability of remanufacturing
7	Liang et al.	2018	Insight into the cognitive, social, and organizational aspects that can determine the social contagion effect of safety violations within a construction crew
8	Khanzadi et al.	2018	Predicting the value of labor productivity taking account of various continuous influencing factors and the interactions between different agents involved in the project
9	Fahhama et al.	2017	Reflecting in detail the network structure of automotive supply chain for supporting decision-making
10	Manisri and Pichitlamken	2017	Supporting decisions concerning natural rubber supply chain
11	Block	2016	Enabling practitioners to analyze the evolution of the staff as well as organizational performance under various and changing conditions

(continued)

Table 17.6 (continued)

	Authors	Year	Purpose of hybrid model
12	Ebrahimi	2017	Analyzing the market share of redesigned and independent designed technologies compared to the acquired ones
13	Asif et al.	2016	Evaluating economic and environmental performance of changing business model
14	Sitepu et al.	2016	Predicting the sustainability impacts of alternative replanting scenarios
15	Kolominsky-Rabas et al.	2015	Assessment of innovative health technologies prior to their launch
16	Swinerd and McNaught	2015	Insight into technological diffusion at national level
17	Nikolic and Simonovic	2015	Examining the codependence between water resource system and socioeconomic environment to support decision-making in water resource management
18	Rondini et al.	2014	Analysis of product-service system in the context of three main performance indicator categories: process efficiency, customers' satisfaction and environmental sustainability
19	Wang et al.	2014	Complementing standard lifecycle assessment in the context of social and economic aspects
20	Swinerd and McNaught	2014	Simulating the diffusion of technological innovation
21	Kortelainen and Lättilä	2013	Testing the profitability of the fast strategy concept in different business environment conditions
22	Shafiei et al.	2013	Comprehensive evaluation of the diffusion process of alternative fuel vehicles taking into account the most influencing stakeholders, including car manufacturers, car dealers, consumers, energy supply system, fuel stations, and government
23	Wang and Van Den Heuvel	2011	Analyzing, optimizing, and tuning the performance of service networks and their resources, including software services and human-operated services
24	Kieckhäfer et al.	2009	Analysis of product strategies in the automotive industry with special regard to alternative fuel and powertrain technologies
25	Meza and Dijkema	2008	Analysis of transitions of socio-technical systems to provide support to decision-makers involved in shaping the transition and to explore the feasibility of transition management

Table 17.7 List of papers discussing application of hybrid SD-ABS in economics and business (own elaboration)

The subject of interest	Number of papers	Authors
Innovation management	5	[67] (Scopus); [66] (ScienceDirect); [68] (Scopus); [69] (Scopus); [70] (Scopus)
Project management	3	[71] (Scopus); [72] (Scopus); [73] (Scopus)
Product management	3	[74] (ScienceDirect); [75] (Scopus); [76] (ScienceDirect)
Production management	2	[77] (Scopus); [78] (ScienceDirect)
Supply chain management	2	[79] (Scopus); [80] (Scopus)
Strategic management	2	[81] (Scopus); [82] (Scopus)
Water resource management	2	[83] (Scopus); [84] (Scopus)
Transition management	2	[85] (Scopus); [86] (Scopus)
Waste management	1	[87] (ScienceDirect)
Workforce management	1	[88] (Scopus)
Sustainable development	1	[89] (ScienceDirect)
Service management	1	[90] (Scopus)

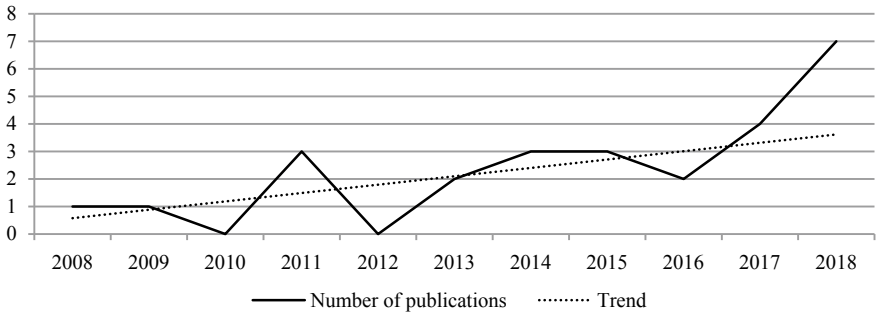


Fig. 17.3 Number of publications presenting the applications of DS-ABS hybrid in economics and business

Acknowledgements The project is financed within the framework of the program of the Minister of Science and Higher Education under the name "Regional Excellence Initiative" in the years 2019–2022, project number 001/RID/2018/19, the amount of financing PLN 10,684,000.00.

References

1. Humphreys, P.: Computer simulations. In: Fine, A., Forbes, M., Wessels, L. (eds.) PSA 1990. Philosophy of Science Association, East Lansing, pp. 497–506 (1991)

2. Humphreys, P.: Extending ourselves: computational science, empiricism, and scientific method. Oxford University Press, New York (2004)
3. Winsberg, E.: Sanctioning Models: the epistemology of simulation. Sci. Context **12**, 275–292 (1999). https://doi.org/10.1017/s0269889700003422
4. Winsberg, E.: Simulations, models, and theories: complex physical systems and their representations. Philos. Sci. **68**(3), 442–454 (2001)
5. Nava Guerrero, G.D.C., Schwarz, P., Slinger, J.H.: A recent overview of the integration of system dynamics and agent-based modelling and simulation. In: Proceedings of the 34th International Conference of the System Dynamics Society, Delft, Netherlands, 17–21 July 2016
6. Meadows, D., Robinson, J.M.: The electronic oracle: computer models and social decisions. Syst. Dyn. Rev. **18**, 271–308 (2002). https://doi.org/10.1002/sdr.239
7. Behdani, B.: Evaluation of paradigms for modeling supply chains as complex socio-technical systems. In: Proceedings of the 2012 Winter Simulation Conference WSC, pp. 1–15, Berlin, Germany, 9–12 Dec 2012. https://doi.org/10.1109/wsc.2012.6465109
8. Borshchev, A.: The big book of simulation modelling. Anylogic, North America (2013)
9. Brailsford, S.C., Eldabi, T., Kunc, M., Mustafee, N., Osorio, A.F.: Hybrid simulation modelling in operational research: a state-of-the-art review. Eur. J. Oper. Res. (In Press): 1–17. (2018). Doi: https://doi.org/10.1016/j.ejor.2018.10.025
10. Mustafee, N., Brailsford, S.C., Djanatliev, A., Eldabi, T., Kunc, M., Tolk, A.: Purpose and benefits of hybrid simulation: contributing to the convergence of its definition. In: Proceedings of the 2017 Winter Simulation Conference (WSC), pp. 1631–1645, Crystal City, VA, 3–6 Dec 2017
11. Djanatliev, A., German, R., Kolominsky-Rabas, P., Hofmann, B.M.: Hybrid simulation with loosely coupled system dynamics and agent-based models for prospective health technology assessments. In: Proceedings of the 2012 Winter Simulation Conference (WSC), pp. 1–12., Berlin, Germany, 09–12 Dec 2012. https://doi.org/10.1109/wsc.2012.6465024
12. Eldabi, T, Balaban, M., Brailsford, S.C., Mustafee, N., Nance, R.E., Onggo, B.S., Sargent, R.G.: Hybrid Simulation: historic.al lessons, present challenges and futures. In: Proceedings of the 2016 Winter Simulation Conference (WSC), pp. 1388–1403, Arlington, Virginia, 11–14 Dec 2016
13. Swinerd, C., McNaught, K.R.: Design classes for hybrid simulations involving agent-based and system dynamics models. Simul. Model. Pract. Theory **25**, 118–133 (2012). https://doi.org/10.1016/j.simpat.2011.09.002
14. Martin, R., Schlüter, M.: Combining system dynamics and agent-based modeling to analyze social-ecological interactions—An example from modeling restoration of a shallow lake. Front. Environ. Sci. **3**, 66 (2015). https://doi.org/10.3389/fenvs.2015.00066
15. Vincenot, C.E., Giannino, F., Rietkerk, M., Moriya, K., Mazzoleni, S.: Theoretical considerations on the combined use of system dynamics and individual-based modeling in ecology. Ecol. Model. **222**, 210–218 (2011). https://doi.org/10.1016/j.ecolmodel.2010.09.029
16. Hamilton, S.H., ElSawah, S., Guillaume, J.H.A., Jakeman, A.J., Pierce, S.A.: Integrated assessment and modelling: overview and synthesis of salient dimensions. Environ. Modell. Softw. **64**, 215–229 (2015)
17. Stemate, L., Pasca, C., Taylor, I.: A comparison between system dynamics and agent based modeling and opportunities for cross-fertilization. In: Proceedings of the 2007 Winter Simulation Conference (WSC), Washington, DC, 9–12 Dec 2007
18. Wallentin, G., Neuwirth, Ch.: Dynamic hybrid modelling: switching between AB and SD designs of a predator-prey model. Ecol. Model. **345**, 165–175 (2017)
19. Forrester, J.W.: Industrial dynamics: a major breakthrough for decision makers. Harv. Bus. Rev. **36**, 37–66 (1958)
20. Phelan, S.E.: A note on the correspondence between complexity and systems theory. Syst. Pract. Action Res. **12**, 237–246 (1999). https://doi.org/10.1023/A:1022495500485
21. Coyle, R.G.: Management System Dynamics. Wiley, Chichester (1979)
22. Sterman, J.D.: Business Dynamics. Systems Thinking and Modeling for a Complex World, McGraw-Hill Education, Boston (2000)

23. Macal, C.M.: To agent-based simulation from system dynamics. In: Proceedings of the 2010 Winter Simulation Conference (WSC), pp. 371–382, Baltimore, Maryland, USA, 5–8 Dec 2010. https://doi.org/10.1109/wsc.2010.5679148
24. Parunak, H.V.D., Savit, R., Riolo, R.L.: Agent-based modeling versus equation-based modeling: a case study and users' guide. In: Sichman, J.S., Conte, R., Gilbert, N. (eds.) Multi-Agent Systems and Agent-Based Simulation, pp. 10–25. Springer, Berlin - Heidelberg (1998)
25. Meadows, D.: Leverage points: places to intervene in a system. Solutions J. **1**(1), 41–49 (2010)
26. Scholl, H.J.: Agent-based and system dynamics modeling: a call for cross study and joint research. In: Proceedings of the 34th Hawaiian International Conference on Systems Sciences, IEEE Computer Society, p. 8, Maui, Hawaii, 3–6 Jan 2001. (2001a). https://doi.org/10.1109/hicss.2001.926296
27. Macal, C.M., North, M.J.: Tutorial on agent-based modeling and simulation PART 2: how to model with agents. In: Proceedings of the 2006 Winter Simulation Conference (WSC), pp. 73–83, Monterey, CA, USA, 3–6 Dec 2006. https://doi.org/10.1109/wsc.2006.323040
28. Frank, A.U., Bittner, S., Raubal, M.: Spatial and cognitive simulation with multi-agent systems. In: Montello, D.R. (ed.) Spatial Information Theory—Foundations of Geographic Information Science. Springer, Berlin, Heidelberg, p 124–139 (2001)
29. Wooldridge, M.: Intelligent Agents. In: Weiss, G. (ed.) Multiagent Systems—A Modern Approach to Distributed Artificial Intelligence, pp. 27–77. MIT Press, Cambridge (1999)
30. Wooldridge, M.: An Introduction to Multi-agent Systems. Wiley, New York (2002)
31. Łatuszyńska, M., Wawrzyniak, A., Furaiji, F.: Multi-agent based simulation in investigating consumer behavior. Iraqi J. Market Res. Consum. Prot. **4**(2), 141–158 (2012)
32. Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. **99**, 7280–7287 (2002). https://doi.org/10.1073/pnas.082080899
33. Epstein, J.M.: Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton University Press, Princeton (2006)
34. Epstein, J.M., Axtel, R.L.: Growing Artificial Societies: Social Science from the Bottom Up. MIT Press, Cambridge (1996)
35. Jennings, N.R., Sycara, K., Wooldridge, M.: A roadmap of agent research and development. Auton. Agents Multi-Agent Syst. **1**, 7–38 (1998). https://doi.org/10.1023/A:1010090405266
36. Holland, J.H.: Emergence: from chaos to order. Addison-Wesley, Redwood City (1998)
37. Cherif, R., Davidsson, P.: Software development process simulation: multi agent-based simulation versus system dynamics. Lect. Notes Comput. Sci. **5683**, 73–85 (2010). https://doi.org/10.1007/978-3-642-13553-8_7
38. Ding, Z., Gong, W., Li, S., Wu, Z.: System dynamics versus agent-based modeling: a review of complexity simulation in construction waste management. Sustainability **10**(7), 1–13 (2018)
39. Fallah-Fini, S., Rahmandad, H., Chen, H.J., Xue, H., Wang, Y.: Connecting micro dynamics and population distributions in system dynamics models. Syst. Dyn. Rev. **29**(4), 197–215 (2013)
40. Lättilä, L., Hilletofth, P., Lin, B.: Hybrid simulation models–when, why, how? Expert Syst. Appl. **37**, 7969–7975 (2010)
41. Osgood, N.D.: Using traditional and agent based toolsets for system dynamics: present tradeoffs and future evolution. In: Proceedings of the 2007 International Conference of the System Dynamics Society, Boston, Massachusetts, USA, 29 July–2Aug 2007
42. Teose, M., Ahmadizadeh, K., O'Mahony, E., Smith, R.L., Lu, Z., Ellner, S.P., Gomes, C., Grohn, Y.: Embedding system dynamics in agent based models for complex adaptive systems. In: Walsh, T. (ed.) Proceedings of The Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI'11), vol. 3, pp. 2531–2538, AAAI Press. (2011). doi:dx.doi.org/https://doi.org/10.5591/978-1-57735-516-8/ijcai11-421
43. Swanson, J.: Business dynamics—Systems thinking and modeling for a complex world. J. Oper. Res. Soc. **53**(4), 472–473 (2002). https://doi.org/10.1057/palgrave.jors.2601336
44. Railsback, S.F., Grimm, V.: Agent-Based and Individual-Based Modeling: A Practical Introduction. Princeton University Press, Princeton (2011)
45. Bobashev, G.V., Goedecke, D.M., Yu, F., Epstein, J.M.: A hybrid epidemic model: combining the advantages of agent-based and equation-based approaches. In: Proceedings of the 2007 Winter Simulation Conference (WSC), p 1532–1537, Washington, DC, 9–12 Dec 2007

46. Dyke, V., Savit, R., Riolo, R.L.: Agent-based modeling versus equation-based modeling: a case study and users' guide. Lect. Notes Comput. Sci. **1534**, 10–25 (1998)
47. Marin, M., Zhu, Y., Meade, P.T., Sargent, M., Warren, J.: System dynamics and agent-based simulations for workforce climate. In: Proceedings of the 2006 Winter Simulation Conference (WSC), pp. 667–671, Monterey, CA, USA, 3–6 Dec 2006
48. Norling, E.: Contrasting a system dynamics model and an agent-based model of food web evolution. Lect. Notes Comput. Sci. **4442**, 57–68 (2006)
49. Rahmandad, H., Sterman, J.: Heterogeneity and network structure in the dynamics of diffusion: comparing agent-based and differential equation models. Manage. Sci. **54**, 998–1014 (2008)
50. Wakeland, W., Gallaher, E.J., Macovsky, L.M., Aktipis, C.A.: A comparison of system dynamics and agent-based simulation applied to the study of cellular receptor dynamics. In: Proceedings of the 37th Hawaiian International Conference on Systems Science, Big Island, Hawaii, 5–8 Jan 2004
51. Wilson, W.G.: Resolving discrepancies between deterministic population models and Individual-Based simulations. Am. Nat. **151**, 116–134 (1998)
52. Akkermans, H.A.: Emergent supply networks: system dynamics simulation of adaptive supply agents. In: Proceedings of the 34th Hawaiian International Conference on Systems Sciences, IEEE Computer Society, Maui, Hawaii, 3–6 Jan 2001
53. Scholl, H.J.: Looking across the fence: comparing findings from SD modeling efforts with those of other modeling techniques. In: Proceedings of the 19th International Conference of the System Dynamics Society, Atlanta, Georgia, 23–27 July 2001 (2001b)
54. Pourdehnad, J., Maani, K.E., Sedehi, H.: System dynamics and intelligent agent-based simulation: where is the synergy? In: Proceedings of the 20th International Conference of the System Dynamics Society, Palermo, Italy, 28 July–1 Aug 2002
55. Schieritz, N.: Integrating system dynamics and agent-based modeling. In: Proceedings of the 20th International Conference of the System Dynamics Society, Palermo, Italy, 28 July–1 Aug 2002
56. Größler, A., Stotz, M., Schieritz, N.: A software interface between system dynamics and agent-based simulations: linking Vensim® and RePast®. In: Proceedings of the 21st International Conference, System Dynamics Society, pp. 20–24, New York, NY, USA, 20–24 July 2003
57. Swinerd, C.: On the Design of Hybrid Simulation Models: Focusing on the Agent-Based System Dynamics Combination. Cranfield University, PhD (2014)
58. Shanthikumar, J.G., Sargent, R.G.: A unifying view of hybrid simulation/analytic models and modeling. Oper. Res. **31**(6), 1030–1052 (1983)
59. Kim, Y., Son, J., Lee, Y.-S., Lee, M., Hong, J., Cho, K.: Integration of an individual-oriented model into a system dynamics model: an application to a multi-species system. Environ. Model Softw. **112**, 23–35 (2019)
60. BenDor, T., Scheffran, J., Hannon, B.: Ecological and economic sustainability in fishery management: a multi-agent model for understanding competition and cooperation. Ecol. Econ. **68**, 1061–1073 (2009)
61. Martinez-Moyano, I.J., Sallach, D.L., Bragen, M.J., Thimmapuram, P.R.: Design for a multilayer model of financial stability: exploring the integration of system dynamics and agent-based models. In: Proceedings of the 25th International Conference of System Dynamics Society, Boston, Mass, 29 July –2 August 2007
62. Schieritz, N., Größler, A.: Emergent structures in supply chains: a study integrating agent-based and system dynamics modeling. In: Proceedings of the 36th Hawaiian International Conference on Systems Science, Waikoloa, Hawaii, 06–09 Jan 2003
63. Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L.K., Young, T.: Simulation in manufacturing and business: a review. Eur. J. Oper. Res. **203**(1), 1–13 (2010)
64. Barbosa, C., Azevedo, A.: Hybrid simulation for complex manufacturing value-chain environments. Procedia Manuf. **11**, 1404–1412 (2017). https://doi.org/10.1016/j.promfg.2017.07.270
65. Scheidegger, A.P., Pereira, T.F., Oliveira, M.L., Banerjee, A., Montevechi, J.A.: An introductory guide for hybrid simulation modelers on the primary simulation methods in industrial engineering identified through a systematic review of the literature. Comput. Ind. Eng. **124**, 474–492 (2018). https://doi.org/10.1016/j.cie.2018.07.046

66. Kolominsky-Rabas, P.L., Djanatliev, A., Wahlster, P., Gantner-Bär, M., Hofmann, B., German, R., Sedlmayr, M., Reinhardt, E., Schüttler, J., Kriza, Ch.: Technology foresight for medical device development through hybrid simulation: the ProHTA project. Technol. Forecast. Soc. Chang. **97**, 105–114 (2015)
67. Ebrahimi, M.: Hybrid simulation approach for technological innovation policy making in developing countries. Adv. Intell. Syst. Comput. **528**, 109–119 (2017). https://doi.org/10.1007/978-3-319-47253-9_9
68. Swinerd, C., McNaught, K.R.: Simulating the diffusion of technological innovation with an integrated hybrid agent-based system dynamics model. J. Simul. **8**(3), 231–240 (2014)
69. Shafiei, E., Stefansson, H., Asgeirsson, E.I., Davidsdottir, B., Raberto, M.: Integrated agent-based and system dynamics modelling for simulation of sustainable mobility. Trans. Rev. **33**(1), 44–70 (2013). https://doi.org/10.1080/01441647.2012.745632
70. Swinerd, C., McNaught, K.R.: Comparing a simulation model with various analytic models of the international diffusion of consumer technology. Technol. Forecast. Soc. Chang. **100**, 330–343 (2015)
71. Nasirzadeh, F., Khanzadi, M., Mir, M.: A hybrid simulation framework for modelling construction projects using agent-based modelling and system dynamics: an application to model construction workers' safety behavior. Int. J. Constr. Manage. **18**(2), 132–143 (2018). https://doi.org/10.1080/15623599.2017.1285485
72. Liang, H., Lin, K.-Y., Zhang, S.: Understanding the social contagion effect of safety violations within a construction crew: a hybrid approach using system dynamics and agent-based modeling. Int. J. Environ. Res. Public Health **15**(12), 2696 (2018). https://doi.org/10.3390/ijerph15122696
73. Khanzadi, M., Nasirzadeh, F., Mir, M., Nojedehi, P.: Prediction and improvement of labor productivity using hybrid system dynamics and agent-based modeling approach. Constr. Innov. **18**(1), 2–19 (2018). https://doi.org/10.1108/CI-06-2015-0034
74. Nassehi, A., Colledani, M.: A multi-method simulation approach for evaluating the effect of the interaction of customer behaviour and enterprise strategy on economic viability of remanufacturing. CIRP Ann. Manuf. Technol. **67**, 33–36 (2018)
75. Rondini, A., Tornese, F., Gnoni, M.G., Pezzotta, G., Pinto, R.: Comparing simulation approaches for designing sustainable PSS. In: Proceedings of the Summer School Francesco Turco, pp. 108–113, Senigallia, 09–12 Sep 2014
76. Wang, B., Brême, S., Moon, Y.B.: Hybrid modeling and simulation for complementing lifecycle assessment. Comput. Ind. Eng. **69**(1), 77–88 (2014)
77. Barbosa, C., Azevedo, A.: Hybrid modelling of MTO/ETO manufacturing environments for performance assessment. Int. J. Prod. Res. **56**(15), 5147–5171 (2018). https://doi.org/10.1080/00207543.2017.1421788
78. Barbosa, C., Azevedo, A.: Towards a hybrid multi-dimensional simulation approach for performance assessment of MTO and ETO manufacturing environments. Procedia Manuf. **17**, 852–859 (2018)
79. Fahhama, L., Zamma, A., Mansouri, K., Elmajid, Z.: Towards a mixed method model and simulation of the automotive supply chain network connectivity. In: 2017 International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA), p 13–18, Rabat, 27–28 Apr 2017
80. Manisri, C., Pichitlamken, J.: Hybrid Simulation Model for the Upstream Rubber Supply Chain. Industrial Engineering & Management Systems **16**(9), 330–341 (2017). https://doi.org/10.7232/iems.2017.16.3.330
81. Kortelainen, S., Lättilä, L.: Hybrid modeling approach to competitiveness through fast strategy. Int. J. Innovation Technol. Manage. **10**(5), 1340016 (2013). https://doi.org/10.1142/S0219877013400166
82. Kieckhäfer, K., Walther, G., Axmann, J., Spengler, T.: Integrating agent-based simulation and system dynamics to support product strategy decisions in the automotive industry. In: Proceedings of the 2009 Winter Simulation Conference (WSC), p 1433–1443, Austin, Texas, USA, 13–16 Dec 2009

83. Alvi, M.S.Q., Mahmood, I., Javed, F., Malik, A.W., Sarjoughian, H.: Dynamic behavioural modeling, simulation and analysis of household water consumption in an urban area: a hybrid approach. In: Proceedings of the 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden, 09–12 Dec 2018
84. Nikolic, V.V., Simonovic, S.P.: Multi-method modeling framework for support of integrated water resources management. Environ. Process. **2**(3), 461–483 (2015). https://doi.org/10.1007/s40710-015-0082-6
85. Meza, C.M.C., Dijkema, G.P.J.: Modelling infrastructure systems: a hybrid approach for system transition. In: 1st International Conference on Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA), pp. 1–6, Rotterdam, Netherlands, 10–12 Nov 2008. https://doi.org/10.1109/infra.2008.5439677
86. Asif, F.M.A., Lieder, M., Rashid, A.: Multi-method simulation based tool to evaluate economic and environmental performance of circular product systems. J. Clean. Prod. **139**, 1261–1281 (2016). https://doi.org/10.1016/j.jclepro.2016.08.122
87. Elia, V., Gnoni, M.G., Tornese, F.: Improving logistic efficiency of WEEE collection through dynamic scheduling using simulation modeling. Waste Manage **72**, 78–86 (2018)
88. Block, J.: A hybrid modeling approach for incorporating behavioral issues into workforce planning. In: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp. 326–331, Budapest, Hungary, 9–12 Oct 2016. https://doi.org/10.1109/smc.2016.7844261
89. Sitepu, M.H., McKay, A., Holt, R.J.: Towards a framework for sustainable development planning in the Indonesian natural rubber industry supply network. Procedia CIRP **48**, 164–169 (2016)
90. Wang, Y., Van Den Heuvel, W.-J.: Towards a hybrid simulation modelling framework for service networks. Lect. Notes Comput. Sci. **6569**, 119–128 (2011). https://doi.org/10.1007/978-3-642-22760-8_13

Part III
Practical Issues—Case Studies

Chapter 18
Measuring Neurophysiological Signals, Fixations and Self-report Data for Product Placement Effectiveness Assessment in Music Videos

Ana C. Martinez-Levy, Giulia Cartocci, Enrica Modica, Dario Rossi, Marco Mancini, Arianna Trettel, Fabio Babiloni and Patrizia Cherubino

Abstract Product placement is a marketing technique that, by inserting products into a narrative structure, constitutes a likely effective tool to increase the visibility and notoriety of a brand. For years, the opportunities for product placement in music videos were limited. Recently, there has been a growth of interest for this tool/advertising modality since the digital community allowed the possibility to move videos from television to the Internet. The scope of the present study is to investigate the effectiveness of the product placement in music videos. An electroencephalographic (EEG) index called mental effort (ME) has been analyzed, in addition to the emotional index (EI), calculated by the combination of galvanic skin response (GSR) and heart rate (HR) signals. Self-report responses have also been collected through an online questionnaire and interviews, since one experimental question was to investigate whether viewing a video containing a commercial product could influence the declared recall of the product inserted in it and the spontaneous recall of the video itself. Furthermore, fixations related to the product inserted in videos have been obtained by the eye-tracking technique (ET). Higher values of the ME ($p = 0.016$) and EI ($p = 0.033$) have been found for videos with product placement in comparison to videos without it. In addition, results show that the number of fixations affects the recall of the showed products ($p < 0.001$). These findings highlight that using product placement in famous singers' music videos is an effective technique for prompting product recall and how it helps to focus the visual attention on them.

Keywords Product placement · EEG · Fixations · Emotion · Mental effort · Recall

A. C. Martinez-Levy (✉)
Department of Communication and Social Science, Sapienza University of Rome, Rome, Italy
e-mail: ana.martinezlevy@uniroma1.it

G. Cartocci · E. Modica · D. Rossi · F. Babiloni · P. Cherubino
Department of Molecular Medicine, Sapienza University of Rome, Rome, Italy

A. C. Martinez-Levy · G. Cartocci · M. Mancini · A. Trettel · F. Babiloni · P. Cherubino
BrainSigns Srl, Rome, Italy

F. Babiloni
College of Computer Science and Technology, University Hangzhou Dianzi, Hangzhou, China

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_18

18.1 Introduction

Product placement is a form of ever-expanding advertising, as it has different characteristics and advantages compared to traditional advertising. It is indirect advertising and, therefore, often perceived as less intrusive by the viewers [22]. The American Federal Trade Commission defines product placement as "*a form of advertising whereby advertisers place branded products in TV programs for a certain fee or other consideration.*" This technique has its beginning in a natural way with the appearance of the cinematographer in 1895 [16]. The Lumière brothers were the first to put into practice the technique of product placement and quickly commercial brands began a relationship with different media communication. In 1926, in the western movie "The Texas Streak," the Coca-Cola brand appeared for the first time in the film industry. Many other brands followed this example with appearances in cinematic productions. The interest in product placement increased dramatically in the 1980s when, with the release of the film "ET," the Reese's Pieces sweets from Hershey foods increased their sales by 65% and became M & M's competitors [23]. Gradually, the spot began to break away and to differentiate itself from the film, even if it continued to be influenced in some ways. Currently, product placement is in fact widely used as a marketing technique within films and television programs, thus allowing the product to reach, and therefore be seen by, an extremely wide audience that is difficult to achieve through traditional advertising. Most research to date on product placement has been focused on movies and television, but not much research has looked more in-depth at product placement in music videos. While it has been studied how product placement during TV exposure increases the memorization of products [17], the few studies focusing on music videos have been conducted with traditional marketing techniques such as questionnaires or interviews. The music industry's revenue from brand or product placement in music videos more than doubled between 2000 and 2010 [25]. Furthermore, in 2017, product placement in online and mobile media jumped 25.3% to over $205 million, according to PQ Media. It has been the fastest growing product placement channel until the date. Product placement appears in music videos on YouTube's networks [27]. The interest of marketers in brand/product placement lies in knowing that today's multi-screen world reduces the effectiveness of television breaks since people do not focus on ads anymore. Fans of music videos, on the other hand, are very engaged in watching the videos, and typically viewing them many times suggesting that brand or product placement within popular music videos are likely to be more effective than the promotional messages of commercial advertisements. The traditional measurements used in marketing research include recall, which is a direct measure that suggests evidence of conscious processing. It relies on an audience ability to identify the memory trace consciously [30]. During product placement exposure, unconscious processing takes place on the audience mind, and so implicit measures should determine the conscious effects [2, 17, 32]. With neuromarketing research, it is possible to obtain consumer insights without a verbal declaration of participants through different neurophysiological tools. Different indexes or psychologic constructs can be analyzed with those

techniques and different stimuli can be tested, not only product placement (please see [9, 12, 13, 19, 21]). The aim of this chapter is thus to contribute to the as of yet sparse literature for product placement studies relative to music videos using neuromarketing techniques. Specially, the levels of two different neurophysiological indexes such as ME index and EI of participants during the visualization of music videos were measured. Also eye fixations on products and self-report answers of participants were analyzed, in order to see how products inserted in music videos are watched and if they are recalled.

18.2 Methodology and Sample

The research involved 22 healthy volunteers (12 Males, 10 Females; average age $= 25.18 \pm 2.28$). During the study, brain activity, physiological reactions, and eye movements of each participant were measured. Particularly, the procedure of the experimental task consisted in watching a film in which a series of music videos were inserted. The participants were told to pay attention to the film. It contained eight music videos: four of them with a commercial product inserted and the other four without. All music videos were randomized, and the total duration of the film was ten minutes. After watching the film, each participant saw twenty images of different products, with eight of them being present in the previously observed videos and 12 of them, the distractors, not. After each image, an online questionnaire was administered in order to collect the product's recall, with the following questions: *How much do you like this product from 0 to 10?/Do you remember if these products were on the music videos that you saw before*? At the end of the experience, a short interview took place, with the aim to collect the spontaneous recall of the videos, the products, and the individual music preferences. The music videos with product placement (PPV) were: Coldplay "Adventure Of A Lifetime"; Avicii "Wake Me Up"; Jennifer Lopez ft. Pitbull "Live It Up"; and Lady Gaga and Beyoncé "Telephone." On the other hand, the music videos without product placement (NPPV) were: Pharrel Williams "Happy"; U2 "Ordinary Love"; Adele "Hello"; and Mark Ronson ft. Bruno Mars "Uptown Funk." The products that appeared in the PPV were "Polaroid" photographic camera, "Nokia" and "Sony" smartphones; "ICE" watch; "Ralph Lauren" brand clothing; "Chanel" and "Swarovski" sunglasses; "Beats" speakers. The products inserted as distractors were "Fujifilm" and "Kodak" photographic cameras; "Htc" and "Lg" smartphones; "Liu Jo" and "Swatch" watches; "Calvin Klein" and "MCS" brand clothing; "Gucci" and "Rayban" sunglasses; and "Bose" and "Jbl" speakers. Figure 18.1 shows the experimental protocol.

During the experimental procedure with videos and images, the neurophysiological data was recorded. Particularly, the biosignals recorded during the entire experimental protocol were electroencephalography (EEG), heart rate (HR), galvanic skin response (GSR) and eye-tracking data (ET).

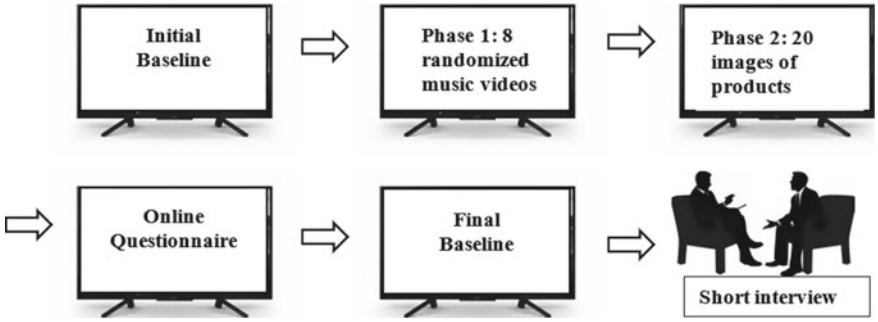


Fig. 18.1 Picture shows the experimental protocol for the study

18.2.1 EEG Recording and Signal Processing

The EEG activity was recorded using ten electrodes (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF5, AF6, AF7, and AF8) placed on the frontal portion of the scalp of participants by means of a portable 21-channels system (BEmicro, EBneuro, and Italy). Although the system allowed to record up to 21 channels, a ready-made headband with ten electrodes placed over the prefrontal and frontal cortex was used, since only this cortical area (prefrontal and frontal) was of interest in the present study and reduced the system's invasiveness and increased the comfort of the participant when compared with traditional EEG caps. The reference and the ground electrodes have been placed, respectively, on the left and right earlobes. The signals have been acquired at a sampling rate of 256 Hz and the impedances were kept below 10 kΩ. After the acquisition phase, the raw EEG signal has been digitally pre-processed by using the EEGLAB [14] MATLAB toolbox. Firstly, a notch filter (50 Hz) was applied in order to reject the main current interference. Secondly, the gathered signal has been band-pass filtered by a 5th order Butterworth filter (2–30 Hz), in order to reject the continuous component as well as high-frequencies interferences, such as muscular artefacts. Then, the independent component analysis (ICA), in particular, the SOBI algorithm [4] has been applied to EEG data in order to identify and remove the component related to eye blinks and eye movements, since their contribution overlaps the EEG bands of interest [15]. The component has been manually selected in order to be removed, and after that the EEG signal has been reconstructed. Furthermore, in order to clean the EEG signal as much as possible, after these conservative steps (until now no EEG data has been lost), the EEG signal segments still affected by artefacts have been automatically detected and rejected. To compute the activity of the cortical areas of interest in a specific frequency band, the global field power (GFP) was then computed. This is a measurement to summarize the synchronization level of the brain activity over the scalp surface [18]. GFP is computed from a specific set of electrodes by performing the sum of squared values of EEG potential at each electrode, averaged for the number of involved electrodes, resulting in a time-varying waveform related to the increase or decrease of the global power in the analyzed EEG.

The GFP formula is presented in the following:

$$\mathrm{GFP} = \frac{1}{N}\sum_{i=1}^{N} X_{\theta_i}(t)^2$$

where θ is the considered EEG band on the frontal cortical area, N is the number of electrodes included in the area of interest, and i is the electrodes' index. To evaluate the ME, EEG activity in the theta band of all the frontal electrodes has been considered for the GFP computation. An increase in the frontal theta would imply an increase in the task difficulty or a higher decodification of information [34]. ME index was estimated for each second, and then normalized with respect to the index of the baseline (1 min of open eyes).

18.2.2 The Autonomic Data Recordings and Signal Processing

The blood volume pulse (BVP) and galvanic skin response (GSR) were recorded with the Shimmer System (Shimmer Sensing, Ireland) with a sampling rate of 64 Hz. For the recording of these signals, two electrodes are placed on the palmar side of the middle phalanges of the second and third fingers on the non-dominant hand of the participant in order to acquire the GSR signal according to published procedures [8] and a photoplethysmographic sensor is placed on the thumb of the same hand for the BVP recording. In order to obtain the heart rate (HR) signal from the BVP, it has been used the Pan–Tompkins algorithm [24]. The constant voltage method (0.5 V) is employed for the acquisition of the GSR, then, by using the LEDAlab software [5], the tonic component of the skin conductance (skin conductance level, SCL) is estimated. Then the combination of HR and GSR signal produces a monodimensional variable which returns the emotional state of subjects [33]. So, for the EI, we refer to effects plane [26, 28] where the coordinates of a point in this space are defined by the horizontal axis (HR) and the vertical axis (GSR). Several studies have highlighted that these two autonomic parameters correlate with valence and arousal, respectively [20]. The interpretation of the EI implies that the higher the value the more emotional engagement experienced by the subject is, and vice versa.

18.2.3 Eye Tracker Recordings and Signal Processing

Eye-tracking data have been acquired by Tobii Pro X2-30 screen-based eye tracker with a sampling frequency of 30 Hz, in order to identify eye fixations on the proposed stimuli. Firstly, all the artifactual or not physiological point of gaze was automatically removed. Secondly, eye-tracking data was analyzed with Tobii studio 3.4.8 for

the extraction of information about fixations in each area of interest (AOI), such as a number of fixations on each AOI; in this case, the AOI was the products inserted in music videos. Basing upon the total number of fixations recorded on the screen throughout the movie, the percentage of fixations of eye-tracking metric has been performed for each subject in order to evaluate the visual attention elicited by each specific AOI. It has been obtained by dividing the number of fixations of each participant for each AOI by the total number of fixations recorded on the screen throughout the movie for each participant.

18.2.4 Data Analysis

For the product's recall analysis, a *t*-test was performed for evaluating the fixation percentage on recalled and non-recalled products during the online questionnaire.

A *t*-test was performed for evaluating the differences between both kinds of videos (PPV and NPPV) with the data from the spontaneous declared recall of products during the final interview, in order to compare the videos' differences.

A paired *t*-test was performed for evaluating the differences between PPV and NPPV with the data from the mental effort index. Another paired *t*-test was performed for evaluating the differences between PPV and NPPV with the data from the emotional engagement index.

18.3 Results

18.3.1 Eye Fixations and Recall of Products

The first evidence of the present research is the relation between self-reported data and eye-tracking data. This relation is between the recall of products obtained during the online questionnaire, where participants answered to the question "*Do you remember if this product was present in the music videos you just saw?,*" and the percentage of fixations that was dedicated to each product contained in each AOI during the videos' visualization. Figure 18.2 shows that the participants that remembered the products included in the music videos are also those who have dedicated, on average, a higher percentage of fixations on the products ($t = -3.735$; $p < 0.001$). An image representation of the heatmap for the fixation during a PPV is shown in Fig. 18.3.

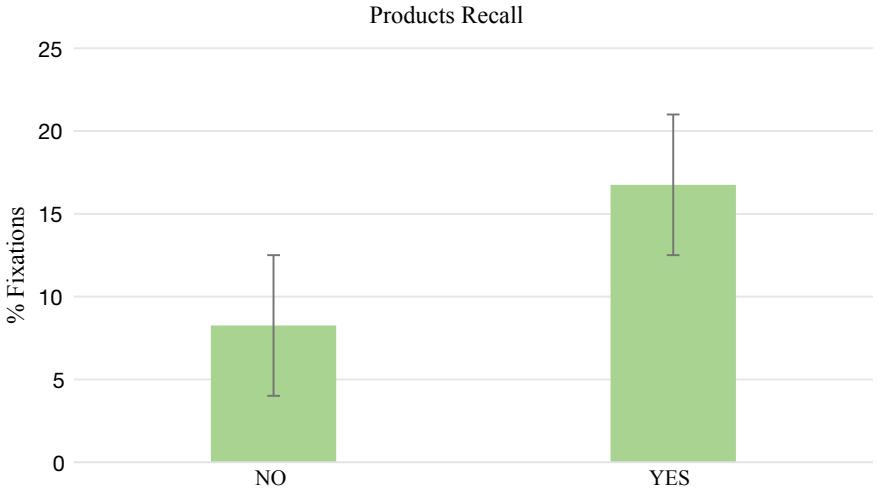


Fig. 18.2 Graph represents the statistical results from the *t*-test considering the fixation percentage on recalled (YES) and non-recalled (NO) products during the online questionnaire. Error bars represent standard error

Fig. 18.3 Picture shows a heatmap image detected during a PPV visualization

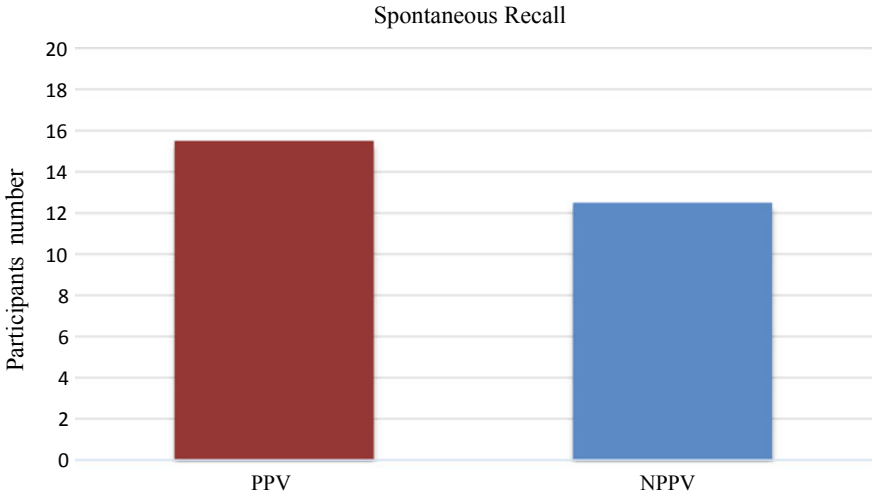


Fig. 18.4 Graph shows the different answers of participants for the spontaneous recall of videos during the interviews

18.3.2 Spontaneous Recall of Videos: Product Placement Videos Versus Non-product Placement Videos

There is a higher spontaneous recall of PPV than NPPV within the interview results. The difference is not statistically significant ($p > 0.05$). Notwithstanding, Fig. 18.4 demonstrates the difference in the participants' answers, which confirms that the placement of products, at least in this case, has led to greater remembrance not only of the products, but of the video itself.

18.3.3 Mental Effort: Product Placement Videos Versus Non-product Placement Videos

Regarding the cognitive index of mental effort, from the comparison between PPV and NPPV, a higher cognitive effort emerged for the first ones (see Fig. 18.5). This difference was statistically significant ($t = 2.786$; $p = 0.016$).

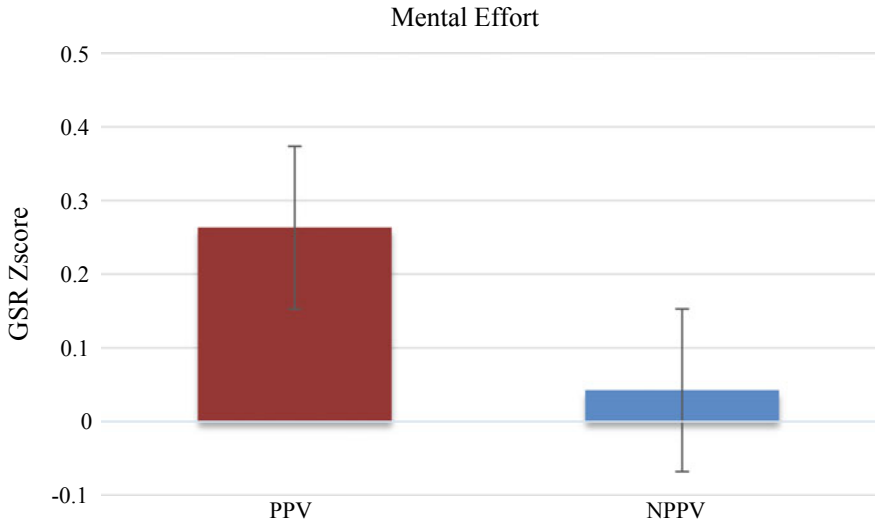


Fig. 18.5 Graph represents the statistical results from the *t*-test conducted for the ME index values concerning the comparison among PPV and NPPV. Error bars represent standard error

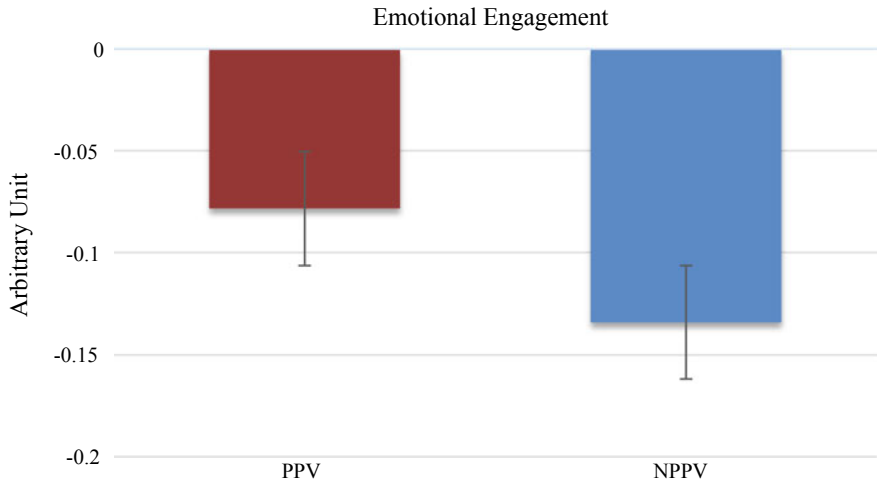


Fig. 18.6 Graph represents the statistical results from the *t*-test conducted for the emotional index values concerning the comparison between PPV and NPPV. Error bars represent standard error

18.3.4 Emotion Engagement: Product Placement Videos Versus Non-product Placement Videos

Both kinds of videos evoked a negative emotional reaction. Nevertheless, in comparison to the NPPV, the PPV achieved a less negative emotional engagement. This statistically significant difference is demonstrated in Fig. 18.6 ($t = 2.405; p = 0.033$).

18.4 Discussion

In accordance with the analyzed indexes of ME index and EI, it has been confirmed that music videos with a product or brand inserted have higher levels of both than music videos without product placement. Regarding the ME index, different researches have previously identified that changes in the EEG spectra reflect levels of mental effort on different tasks [1, 6, 7, 9–12]. During the visualization of a music video, this index could be connected to the presence of commercial products, which require a higher information processing. In addition, the emotional engagement during the visualization of the videos differed between PPV and NPPV. In particular, the former had higher EI values compared to the latter. Since the PPV had higher content decoding, this could be explained with the context of the video, which is being more comprehensible and thus generating a higher level of emotional engagement. The results of self-report data collected through the online questionnaire reveal that for those who remembered the products included in the music videos, there was, on average, a higher percentage of fixations on the products. Therefore, the relation between the number of fixations dedicated to a product and the recall gives us information to confirm a significative, positive relation between both measures. Connectedness research shows that the intensity of the relationship that viewers develop with television programs and the characters in those programs affects their memory of placements [29]. Recall is a direct measure that relies on an audience member's ability to identify the memory trace consciously, which suggests evidence of conscious processing. Although recognition also provides a direct measure, its relationship with unconscious processing is not clear, in that recognition often serves to identify sensory memory content [3]. The relationship between memorization and attention has been widely studied in marketing research, with the consensus being that greater allocation of attention to a stimulus facilitates its memorization [31]. In marketing and particularly in product placement, it is crucial to understand how visual attention influences product and brand recall. During this study, we utilized different videos with and without product placement to analyze and compare the perception of real top ten music videos. An important limitation of the study is that the potential familiarity of participants with the inserted products as well as the type of video might influence results by reflecting individual differences in terms of music preferences and liking of the video content.

18.5 Conclusion

This study sought to determine whether the insertion of commercial products in music videos affects on implicit responses of consumers to the videos as well as to the products themselves and consequently also on the conscious recall. Connecting the number of fixations with visual attention and products' recall answers with their memorization, results denote that products with higher visual attention have higher memorization. Likewise, PPV shows higher values for mental effort, emotional engagement, and spontaneous recall than NPPV. These results highlight the importance of a selected product positioning during a video exposure, as well as, of the time of the products' appearance during their insertion in different communication media. This can help marketers ensuring that visual attention to the products will lead to a higher level of recall.

For the future studies in this field, the context of the video in which the product or brand is placed should be considered. An in-depth analysis could denote the eventual perfect match between the type of video and the product promotion suiting the company's strategy. Likewise, in the future studies, videos could be "laboratory processed" in order to test the same videos with and without product placement. Results could shed light on the influence of product appearance and less on the perception of the video itself, reducing the bias of the context in videos.

References

1. Aricò, P., Borghini, G., Graziani, I., et al.: Towards a multimodal bioelectrical framework for the online mental workload evaluation. In: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, pp. 3001–3004 (2014)
2. Auty, S., Lewis, C.: The 'delicious paradox': preconscious processing of product placements by children. In: The Psychology of Entertainment Media: Blurring the Lines Between Entertainment and Persuasion, pp. 117–133 (2004)
3. Baddeley, A.: Working Memory, Thought, and Action, vol. 45. OUP, Oxford (2007)
4. Belouchrani, A., Abed-Meraim, K., Cardoso, J.F., et al.: A blind source separation technique using second-order statistics. IEEE Trans. Sig. Process. **45**(2), 434–444 (1997)
5. Benedek, M., Kaernbach, C.: A continuous measure of phasic electrodermal activity. J. Neurosci. Methods **190**, 80–91 (2010)
6. Berka, C., Levendowski, D.J., Lumicao, M.N., et al.: EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. Aviat. Space Environ. Med. **78**(5), B231–B244 (2007)
7. Borghini, G., Astolfi, L., Vecchiato, G., et al.: Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci. Biobehav. Rev. **44**, 58–75 (2014)
8. Boucsein, W.: Electrodermal Activity. Springer Science & Business Media (2012)
9. Cartocci, G., Caratù, M., Modica, E., et al.: Electroencephalographic, heart rate, and galvanic skin response assessment for an advertising perception study: application to antismoking public service announcements. J. Vis. Exp. JoVE (126) (2017)
10. Cartocci, G., Maglione, A.G., Vecchiato, G., et al.: Mental workload estimations in unilateral deafened children. In: Engineering in Medicine and Biology Society (EMBC), 37th Annual International Conference of the IEEE, pp. 1654–1657, Aug 2015

11. Cartocci, G., Maglione, A.G., Rossi, D., et al.: The influence of different cochlear implant features use on the mental workload index during a word in noise recognition task. Int. J. Bioelectromagn. **18**, 60–66 (2016)
12. Cartocci, G., Maglione, A.G., Modica, E., et al.: Is the younger the less effortful? An electroencephalographic comparison among consecutive generations of cochlear implant sound processors. Int. J. Bioelectromagn. **19**(1), 11–17 (2017)
13. Cherubino, P., Cartocci, G., Modica, E., et al.: Wine tasting: how much is the contribution of the olfaction? In: International Conference on Computational Methods in Experimental Economics, pp. 199–209. Springer, Cham (2017)
14. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods **134**(1), 9–21 (2004)
15. Di Flumeri, G., Aricó, P., Borghini, G., et al.: A new regression-based method for the eye blinks artifacts correction in the EEG signal, without using any EOG channel. In: Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference, Aug 2016
16. González, M.B., García, T.R.: Imagen de marca y product placement, pp. 120–121. ESIC Editorial (2012)
17. Law, S., Braun, K.A.: I'll have what she's having: gauging the impact of product placements on viewers. Psychol. Mark. **17**(12), 1059–1075 (2000)
18. Lehmann, D., Michel, C.M.: Intracerebral dipole source localization for FFT power maps. Clin. Neurophysiol. **76**(3), 271–276 (1990)
19. Martinez-Levy, A., Cherubino, P., Cartocci, G.: Gender differences evaluation in charity campaigns perception by measuring neurophysiological signals and behavioural data. Int. J. Bioelectromagn. **19**(1), 25–36 (2017)
20. Mauss, I.B., Robinson, M.D.: Measures of emotion: a review. Cogn. Emot. **23**, 209–237 (2009)
21. Modica, E., Cartocci, G., Rossi, D., et al.: Neurophysiological responses to different product experiences. Comput. Intell. Neurosci. **2018** (2018)
22. Nebenzahl, I.D., Secunda, E.: Consumers' attitudes toward product placement in movies. Int. J. Advert. **12**(1), 1–11 (2015)
23. Nelli, R.P., Bensi, P.: Il product placement nelle strategie di convergenza della marca nel settore dell'intrattenimento, vol. 23. Vita e Pensiero (2007)
24. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. **32**(3), 230–236 (1985)
25. Plambeck, J.: Product placement grows in music videos. N. Y. Times **5**, B8 (2010)
26. Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev. Psychopathol. **17**, 715–734 (2005)
27. PQ Media.: Global Branded Entertainment Marketing Forecast 2018. (2018)
28. Russell, J.A., Barrett, L.F.: Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. J. Pers. Soc. Psychol. **76**, 805–819 (1999)
29. Russell, C.A., Norman, A.T., Heckler, S.E.: People and "their" television shows: an overview of television connectedness. In: The Psychology of Entertainment Media: Blurring the Lines Between Entertainment and Persuasion, pp. 275–290 (2004)
30. Scott, J., Craig-Lees, M.: Audience engagement and its effects on product placement recognition. J. Promot. Manag. **16**(1–2), 39–58 (2010)
31. Unsworth, N., Spillers, G.: Working memory capacity: attention control, secondary memory, or both? A direct test of the dual-component model. J. Mem. Lang. **62**(4), 392–406 (2010)
32. Van Reijmersdal, E.A., Neijens, P.C., Smit, E.G.: Effects of television brand placement on brand image. Psychol. Mark. **24**(5), 403–420 (2007)

33. Vecchiato, G., Maglione, A.G., Cherubino, P., et al.: Neurophysiological tools to investigate consumer's gender differences during the observation of TV commercials. Comput. Math. Methods Med. **2014** (2014)
34. Wisniewski, M.G., Thompson, E.R., Iyer, N., et al.: Frontal midline θ power as an index of listening effort. NeuroReport **26**(2), 94–99 (2015)

Chapter 19
Objective and Subjective Perception of Real Estate Features in the Light of an Experimental Study

Iwona Foryś and Radosław Gaca

Abstract Issues related to the objectification of the assessment of real estate as a subject of market turnover or valuation constitutes one of the basic problems of real estate management. The problems described are not limited only to the real estate market. It is universal in relation to various types of consumer and capital goods, especially in the case of the assessment of qualitative characteristics of these goods. The research problem concerns the issue of objectification of the assessment of the quality characteristics of real estate using standardized methods for measurement. For this purpose, an experimental study was carried out on the perception of the characteristics of real estate and the assessments made by real estate appraisers, using the modified method of semantic differential. The obtained results confirmed that the use of this standardized measurement tool leads to an increase in the objectivity of measurement. In the following stages, the results obtained should be generalized for the entire population of real estate appraisers, ensuring that the research is representative.

Keywords Intuitive methods · Real estate valuation · Semantic differential method · Market value

19.1 Introduction

The objectification of the assessment of the characteristics of marketable items, including in particular those of a qualitative nature, occupies a very important place in both research and practical decisions taken by manufacturers, consumers, investors, and appraisers. This assessment is carried out by marketing specialists as well as other management bodies, basing their decisions on models that take into account

I. Foryś (✉)
Faculty of Economics and Management, University of Szczecin, Szczecin, Poland
e-mail: iwona.forys@usz.edu.pl

R. Gaca
Kancelaria Rzeczoznawcy Majątkowego, Bydgoszcz, Poland
e-mail: radoslaw.gaca@gmail.com

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_19

quality assessments [17]. The problems described are not limited only to the real estate market. It is universal in relation to various types of consumer and capital goods. However, these issues are one of the basic problems of real estate management economics, as the assessment (measurement) of the characteristics of real estate along with the premises deciding about the purchase is the basis for estimating its value [2].

In classical terms, the analysis of real estate features, understood as their measurement, is nowadays most often carried out using various types of methods and rank strategies, operating on individually defined scales. Such measurement strategies lead to significant discrepancies not only in terms of semantic scope (level naming) but also in terms of the assessment criteria themselves [9].

Therefore, can the use of standardized methods to measure the quality characteristics of real estate lead to an increase in the objectivity of assessments? In order to answer this question, an experimental study was carried out on the perception of the characteristics of real estate and real estate appraisers' assessments using an adapted method of semantic differential.

According to research [9, 11], currently, most often, for the measurement of quality characteristics of real estate, appraisers use scales for which they introduce their own individual characteristics of facilities, corresponding to individual assessments of the adopted scale. Taking into account the above, assessments made by different appraisers for the same property differ significantly. This quite natural phenomenon is related to the above-described variety of scales and the lack of unambiguous levels, reference criteria.

In this context, a fundamental question arises as to the level of the subjectivity of assessment, i.e., whether the discrepancy in assessments is more related to the individual perception or to the distinctness of the criteria adopted for measurement scales. Providing answers to such questions will also allow us to answer the question whether the divergence of assessments results mainly from individual ways of processing information by individual experts or whether it is much more related to the diversity of assessments concerning the formulation of different criteria for different levels of measurement scales.

The research was preceded by a review of literature in the field of economics and psychology. As part of a critical analysis of the results obtained, a discussion was also held on the further development of methods for measuring quality characteristics.

19.2 Literature Review

Research on cognitive strategies has been carried out for many years in the field of psychology and economics [15]. In economics, for many years, research on the subjective assessment of characteristics and attributes of consumer goods has been conducted mainly in the area of marketing [4, 5, 16]. They referred to various issues, including subjective assessment or attitudes [10]. For example, from the perspective of consumption behavior analysis, Lai's [13] paper constructs a framework of product

valuation for consumers and its typology of product benefits and the implications of holistic consumption behavior analysis for marketing strategy are discussed.

This research has led to the possibility of formulating certain laws (Weber's law and Stevens' power law) or a theory which identifies five consumption values influencing consumer choice behavior [18]. These rights relate to the context of physical measurement and the ability to estimate the correct result on the basis of an individual assessment. However, they are related to the imaging process itself and not directly to the individual experience of the assessor.

The main problem in the real estate market is its low information efficiency, especially in relation to the description of the quality characteristics of real estate [24]. Hence, decisions of market participants result from intuitive actions and not from decisions taken on the basis of reliable and unambiguous information [21].

In the previous research, some authors provided the response on the appraisers' preference for property valuation approaches. The authors, on the basis of the collected results, presented recommendations regarding possible changes to legal regulations and professional norms, and a critical assessment of the causes influencing the range of applied analytical tools in valuation process was also carried out [8].

Observation and measurement are the basic tools of scientific cognition [12]. While in the case of exact sciences, measurement has never been questioned, in the case of the social sciences, the mere possibility of measuring was the subject of numerous doubts. These doubts have not been partially eliminated to this day. In particular, they have concerned the possibility of measuring various types of phenomena and information of a qualitative nature [6]. From the mathematical point of view, the theory of measurement considers the system of objects (features, events) which are subject to measurement and the system of numbers through which the mapping is performed. The essential element connecting the systems described above is the possibility of performing two empirical operations:

- a comparison operation according to this value, corresponding to mathematical relation ">" between numbers;
- an object merging operation that would correspond to an add operation performed on numbers.

In the context of measurement in social sciences, the following counterarguments can be cited (Stevens). The empirical operation corresponding to an addition does not have to be a variation of physical joining of objects, such as the addition of segments for length measurement; it can be a hypothetical operation, e.g., carried out in the mind of the person tested. Assigning numbers to measured objects does not require an operation corresponding to addition. In this sense, the measurement is carried out at weaker levels, for which the subsequent addition of assigned numbers will not be allowed.

In order to increase the objectivity of the measurement of the quality characteristics of real estate, the use of a semantic differential method adapted to this end [9] was proposed. The scale of the semantic differential was proposed for the first time

by Osgood [14]. The concept of scale focuses on the assessment of the level of intensity of a phenomenon or the state of an object which they cause in the respondent, by determining its connotative meaning. The semantic differential is a multidimensional scaling method used to study the meaning of specific concepts or perceptions of objects (people, events, items, terms, and phenomena) in semantic space [19, 20, 22].

As part of the adaptation of the method for the purpose of measuring the quality characteristics of real estate, the appropriate transformation, and determination of individual scales was made with the use of pairs of adjectives corresponding to the characteristics of a phenomenon under examination [7, 9]. According to the assumption, the results of a measurement carried out with the use of the method of semantic differential can be treated as a measurement carried out at least at the interval level [23].

19.3 Scope of Data and Organization of the Study

The study was conducted in the form of a survey in which respondents assessed the technical and usable condition of real estate on the basis of photographs presented to them for sets appropriately similar to each other, in the generic sense of real estate. The assessment of condition on individual scales was carried out using the method of a semantic differential in a version adapted to the analyzed characteristic of real estate.

The study used two sets of real estate related to real estate developed with single-family residential buildings. In the case of the first set (Z_1), the subject of analysis was photographic documentation including photographs of properties showing their depicted exterior (a single photograph of a facility). In the case of the second set (Z_2), the documentation presented for analysis included both the external and internal picture (two exterior photographs and three interior photographs). For set (Z_1) respondents assessed seven properties (N_i, $i = 1, \ldots, 7$), and for the set (Z_2) six properties ($N_j, j = 1, \ldots, 6$). Before the study itself, the participants were informed about the formula according to which the measuring scales provided for in the study should be used. Examples of photographic documentation used as a basis for the assessment of the study are shown in Figs. 19.1 and 19.2.

The study was conducted in the period from September 2016 to March 2017 on various groups of real estate appraisers, as part of training on the use of econometric methods in the valuation of real estate (group A–E). As a result, 55 responses were obtained for the set (Z_1) and 35 for the set (Z_2). Detailed dates of the study and the number of study groups are listed (Table 19.1).

In the study for both sets, scales were used, covering five levels and the following pairs of adjectives: neglected–well maintained, dry–damp, undamaged–damaged, without cracks–cracked, damaged–renovated, old–new, mold-infested–non-mold-infested, worn–unworn, modern–obsolete, esthetic–unesthetic, defective–no defects,

Fig. 19.1 Examples of photographic documentation for set Z_1

Fig. 19.2 Examples of photographic documentation for set Z_2

Table 19.1 Number of groups of respondents and dates of the experiment

Indication of the group of respondents	Size of set Z_1	Size of set Z_2	Study period
A	5	5	From 09/26/2016 to 09/27/2016
B	12	11	From 10/1/2016 to 10/10/2016
C	6	8	From 10/14/2016 to 11/17/2016
D	20	5	From 11/25/2016 to 01/17/2017
E	12	6	From 3/11/2017 to 03/15/2017

Table 19.2 Statistics of assessments for sets (Z_1) and (Z_2)

Property	Min	Max	Average	Number of all responses	Number of the medium range
Set (Z_1)					
N_1	1	5	3.81	55	38
N_2	1	5	2.11	55	39
N_3	1	5	2.16	55	35
N_4	1	5	2.86	55	41
N_5	1	5	2.58	55	31
N_6	1	5	1.98	55	27
N_7	1	5	4.33	55	28
Set (Z_2)					
N_1	1	5	3.68	35	22
N_2	2	3	4.11	35	16
N_3	1	4	2.67	35	18
N_4	2	5	3.41	35	22
N_5	1	5	3.04	35	22
N_6	1	5	3.79	35	18

modernized–no modernization, energy-intensive–energy-saving, whole–leaky, and dark–light.

During the survey, there was generally no time limit for respondents to respond, but in no case did the response per respondent exceed 10 min for each of the surveyed groups. The obtained results of average assessments for particular properties (N_i) and (N_j) and for studied sets (Z_1) and (Z_2) are presented in Table 19.2.

In addition, distributions of the obtained responses for individual scales for both sets of properties (Figs. 19.3 and 19.4) are presented.

19.4 Test Results

Validity and reliability are the key measures for assessing the objectivity of a measurement tool in social sciences [1]. Validity refers in this case to the level at which the measuring tool adequately reflects the measured parameter (condition, feature, phenomenon, and significance). Reliability refers to the level of repeatability of the result.

Within the framework of validity, the following are distinguished: face validity, criterion validity, construct validity, and content validity. Analysis of the validity and reliability of a research tool is essential for its further use. It allows determining whether the tool used actually allows for the measurement of a specific factor and whether such results can be applied in practice. To sum up, the assessment of the

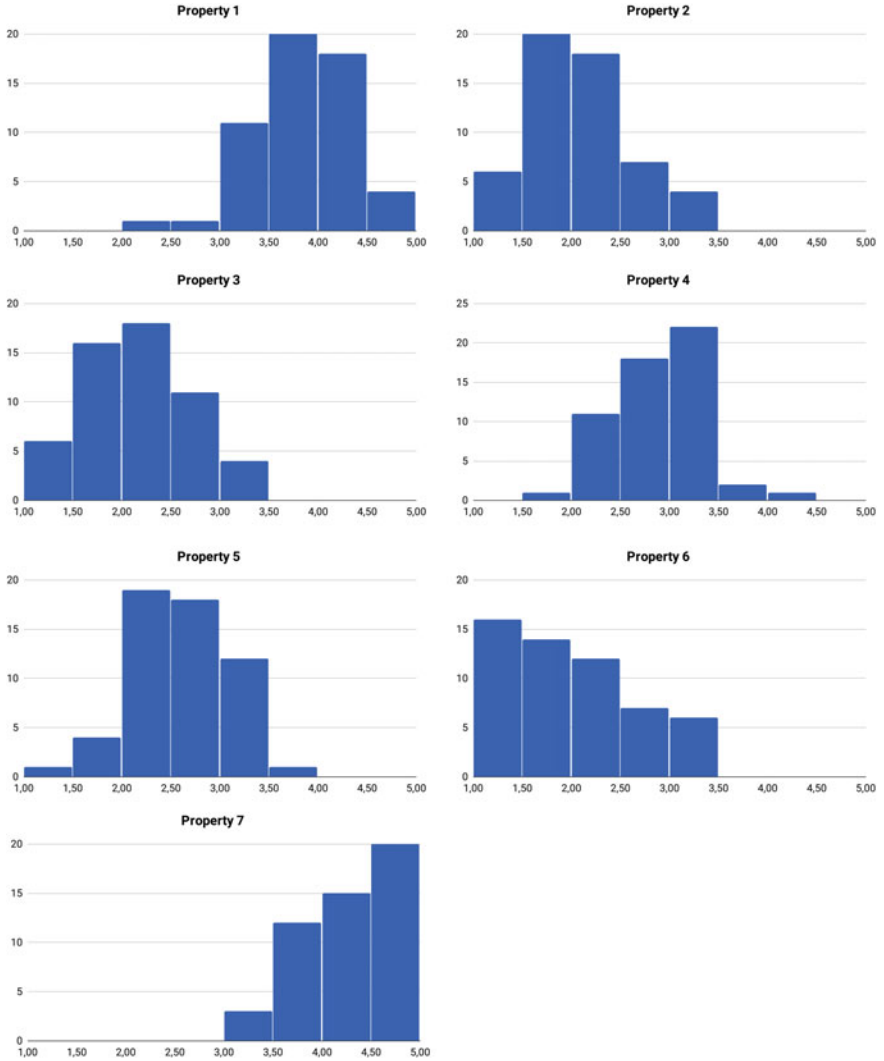


Fig. 19.3 Distribution of respondents' assessments by scales (set Z_1)

above-mentioned characteristics allows us to conclude on the level of objectivity of a measurement.

In the study, the method of testing the internal structure of the test was used to assess its validity. The reliability was assessed using a retest [3]. As part of studying the internal structure of the test, Pearson's correlation coefficients were calculated for individual scales and mean values. In the case of reliability assessment, the same correlation coefficient was calculated between the results of the averages for different

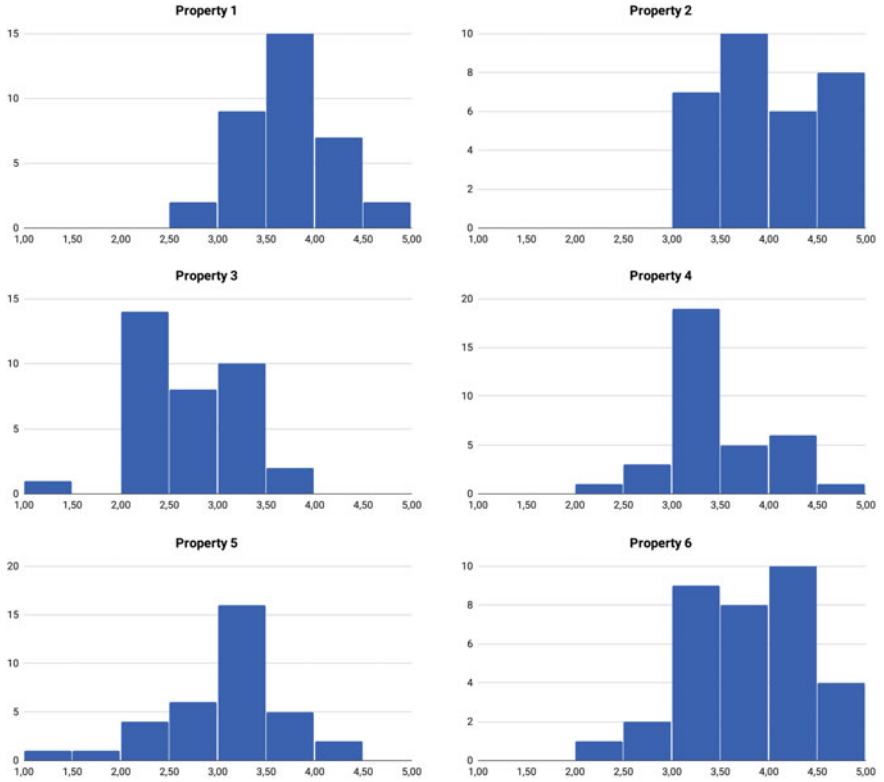


Fig. 19.4 Distribution of respondents' assessments by scales (set Z_2)

scales for different groups. The results for each set of responses are presented in Table 19.3.

In the next step, the Pearson's correlation of assessments for the retest (Table 19.4) was determined.

The obtained results indicate a high level of correlation both in the case of validity testing (above 0.5 except for one case) and reliability of the tested measuring tool (above 0.7). In the tool reliability study, the majority of correlation dependencies were even close to 0.9. Weaker but satisfactory results were obtained in the case of testing the validity of the tested tool.

Table 19.3 Results of Pearson's correlation of the study on the internal structure of the test for sets (Z_1) and (Z_2)

Scale number	*r*	Scale number	*r*	Scale number	*r*
Set (Z_1)					
1	0.6178	6	0.5452	11	0.7197
2	0.6449	7	0.4927	12	0.5691
3	0.6813	8	0.7127	13	0.5942
4	0.7040	9	0.6310	14	0.7461
5	0.7636	10	0.6108	15	0.5295
Set (Z_2)					
1	0.7522	6	0.6129	11	0.7061
2	0.6066	7	0.6450	12	0.7864
3	0.8291	8	0.7538	13	0.7184
4	0.6297	9	0.6165	14	0.6957
5	0.7292	10	0.7371	15	0.5171

Table 19.4 Results of Pearson's correlation of assessments for the retest for sets (Z_1) and (Z_2)

Group	A	B	C	D	E
Set (Z_1)					
A	1				
B	0.8397	1			
C	0.8729	0.8328	1		
D	0.9202	0.8493	0.8630	1	
E	0.9168	0.8132	0.8720	0.9156	1
Set (Z_2)					
A	1				
B	0.9111	1			
C	0.7156	0.7657	1		
D	0.7143	0.7415	0.73468	1	
E	0.8088	0.9088	0.7581	0.7054	1

19.5 Conclusions

The use of a standardized measurement tool in the form of a modified method of semantic differential to measure the quality characteristics of real estate leads to an increase in the objectivity of measurement. The conducted study indicates a high level of both validity and reliability of the tool used. It should be remembered, however, that the set of qualitative variables completed took place in relation to sets of similar properties, characterized by a very wide range of similarities. Thus, it

can be assumed that for real estate sets differing in quality features, it is possible to significantly increase the objectivity of assessments if a modified method of the semantic differential is used to measure them.

This objective was achieved by measuring with the use of fifteen scales allowing for the measurement (mapping) of the impact of the technical and usable condition of real estate on the assessment of its usability. The proposed method of measurement in the described approach significantly improves the objectivity of the assessment of the quality characteristics of real estate. It should be emphasized that the survey is still of a pilot nature, therefore in subsequent stages both the substantive scope of the survey and the respondent base will be extended, in the direction of representative surveys, allowing for far-reaching generalizations.

References

1. Babbie, E.: Podstawy badań społecznych. Wydawnictwo Naukowe PWN, Warszawa (2008)
2. Betts, R.M., Ely, S.J.: Basic Real Estate Appraisal: Principles and Procedures, 7th edn. Thomson South-Western, USA (2008)
3. Cronbach, L.J., Meehl, P.E.: Construct validity in psychological tests. Psychol. Bull. **52**(4), 281–302 (1955). https://doi.org/10.1037/h0040957
4. Day, G.S.: Market Driven Strategy. Free Press, New York (1990)
5. Engel, J.F., Blackwell, R.D., Miniard, P.W.: Consumer Behavior and Marketing Strategy, 6th edn. Dryden Press, Chicago IL (1990)
6. Foryś, I., Gaca, R.: Theoretical and practical aspects of qualitative variable descriptions of residential property valuation multiple regression models. In: Proceedings Foundation of the Cracow University of Economics, 978-83-65173-48-5 (HTML), pp. 36–44 (2016)
7. Foryś, I., Gaca, R.: Application of the Likert and Osgood scales to quantify the qualitative features of real estate properties. Folia Oeconomica Stetinensia **16**(2), 7–16 (2017). https://doi.org/10.1515/foli-2016-0021
8. Foryś, I., Gaca, R.: Intuitive methods versus analytical methods in real estate valuation: preferences of Polish real estate appraisers. In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics, pp. 79–87. Springer International Publishing (2018). https://doi.org/10.1007/978-3-319-99187-0_6
9. Gaca, R.: Zmienne jakościowe w procesie szacowania wartości rynkowej nieruchomości. Rozprawa doktorska. Uniwersytet Szczeciński - maszynopis (2017)
10. Himmelfarb, S.: The measurement of attitudes. In: Eagly, A.H., Chaiken, S. (eds.) Psychology of Attitudes, pp. 23–88. Thomson/Wadsworth (1993)
11. Kane, M.S., Linne, M.R., Johnson, J.A.: Practical Applications in Appraisal Valuation Modeling. Statistical Methods for Real Estate Practitioners, Chicago, IL (2004)
12. Kuhn, T.S.: Dwa bieguny: tradycja i nowatorstwo w badaniach naukowych. Państwowy Instytut Wydawniczy, Warszawa (1985)
13. Lai, A.W.: Consumer values, product benefits and customer value: a consumption behavior approach. Adv. Consum. Res. **22**, 381–388 (1995)
14. Osgood, C.E.: The nature and measurement of meaning. Psychol. Bull. **49**(3), 197 (1952)
15. Osgood, C.E., Suci, G., Tannenbaum, P.: The Measurement of Meaning. University of Illinois Press, Urbana IL (1957)
16. Paul, P.J., Olson, J.C.: Consumer Behavior and Marketing Strategy, 2nd edn. Irwin Homewood, IL (1990)
17. Schinzler, B.: A multiattribute combinatorial exchange. In: Negotiation, Auctions, and Market Engineering: International Seminar, Dagstuhl Castle, Germany, 12–17 Nov 2006, pp. 84–100. Springer Science & Business Media (2008)

18. Sheth, J.N., Newman, B.I., Gross, B.L.: Why we buy what we buy: a theory of consumption values. J. Bus. Res. **22**, 159–170 (1991). https://doi.org/10.1016/0148-2963(91)90050-8
19. Snider, J.G., Osgood, C.E.: Semantic Differential Technique: A Sourcebook. Aldine, Chicago (1969)
20. Smith, G.T.: On construct validity: issues of method and measurement. Psychol. Assess. **17**(4), 396–408 (2005). https://doi.org/10.1037/1040-3590.17.4.396
21. Thaler, R.H.: Misbehaving: The Making of Behavioral Economics. W.W. Norton, New York, London (2016)
22. Verhagen, T., Hooff, B., Meents, S.: Toward a better use of the semantic differential in IS research: an integrative framework of suggested action. J. Assoc. Inf. Syst. **16**(2), 1 (2015)
23. Walesiak, M., Dudek, A.: Symulacyjna optymalizacja wyboru procedury klasyfikacyjnej dla danego typu danych–charakterystyka problemu. Zeszyty Naukowe Uniwersytetu Szczecińskiego **450**, 635–646 (2006)
24. Wolverton, M.L.: Comments on "Common statistical errors and mistakes: valuation and reliability". Letters to the Editor. Apprais. J. **176** (2014)

Chapter 20
Optimization of Location Attractiveness Zones for the Purpose of Property Mass Appraisal

Sebastian Kokot and Sebastian Gnat

Abstract Mass land appraisal is a specific type of appraisal. A suitable algorithm used for mass appraisal constitutes one of the main determinants of the quality of achieved results. Many of the algorithms cluster real estate in accordance with specific criteria, such as a real estate type, its characteristics or location. Assuming a "real estate location" as a clustering basis, we use terms such as a taxation zone (which originates from a basic application of mass appraisal, which involves ad valorem real estate taxation), along with an elementary terrain (which originates from the nomenclature used in spatial planning). The main goal of this chapter is to find the solution to the designation of such zones, or real estate clusters for the purpose of mass land appraisal conducted with the use of Szczecin Algorithm of Real estate Mass Appraisal (*SAREMA*). Each mass appraisal algorithm has its own specificity, and consequently, it may pose various requirements to designated zones or real estate clusters (here referred to as location attractiveness zones—*LAZ*'s). The chapter presents the methods of *LAZ*'s designation formed on the basis of the existing cadastral districts, real estate market analysis, and hierarchical clustering. The chapter is additionally focused on the manner of determining *LAZ* optimal number and consequently the number of properties that will be included in an automated mass appraisal process in a given zone. The uniformity of isolated *LAZ*'s was subject to verification with the use of an entropy ratio, especially constructed for that purpose. The study was conducted on the land plots located in the city of Szczecin.

Keywords Mass appraisal · Real estate appraisal algorithms · Hierarchical clustering

S. Kokot · S. Gnat (✉)
Faculty of Economics and Management, University of Szczecin, ul. Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: sebastian.gnat@usz.edu.pl

S. Kokot
e-mail: sebastian.kokot@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_20

20.1 Introduction

Contemporary methods of mass appraisal use various algorithms of machine learning, including artificial neural networks, and many more (e.g., [12, 20, 30]). Real estate mass appraisal is a completely different procedure from individual appraisal in terms of the organization of the entire process of valuation. The subject matter of individual appraisal involves one property or a relatively small group of properties. Valued properties typically differ in terms of their location, type of market (segment or sub-segment), purpose, and the scope of appraisal, dates for which the real estate conditions and price levels are taken into consideration, etc. Mass appraisal is carried out in an entirely different fashion and it takes place, when:

1. the subject of an appraisal involves a large number of properties of one type,
2. an appraisal is to be conducted with a uniform approach and consequently yield cohesive results,
3. all the properties subject to appraisal are evaluated "at the same time" with the same dates for which the property condition and price level is taken into consideration (e.g., [11]).

As a result, the organization of valuation process relative to mass appraisal is also different in comparison to individual appraisal. Two main stages of this process can be distinguished. During the first stage, all the necessary information and data concerning the valuated properties and a relevant market are gathered, whereas during the second stage the values of all the real estate subject to appraisal are calculated through the application of the proper algorithm.

The use of mass appraisal is justified, since in certain areas it occurs to be quicker and less costly than individual valuation. Typical areas in which mass appraisal appears to be useful include:

- updating the charges calculated on a property value (e.g., in Poland, annual perpetual usufruct charges), including universal real estate taxation [3],
- estimating the economic results of adoption or changes of local zoning plans,
- monitoring the values of property sets, constituting a security of credit exposures held by a bank in order to calculate a LtV ratio for a bank's credit portfolio [16, 26].

The purpose and the conditions of a mass appraisal indicate whether the application of logarithm-based procedures is justified. When attempting to classify the methods used in practice or subjected to scientific study as a possible way of conducting the process, one can distinguish groups of methods which use econometric models of multiple regression, methods based on neural networks or machine learning, and the so-called introspective methods (e.g., [6, 9, 14]).

The methods based on econometric models of regression are the most frequently applied as a method of a statistical analysis of the market, particularly in linear approach. However, they usually produce unsatisfactory results [21]. The reason for this is the difficulties in fulfilling a number of formal requirements that are posed already at the stage of constructing a model [9].

Artificial neural networks operate on the principle based on the functioning of the human brain. They are trained and taught on example data sets. Literature reports indicate that despite certain problems regarding the optimization of networks construction, the results obtained through their application in real estate appraisal are satisfactory [18, 28]. The main barrier for their application in practice seems to be the issue of justifying valuation results, since a part of the real estate appraisal process occurs beyond the control of a property appraiser.

Introspective methods should not feature the above-described drawbacks. Those methods are based on models devised specifically for real estate valuation, which on the one hand enable the specificity a local market operation to be taken into account (spatial layout, local fashions, etc.), and on the other hand they manage (within a specified scope) to cope with the imperfections of the market in the form of a relatively low number of transactions far better than the methods listed above.

As was previously mentioned, some real estate appraisal methods require the appraised area to be divided into smaller zones. The objective of the study is an analysis of the effects of applying various methods for determining *LAZ*'s. It will be examined how individual methods affect the degree of uniformity of the properties qualified into particular zones as well as how such a number of zones influences the number of representative properties designated for individual valuation. On the basis of these two criteria, conclusions will be drawn as to which of the methods in the presented problem is the best. A modification of a classic entropy method will be proposed, which will enable a better presentation of the specificity of a real estate market. The geographical area that the study encompasses includes the northern part of Szczecin (the largest city in the northwest of Poland). The subject of the appraisal, within the scope of which *LAZ*'s are created, entails 1500 land plots.

20.2 Location Attractiveness Zones

The issue of a proper designation of location attractiveness zones ought to be considered as one of the key problems, from the standpoint of the accuracy of appraisal results obtained with the use of many algorithms. An area in which a certain number of appraised properties, which are most importantly characterized by the same impact of their location on their value and, as far as possible, characterized by a similar impact of other market properties, constitutes such a zone. It means that all properties located in a given zone formally should not differ in terms of their location. The assumption takes on a particular significance in the case of certain algorithms used for mass appraisal, including the Szczecin Algorithm of Real Estate Mass Appraisal. *LAZ*'s ought to be created on the grounds of the so-called pricing factors of a property, the basic one of which includes a property function in a zoning plan [23]. The existing approaches to real estate zone designation use also, apart from the provisions of local zoning plans, aerial photographs [25], or additionally district and housing estates boundaries, land plot lines, streets, roads, rivers, railroads, as well as other artificial and natural features as auxiliary information [5].

Expert intervention in zone boundaries is not precluded, on the basis of those criteria or other methods. Furthermore, there are postulates put forward to take into account the so-called physiognomic features of terrains, or even legal properties (real estate ownership structure) apart from terrain functions [17]. A separate group of methods that can be employed for the designation of *LAZ*'s involves various methods of grouping. One of the available approaches is hierarchical clustering (e.g., [13]). This method enables clustering of objects (in this case properties) that are similar to one another on the basis of many characteristics that describe them. Clustering is an issue frequently found in scientific research. The question of determining an optimal number of clusters is one of particular interest to researchers [7, 15]. What is more, researcher's present possible ways of improving the existing methods (e.g., [1]) or they propose to use ensemble of clustering methods [2, 29]. From the perspective of a real estate market, the possibility of introducing spatial constraints to the algorithm is a particularly significant element of this clustering method [4, 8]. Such constraints will enable taking into consideration the contiguity of certain objects, not only on account of their descriptive characteristics, but also their neighboring location in space. Spatial weight matrices constitute a dominant form of accounting for spatial relations between objects (see Suchecka (ed.) [24]).

From the point of view of mass appraisal, it is important that location attractiveness zones should contain properties similar to one another. A modified entropy measure was employed in order to evaluate the uniformity of location attractiveness zones. The distribution entropy of the analyzed variable enables calculating the degree of determinateness (definitiveness) of that distribution on account of the analyzed variable (e.g., [22, p. 90]). Shannon entropy ratio is normalized and it reaches the values from the range of $\langle 0, 1\rangle$. A high value of the ratio indicates a significant degree of definiteness of the analyzed system (the system does not demonstrate an inclination). In order to assess the uniformity of achieved *LAZ*'s a classic entropy method was employed, which was subject to a modification, owing to a different number of possible land plot types (plot classes) that may occur. The entropy of *LAZ*'s was calculated with the following measure:

$$H_z = \frac{\left(-\sum_{i=1}^{k} p_i \cdot \log_k p_i\right) \cdot k}{L}, \tag{1}$$

where

p_i—percentage share of the properties belonging to ith class,
k—number of classes (combination of variants of the market properties found in a given *LAZ*),
L—number of land plot classes found in the analyzed set.

The modification of a classic entropy measure involves changing the assessment depending on the number of classes found in particular zones. For instance, in a classic approach, *LAZ*'s with two and ten classes and uniform percentage share of the classes will feature complete entropy. In turn, in the case of a real estate market analysis,

these two situations need to be assessed differently. Two classes of combinations of market features, in a situation, when a designated terrain contains several dozen or more properties it translates into high uniformity, even if the percentage shares of both classes are equal to 50%. H_z measure will differentiate between entropy levels depending on the percentage shares of classes as well as their number.

Incorrect designation of taxation zones, i.e., their excessive or insufficient number or incorrectly devised borders, may have far-reaching consequences. On the one hand, in order to ensure a lower degree of labor-intensiveness, time demands, and costs we try to create the lowest possible number of taxation zones. In the case of *SAREMA*, it means that fewer properties will need to be appraised individually as representative properties. The negative effect of an insufficient number of zones and consequently too expansive zones will result in qualifying distinctly different properties in terms of market location attractiveness. In other words, properties in different locations will be treated as properties having identical location attractiveness, which will naturally affect appraisal results with excessive errors. In turn, if too many zones are formed, thereby zones of relatively smaller surface, we might not be undermining the results accuracy, but we unnecessarily lengthen the process and raise the cost of such an appraisal, because in that situation more representative properties need to be valued individually. Moreover, a greater number of zones may be created, in which few or only several or even one property will be subject to evaluation. In such cases, the number of representative properties will constitute a very high percentage of all appraised properties, at times rendering the point of mass appraisal questionable.

The problem with designating *LAZ*'s in this chapter was presented with an example of one of algorithm-based methods of appraisals classified into a group of introspective methods, namely Szczecin Algorithm of Real Estate Mass Appraisal. The procedure of that algorithm entails the designation of *LAZ*'s (in other publications concerning the algorithm also termed as elementary terrains). Szczecin Algorithm of Real Estate Mass Appraisal was described in the paper by Hozer et al. [10] "Econometric algorithm of land real estate mass appraisal."

20.3 Source Data and Research Assumptions

In the presented example, the appraisal includes 1512 land plots located in region No. 3—Nad Odrą in the city of Szczecin. In order to designate taxation zones in an optimum manner, it is necessary to adopt research assumptions regarding:

1. the number of representatives, i.e., properties intended for individual appraisal out of each taxation zone that is to be designated,
2. the objective scope of *LAZ*'s.

The number of properties intended for individual appraisal in each taxation zone ought to ensure representativeness of such properties for all the remaining ones located in a given zone of appraised properties. The greater the number of representatives will be selected, the higher the representativeness will be ensured. On the

other hand, too high a number of representatives will naturally result in an increased workload required for individual appraisals, thereby—an increase in costs. Practical experience of the algorithm creators indicates that the following sampling selection ensures a suitable representativeness of a location:

- from zones where up to 10 properties are appraised—one representative property,
- from zones where from 10 to 50 properties are appraised—two properties,
- from zones where from 51 to 100 properties are appraised—three properties,
- from zones where from 101 to 500 properties are appraised—four properties,
- from zones where from 501 to 1000 properties are appraised—five properties,
- from zones where over 1000 properties are appraised—six properties.

On the other hand, when discussing the objective scope, we take into account whether individual elementary terrains should include only properties of similar, comparable purpose, or whether we should allow the possibility of properties of different purposes located in one zone. If the first option is assumed, numerous small elementary terrains will need to be designated, comprising a small number, or even individual properties. It will be the case when, e.g., a property with a commercial facility erected on it is located in a housing real estate, which is in fact a commonplace occurrence. In such circumstances, that property will constitute a separate zone. The consequence of adopting the second solution may also be varied. There are, however, ways in which the situation may be resolved, which has been discussed in a separate paper. In the presented study, in order to avoid complicating the nature of the analyzed problem, the analysis was limited to residential real estate.

20.4 Research and Results

When attempting to designate *LAZ*'s it is helpful if one has got a certain preliminary division of the area in which appraised properties are located. The designation of *LAZ*'s boundaries can then be conducted through an analysis of the accuracy of such a division in relation to the desired *LAZ*'s boundaries. The division into cadastral districts or boundaries of local zoning plans may serve as such a preliminary division. *LAZ*'s may also be designated without any preliminary division. Irrespectively of the fact whether and what preliminary division we have at our disposal, we need to know the specificity of the appraised area. Mass appraisal, no matter what method is used, should be conducted with the participation of individuals who have such knowledge of the terrain.

In this study, an attempt was made at comparing several approaches to the creation of location attractiveness zones. The first approach assumes that boundaries of cadastral districts constitute *LAZ* borders; therefore, the process of their designation does not need to be carried out. Existing cadastral districts will be evaluated as zones. A fragment of the area under study, taking into account the boundaries of a geodetic area, was presented in Fig. 20.1.

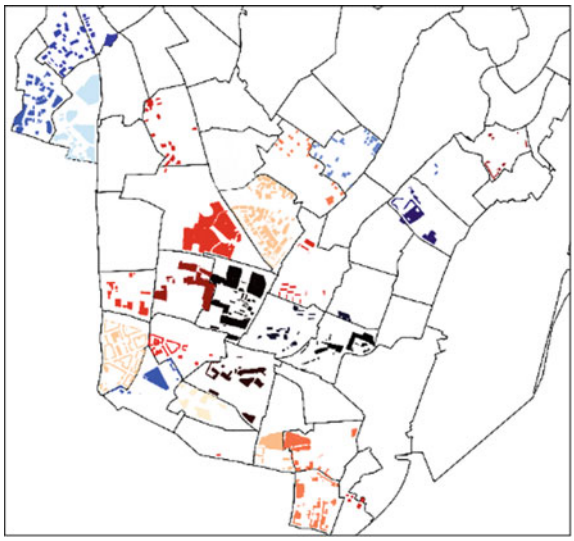


Fig. 20.1 Location of appraised properties with marked geodetic area borders (own elaboration)

The second approach is based on experts' involvement. A team of property appraisers conducted an analysis of the local market of undeveloped land real estate and using the preliminary division of cadastral districts the team designated the boundaries of location attractiveness zones. As a result, the appraised area was divided into 17 location attractiveness zones, which according to the experts' intention, are meant to cluster properties of the same location impact on value. The selected *LAZ*'s created with this expert approach are presented in Fig. 20.2.

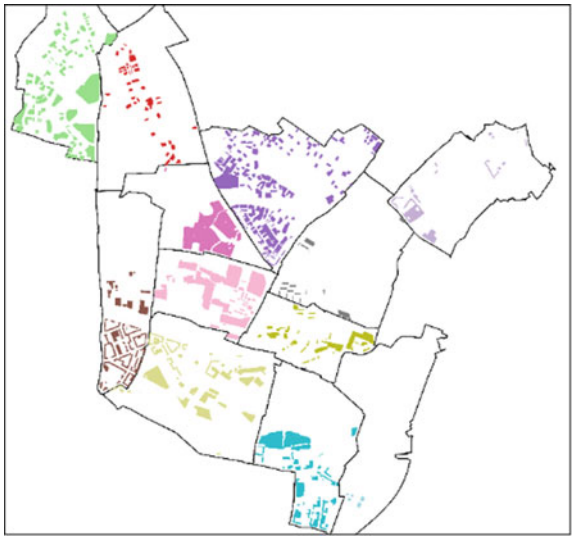
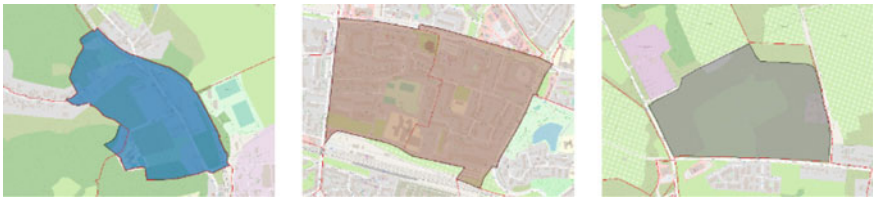


Fig. 20.2 Location of selected *LAZ* created with the expert approach (own elaboration)



Cadastral district with a fragment of another district

Two cadastral districts

Fragment of a cadastral district

Fig. 20.3 Examples of various types of location attractiveness zone borders designated by experts (own elaboration)

During the conducted works it sometimes happened that one geodetic area was qualified as one *LAZ*, different *LAZ* encompassed one geodetic area and a fragment of another one, another *LAZ* was composed of several geodetic areas and in some cases *LAZ*'s encompassed only a part of one geodetic area. The examples of those situations are presented in Fig. 20.3.

The third method of *LAZ*'s designation was based on a statistical approach. With the use of hierarchical clustering with spatial constraints, land plots constituting the subject of the study were qualified to individual sub-areas. A spatial weight matrix using a distance band was used as a spatial constraint. As the band length constituted distance ensuring that each analyzed land plot will have at least one neighbor. In order to evaluate the uniformity of obtained *LAZ*'s, a classic entropy method was used, having been modified on account of a various number of possible land plots (plot classes) that may occur. In order to evaluate entropy, converted information was used concerning the variants of four market characteristics: utilities, surroundings, transport availability, and location.

The variants of those properties were coded into the form of natural numbers (the worst condition 0, medium 1, the best 2 and in the case of transport unavailability, which was a property of four states—3) and they were combined into a four-digit code. Each code value constitutes a combination of the value of market properties, which was understood as a class. Calculations were performed with Python computer language statistics packages [19, 27].

Hierarchical clustering was conducted assuming a different number of clusters as an initial algorithm parameter. There were between 2 and 40 clusters. Two clusters constituted the smallest possible number. While the upper limit resulted from the fact that an original division comprised 39 cadastral districts. An example of hierarchical clustering with two *LAZ*'s is presented in Fig. 20.4. The fact that the method does not permit designating zone boundaries, but it only groups the objects into a predetermined number of clusters, needs to be considered a disadvantage of this method.

For each of instances of hierarchical clustering, on the basis of the size of individual *LAZ*'s, a required number of representative properties and an average entropy ($\bar{H}_z$) were determined. Thus, obtained parameters of each clustering will be compared between one another and with both previously employed methods of *LAZ*'s

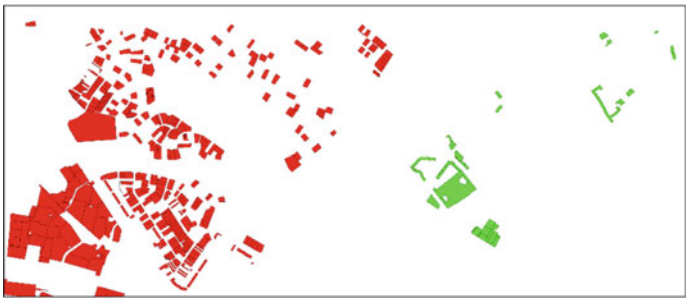


Fig. 20.4 Fragment of assigning land plots into two clusters through hierarchical clustering (own elaboration)

designation. On the grounds of a required number of representative properties and an average entropy, it will be determined, which method or methods of land plot clustering ought to be considered as being the most favorable. Average *LAZ*'s entropies created with all the specified methods are compiled in Fig. 20.5.

As can be observed, the average entropy of the zones formed through hierarchical clustering initially, with a small number of clusters, is significantly higher than the average entropy of cadastral districts and the expert approach. Hierarchical clustering reaches the level of entropy of the latter with 7 *LAZ*'s, while the level of cadastral districts entropy with 10 *LAZ*'s. With a cluster number higher than 10, hierarchical clustering enables achieving zones of lower average entropy than the two previous approaches. However, an inevitable consequence of the increase in the number of zones entails an increase in the number of representative properties. Each representative property needs to be appraised individually, which involves the use of time and monetary resources. Therefore, it becomes necessary to arrive at a compromise between entropy and the number of representative properties. In the case of an approach based on cadastral districts, the number of representative properties was, in line with previously set guidelines, defined at the level of 83. In turn, in the case

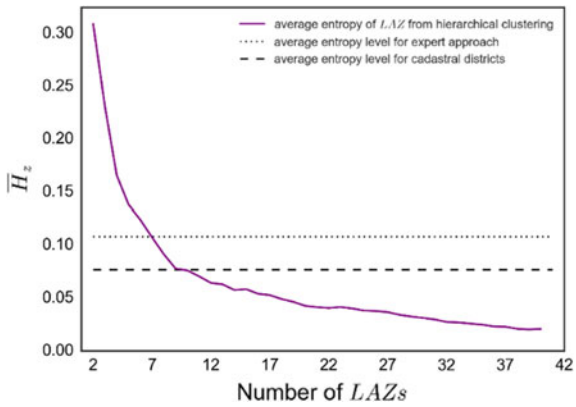


Fig. 20.5 Average entropies of *LAZ*'s designation (own elaboration)

of the expert approach—it was 47. Figures 20.6 and 20.7 present average entropies of individual hierarchical clusterings along with a number of representatives in a cluster resulting from the number of land plots in it. These figures were confronted with average entropies and the number of representative plots for the two remaining approaches. On the basis of the data presented in Fig. 20.6, it was demonstrated that for 10 clusters average *LAZ* entropy is lower that for a division based on geodetic boundaries. What is more, such a number of zones require a far lower number of representative plots than is the case for cadastral districts. When analyzing the results presented in Fig. 20.6, it becomes evident that all the specified numbers of clusters exceeding 10 require an individual appraisal of a lower number of plots. It means that a division based on cadastral districts is not the best approach, because those districts were not created on the grounds of a location or other pricing factors impact on real estate value.

The situation is slightly different when we compare hierarchical clustering with a division conducted through expert approach (see Fig. 20.7). Only the number of

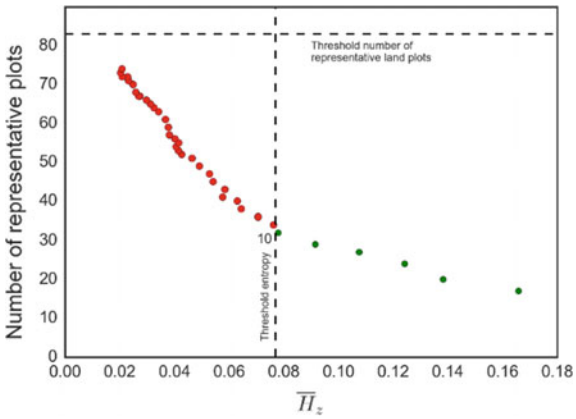


Fig. 20.6 Average entropies of hierarchical clustering and the numbers of representative land plots required for them in comparison to cadastral districts (own elaboration)

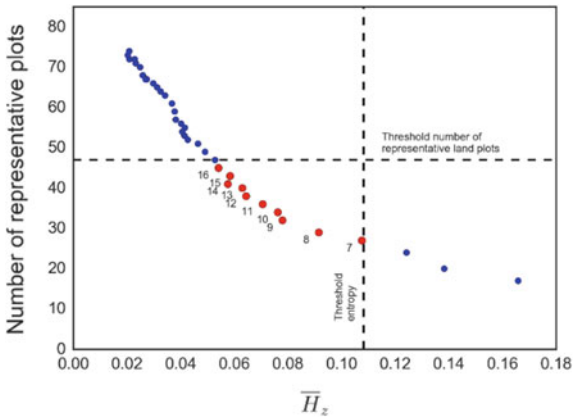


Fig. 20.7 Average entropies of hierarchical clustering and the numbers of representative plots required for them in comparison to the expert method (own elaboration)

clusters from 7 to 16 yields lower average entropy and requires a smaller number of representative plots than the *LAZ*'s created by experts on the basis of a real estate market analysis. Starting from 17 clusters, average entropy is still lower than for the zones designated by experts, but it requires a greater number of representative plots.

20.5 Conclusions

The chapter analyses the issue of the division of areas subject to mass appraisal into smaller and uniform subzones. It was demonstrated that there are many possibilities of devising such zones. Each of the approaches has its advantages and disadvantages. Out of the existing possibilities, three were presented in the study. The first of them is based on using an existing cadastral districts. The benefit of this method is that the division has already been performed. However, if such a division is not confronted with the conditions of a real estate market it results in low uniformity of zones and the number of representative properties required (for individual appraisal) is high. In turn, an approach based on hierarchical clustering is relatively easy to implement, but it creates problems related to the specification of spatial weight matrices. It is worth emphasizing that as a result of the clustering process no boundaries of location attractiveness zones can be determined, but only the appraised properties are classified into the zones. If a property has not been assigned into a set of properties to be valued, there is no information as to which of the *LAZ*'s it ought to be assigned to. As far as the expert approach is concerned, extensive knowledge regarding the specificity of an appraised area is required in order to successfully designate zone boundaries.

The question of choosing an optimal number of *LAZ*'s is linked to both their uniformity and the costs of appraising representative properties. Therefore, a decision needs to be taken, which is problematic, since a greater number of zones translates into their lower average entropy, but a greater number of representative properties. The conducted comparisons demonstrated that a division based on cadastral districts yields poorer results in relation to hierarchical clustering, owing to the cost of individual appraisals. In turn, when the expert method was compared with hierarchical clustering method it was observed that the latter approach is more favorable only for a portion of clustering (number of clusters between 7 and 16). Considering the advantages and disadvantages of the applied approaches, it seems that the expert method is the best one of all those presented.

References

1. Arguelles, M., Benavides, C., Fernandez, I.: A new approach to the identification of regional clusters: hierarchical clustering on principal components. Appl. Econ. **46**(21), 2511–2519 (2014). https://doi.org/10.1080/00036846.2014.904491

2. Boongoen, T., Iam-On, N.: Cluster ensembles: a survey of approaches with recent extensions and applications. Comput. Sci. Rev. **28**, 1–25 (2018). https://doi.org/10.1016/j.cosrev.2018.01.003
3. Bradbury, K.L., Mayer, C.J., Case, K.E.: Property tax limits, local fiscal behavior, and property values: evidence from Massachusetts under Proposition 2,5. J. Public Econ. **80**, 287–311 (2001)
4. Davidson, I., Ravi, S.S.: Agglomerative hierarchical clustering with constraints: theoretical and empirical results. In: Jorge, A.M., Torgo, L., Brazdil, P., Camacho, R., Gama, J. (eds.) Knowledge Discovery in Databases: PKDD 2005. Lecture Notes in Computer Science vol. 3721, pp. 59–70. Springer, Berlin, Heidelberg (2005). https://doi.org/10.1007/11564126_1
5. Dedkova, O., Polyakova, I.: Development of mass valuation in Republic of Belarus. Geomatics Environ. Eng. **12**(3), 29–39 (2018). https://doi.org/10.7494/geom.2018.12.3.29
6. Doszyń, M.: Ekonometryczna wycena nieruchomości. Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania **26**, 41–52 (2012)
7. Fang, Y.X., Wang, Y.H.: Selection of the number of clusters via the bootstrap method. Comput. Stat. Data Anal. **56**, 468–477 (2012). https://doi.org/10.1016/j.csda.2011.09.003
8. Guo, G.: Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). Int. J. Geogr. Inf. Sci. **22**(7), 801–823 (2008). https://doi.org/10.1080/13658810701674970
9. Hozer, J.: Regresja wieloraka a wycena nieruchomości. Rzeczoznawca Majątkowy **2**, 13–14 (2001)
10. Hozer, J., Kokot, S., Foryś, I., Zwolankowska, M., Kuźmiński, W.: Ekonometryczny algorytm masowej wyceny nieruchomości gruntowych. Uniwersytet Szczeciński, Stowarzyszenie Pomoc i Rozwój, Szczecin (1999)
11. Hozer, J., Kokot, S., Kuźmiński, W.: Metody analizy statystycznej rynku w wycenie nieruchomości. PFSRM, Warszawa (2002)
12. Jahanshiri, E., Buyong, T., Shariff, A.R.M.: A review of property mass valuation models. Pertanika J. Sci. Technol. **19**, 23–30 (2011)
13. Kantardzic, M.: Data Mining. Concepts, Models, Methods, and Algorithms. Wiley-IEEE Press (2003)
14. Kokot, S.: Model wielu regresji pojedynczych w wycenie nieruchomości. Studia i Materiały Towarzystwa Naukowego Nieruchomości **12**(1), 106–122 (2004)
15. Kolesnikov, A., Trichina, E., Kauranne, T.: Estimating the number of clusters in a numerical data set via quantization error modeling. Pattern Recogn. **48**(3), 941–952 (2015). https://doi.org/10.1016/j.patcog.2014.09.017
16. Korteweg, A., Sorensen, M.: Estimating loan-to-value distributions. Real Estate Econ. **44**(1), 41–86 (2016). https://doi.org/10.1111/1540-6229.12086
17. Kotkowski, P.: Propozycja nowej klasyfikacji terenów miejskich. Acta Universitatis Lodziensis, Folia Geographica Socio-Oeconomica **2**, 115–124 (1999)
18. Lis, C.: Sieci neuronowe a masowa wycena nieruchomości, vol. 318, pp. 331–338. Zeszyty Naukowe Uniwersytetu Szczecińskiego, Prace Katedry Ekonometrii i Statystyki, "Mikroekonometria w Teorii i Praktyce" (2001)
19. Müller, A.C., Guido, S.: Introduction to Machine Learning with Python. O'Reilly, Sebastopol (2016)
20. Park, B., Bae, J.K.: Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data. Expert Syst. Appl. **42**, 2928–2934 (2015)
21. Prystupa, M.: O potrzebie dalszych prac nad zastosowaniem regresji wielorakiej w wycenie nieruchomości. Rzeczoznawca Majątkowy **4**, 16–17 (2000)
22. Raschka, S., Mirjalili, V.: Python Machine Learning. Packt Publishing, Birmingham, Mumbai (2017)
23. Sawiłow, E.: Analiza wybranych metod modelowania wartości katastralnych nieruchomości. Acta Scientiarum Polonorum Geodesia et Descriptio Terrarum **8**(2), 27–38 (2009)
24. Suchecka, J. (ed.): Statystyka przestrzenna. C.H. Beck, Warsaw (2014)
25. Surowiec, G., Malczewska, A.: Automatyzacja procesu wyodrębniania stref izowartościowych nieruchomości w oparciu o zdjęcia lotnicze. Archiwum Fotogrametrii, Kartografii i Teledetekcji **11**, 21–35 (2001)

26. Tzioumis, K.: Appraisers and valuation bias: an empirical analysis. Real Estate Econ. **45**(3), 679–712 (2017)
27. Unpingco, J.: Python for Probability, Statistics, and Machine Learning. Springer International Publishing (2016)
28. Wiśniewski, R.: Zastosowanie sztucznych sieci neuronowych do wyceny masowej. Wycena **1**, 15–20 (1998)
29. Wu, X., Ma, T., Cao, J., Tian, Y., Alabdulkarim, A.: A comparative study of clustering ensemble algorithms. Comput. Electr. Eng. **68**, 603–615 (2018). https://doi.org/10.1016/j.compeleceng.2018.05.005
30. Zurada, J., Levitan, A.S., Guan, J.: A comparison of regression and artificial intelligence methods in a mass appraisal context. J. Real Estate Res. **33**(3), 349–387 (2011)

Chapter 21
Market Timing Models for Equity Funds Operating on the Polish Market in the Years 2003–2017

Dorota Żebrowska-Suchodolska and Andrzej Karpio

Abstract The use of a proper investment policy by fund managers has a huge impact on the results achieved. Asset selection for the portfolio is of key importance, in line with the investment policy declared by the fund, and the appropriate timing of changes in the assets in the portfolio. This research aims to assess whether managers in the Polish market have the ability to select assets with a sense of market timing and selectivity. The research concerns domestic equity funds operating in between 2003 and 2017, from which 15 were selected. The entire research period was divided into three five-year sub-periods, to some extent coinciding with changing market conditions, and seven two-year periods, covering the years 2003–2016. It was assumed that the results for the entire period describe long-term strategies of funds, five-year and two-year periods, being respectively the medium and short term. The basis for calculations was the monthly rates of return on participation units. The research was carried out using the Treynor–Mazuy, Henriksson–Merton, and Busse models. WIBOR 1M was accepted as a risk-free rate, and the WIG index was the market factor. The summary of the research is a comparison of the results regarding the selectivity of asset selection, market timing, and the ability of managers of open investment fund portfolios to respond to the market situation. The final conclusions are not very favorable for the managers, and the clients require active involvement in the selection of funds.

Keywords Investment fund · Market timing · Market sensitivity · Selectivity

JEL Codes C51 · G11 · G23

D. Żebrowska-Suchodolska · A. Karpio (✉)
Division of Financial Engineering, Department of Econometrics and Statistics, Faculty of Applied Informatics and Mathematics, University of Life Sciences—SGGW, Nowoursynowska 159, 02-787 Warsaw, Poland
e-mail: andrzej_karpio@sggw.pl

D. Żebrowska-Suchodolska
e-mail: dorota_zebrowska_suchodolska@sggw.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_21

21.1 Introduction

Clients of open-end mutual funds entrust their savings to professionals hoping for better investment results than they would be able to achieve themselves. Their motivations also include access to a market that requires appropriate knowledge and skills, namely the stock market. Lacking adequate knowledge, an individual investor is exposed to many forms of risk that accompany investments on the stock exchange. An additional barrier is access to information and the amount of work that is required to analyze financial statements, market trends, macroeconomic indicators, etc., on an ongoing basis. Therefore, clients of open-ended investment funds in shares decide to bear the costs of which are to be the payment for the professionalism and skills of those managing the investment portfolios. In the financial press and on specialized sites dedicated to capital markets, one can find rankings based on the results of funds achieved over different periods of time: a month, a year, from the beginning of activity, etc. Unfortunately, these are usually based on appropriate rates of return without taking into account the risk. Sometimes, an information indicator appears (e.g., on the Analizy Online portal), which takes into account the ratio of the excess rate of return to the risk measured by standard deviation. It is included in the group of standard measures such as Sharpe, Treynor, or Jensen ratios and many others. Their importance for the assessment of investment effectiveness is quite significant, but they do not refer directly to management skills, only to the final result of the investment. This is why these rankings differ significantly depending on the market situation; see Karpio and Żebrowska-Suchodolska [12]. Thus, they show that managers are not able to adapt their investment policy to the changing market situation. Therefore, investing savings in units of a particular fund does not guarantee profits in the long term, or even in the medium term. It exposes customers to losses, forcing them to actively manage deposits through, for example, changes in funds depending on the anticipated market situation. This, in turn, requires the client to undertake actions that are within the competence of fund managers and what clients pay them to do.

The aim of this work, therefore, is an econometric analysis of the ability to use market trends, market timing, and the ability to carefully select assets for open-end mutual fund portfolios. For this purpose, three models were used: Treynor–Mazuy, Henriksson–Merton, and Busse. The first two take into account the market trend measured by the excess of the percentage changes in the market index above the risk-free rate. They differ in the definition of the variable related to market timing skill. In contrast, the Busse model in addition includes market risk. The aim of the work is to compare management skills as described by these three different models and thus the measures of investment efficiency. Next, the stability of the ranking positions is examined, which provides information on whether there are leaders among investment funds or not.

The studies took into account entities operating continuously on the Polish capital market between the years 2003–2017. The models are constructed for the entire

period and for two- and five-year sub-periods. This is to examine the skills of managing fund portfolios at different investment time horizons. It was assumed that the entire 15-year period corresponds to long-term investments, and the five-year and two-year periods to the medium- and short-term, respectively. In addition, the unit of participation was taken into account without consideration of management fees and commissions.

21.2 Literature Review

The first research on market sensitivity dates back to the 1960s. At that time, Jensen formulated a single-index model and an alpha coefficient, now called Jensen's alpha coefficient. His model was subsequently applied in research by Carlson [3], Ippolito [10] or Grinblatt and Titman [7], among others. Both Ippolito, and Grinblatt and Titman pointed to managers obtaining above-average rates of return. In addition, they suggested that higher growth rates or the smallest net asset value is more likely to have above-average rates of return. In turn, Carlson [3] indicated that the results obtained depend on the selection of the period and the market index.

The two most used models of market timing and selectivity are the models created by Treynor–Mazuy [18] and Henriksson–Merton [9]. Studies of market sensitivity using these models have been performed for a number of markets. Noteworthy here is the work of Chang and Lewellen [4], or Gallo and Swanson [6], who studied the US market. Gallo and Swanson, using the Treynor–Mazuya model, indicate in their research the absence of market timing, and Chang and Lewellen [4] point to its existence to a small extent using the Henriksson–Merton model (for 67 mutual funds in 1971–1979). More recent studies include the work of Sheikh and Noreen [16], in which the authors analyze market timing using the Treynor–Mazuya model. Investigating 50 randomly selected funds on the American market between 1990 and 2008, they conclude that managers do not demonstrate the ability to sense the market and those funds are not able to achieve better results than the market. Research has been carried out on the Indian market by, among others, Rao and Venkanteswarlu [15]. Their research has shown that managers do not have the skill of market timing. Gudimetla [8] has also conducted research on the Indian market. When examining sectoral funds in 2012–2015, he came to the conclusion that they do not show market timing. On the other hand, out of 41 Indian market funds surveyed in 2008–2013 only a few of them showed selectivity. A study of 15 Greek equity funds in the years 2000–2008 using the Treynor–Mazuy model indicates a lack of selectivity and to a small extent market timing Koulis et al. [13]. Research on the Greek market was also performed by Philippas and Tsionas [14] or Thanou [17]. In turn, of the 28 Pakistani market funds examined in 2005–2011 using the Treynor–Mazuya model, two showed characteristics of timing and only one any of selectivity Ahmat and Sattar [1]. Busse [2] pointed out in his research that, taking into account the market situation, managers can change the exposure to risk. He thus proposed a model that takes into account risk variability.

In the case of the Polish market, research on investment funds has been conducted using the Treynor–Mazuy or Henriksson–Merton models by Jamróz [11], Węgrzyn [20], Żelazowska [22], Ważna [19], Witkowska et al. [21], among others. These authors' research on for various types of funds and various periods indicates the existence of funds that demonstrate the ability to select stocks appropriately.

21.3 Models of Portfolio Investment Management Skills

Among the econometric models that are applied in research in investment portfolio management skills, three were used in this work. Two of them, namely the Treynor—Mazuy and the Henriksson–Merton models, concentrate on excess rate of return over a risk-free rate. The third model, the Busse model, takes into account the risk measured by variability, i.e., the standard deviation. It therefore takes into account a factor that is very important from the point of view of active management of the investment portfolio. Considering risk when making investment decisions, rather than just using return rates, indicates the ability, or lack thereof, to adjust investment policy to market risk. Managers should take into account both the rates of return and risk if they are to actively approach the work funds charge their client's fees for. As a consequence, these three models are aimed at singling out potential leaders among investment funds, both from the point of view of the ability to choose the right assets, to adapt to changing market conditions, and to take advantage of the opportunities offered by market volatility.

In all models, the symbols $r_{A,t}$, $r_{f,t}$ and $r_{M,t}$ mean, respectively, the return on the fund A portfolio, the risk-free rate and the percentage change in the market index in the period t. All interest rates are calculated in a "traditional" way not logarithmic. The upper letters in the alpha and beta coefficients will identify them in all models. In the following, they will be HM and BU, respectively, for the Henriksson–Merton and Busse models. The following equation describing the Treynor–Mazuy model [18] is adopted:

$$r_{A,t} - r_{f,t} = \alpha_A^{\mathrm{TM}} + \beta_A^{\mathrm{TM}}\left(r_{M,t} - r_{f,t}\right) + \delta_A\left(r_{M,t} - r_{f,t}\right)^2 + \varepsilon_{A,t} \tag{21.1}$$

The constant term α_A^{TM} is a measure of the ability to select assets for the portfolio regardless of the market situation, and β_A^{TM} describes the manager's ability to use market trends. The δ_A coefficient is a measure of market timing skills. It should be noted that the variable at the δ_A is raised to a square; therefore, a positive value of the coefficient indicates the skill of market timing regardless of whether the variable is positive or negative. The expression of the professionalism of managers is the management of both a rising market $\left(r_{M,t} - r_{f,t} > 0\right)$ and a falling one $\left(r_{M,t} - r_{f,t} < 0\right)$. It is slightly different with the coefficient β_A^{TM}. In order to earn profits from using market trends, the portfolio should have a positive β_A^{TM} value when the market is up and down when it is in decline. Another way to describe management skills is the

Henriksson–Merton model [9], which can be described by the equation:

$$r_{A,t} - r_{f,t} = \alpha_A^{\mathrm{HM}} + \beta_A^{\mathrm{HM}}\left(r_{M,t} - r_{f,t}\right) + \gamma_A y_{M,t} + \varepsilon_{A,t} \tag{21.2}$$

The variable γ_A is equal to $y_{M,t} = \max\left\{0, r_{M,t} - r_{f,t}\right\}$. As a consequence, the coefficient γ_A, similar to the Treynor–Mazuy model, is a measure of market timing, but this time only a raising market is taken into account $\left(r_{M,t} - r_{f,t} > 0\right)$. The interpretation of the remaining coefficients is the same as before. Indeed, another model of describing management skills was proposed by Busse [2]. In addition to the return rates, it takes into account the risk and as a consequence has the form of the equation:

$$r_{A,t} - r_{f,t} = \alpha_A^{\mathrm{BU}} + \beta_A^{\mathrm{BU}}\left(r_{M,t} - r_{f,t}\right) + \lambda_A\left(\sigma_{M,t} - \bar{\sigma}_M\right)\left(r_{M,t} - r_{f,t}\right) + \varepsilon_{A,t} \tag{21.3}$$

where $\sigma_{M,t}$ is the volatility of the market in the period t measured by the standard deviation, and $\bar{\sigma}_M$ is the average volatility Foran and O'Sullivan [5]. The above equation results from the standard formulation of the CAPM model and the assumption by Busse that the market beta coefficient is a linear function of volatility. The coefficient λ_A, depending on the sign, describes the ability to adapt investment policy to changes in market risk in conjunction with the market situation. A negative value means that with high market volatility, the return on the portfolio should change against the market. Thus, in this case, the portfolio has less risk exposure when the anticipated scenario assumes the shape of a rising market. Consequently, it is desirable that the fund's portfolio have a negative value of the λ_A coefficient.

In all models, the significance of the estimators of structural parameters was tested at 5%, testing the following two-tailed hypotheses: H_0: parameter $= 0$ and H_1: parameter $\neq 0$. In order not to complicate the determinations in the further part of the work, the parameter estimators will be marked with the same symbols as the structural parameters.

21.4 Methodical Assumptions

As mentioned above, the research covered open-end equity mutual funds operating on the Polish capital market between 2003 and 2017. They are as follows (in brackets their short names are given, under which they will appear in the further part of the work): Arka BZ WBK Akcji Polskich (Arka), Aviva Investors Polskich Akcji (Aviva), Investor Top 25 Małych Spółek (AXA), Esaliens Akcji (Esaliens), Investor Akcji (Investor), Investor Akcji Spółek Dywidendowych (Investor D), Millennium Akcji (Millennium), NN Akcji (NN), Novo Akcji (Novo), Pioneer Akcji Polskich (Pioneer), PZU Akcji Krakowiak (PZU), Rockbridge Akcji (Rock), Rockbridge Akcji Dynamicznych Spółek (Rock D), Skarbiec Akcja (Skarbiec), UniKorona Akcje (UniKorona).

In addition to the entire period from 2003 to 2017, the studies were repeated for the following two-year sub-periods: 2003–2004, 2005–2006, 2007–2008, 2009–2010, 2011–2012, 2013–2014, and 2015–2016, and the five-year periods 2003–2007, 2008–2012, 2013–2017. The choice of various sub-periods was aimed at accounting for the investment time horizon from the point of view of potential customers. The authors found it interesting to explore the problem of careful asset selection, the use of market conditions, and the adjustment of investment policy to market volatility in different periods of time. This will make it possible to additionally link the given elements of management strategies to changing market conditions. As a risk-free rate, the WIBOR 1M from the interbank market was adopted, i.e., in the period for which the changes in share units were calculated. The benchmark describing the situation on the Polish stock market was the WIG index. This indicator was chosen for two reasons. It is strongly correlated with the WIG20 index and thus similarly describes the market situation. In addition, many funds have shares in portfolios that are not part of the WIG20 index but are included in the WIG. All percentage changes, both fund participation units and the market factor and risk-free rate, were calculated on a monthly basis. It was assumed that this is the period corresponding to possible changes in the funds' investment portfolios.

21.4.1 Research Results in 2003–2017

The results obtained over the entire research period will be presented first. It turns out that some aspects are shared by the sub-periods. Table 21.1 presents the results of parameter estimation of all the models considered.

It is easy to see that only the beta factors, which in all models describe the use of market trends, are different from zero. Their average values are equal to $\overline{\beta}^{\mathrm{TM}} = 0.8513$, $\overline{\beta}^{\mathrm{HM}} = 0.9236$, $\overline{\beta}^{\mathrm{BU}} = 0.8204$. It should be noted that they are less than one, which means that managers fail to build "better" portfolios than the market measured by the surplus of the WIG index above the risk-free rate. The table shows that only the portfolios of the Pioneer fund had a beta coefficient slightly higher than one and only in the Treynor–Mazuy and Henriksson–Merton models. In the Busse model, it is less than one, as for all other funds. The coefficients of variation of the beta coefficients of the models indicate a relatively small dispersion around the mean values and have the values $V_{\mathrm{TM}} = 7.9\%$, $V_{\mathrm{HM}} = 6.7\%$, $V_{\mathrm{BU}} = 9\%$, which proves that the values of the beta coefficients for various investment funds show a relatively small dispersion.

The alpha coefficients that indicate the ability to carefully select assets for portfolios in most cases are statistically insignificant. In particular, in the Treynor–Mazuy model, only one factor is different from zero and negative. In the Henriksson–Merton model, there are four (three positive and one negative), and in the Busse model seven (six positive and one negative). Moreover, the coefficients that are significantly

Table 21.1 Results of the estimation of structural parameters of the model between 2003 and 2017

Funds	Treynor–Mazuy model			Henriksson–Merton model			Busse model		
	Alpha	Beta	Delta	Alpha	Beta	Gamma	Alpha	Beta	Lambda
Arka	**0.0007**	0.9281	−0.3754	**0.0021**	0.9914	**−0.1206**	**0.0008**	0.8786	9.6435
Aviva	**0.0024**	0.8919	−0.5288	0.0049	0.9942	−0.1949	**0.0019**	0.8398	9.2554
AXA	**0.0006**	0.8254	**−0.6810**	**0.0030**	0.9364	**−0.2115**	**0.0001**	0.7630	12.6853
Esaliens	**0.0015**	0.8188	−0.5765	0.0032	0.9055	−0.1651	**0.0003**	0.7886	6.4160
Investor	**0.0023**	0.8046	−0.7521	0.0063	0.9597	−0.2955	**0.0012**	0.7486	9.8766
Investor D	**−0.0010**	0.8299	−0.5357	**0.0016**	0.9346	−0.1994	**−0.0020**	0.7902	5.5049
Millennium	**−0.0019**	0.8147	−0.3877	**−0.0010**	0.8679	**−0.1014**	−0.0029	0.7992	**3.4279**
NN	**−0.0016**	0.8829	**0.0049**	**−0.0012**	0.8920	**−0.0172**	−0.0020	0.8995	**−3.3803**
Novo	**−0.0025**	0.9006	**−0.3861**	**−0.0018**	0.9492	**−0.0927**	−0.0029	0.8708	**5.9944**
Pioneer	**−0.0061**	1.0100	**−0.2053**	−0.0045	1.0636	**−0.1021**	−0.0054	0.9515	**9.0578**
PZU	−0.0024	0.8496	−0.4630	**−0.0001**	0.9425	−0.1769	−0.0035	0.8316	**3.7354**
Rock	**−0.0020**	0.8471	**−0.2152**	**−0.0011**	0.8878	−0.0776	−0.0028	0.8465	**0.0983**
Rock D	**−0.0033**	0.6948	−0.8027	**−0.0004**	0.8269	**−0.2516**	**−0.0040**	0.6232	18.0904
Skarbiec	**0.0000**	0.8251	**−0.1828**	**0.0005**	0.8504	**−0.0484**	**−0.0006**	0.8231	**0.3939**
UniKorona	**0.0006**	0.8462	**−0.0774**	**0.0006**	0.8521	**−0.0112**	0.0002	0.8509	−0.8409

The numbers in boldface indicate the statistically insignificant factors
Source Own study

different from zero have very low values. Therefore, it can be concluded that the managers show a negligible degree of ability to carefully select assets for the portfolio in a way that increases the investment performance of the funds (rate of return on participation units).

The analysis of market timing skills is also not encouraging. The first two models, which in this case do not take risk into account, show some similarity. Of the 15 coefficients in the Treynor–Mazuy model, eight are significantly different from zero and five in the Henriksson–Merton model. However, all of them are negative, which means that managers react to changes in the market in a manner opposite to what they should do. The third Busse model confirms this conclusion, as nine lambda coefficients are statistically significant and positive. One lambda coefficient in the case of the UniKorona fund is significant and negative, which means that only one manager reacts to changes in the market situation and risk in a way that leads to an increase in the rate of return on participation units.

The results presented here justify the creation of fund rankings only on the basis of beta coefficients. From the point of view of the analysis of the equity mutual funds market, the answer to the question is whether the ranking items created on the basis of different models are compatible with each other. Spearman's rank correlation coefficient can be a measure. Its values are as follows: $\rho_{\mathrm{TM,HM}} = 0.6357$, $\rho_{\mathrm{TM,BU}} = 0.8929$, $\rho_{\mathrm{HM,BU}} = 0.3464$,. The first two are statistically significant, and the third one is not. As a consequence, it can be concluded that following the trend measured by beta factors indicates a relatively high consistency of rankings created on the basis of the Treynor–Mazuy model and the other two models. However, the rankings based on the Henriksson–Merton and Busse model are characterized by a lack of correlation. In this case, the rankings differ from each other considerably. It is difficult to explain this result, but the answer may lie in the definitions of risk in Busse and Henriksson–Merton models. However, the Treynor–Mazuy model does not take risk into account.

21.4.2 Research Results in Five-Year Periods

Divided into five-year periods, it can be seen that in the first period, there was a bull market, with the exception of the end of year 2007. The years 2008–2012 are a period of changing economic conditions, including a decisive decline, as well as a rise and another fall. The third period is also characterized by economic volatility, but is definitely less dynamic than in the previous period.

As mentioned in Sect. 21.4.1, some aspects of the test results for the entire period remain valid in this case as well as the next two-year periods. One can speak here about the significance of the structural parameter estimators of the models. Beta coefficients are statistically significant in most cases. Unfortunately, this can not be said about the other coefficients. The number of significant estimators, apart from the beta coefficient, is given in Table 21.2.

Table 21.2 Number of statistically significant estimators of structural coefficients of models in five-year periods

Coefficients	2003–2007			2008–2012			2013–2017		
	TM	HM	BU	TM	HM	BU	TM	HM	BU
Alpha	1	1	1	8	3	8	2	3	2
Delta, gamma, lambda	0	0	2	3	4	3	1	2	10

Source Own study

In the case of the alpha coefficients, the second period is the best. In that period, eight of them are statistically significant in the case of the Treynor–Mazuy and Busse models. However, it should be added that all of them are negative. This results in a decrease in the rate of return from fund participation units. In the Busse model, between the years 2013–2017, ten lambda coefficients were statistically significant, but all were positive. This indicates a mismatch of portfolios to risk, leading to adverse changes in the rate of return. As a consequence, just as before, only the beta coefficients provide the basis for evaluating managers. These were gathered in Table 21.3.

Beta coefficients are only sporadically larger than one which means that managers rarely manage to build a portfolio that yields a return rate greater than the WIG surplus above the risk-free rate. This conclusion is confirmed by the average value, which is lower than one for all models and for all sub-periods. Moreover, the coefficients of variation are relatively small. One can therefore risk saying that managers have similar strategies regarding the use of market trends. It is also worth noting that these conclusions do not depend on the market situation. In the years 2003–2007, aside from 2007, there was a bull market. In the remaining sub-periods, the economic situation was volatile, but the upward trend prevailed. This should be used by managers to increase profits, but from the point of view of beta coefficients, this did not happen. Only the volatility of this measure of effectiveness is significantly lower in the years 2008–2012 compared to the remaining periods.

The analysis of ranking positions based on beta coefficients and occupied in subsequent periods in the analyzed models additionally confirms the above conclusions. It turns out that for all models, Spearman's rank correlation coefficients between rankings created for subsequent sub-periods within each model are statistically insignificant. It should be noted that the rankings were created without taking the Rock D fund into account, because in the first period the beta factor was insignificant. Spearman's coefficients calculated for rankings based on different models are given in Table 21.4.

The table shows that there is a very large correlation between the ranking items created on the basis of the Henriksson–Marton and Busse models in 2003–2007 and 2013–2017. It can be assumed that the ranking of funds is then practically the same. A similar situation takes place for the Treynor–Mazuy and Henriksson–Merton models in 2008–2012 and 2013–2017, but at a slightly lower level. However, the insignificance of coefficients means that the ranking positions differ significantly.

Table 21.3 Values of beta factor estimators

Funds	2003–2007			2008–2012			2013–2017		
	TM	HM	BU	TM	HM	BU	TM	HM	BU
Arka	0.8716	0.8670	0.8563	0.9289	1.0124	0.8707	0.8791	0.9637	0.8513
Aviva	0.8805	0.9612	0.8290	0.9324	1.0073	0.8347	0.7676	0.8732	0.7269
AXA	0.8000	0.8449	0.7698	0.8725	0.9295	0.8800	0.6295	0.7032	0.6293
Esaliens	0.8278	0.8977	0.8022	0.8048	0.9317	0.7047	0.7752	0.6534	0.7922
Investor	0.7785	0.8243	0.7146	0.8720	0.9531	0.8733	0.6605	1.0549	0.5776
Investor D	0.7631	0.8002	0.7353	0.9000	0.9883	0.9001	0.7226	0.8048	0.7378
Millennium	0.8038	0.8577	0.7454	0.8292	0.8554	0.8315	0.8356	0.8598	0.8242
NN	0.8518	0.7875	0.8792	0.8761	0.8822	0.9066	0.9536	1.0657	0.9284
Novo	0.7749	0.7939	0.7247	0.9121	0.9649	0.8805	1.1674	1.2131	1.2336
Pioneer	0.8925	0.8799	0.9083	1.1051	1.1260	1.0671	0.8933	0.8709	0.8988
PZU	0.7542	0.8374	0.7195	0.8687	0.9368	0.8721	1.0240	1.1407	1.0310
Rock	0.7976	0.8435	0.7435	0.8720	0.8616	0.9122	0.8824	0.9632	0.8880
Rock D	0.3855	**0.4016**	0.2857	0.9104	0.8712	0.9129	0.7353	0.8075	0.7688
Skarbiec	0.7802	0.8069	0.7519	0.8246	0.8241	0.8315	0.9467	1.0627	0.9448
UniKorona	0.8178	0.7808	0.8223	0.8506	0.8453	0.8731	0.8768	0.9545	0.8956
Average	0.7853	0.8123	0.7525	0.8906	0.9327	0.8767	0.8500	0.9328	0.8485
Volatility coefficients (%)	14.61	14.66	18.36	7.61	8.33	8.11	16.13	16.36	18.31

A number in boldface indicates a statistically insignificant factor
Source Own study

Table 21.4 Spearman's rank correlation coefficients for rankings based on different models

Periods	TM/HM	TM/BU	HM/BU
2003–2007	**0.5253**	**0.3143**	0.9253
2008–2012	0.7514	**0.1821**	0.5536
2013–2017	0.7464	0.6750	0.9714

The numbers in boldface indicate the statistically insignificant factors
Source Own study

This is the case for fully two periods in the Treynor–Mazuy model and in the Busse models for the first two sub-periods. This is evidence of instability in the ranking positions. This is a very unfavorable situation for fund clients.

21.4.3 Research Results in Two-Year Periods

In this case, the entire research period does not include 2017 and covers the years 2013–2016. This is related to the need to maintain identical two-year periods. In two-year periods, corresponding to investing in the short term, the situation is analogous to the previous ones. In the vast majority of cases, beta factor estimates are statistically significant. However, estimators of other structural parameters are not. In the years 2003–2004, 2005–2006, 2009–2010, there was an almost uniform increase in the index. The years 2013–2014 are also characterized by an increase in the index, but its fluctuations are clearly greater than in other periods of the boom. However, the years 2007–2008, coinciding with the period of the global financial crisis, show a sharp drop in WIG values (Table 21.5).

Ten lambda coefficients in the Busse model were statistically significant in the years 2015–2017. However, all the values were positive, which indicates that risk management is incompatible with the market. In all other periods and models, only a small number of factors related to market timing were statistically significant. Beta coefficients were collected in Tables 21.6, 21.7, and 21.8.

In the first two periods and in all models, we have to do with relatively high volatility exceeding 20%. Bear in mind that these are periods of boom. Thus, the high volatility shows that the efficiency of managers following the market trend was characterized by a widespread. At the same time, the average beta factor has a relatively low value, much less than one. Therefore, the managers were not able to beat the market, despite the fact that the market situation offered opportunities to do so. Taking all the models into account, only the Arka fund portfolio in 2005–2006 had a beta factor greater than one. In the next boom period, 2009–2010, the coefficient of variation of the beta parameter has a much lower value than in other periods and its average value of 0.8 is higher than in previous periods of market growth. In the case of the Henriksson–Merton model in periods of changing market conditions (2011–2012 and 2015–2016), the average value of the beta coefficient exceeded one. This value

Table 21.5 Number of statistically significant estimators of structural coefficients of models in two-year periods

Coefficients	2003–2004			2005–2006			2007–2008			2009–2010			2011–2012			2013–2014			2015–2016		
	TM	HM	BU	TM	HM	BU	TM	HM	BU	TM	HM	BU	TM	HM	BU	TM	HM	BU	TM	HM	BU
Alpha	1	1	1	2	2	2	4	2	3	0	0	1	3	1	5	2	2	4	4	4	5
Delta, Gamma, Lambda	1	0	0	1	1	0	2	3	2	2	2	5	2	1	3	0	0	0	3	5	10

Source Own study

Table 21.6 Beta coefficients in the Treynor–Mazuy model in two-year periods

Funds	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016
Arka	0.8223	0.8971	0.8603	0.8674	0.8851	0.8614	0.8954
Aviva	0.7524	0.8810	1.0114	0.9412	0.8289	0.7855	0.8259
AXA	0.6986	0.6050	1.1344	0.8609	0.8843	0.7236	0.6909
Esaliens	0.5907	0.9075	0.9752	0.8088	0.8140	0.8685	0.7788
Investor	0.5811	0.8390	0.9161	0.7877	0.9677	0.7928	0.7116
Investor D	0.6332	0.8801	0.8096	0.8411	0.9034	0.7761	0.7576
Millennium	0.7524	0.7230	0.9601	0.7660	0.9074	0.8503	0.8647
NN	0.7645	0.9323	0.8498	0.8801	0.9623	0.9399	0.9272
Novo	0.6285	0.8084	0.8891	0.8928	0.8527	0.9256	1.3332
Pioneer	0.7652	0.9135	0.9947	0.9495	1.2238	0.9485	0.8734
PZU	0.6268	0.7783	0.8590	0.8657	0.9232	1.0851	1.1013
Rock	0.6235	0.8261	1.0116	0.9165	0.9946	0.8948	0.9632
Rock D	**0.0991**	**0.1534**	1.1414	0.8344	0.8813	0.8330	0.8132
Skarbiec	0.7383	0.8208	0.7507	0.7895	0.8893	1.0037	0.9698
UniKorona	0.7284	0.8341	0.9061	0.9031	0.8578	0.8841	0.9247
Average	0.6537	0.7866	0.9389	0.8603	0.9184	0.8782	0.8954
Volatility coefficients (%)	25.20	23.85	11.39	6.29	10.34	10.32	17.47

The numbers in boldface indicate the statistically insignificant factors
Source Own study

Table 21.7 Beta coefficients in the Henriksson–Merton model in two-year periods

Funds	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016
Arka	0.6290	1.1278	1.1094	0.9033	0.9135	0.9525	1.0590
Aviva	0.8057	0.8787	0.9954	0.9989	1.0418	0.8192	1.2319
AXA	0.5727	**0.7924**	0.8165	0.8990	1.1250	**0.3538**	1.5025
Esaliens	0.6223	0.9617	0.9728	0.9402	0.8134	0.7272	0.6828
Investor	0.6160	0.7123	0.9124	0.7765	1.1650	0.8656	1.4244
Investor D	0.6373	0.7913	1.0321	0.8481	0.9471	0.8119	0.9215
Millennium	0.8732	0.5914	0.8555	0.7261	0.9488	0.8268	1.0683
NN	0.6736	0.9132	0.7905	0.9351	1.1322	1.0593	1.0998
Novo	0.5777	0.7917	1.0184	0.9209	0.7892	1.0880	1.2621
Pioneer	0.6994	0.8706	1.0628	0.8736	1.4889	0.9005	0.9398
PZU	0.6392	0.8464	0.9422	0.9696	0.9767	1.0215	1.4243
Rock	0.6533	0.8295	0.8423	1.0152	0.8911	0.9443	1.2056
Rock D	**0.0197**	**0.2800**	0.8483	0.7118	1.0385	0.6985	1.4167
Skarbiec	0.7725	0.7153	0.8498	0.7206	0.9548	0.9931	1.3423
UniKorona	0.6448	0.6760	0.8632	0.9789	0.7905	0.9832	0.9504
Average	0.6291	0.7852	0.9274	0.8812	1.0011	0.8697	1.1688
Volatility coefficients (%)	28.92	23.29	10.33	11.30	17.36	20.39	19.41

The numbers in boldface indicate the statistically insignificant factors
Source Own study

Table 21.8 Beta coefficients in the Busse model in two-year periods

Funds	2003–2004	2005–2006	2007–2008	2009–2010	2011–2012	2013–2014	2015–2016
Arka	0.9133	0.8154	0.9029	0.8356	0.8649	0.8339	0.8635
Aviva	0.7414	0.8627	1.0094	0.9080	0.7624	0.7678	0.8071
AXA	0.7407	0.5625	1.0404	0.8110	0.8411	0.7869	0.5691
Esaliens	0.6244	0.8781	0.9747	0.7388	0.8385	0.9103	0.7929
Investor	0.5706	0.8531	0.9122	0.8302	0.9400	0.8039	0.6006
Investor D	0.6500	0.8869	0.8710	0.9349	0.8821	0.7661	0.7502
Millennium	0.6547	0.7198	0.9176	0.8397	0.9147	0.8494	0.8698
NN	0.8722	0.9257	0.8289	0.9245	0.9500	0.9116	0.9018
Novo	0.6185	0.8273	0.9234	0.8020	0.8405	0.9174	1.4077
Pioneer	0.8610	0.9319	1.0141	0.9016	1.0820	0.9611	0.8649
PZU	0.6162	0.7716	0.8735	0.8474	0.9297	1.1031	1.0535
Rock	0.5802	0.8195	0.9741	0.8873	1.0210	0.8827	0.9320
Rock D	**0.1132**	**0.1002**	1.1002	0.8554	0.8486	0.8528	0.7912
Skarbiec	0.7240	0.8451	0.7580	0.8935	0.8877	1.0034	0.9559
UniKorona	0.7449	0.8857	0.9047	0.9185	0.8379	0.8721	0.9444
Average	0.6683	0.7790	0.9337	0.8619	0.8961	0.8815	0.8736
Volatility coefficients (%)	27.08	25.91	9.03	6.11	8.68	10.09	21.60

The numbers in boldface indicate the statistically insignificant factors
Source Own study

Table 21.9 Spearman's rank correlation coefficients between rankings created on the basis of different models in two-year periods

Periods	TM/HM	TM/BU	HM/BU
2003–2004	**0.5253**	**0.3143**	0.9165
2005–2006	0.6220	**0.1736**	0.8022
2007–2008	**−0.2250**	**−0.0250**	0.9500
2009–2010	0.6750	**0.1036**	**0.3857**
2011–2012	**0.4964**	**0.4643**	0.9715
2013–2014	0.7407	0.6132	0.9648
2015–2016	**0.0214**	**−0.0357**	0.9714

The numbers in boldface indicate the statistically insignificant factors
Source Own study

cannot be interpreted unequivocally positively, because following the market trend when the economic downturn is a negative activity. Note that in the years 2007–2008, there was a bear market, and in 2009–2010 a bull market. In the first case, the average beta coefficients, in all models, exceeds 0.9. In the second, it is greater than 0.6. The coefficients of variation are relatively low. Thus, the conclusion to be drawn states that managers follow the market trend independently of the market situation and are not able to use the bull market to improve their investment performance. This conclusion is confirmed by the analysis of Spearman's rank correlation coefficients. In the case of rankings created for each model separately in subsequent sub-periods, almost all ratios are statistically insignificant. The situation is different in the case of correlations between rankings based on different models.

Table 21.9 shows that comparison of rankings created on the basis of the Treynor—Mazuy and Henriksson–Merton and Treynor–Mazuy and Busse models produces completely different results. Of the 14 coefficients, only four are statistically significant. In the years 2013–2014, when there was an upward trend, but with the relatively high volatility on the stock exchange index, both ratios are significant and positive. Thus, during this period, regardless of the model, the majority of market leaders remained leaders. An interesting result is obtained for the rankings based on Henriksson–Merton and Busse models. Only one correlation coefficient (2009–2010) is zero from a statistical point of view. The remaining ones are positive with high values. It can be concluded that the rankings created on the basis of these two models carry the same information about the effectiveness of managers from the point of view of beta coefficients.

21.5 Conclusion

The research carried out on 15 open investment funds of shares operating on the Polish capital market does not lead to optimistic conclusions. The 15-year period considered is long enough to cover changing market conditions and gives the managers the

opportunity to demonstrate their skills. The fee is paid by clients regardless of whether the fund made a profit or suffered a loss. In return, the client has the right to expect professional management of the fund's assets and thus also her or his savings. This professionalism is the careful selection of portfolio components that should yield a profit regardless of the market situation. Skillful use of changing market conditions is measured by an appropriate correlation of changes in participation units with the factor of the market. Finally, there are managers having market timing skills, i.e., predicting changes in the economic situation and making appropriate changes in the composition of portfolios. Returning to the elements mentioned above used in the assessment of the effectiveness of fund portfolio management, it can be concluded that the ability to carefully select assets is at best unsatisfactory. The measure of this is the constant terms in the considered models. In nearly all periods and in nearly all models, the vast majority of alpha coefficients were statistically insignificant. The exceptions were the years 2008–2012, when eight of fifteen coefficients were statistically significant in the Treynor–Mazuy and Busse models. These, however, were negative, which can be said to have been particularly disadvantageous for fund clients.

The analysis of managers' use of market timing, unfortunately, does not offer flattering portrait. The delta, gamma, and lambda coefficients in the models under consideration show statistical properties similar to those of the alpha coefficients. The vast majority of them are either irrelevant or statistically significant but with the undesired sign. The Busse model is worth highlighting here. In the entire research period, eight lambda coefficients were statistically significant, but only one was negative. However, in the years 2013–2017 and 2015–2016, ten of them were different from zero, but had positive values. This means that the linking of the anticipated market situation with the risk was almost always insignificant.

The situation is different from the one described above, when we consider beta factors as a measure of the managers' ability to use market trends. In all periods and models, only in a few cases are they statistically insignificant and these almost always apply to the same fund. In addition, they always have positive values. It is worth mentioning that this is not an unambiguously positive situation from the point of view of customers. The studies also included bear market periods, which is when managers should build portfolios leading to negative beta factors. This conclusion is confirmed by the statistical characteristics of these coefficients in almost all models and all sub-periods under consideration. Their average values are greater than 0.6 and not higher than 0.94, except for the Henriksson–Merton model in 2011–2012 and the Busse model in 2015–2016. In those periods, the average is slightly larger than one. The calculation of variation coefficients for beta parameters shows that in all models and all sub-periods, the distribution of their values around the averages is small. It can be concluded that the managers of all funds follow the market in a very similar way. Moreover, their approach is more reminiscent of index funds rather than aggressive stock strategies.

Another element of the research conducted was the persistence of ranking positions. Due to the fact that the beta coefficients were statistically significant, the rankings could only be created based on this parameter. In the entire 2003–2017 period,

there is relatively high compatibility of ranking items based on the Treynor–Mazuy model and the other two models. Therefore, relatively low- or high-ranking positions in this model remain the same for both. However, the situation is quite different when we consider the Henriksson–Merton and Busse models, where the Spearman's coefficient is equal to zero from the statistical point of view. Thus, leaders become "marauders" and vice versa. The situation changes radically when we repeat the calculations for sub-periods. In the case of five-year periods, all three correlation coefficients between the rankings created on the basis of the Henriksson–Merton and Busse models are important. In 2003–2007 and 2013–2017, they have values greater than 0.9, and in 2008–2012, the coefficient is slightly greater than 0.5. In other cases, three coefficients are insignificant and three significant, but they have values close to 0.7. A similar situation occurs for two-year sub-periods. In the case of the Henriksson–Merton and Busse models for seven sub-periods, only one factor is statistically insignificant (2009–2010), while in the others they have high values, often significantly above 0.9. The correlation coefficients between rankings based on the Treynor–Mazuy and Henriksson–Merton models are insignificant in four cases and in three cases statistically significant, but they show values much lower than in the previous pair of models. Consequently, it can be stated that depending on the pair of models and the period of selected for study, rankings based on beta factors lead to very different results from the point of view of market persistence. Relatively high agreement exists between the Henriksson–Merton and Busse models.

One might be tempted to say that over the past 15 years, the market for open-end mutual funds has been characterized by low management efficiency. Managers paid the most attention to following the market trend, without even trying to predict it. The ability to carefully select assets was practically non-existent, and probably not considered. On the other hand, the effects of the active use of market timing were counterproductive. As a consequence, the portfolios of funds that should be actively managed, not to say aggressively, rather resemble the portfolios of index funds. As a consequence, clients of funds are exposed to not very encouraging profits from the entrusted money. One solution is systematic payment of small amounts with active selection and possible change of funds depending on the market situation.

References

1. Ahmat, J., Sattar, A.: Estimation of fund selectivity and market timing: performance of Pakistani mutual funds. Glob. Manage. J. Acad. Corp. Stud. **6**(2), 162–169 (2016)
2. Busse, J.A.: Another Look at Mutual Fund Tournaments. Working paper Emory University (1999)
3. Carlson, R.: Aggregate performance of mutual funds, 1948–1967. J. Fin. Quant. Anal. **5**(1), 1–32 (1970)
4. Chang, E.C., Lewellen, W.G.: Market timing and mutual fund performance. J. Bus. **57**, 57–72 (1984)
5. Foran, J., O'Sullivan, N.: Mutual fund skill in timing market volatility and liquidity. Forthcoming, Int. J. Fin. Econ. (2017). Available at SSRN: https://ssrn.com/abstract=2977795

6. Gallo, J., Swanson, P.: Comparative measures of performance for US based international equity mutual funds. J. Bank. Fin. **20**(10), 1635–1650 (1996)
7. Grinblatt, M., Titman, S.: Mutual fund performance: an analysis of quarterly portfolio holdings. J. Bus. **62**(3), 393–416 (1989)
8. Gudimetla, A.: Timing and Selectivity in Indian Sector Mutual Funds Performance. Jawaharlal Nehru Technological University Hyderabad, India (2015)
9. Henriksson, R., Merton, R.: On market timing and investment performance, II statistical procedures for evaluating forecasting skills. J. Bus. **54**(4), 513–533 (1981)
10. Ippolito, R.A.: Efficiency with costly information: a study of mutual fund performance, 1965–1984. Quart. J. Econ. **104**(1), 1–23 (1989)
11. Jamróz, P.: Parametryczna ocena umiejętności selektywności i wyczucia rynku zarządzających OFI akcji. Zeszyty Naukowe Uniwersytetu Szczecińskiego 639, Finanse, Rynki finansowe, Ubezpieczenia nr 37, pp. 221–231 (2011)
12. Karpio, A., Żebrowska-Suchodolska, D.: Dostosowywanie polityki inwestycyjnej FIO do zmiennej koniunktury rynkowej w latach 2003–2010. Metody i zastosowania ekonometrii współczesnej, vol. VII, pp. 105–115 (2013)
13. Koulis, A., Beneki, Ch., Adman, M., Botsaris, Ch.: An assessment of the performance of Greek mutual equity funds selectivity and market timing. Appl. Math. Sci. **5**(4), 159–171 (2011)
14. Philippas, N., Tsionas, E.: Performance evaluation: an review article and an empirical investigation of Greek mutual fund managers. Int. Bus. Econ. Res. J. **1**(6), 31–44 (2002)
15. Rao, K.V., Venkateswarlu, K.: Market Timing Abilities of Fund Managers—A Case Study of Unit Trust of India. UTI Institute of Capital Markets and Tata Mcgraw Hill Publications, Indian Capital Markets—Trends and Dimensions (2000)
16. Sheikh, M.J., Noreen, U.: Validity of efficient market hypothesis: evidence from UK mutual funds. Afr. J. Bus. Manage. **6**(2), 514–520 (2012)
17. Thanou, E.: Mutual fund evaluation during up and down market conditions: the case of Greek equity mutual funds. Int. Res. J. Fin. Econ. **13**, 84–93 (2008)
18. Treynor, J.L., Mazuy, K.: Can mutual funds outguess the market? Harvard Bus. Rev. **44**, 131–136 (1966)
19. Ważna, E.: Efektywność inwestowania otwartych funduszy inwestycyjnych w Polsce w latach 2007–2014. Zeszyty Naukowe Uczelni Vistula **56**(5), 64–83 (2017)
20. Węgrzyn, T.: Efektywność funduszy inwestycyjnych stosujących aktywne strategie zarządzania portfelem. Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach nr 239, pp. 141–152 (2015)
21. Witkowska, D., Kompa, K., Grabska, M.: Badanie informacyjnej efektywności rynku w formie silnej na przykładzie wybranych funduszy inwestycyjnych. Metody Ilościowe w Badaniach Ekonomicznych **X**, 265–285 (2009)
22. Żelazowska, I.: Wyczucie rynku w działalności funduszy inwestycyjnych akcji w Polsce-badanie empiryczne. Annales, LI **1**, 125–134 (2017)

Chapter 22
Identification of EU Countries Due to the Quality of Data on Intra-Community Supplies of Goods Shipped from Poland in the Years 2005–2017

Iwona Markowicz and Paweł Baran

Abstract Public statistics data regarding trade in goods between European Union Member States are collected and disseminated by Eurostat. They are aggregated data collected from national statistical offices from all Member States. In Poland, as in all other countries within EU, Intrastat system was introduced. Its main goal is collecting data from declarations submitted by businesses shipping and acquiring goods from partners in other EU Member States. Data from Intrastat (intra-Community trade) as well as Extrastat (foreign trade with third parties, i.e. partners residing outside of EU) serve as a base in research of foreign trade or economic development as well as establishing many macroeconomic measures. In the survey presented, data from Eurostat's COMEXT database was used. We utilised individual as well as aggregated indices of mirror data divergence. The aim of the paper was to analyse intra-Community supplies from Poland to the whole of EU as well as to individual countries over the period 2005–2017. Then, a ranking of Poland's trade-partner countries was created for every year under consideration. Those rankings had been compiled together and further conclusions about dynamics were drawn, particularly we show changes in data quality of Poland's intra-Community supplies.

Keywords Trade in goods shipped from Poland · Eurostat's Comext database · Mirror data

I. Markowicz (✉)
Department of Statistics, Faculty of Economics and Management, Econometrics and Statistics Institute, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: iwona.markowicz@usz.edu.pl

P. Baran
Department of Operations Research and Applied Mathematics in Economics, Faculty of Economics and Management, Econometrics and Statistics Institute, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: pawel.baran@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_22

22.1 Introduction

One can argue that research into economic phenomena is based on statistical data. They are the starting point for detecting regularities in the economy. The quality of data characterising the economy of a country or group of countries (e.g. the European Union) is an important issue. It should, therefore, be measured, and efforts should be made to improve it.

The authors of this paper focus in their research on the data characterising Poland's intra-Community trade with EU countries. The necessary data are provided by public statistics. The Central Statistical Office of Poland on its website provides data in the form of a SWAID database, where exports and imports are presented by sections of CN/SITC/CPA/BEC and groups of countries. At EU level, intra-Community trade data are presented by Eurostat in the COMEXT database. They are aggregated data collected from national statistical offices from all Member States. Data in the Member States are collected in the form of statistical declarations of entities engaged in intra-Community trade. Declarations are collected in the Intrastat system, which was launched on 1 January 1993 [6]. Countries that joined the EU after that date introduced this system at the time of their accession to the EU. The Intrastat statistical declaration replaced an earlier customs document due to the removal of customs borders between Member States.

In foreign trade statistics, the comparability of data derived from public statistics of different countries is a priority, as emphasised in EU legal acts [19, 20].

Baran and Markowicz [2, 3] point out some of possible irregularities in economic activities' evidence of Polish tax payers whose businesses depend on intra-Community trade (both supplies and acquisitions). These anomalies may be the result of hiding transactions by avoiding statistical declarations or late declaring, although there exist other possible causes of differences between intra-Community trade figures.

The specific details of data collection procedures differ between Member States; hence, mirror data are not fully compatible with each other. Mirror data are information on the same transactions recorded in public statistics of both countries of dispatch and acquisition (Fig. 22.1). Differences in declared values of transactions between partner states measure the quality of data collected.

The aim of the study was to analyse intra-Community supplies from Poland to the whole of EU as well as to individual countries over the period 2005–2017. In the first part, we utilised individual as well as aggregated indices of mirror data asymmetries. In subsequent part of the study, rankings of Poland's trade-partner countries were created for years 2005, 2008, 2011, 2014 and 2017. Those rankings had been compiled together, and further conclusions about dynamics were drawn. We show changes in data quality of Poland's intra-Community supplies and try to explain some of them.

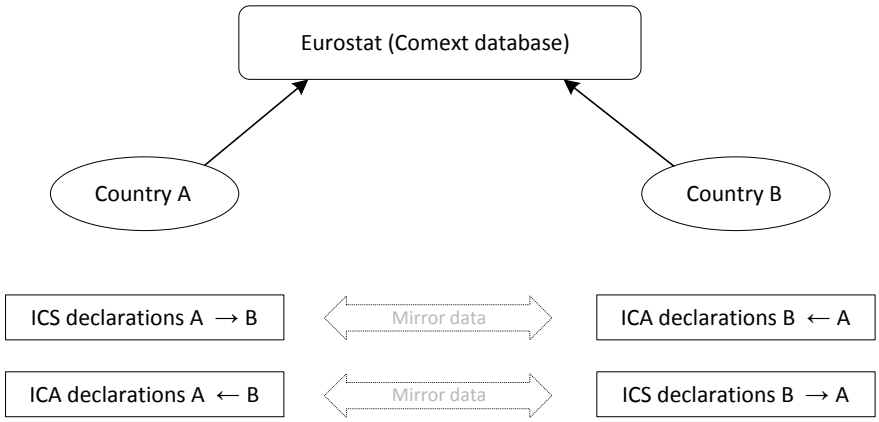


Fig. 22.1 Mirror data in intra-Community trade

22.2 The Need and Ways of Identifying Irregularities in Mirror Data—Literature Review

According to Parniczky [17], discrepancies in mirror data on trade have been present in economic literature at least since the 1920s. Tsigas et al. [21] argue that discussion on that issue is older, dating it back to the 1880s. Modern approaches to this issue have emerged in the 1960s.

The reasons for errors or irregularities in intra-Community trade mirror data are numerous, including: established values of statistical thresholds for exports and imports, late submission of declarations by obligated entities, misclassification of goods (CN nomenclature), concealment of transactions, declaration of fictitious transactions and misrepresentation of the country of destination or origin of goods. Also relevant are transportation costs, which are generally only recorded on the import side and exchange rate differences. An extensive review and discussion were carried out by Hamanaka [11], who, after Federico and Tena [7], divides the reasons for the differences between the mirror data into unavoidable differences between CIF-based and FOB-based reporting, structural differences between customs offices' approach to classifying transactions and human errors and deliberate misclassification. Hamanaka suggests that export data are generally less accurate than import data because governments are interested in proper registering imported goods in the first place. But it is disputable whether the same condition holds true in case of a customs union and a loosened regime of declaring foreign trade. Nevertheless, most of the above causes are responsible for differences between intra-Community trade data as well. Guidelines for data collection units at national level, including typology, sources and means of tackling irregularities in the foreign trade datasets, are also provided by Eurostat [6]. However, even their full implementation does not eliminate all sources of errors and data asymmetries. As a result, there are huge differences between datasets containing the same data reported by both sides of transactions, i.e.

as intra-Community supplies (ICS) by Polish firms and as intra-Community acquisitions (ICA) from Polish firms [3].

Many authors emphasise the fact that discrepancies in foreign trade data result from errors in the data entered or from deliberate concealment of economic fraud. As exports of goods and services to another Member State continue to be VAT-exempt, this has created a risk that these goods and services remain untaxed in both the supplying state and in the state of consumption [5]. In addition to the revenue loss for Member States, uncollected VAT has an effect on the European Union's own resources which in effect affects all Member States. For intra-Community VAT purposes, suppliers must be identified with a specific VAT identification and must regularly report their intra-Community supplies or transfers in a recapitulative statement submitted to the tax authorities of the supplying Member State. The EU has set up an electronic system (VIES), through which Member States exchange information on traders registered for VAT purposes and on intra-Community supplies.

Keen and Smith [14] state that like any tax, the VAT is vulnerable to evasion and fraud, and abuse of the weaknesses of VAT system is a serious problem in the European Union. They describe the main forms of non-compliance distinctive to VAT, consider how they can be addressed, and assess evidence on their extent in high-income countries. According to the same authors, while the practical significance of current difficulties in the EU should not be over-stated, administrative measures alone may prove insufficient to deal with them, and a fundamental redesign of the VAT treatment of intra-Community trade is required.

Pope and Stone [18] already in 2009 concluded that missing trader intra-Community (MTIC) fraud had been a problem across the European Union for many years, and much had been written about its effects and how best to tackle it. According to them, MTIC fraud exploits the zero-rated supply across national boundaries as a means for stealing revenues from national states (carousel fraud) or creating a VAT debt to be used as a subsidy for undercutting legitimate supplies (acquisition fraud).

MTIC fraud has been slowly morphing from cell phones and computer chips to other commodities [4]. Borselli states that MTIC can involve any type of goods but had favoured mobile phones and computer chips and later was quite common in sales of meat, cars, mineral oils, soft drinks, cosmetics and precious metals. According to Ainsworth [1], however, in the last few months of 2009, MTIC frauds made a dramatic appearance in tradable CO2 permits. In recent years, evidence of MTIC and MTEC (extra-Community) fraud involving fictitious trading in electricity and gas has also emerged [15].

22.3 Analysis Framework

Poland's place in intra-Community trade is shown in Fig. 22.2. Figures 22.2 and 22.3 were compiled using the `circular` library by Gu et al. [9]. The outer ring shows the share of each country in intra-Community trade in goods (exports and

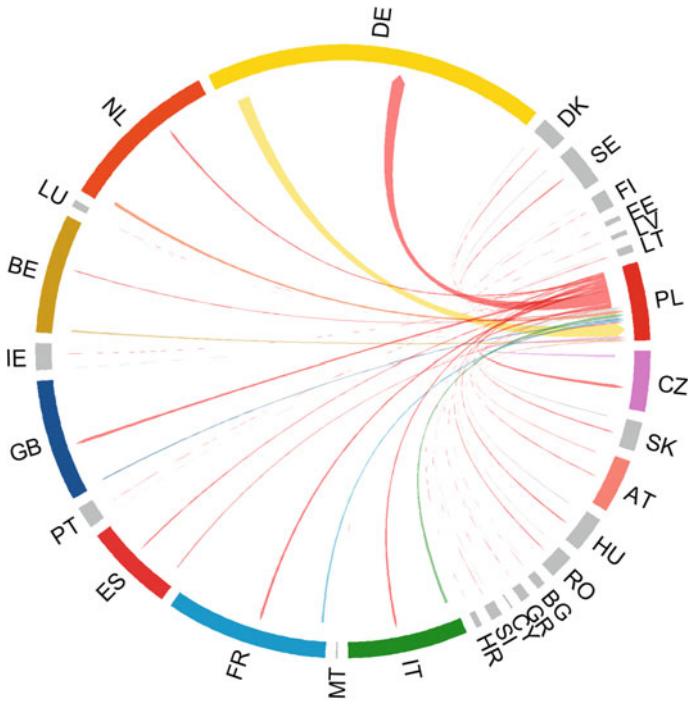


Fig. 22.2 Poland's ICS and ICA as a part of overall EU intra-Community trade in goods (2017 data)

imports combined) in value terms in 2017. The ten countries with the highest share in intra-Community trade in goods are marked in colours. The high position of the Netherlands and Belgium is, of course, due to the fact that they are often indicated as an importing country when imports from third countries have taken place, and the status of EU goods has been acquired in Dutch or Belgian ports (mainly Rotterdam and Antwerp). Poland ranks eighth in terms of intra-Community trade turnover. Polish trade (both exports under ICS and imports under ICA) is marked with arrows. The thickness of the arrows corresponds to the value of turnover. Particularly noteworthy is the trade in goods with Germany, Poland's largest trading partner.

The position of individual countries in intra-Community trade of Polish entities is presented in Fig. 22.3. Now, only Polish exports within the ICS are presented, while the proportions of the figure correspond to the share of the value of exports to individual countries. The ten countries to which the combined value of intra-Community supplies from Poland is the highest were marked with colours. This list does not include Belgium and Austria, but there are Sweden, Slovakia and Hungary instead, i.e. neighbouring countries, with which Polish companies trade relatively more often.

In the study, aggregate data on intra-Community supplies and acquisitions for years 2005, 2008, 2011, 2014 and 2017 have been utilised. The survey period was

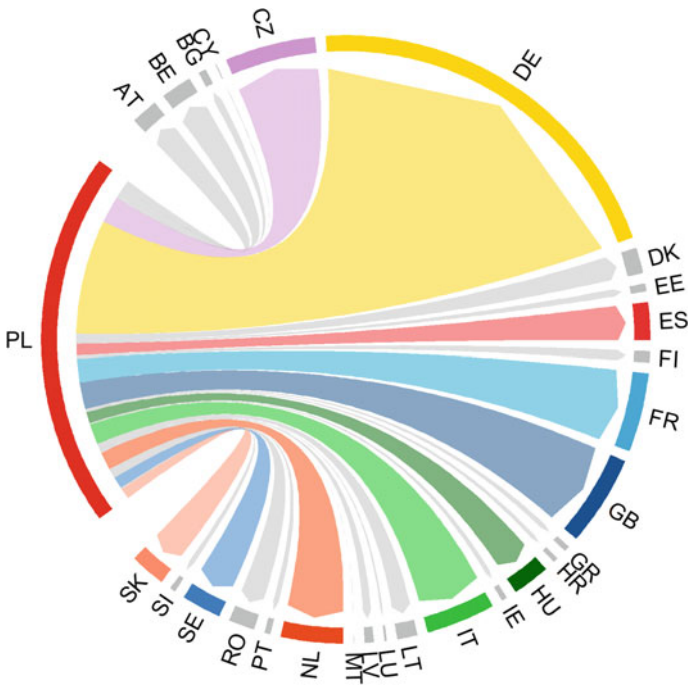


Fig. 22.3 Distribution of ICS from Poland among EU Member States (2017 data)

selected in such a way as to include data for full years and the entire period of Poland's membership in the EU, thus the first audited year was 2005 (the first full year of membership). The period closes in 2017, the last year for which full, verified data were available. We deliberately resigned from the survey covering each year in the period under consideration. Such a data range, according to the authors, provides a proper representation of the studied phenomena and at the same time reduces information noise to acceptable sizes. The data quality survey covered several stages, as listed below:

1. Comparison of the values of ICS reported in Poland to individual EU Member States with the values of ICA reported in each of the countries as originated in Poland.
2. Calculation of aggregated data quality indices for each trading partner country and for the EU as a whole.
3. On the basis of the calculated indices, countries (export directions) were ranked from the highest to the lowest quality of data.
4. The above points have been repeated for each of the years under consideration, the results obtained have been compiled and changes in the quality of the data over time have been examined.

22.4 Data and Methodology

Official statistics on intra-Community trade in goods depend on data gathered within the national Intrastat systems. The database consists of declarations made by businesses dispatching goods to other EU Member States or receiving goods from them. Due to the specifics of data, it is possible to assess its quality. Mirror data, i.e. data gathered by both sides of a transaction, should be reported twice—by the country of the dispatcher and at the same time by the country of the receiver of traded goods. Thus, for every pair of countries, the Intrastat system should contain the same values (and quantities) of trade in goods reported twice for the same period. Slight differences are permitted for various reasons, mostly due to different statistical thresholds used in individual countries. Generally, however, discrepancies in mirror data are due to errors in reporting or concealment of transactions. And this affects the quality of collected data.

The discrepancy between mirror data on imports and exports can be measured in different ways. In the literature, there are approaches based on the determination of the difference between the two figures [6, 12, 17] or the relative difference. In the latter case, the denominator is either the total volume of exports or imports [10, 13], their sum or average [6, 8, 12] or difference between them [12]. Some researchers use the quotient of imports and exports [7, 17]. The discrepancies are aggregated by goods or by countries using their sum or sum of absolute values [7]. Ferrantino and Wang [8] propose that the aggregated absolute asymmetry values should be weighted by the shares of individual product groups in the total trade volume, but we consider this approach disputable.

We used Comext data on intra-Community supplies for 2005, 2008, 2011, 2014 and 2017 aggregated on country and chapter level. The goods listed in the database are grouped into 98 chapters within the Harmonized System (HS). To ensure data consistency, we omitted last two chapters ('complete industrial plant' and 'special codes'), thorough explanation of such an approach is given by Ferrantino and Wang [8]. The analysis started with calculating data quality indices for each EU-27 country being an export direction from Poland and each HS chapter.

The quality of data on ICS from Poland to country A combined by HS chapters was calculated with the use of aggregated index of data quality:

$$ {}_{Z}W_{E}^{\mathrm{PL},A} = \frac{\sum_{i=1}^{n} \left| E_{d_i}^{\mathrm{PL},A} - I_{d_i}^{A,\mathrm{PL}} \right|}{K} \tag{22.1} $$

where

$E_{d_i}^{\mathrm{PL},A}$ declared value of dispatches (supply) from Poland to country A by HS section d_i,

$I_{d_i}^{A,\mathrm{PL}}$ declared value of acquisitions by country A delivered from Poland by HS section d_i (mirror data),

$K = \sum_{i=1}^{n} \frac{E_{d_i}^{\mathrm{PL},A} + I_{d_i}^{A,\mathrm{PL}}}{2}$ hypothetical average turnover,

$d_i = 1, \ldots, 76, 78, \ldots, 97$ HS section number.

The aggregate index takes values from the range from 0 to 2. The higher its value, the lower the quality of the analysed data.

On the basis of data quality indices, a ranking of the EU countries was established and compared. This made it possible to identify the directions of Poland's exports with the highest and lowest data quality.

Comparison of the quality of data on Polish ICS dispatched to the whole of the EU in the years under consideration was done using similar index with aggregation by countries and chapters of HS of the form

$$ {}_{Z}W_{E}^{\mathrm{PL,EU}} = \frac{\sum_{i=1,j=1}^{n} \left| E_{d_i}^{\mathrm{PL},A_j} - I_{d_i}^{A_j,\mathrm{PL}} \right|}{K} \tag{22.2} $$

where

$A_j = 1, \ldots, 27$ EU Member State, other symbols as in the Formula (22.1).

22.5 Quality of Data on Poland's Intra-Community Supplies by EU Member States—Research Results

The ranking of EU countries according to the quality of data on intra-Community trade and its changes over time are presented in Fig. 22.4 (made in *R* using the `slopegraph` library [16]).

Figure 22.4 shows groups of countries characterised by specific changes of position in such a ranking over time. The countries with the highest increase in the ranking (data quality improved significantly compared to other countries) are shown in green. The red colour indicates countries where the quality of data has deteriorated significantly compared to other countries (the most significant drops in the ranking). The blue colour indicates the countries for which the greatest fluctuations were observed during the period considered, but the total change of position between 2005 and 2017 was not as huge. Italy, the Czech Republic, Romania and Croatia are the countries with the best quality data on trade with Poland, but they are the countries with the lowest quality of mirror data that should be paid special attention in order to ensure a lasting increase in the quality of data in the whole set. This group of countries includes Cyprus and Malta as well as Ireland and Luxembourg.

The Netherlands, Spain and Germany were the countries that made the biggest leaps in the quality of data on Polish ICS. Note that Germany is also Poland's largest trading partner, so improving the quality of data on trade with this country is particularly important. In absolute figures, this improvement means a reduction in the mirror data asymmetry (aggregated at HS chapter level) from 4.64 EUR billion in 2005 to 1.33 EUR billion in 2017, i.e. by 71.4%, which should be considered as a

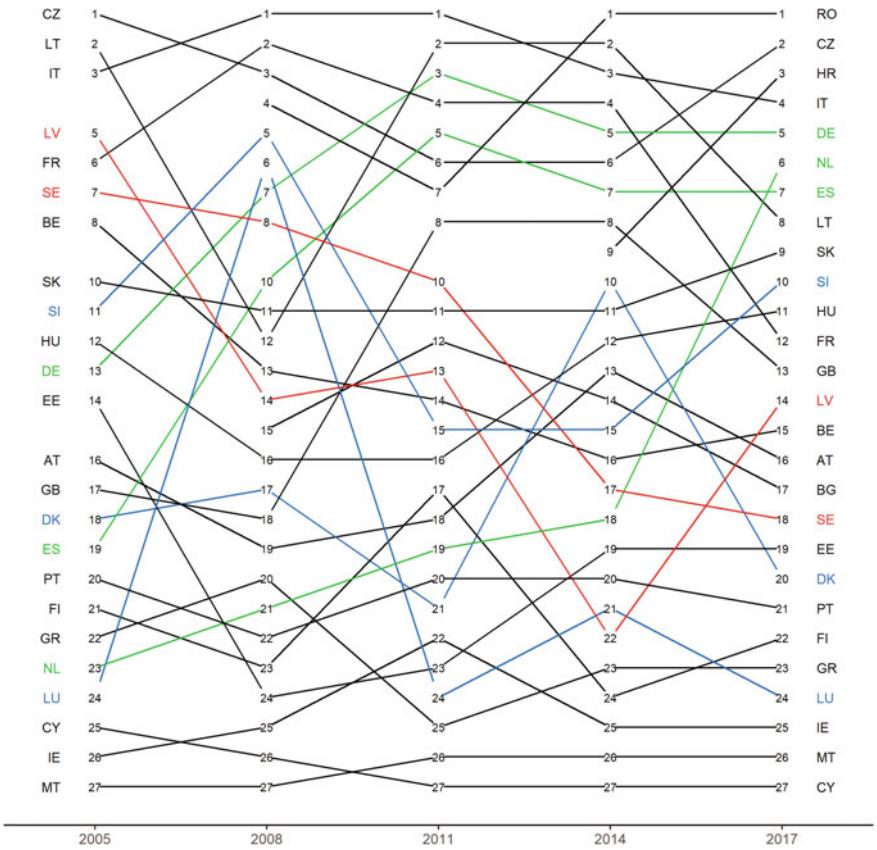


Fig. 22.4 Ranking of EU Member States according to quality of data on ICS and its changes over the period of 2005–2017

satisfactory result, all the more so because at the same time the total value (according to Polish declarations) increased from 20.3 EUR billion to 57.0 EUR billion, i.e. by 181.2%. Latvia and Sweden are of particular concern among the countries for which we record data of ever poorer quality. The countries with the highest fluctuations in the ranking in the period under consideration are Luxembourg, Denmark and Slovenia.

The changes in the ranking position are partly due to the fact that the values of indices on which the ranking is based are similar for many countries and are subject to significant fluctuations in the analysed period. Mirror data from Intrastat declarations for most countries differ relatively little, with median value of the index close to 0.2 while third quartile of the measure decreased from around 0.4–0.29. This means that in 2017, the aggregate differences between collected data on ICS and mirror ICA did not exceed 30% of the hypothetical average turnover for 21 countries, and for 12 countries, the upper limit is 20% of the hypothetical average turnover.

Figure 22.5 shows the values of the individual data quality indices for the selected countries and for the EU as a whole. The chart shows the four countries that ranked the highest in the whole period—Italy, Romania, the Czech Republic and additionally Germany, as well as the three countries with the lowest data quality, which were ranked lowest in the created rankings: Cyprus, Malta and Ireland. In addition, the level of discrepancy established for the EU as a whole is also shown, calculated with (22.2). The nature of the procedure for aggregation of discrepancies means that the value of index for EU is not close to the average level in the whole group but rather to the level of the indices for countries with the highest quality of data.

Table 22.1 shows the ICS indices for 2017. It is primarily possible to indicate a group of countries with data quality different from the others—Malta and Cyprus—and two other countries with data quality significantly lower than the others—Ireland and Luxembourg. However, it is not the case that trade in goods with small countries results in lower data quality. A counterevidence to such a statement is good quality data on Poland's trade with Croatia or Lithuania.

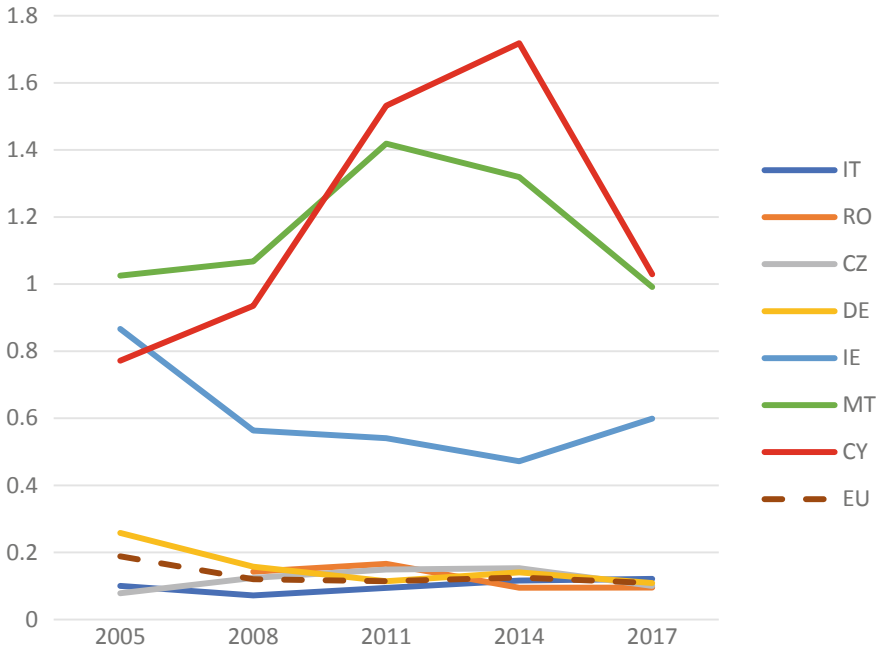


Fig. 22.5 Values of data quality indices for selected countries (four with the best data quality, three with the worst data quality) and for the whole of EU aggregate

Table 22.1 Values of individual countries' quality indices of Comext data on ICS

Country	Index	Country	Index	Country	Index
RO	0.0951	SI	0.1564	EE	0.2472
CZ	0.1004	HU	0.1583	DK	0.2836
HR	0.1191	FR	0.1663	PT	0.2916
IT	0.1209	GB	0.2073	FI	0.3227
DE	0.1229	LV	0.2180	GR	0.3592
NL	0.1407	BE	0.2220	LU	0.5311
ES	0.1468	AT	0.2279	IE	0.5989
LT	0.1477	BG	0.2282	MT	0.9916
SK	0.1526	SE	0.2339	CY	1.0295

22.6 Conclusions

Poland ranks eighth in 28 countries in terms of intra-Community trade in goods turnover. The quality of data on Polish trade—important from the point of view of Polish users—therefore, makes a significant contribution to the quality of aggregate data for the entire EU.

The quality of data on Polish trade in goods with other Member States is high. In particular, it is worth noting that the quality of data on the trade with the largest partner, Germany rose. While the total turnover almost tripled, the size of the discrepancy (absolute difference) decreased by more than 70%, which resulted in a significant decrease (improvement) in the value of the data quality index. Unfortunately, as for the data on trade with several countries—Cyprus, Malta and to a lesser extent Luxembourg and Ireland—the observed quality is low and is not improving. A relatively small turnover as well as varying reporting regimes are among the main reasons.

Poor data quality in trade with some partner countries may also be related to deliberate, unwanted actions of Intrastat reporting entities. Such a hypothesis could be confirmed by carrying out extended analyses at the level of individual transactions. Such a study is currently not possible because it would require a comparison of individual data on transactions from two parties (such data are subject to statistical and/or fiscal confidentiality rules, and their anonymisation would not allow to link mirror data pairs). It is only possible to indicate which goods accounted for the largest part of the generation of asymmetries and to determine this at the level of individual classification codes. Such research is being carried out by the authors of this study and will be continued.

References

1. Ainsworth, R.T.: The Morphing of MTIC Fraud: VAT Fraud Infects Tradable CO_2 Permits. Boston University School of Law Working Paper No. 09-35 (2009)
2. Baran, P., Markowicz, I.: Analysis of intra-community supply of goods shipped from Poland. In: Papież, M., Śmiech, S. (eds.) The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings, Zakopane. Socio-Economic Modelling and Forecasting, vol. 1, pp. 12–21 (2018a)
3. Baran, P., Markowicz, I.: Behavioral economics and rationality of certain economic activities. The case of intra-community supplies. In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics. Proceedings of Computational Methods in Experimental Economics (CMEE) 2017 Conference, pp 285–299. Springer, Cham (2018b)
4. Borselli, F.: Pragmatic policies to tackle VAT fraud in the European union. Int. Vat Monit. **5**(19), 333–342 (2008)
5. European Court of Auditors: Tackling intra-community VAT fraud: more action needed. Special Report, No. 24. Publications Office of the European Union, Luxembourg (2016)
6. Eurostat: Compilers guide on European statistics on international trade in goods, 2017 edition. Manuals and guidelines. Publications Office of the European Union, Luxembourg (2017)
7. Federico, G., Tena, A.: On the accuracy of foreign trade statistics (1909–l935). Morgenstern Revisited. Explor. Econ. Hist. **28**(3), 259–273 (1991)
8. Ferrantino, M.J., Wang, Z.: Accounting for discrepancies in bilateral trade: the case of China, Hong Kong, and The United States. China Econ. Rev. **19**(3), 502–520 (2008)
9. Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B.: Circlize implements and enhances circular visualization in R. Bioinformatics **30**(19), 2811–2812 (2014)
10. Guo, D.: Mirror Statistics of International Trade in Manufacturing Goods: The Case of China. UNIDO. Research and Statistics Branch Working Paper 19/2009 (2009)
11. Hamanaka, S.: Whose trade statistics are correct? Multiple mirror comparison techniques: a test of Cambodia. J. Econ. Policy Reform **15**(1), 33–56 (2012)
12. HMRC Trade Statistics: A Reconciliation of Asymmetries in Trade-in-Goods Statistics. Published by the UK and Other European Union Member States. Southend-on-Sea (2014)
13. Javorsek, M.: Asymmetries in International Merchandise Trade Statistics. UN ESCAP Statistics Division Working Paper 02 (2016)
14. Keen, M., Smith, S.: VAT Fraud and Evasion: What Do We Know, and What Can Be Done? IMF Working Papers, vol. 31, issues 7, pp. 1–33 (2007)
15. Kim, S.Y.: An Analytical Approach to Detect MTIC/MTEC VAT Fraud Involving Energy Trading. IOTA Papers (2017)
16. Leeper, T.J.: Slopegraph: Edward Tufte-inspired Slopegraphs. R package version 0.1.14 (2017)
17. Parniczky, G.: On the inconsistency of world trade statistics. Int. Stat. Rev. **48**(1), 43–48 (1980)
18. Pope, S., Stone, R.: Civil interventions for tackling MTIC fraud: a UK perspective. Amicus Curiae **80**, 8–13 (2009)
19. Regulation (EC) No 638/2004 of the European Parliament and of the Council of 31 March 2004 on Community statistics relating to the trading of goods between Member States. https://eur-lex.europa.eu (5.06.2018)
20. Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics. https://eur-lex.europa.eu (5.06.2018)
21. Tsigas, M.E., Hertel, T.W., Binkley, J.K.: Estimates of systematic reporting biases in trade statistics. Econ. Syst. Res. **4**(4), 297–310 (1992)

Chapter 23
Measuring the Provision of Access to Sustainable Energy in the EU as a New Task of Modern Economics in Line with Agenda 2030

KRYK Barbara

Abstract The purpose of the chapter is to measure and assess the level of sustainable energy provision in EU countries in the context of Goal 7 of the Agenda 2030. Ensuring access to sustainable energy for people and increasing energy efficiency in all sectors of the economy is one of the most important goals both in the European Union (EU) and globally. It was recognized that access to modern energy services at affordable prices is necessary to achieve the goals of the concept of sustainable development, Agenda 2030, Europe 2020 strategy, and the European energy policy. Precise evaluation of achievements in this area (using modern quantitative methods) is a relatively new issue, both in economic practice and in modern economics; therefore, there is a need to develop ways to measure access to sustainable energy, other than simply analyzing the time series of data in individual years on which the previous research was based. Hence, an attempt was made to use the taxonomic measure of development, to assess the achievements of individual EU countries, and to organize them according to the degree of sustainable energy provision, which is a novelty in this respect and includes the article to the research trend oriented toward the operationalization of the concept of sustainable development. The research used eight core indicators for Sustainable Development Goals (SDGs), seven reported by Eurostat and set by the UN. The research period is 2010 and 2016. The conducted research allowed to formulate conclusions on the current achievements of the member states and the EU as a whole with respect to the achievement of Goal 7 of the Agenda 2030, as well as formulating recommendations for necessary actions.

Keywords Sustainable energy · Agenda 2030 · EU · Modern economy · Taxonomic measure of development

KRYK Barbara (✉)
University of Szczecin, Szczecin, Poland
e-mail: Barbara.Kryk@usz.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_23

23.1 Introduction

In 2010, United Nations in the report "The Global Partnership for Development: Time to Deliver" [28] for the first time drew attention to the difficulties in supplying energy to a significant part of the Earth's population, stressing that providing access to modern energy services at affordable prices in developing countries is necessary to achieve the Millennium Development Goals and the goals of the concept of sustainable development. From that moment, various initiatives were started to implement the goals set by the UN in this regard. The formalization of the goals set at that time took place by making them obligatory tasks,[1] included in Goal 7 of the Agenda 2030 and adopted at the UN General Assembly in September 2015 [29]. The Agenda has 17 goals (quantitative and qualitative) and 169 tasks for the next 15 years, which—together with the Paris climate agreement, known as COP21[2] [4] and the Addis Abeba Action Program [27]—create a global framework for international cooperation in the field of sustainable development, in all its dimensions (environmental, economic, social, and administrative/institutional), and also help guide the convergence process of EU member states, societies, and the rest of the world [18].

The Goal 7 is one of the most important goals of sustainable development,[3] because it determines the achievement of several other goals. To better monitor its implementation, the UN has established indicators that are to be used to determine the level of sustainability in access to energy as well as the development itself. This is a kind of novelty in enforcing the goals set in Agenda 2030. At the same time, the designated indicators form the basis for the search/application of more advanced methods of monitoring the performance of tasks by individual countries/groups of countries, than the usual statistical analysis, which until now was the norm in this respect. Therefore, an attempt was made to use the taxonomic measure of development to achieve the purpose of the article, which is to measure and assess the level of sustainable energy supply in the European Union (EU) countries in the context of Goal 7 of the Agenda 2030. The use of modern econometric methods to measure the provision of sustainable energy also meets the expectations of modern economics, as it contributes not only to broadening the spectrum of analyzed issues and ways of operationalizing sustainable development, but also increasing the accuracy of calculations and formulating conclusions. So we can say that the article is complementary to the research gap that has been spotted. The need of using advanced econometric methods for the operationalization of sustainable development was also mentioned by such researchers as, among others, Radovanovic et al. [22], Mardani and others [20], Kumar and others [19].

The research covers EU Member States, while its subject is the degree of sustainable energy supply, analyzed on the basis of Eurostat data on 8 core indicators for

[1] This was due to their non-fulfillment by 2015.

[2] Paris Agreement—United Nations Framework Convention on Climate Change.

[3] Goal 7—"Provide everyone with access to affordable and clean/sustainable and modern energy at an affordable price."

SDG 7 of Agenda 2030, which is the reference point in the article. The beginning of the research period is 2010, when the discussion and actions aimed at increasing the access of societies to sustainable energy were initiated. The end of the research period (2016) is determined by available data. The description of the research was preceded by a synthetic presentation of the EU's commitment to the achievement of Goal 7 of the Agenda 2030. The conducted research enabled assessment of the achievements of individual EU countries and their ordering by to the level of sustainable energy provision as well as the formulation of conclusions.

23.2 Outline of the Issues

The European Union is known for its high commitment and experience in the implementation of sustainable development, which is anchored in the highest ranking documents, including Treaty on European Union [3], strategy for sustainable development ([5], 400 final), "Europe 2020" strategy [11], and others. EU has also played an important role in shaping Agenda 2030, which is fully in line with the European vision and is a global action program for global sustainable development based on the objectives of this concept. The EU was one of the leading forces that led to the adoption of Agenda 2030. It is fully committed to its implementation and determined to act as a precursor in the implementation of this program and the objectives of sustainable development in cooperation with the member states in accordance with the principle of subsidiarity. That is why the EU adopted in 2016 the action program for sustainable development entitled *Next steps for a sustainable European future European action for sustainability* [6]. It is implemented in two ways. The first direction of works is to completely integrate the goals of sustainable development within European policy with the priorities of the European Commission. The second direction is to develop a long-term vision and the main points of the sectoral policy after 2020 as part of the preparation for the long-term implementation of the sustainable development goals. With regard to this work, a long-term vision was created to establish a prosperous, modern, competitive, and climate neutral EU economy by 2050 [8]. This vision paves the way for a structural change in the European economy, stimulating sustainable growth and job creation. On the other hand, with regard to the first direction, the identification has shown that under the current EU policies, all 17 sustainable development objectives are being implemented. Moreover, they are included in all 10 of the Commission's priorities for 2015–2019.[4] An important role in the process of implementing a number of sustainable development goals is played by the "Europe 2020" strategy and its assumptions regarding sustainable energy and climate to be implemented by 2020: increase of energy efficiency by 20%, increase

[4]These are: (1) employment, economic growth, and investments; (2) the single digital market; (3) energy union and climate; (4) internal market; (5) a deepened and fairer economic and monetary union; (6) a balanced and progressive trade policy to exploit the opportunities of globalization; (7) justice and fundamental rights; (8) migration; (9) stronger position in the international arena; (10) democratic changes.

of up to 20% share of energy from renewable sources in total energy consumption, reduction of greenhouse gas emissions to 20% compared to 1990. In the document *Towards a Sustainable Europe by 2030* [7], this strategy was pointed out as a way for achieving, among others, Goal 7 of the Agenda 2030. Namely,

- The process of **separating economic growth** from energy inputs and related **greenhouse gas emissions**. In 1990–2017, greenhouse gas emissions fell by 22%, while GDP increased by 58%. Since 2000, the energy productivity and the intensity of greenhouse gas emissions have been improving almost uninterruptedly in the EU.
- The EU seeks to achieve its target value for **energy efficiency,** set at 20% by 2020. In 2005–2016, primary energy consumption in the EU decreased by 9.9% and final energy consumption by 7.1%.
- The EU is on track to achieve the 2020 target of a 20% share **of renewable energy** in final energy consumption. In the last decade, the use of renewable energy has grown steadily in the EU—from 9 to 17% of gross final energy consumption (in 2005–2016). The main driver of this growth was the predictable EU regulatory framework, more efficient technologies, falling costs of renewable energy technologies and more market-oriented support.
- The EU continues to meet its demand for energy mainly by **importing fuels from non-EU countries**. The EU dependence on import is 53.6% and has practically remained unchanged between 2006 and 2016, while energy production has decreased by 14% during this period. In the same period, there was a constant decrease in primary energy consumption—by about 10%.
- The EU has made progress in increasing **access to affordable energy**. In recent years, the percentage of households that cannot afford to heat the house to the right temperature has decreased. In 2017, 8.1% of the EU population complained about the lack of access to affordable energy—which is less by 2.8% points than in 2007.

The presented trends, based on the analysis of data in the form of time series, suggest that changes in the provision of sustainable energy are satisfactory. Meanwhile, research carried out using a taxonomic measure of development shows less satisfactory results, as presented later in the chapter.

23.3 Research Methods

Observation of the course of implementation of tasks aimed at achieving the sustainable development goals included in the Agenda 2030, including ensuring access to sustainable energy, is based on the most common tool, namely indicators. According to Borys [2], the indicator refers to a certain state of the phenomenon. The most important feature of each indicator is the comparability of its value, which makes

it possible to position a given object (e.g., a country) against others.[5] Sustainable development indicators are tools whose task is to monitor the implemented changes and provide information on the level of implementation of tasks in the field of sustainability and the current state of the environment. Using indicator as an information medium has become quite popular due to its multidimensionality, accessibility, readability, unambiguity, and comparability. Indicators assigned to particular governance (economic, environmental, social) or to sustainable development objectives create the possibility of constructing economic models, forecasting, monitoring/evaluation of change implementation. The establishment of indicators by the United Nations to achieve Goal 7 of the Agenda 2030 has created the possibility of using econometric methods to measure its implementation—related to the use of methods from the multidimensional comparative analysis of MCA (e.g., Hellwig [15], Shen and Tzeng [23], GDM [16, 30], TMAI [26]). Hence, this article uses a taxonomic measure of development that allows the assessment and arrangement of EU countries according to the degree of sustainable energy provision (distance to the target) and a comparison in 2010 and 2016.

The task of taxonomic analysis (the method of organizing a linear set of objects) is the ordering, that is, determining the order of objects according to a given criterion.[6] A synthetic measure of development is the tool of ordering methods. The use of a taxonomic measure of development makes it possible to classify objects described by many features using one synthetic variable.[7] There are many ways to construct this measure. Generally, they are divided into the model ones and the model-free ones. The study used the Hellwig model formula [1]. The Hellwig method is based on standardized values of diagnostic features. The distance of each test object from the established development pattern is determined by the formula:

$$d_i = \left[\sum_{k=1}^{K} (z_{ik} - z_{0k})^2 \right]^{\frac{1}{2}},$$

where

d_i the distance of the object and from the pattern,
i number of objects (countries), $i = 1, 2, \ldots, N$,
k number of features (variables), $k = 1, 2, \ldots, K$,
S_k standard deviation
Z_{ik} standardized values of features,

[5]The indicator expressed as an absolute or relative number (presented as a percentage ratio of the considered values to the adopted basis), expressing the level of a given phenomenon, is the most readable form of statistical description, because it is a simplified information referring to the description of a complex phenomenon.

[6]The multidimensional comparative (taxonomic) analysis was created by Hellwig [15].

[7]Taxonomic development measures replace the description of the studied objects with the help of a number of features, with a single aggregate size [14, 24].

$$z_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k}$$

z_{0k} development model, taking the maximum value of a feature for the examined objects if the feature is a stimulant or a minimal if it is a destimulant.

The distance d_i can be used to compare the level of development of the examined objects; the smaller the value the higher the level of development has reached the given object. In order to normalize the value of the measure, a relative taxonomic measure of development was established, also called a taxonomic measure of investment attractiveness (TMAI) [21, 25], using the formula:

$$z_i = 1 - \frac{d_i}{d_0},$$

where

z_i taxonomic measure of development for the object, and
d_0 a norm that ensures that variable z_i will take values from the range [0, 1],

$$d_0 = \bar{d} + aS_d$$

where

$\bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$, $S_d = \left[\frac{1}{N}\sum_{i=1}^{N}(d_i - \bar{d})^2\right]^{\frac{1}{2}}$ the a value is assumed as 2 or maximum d_i,

S_d standard deviation of d_i.

The z_i measure is normalized and takes values from the range [0, 1]. The higher its value, the better the position of the object in the set (higher level of the examined objects). Therefore, the closer its value is to 1, the greater the degree of access to sustainable energy by EU countries, as well as the achievement of Goal 7 of the Agenda 2030. On the basis of z_i values, a ranking of EU countries in 2010 and 2016 was established, which made it possible to assess the degree of sustainable energy provision by these countries and to indicate whether the desired changes take place—progress in the implementation of the studied goal.

23.4 Results

A total of 8 indicators were established to measure the achievement of Goal 7 in the UN Agenda 2030. They were adopted in this article as variables in respect of which 28 EU countries were examined (Table 23.1). Variables 4 and 5 were considered stimulants (the higher their value the better), while the other variables are destimulants (the lower their value the better).

Table 23.1 Set of variables

Designation of variable	Name of variable
x_1	Primary energy consumption (million tonnes of oil equivalent—TOE)
x_2	Final energy consumption (million tonnes of oil equivalent—TOE)
x_3	Final energy consumption in households per capita (kg of oil equivalent)
x_4	Energy productivity (Euro per kilogram of oil equivalent—KGOE)
x_5	Share of renewable energy in gross final energy consumption by sector (%)
x_6	Energy dependence by-product (% of imports in total energy consumption)
x_7	Population unable to keep home adequately warm by poverty status (% of the population)
x_8	Greenhouse gas emissions intensity of energy consumption year 2000 = 100

Source The author's own development based on Eurostat

After standardizing the values of the features (variables) and determining the distance of each study object from the established standard, a relative taxonomic measure of development z_i (in accordance with the previously given formula) for 2010 and 2016 was calculated, reflecting the level of sustainable energy provision. Then, the examined countries were sorted according to the value of this measure, indicating the change in the place in the ranking (Table 23.2). The arrow ↑ indicates the improvement of the place in the ranking, and the arrow ↓—deterioration, constant means maintaining the position.

In 2010, the average value of the measure z_i for the EU was 0.2829 and decreased to 0.2735 in 2016, which means—albeit slight but still—the deterioration of the level of sustainable energy provision at the community level. This may have been caused by the adoption of the poorest countries in the group during the analyzed period and the continuation of the global economic crisis until 2014. It should be emphasized that the average relative taxonomic development measure for the EU reflects a slightly more pessimistic trend of changes in this area than the traditional analysis of changes in the time series of individual indicators used to measure the achievement of Goal 7 of the Agenda 2030. The traditional analysis indicated an improvement in the level of all indicators, thus increasing the level of achieving the goal [13, 17], and meanwhile, the measure z_i shows that it is different. This confirms the necessity of using modern methods of measuring the implementation of sustainable development goals in order to be able to precisely determine the actual effects.

In 2010, in all EU countries, the height of the z_i measure was far from the borderline, ideal value of 1, which would mean total energy balancing. The highest measure was reported for Denmark and the lowest for Bulgaria. In fifteen surveyed countries, the indicator was higher than the EU average (Denmark, Sweden, Austria, Romania, Portugal, Croatia, Slovenia, Latvia, Greece, Spain, Estonia, Czechia, Hungary,

Table 23.2 Ranking of countries according to the value of the z_i measure

2010			2016		
Country	z_i	Position in ranking	Country	z_i	Position in ranking
Denmark	0.5002	1	Denmark	0.5192	1 const.
Sweden	0.4656	2	Sweden	0.4900	2 const.
Austria	0.4476	3	Austria	0.3866	3 const.
Romania	0.4288	4	Romania	0.3740	4 const.
Portugal	0.3967	5	Latvia	0.3675	5 ↑
Croatia	0.3962	6	Portugal	0.3647	6 ↓
Slovenia	0.3823	7	Ireland	0.3397	7 ↑
Latvia	0.3662	8	Malta	0.3358	8 ↑
Greece	0.3511	9	Slovenia	0.3265	9 ↓
Spain	0.3446	10	Croatia	0.3264	10 ↓
Estonia	0.3125	11	Slovakia	0.3082	11 ↑
Czechia	0.3108	12	Finland	0.3069	12 ↑
Hungary	0.3053	13	Czechia	0.2950	13 ↓
Netherlands	0.3016	14	Spain	0.2947	14 ↓
Slovakia	0.2956	15	Hungary	0.2773	15 ↓
Poland	0.2811	16	Greece	0.2771	16 ↓
Finland	0.2777	17	Estonia	0.2618	17 ↓
Ireland	0.2725	18	UK	0.2525	18 ↑
Italy	0.2421	19	Italy	0.2451	19 const.
Malta	0.2261	20	Poland	0.2382	20 ↓
UK	0.2036	21	Netherlands	0.2248	21 ↓
France	0.2013	22	France	0.2234	22 const.
Belgium	0.1962	23	Belgium	0.2133	23 const.
Cyprus	0.1722	24	Luxembourg	0.1259	24 ↑
Lithuania	0.1026	25	Cyprus	0.1115	25 ↓
Luxembourg	0.0985	26	Lithuania	0.1082	26 ↓
Germany	0.0413	27	Bulgaria	0.0624	27 ↑
Bulgaria	0.0000	28	Germany	0.0000	28 ↓

Source The author's own study

Netherlands, Slovakia), in the remaining thirteen—lower (Poland, Finland, Ireland, Italy, Malta, UK, France, Belgium, Cyprus, Lithuania, Luxembourg, Germany, Bulgaria). Among the latter, there are even the countries that were set as a model for the development of sustainable energy (e.g., Germany, France), and the results show something different. The difference in the level of measures between Denmark and countries below the EU average ranged from around 2–12 times, suggesting a fairly large variation in the provision of sustainable energy within the community itself.

In 2016, in all EU countries—as in 2010—the height of the z_i measure was far from the borderline, the ideal value of 1. At the same time, Denmark had the highest measure and Germany the lowest. The number of countries in which the z_i measure was higher than the EU average increased to sixteen (Denmark, Sweden, Austria, Romania, Latvia, Portugal, Ireland, Malta, Slovenia, Croatia, Slovakia, Finland, Czechia, Spain, Hungary, Greece), up to twelve has decreased the number of countries with a measure below the average (Estonia, UK, Italy, Poland, Netherlands, France, Belgium, Luxembourg, Cyprus, Lithuania, Bulgaria, Germany). Moreover, there was a change of places in the rankings. Seven countries have kept their place (Denmark, Sweden, Austria, Romania, Italy, France, Belgium), eight have improved their position (Latvia, Ireland, Malta, Slovakia, Finland, UK, Luxembourg, Bulgaria), and 13 have worsened (Portugal, Slovenia, Croatia, Czechia, Spain, Hungary, Greece, Estonia, Poland, Netherlands, Cyprus, Lithuania, Germany). The reason for the decline to lower places in the ranking was a decrease in the value of the z_i measure, caused by the change in conditions (including economic, social, legal) in individual countries. The value of this measure also decreased in two countries that maintained their position (Austria, Romania). Thus, the total number of countries in which the value of the measure decreased was greater than the number of countries with the reverse situation, which determined the reduction of the EU average measure, i.e., the reduction of the level of sustainable energy provision by the EU. This suggests a lack of progress in achieving Goal 7 of the Agenda 2030 on the EU scale. The only satisfactory thing is the reduction of difference in the amount of measures between Denmark (with the highest value of z_i) and countries below the EU average. The difference was at the level of just over 8 times, which is 4 less than in 2010, which indicates a reduction in the disproportion in providing sustainable energy in the community itself.

23.5 Conclusions

The research carried out shows that

- In 6 years, the EU, as an organization, has not made so much progress in ensuring sustainable energy, as suggested by the analysis of time series [17]. The relative taxonomic measure of development z_i for the Union was not only far from the borderline, ideal value of 1, but also almost twice lower than the z_i measure for the

country with its highest value (Denmark). This indicates the need to modify/take actions that will intensify the effects in this area.

- The EU countries are strongly different in terms of providing sustainable energy, as evidenced by differences in the amount of the z_i measure in the analyzed years. The decrease in these disparities is positive.
- The first two places in the ranking are occupied by countries that have for years been a model not only in terms of providing sustainable energy, but also other goals of sustainable development. That is why it is worth spreading their good practices/solutions in a given area.
- In five countries, which maintained their position in the ranking in 2016, the value of the z_i measure increased, which means increasing the degree of achievement of Goal 7 of the Agenda 2030, as well as the existence and creation of new, additional opportunities to increase energy sustainability compared to the currently known. This should encourage other countries to use solutions conducive to the implementation of the target and applied in countries with a higher level of sustainable energy provision.
- In 2016, in 15 countries, the z_i measure was lower than in 2010. The biggest change occurred in the case of Greece, Netherland (by 7 places down), and Estonia (by 6 places), and the smallest in the case of Germany, Lithuania, Portugal (by 1 place). The reasons for these adverse changes should be investigated in order to take action to support Goal 7 of the Agenda 2030.
- In 2016, 8 countries took higher places in the ranking, which means they increase their level of sustainable energy provision. The largest change occurred in the case of Malta (by 12 places) and Ireland (11 places), and the smallest in the case of Luxembourg (by 2 places) and Bulgaria (by 1 place).

In general, EU Member States are characterized by a low and medium degree of sustainable energy provision, which means still unsatisfactory results in the implementation of Goal 7 of the Agenda 2030. Despite some progress in the studied area, previous initiatives have not been able to ensure the full implementation of the UN-designated action program by 2030. Therefore, there was a need for more efficient implementation and further targeted actions in all areas. In the context of the above, the EU has taken new initiatives to facilitate the implementation of Goal 7 (COM (2019) 22 final). The first one is the *Energy Union Strategy* [10], which sets new EU targets for 2030[8]: an increase in energy efficiency of at least 32.5% with the possibility of upward revision in 2023, increasing to least 32% share of energy from renewable sources in total energy consumption. For the electricity market, the goal of interconnectors for 2030 was confirmed, assuming that at least 15% of the installed electricity production capacity can be sent across borders. These ambitious goals are intended to boost Europe's industrial competitiveness, stimulate economic growth

[8]The objectives set out in 2014 have been modified: A reduction of at least 40% of greenhouse gas emissions (compared to 1990 levels), at least 27% of final energy consumed, will come from renewable sources (a binding target at EU level), an indicative achievement the objective of increasing energy efficiency by at least 27% compared to the forecasted future energy consumption. This objective is to be revised in 2020 with a view to increasing the EU indicator to 30% [9].

and employment, reduce energy bills, help fight energy poverty, and improve air quality. The following initiatives complementing the above-mentioned one are also to support the achievement of the plans: The "Europe on the move" packages, the "Clean energy for all Europeans" package, Clean planet for all—a long term vision of creating a thriving, modern, competitive and climate neutral economy to 2050, the policy framework energyclimate and energy by the year 2030, a broad-based initiative "Horizon 2020" for the digital transformation in the energy sector, thanks to the Internet of things, a strategic energy technology plan, the European Alliance for for batteries, "Mission Innovation", support for the mining regions in the transitional period, the Observatory of Energy Poverty, the Initiative Clean Energy for islands in the EU Cohesion policy, a Plan of action for the sustainable finance, Low emission strategy, the European consensus on development, Strategy by providing energy services in Africa, European and global Covenant of mayors for climate and energy, an updated European neighbourhood policy and the EU's enlargement strategy, the Strategy for the Western Balkans, Global strategy for foreign and security policy of the European Union.

The full implementation of all these initiatives is to lead not only to stricter emission reductions across the EU than it was expected—around 45% by 2030 compared to 1990 (compared to the current reduction target of 40%)—but to achieve Goal 7 of the Agenda 2030.

Acknowledgements The project is financed within the framework of the program of the Minister of Science and Higher Education under the name "Regional Excellence Initiative" in the years 2019–2022, project number 001/RID/2018/19, the amount of financing PLN 10,684,000.00.

References

1. Bąk, I., Markowicz, I.: Taksonomiczne mierniki rozwoju w klasyfikacji obiektów (na przykładzie usług kulturalnych). Wiadomości Statystyczne **4**, 32–39 (2003)
2. Borys, T. (ed.): Wskaźniki rozwoju zrównoważonego, p. 62. Wydawnictwo Ekonomia i Środowisko, Warszawa-Białystok (2005)
3. Consolidated version of the Treaty on European Union: JoL C 326 of 26.10.2012 (2012)
4. Council of the EU: Council Decision 590 of 11 April 2016 on the Signing, on Behalf of the European Union, of the Paris Agreement Adopted Under the United Nations Framework Convention on Climate Chang, Brussels (2016)
5. European Commission: Mainstreaming Sustainable Development into EU Policies: 2009 Review of the European Union (2009) Strategy for Sustainable Development. COM (2009) 400 final, Brussels (2009)
6. European Commission: Next Steps for a Sustainable European Future European Action for Sustainability. COM (2016) 739 final, Brussels (2016)
7. European Commission: Towards a Sustainable Europe by 2030. COM (2019) 22 final. ANNEXES 1 to 3, Brussels (2019)
8. European Commission: A Clean Planet for all A European Strategic Long-term Vision for a Prosperous, Modern, Competitive and Climate Neutral Economy. COM 773 final, Brussels (2018)
9. European Commission: A Policy Framework for Climate and Energy in the Period from 2020 to 2030. COM 15 final, Brussels (2014)

10. European Commission: A Framework Strategy for a Resilient Energy Union with a Forward-Looking Climate Change Policy COM 80 final, Brussels (2015)
11. European Commission: Europe 2020: A Strategy for Smart, Sustainable and Inclusive Growth. COM 2020, Brussels (2010)
12. Eurostat. https://ec.europa.eu/eurostat/statistics-explained/index.php/SDG_7_-_Affordable_and_clean_energy_(statistical_annex)
13. Eurostat: Sustainable Development in the European Union, Monitoring Report On Progress Towards The SDGs In An Eu Context, p. 132. Publications Office of the European Union, Luxembourg (2018)
14. Grzegorczyk, E.: Taxonomic analysis of the level private equity venture capital market in Europe. In: Proceedings of IAC in Dresden. Czech Institute of Academic Education z.s., Praque. https://doi.org/10.20944/preprints201810.0749.v1 (2018)
15. Hellwig, Z.: Application of the taxonomic method to the countries typology according to their level of development and the structure of resources and qualified staff (in Polish). Przegląd Statystyczny **15**(4), 307–326 (1968)
16. Jajuga, K., Walesiak, M., Bąk, A.: On the general distance masure. In: Schwaiger, M., Opitz, O. (eds.) Exploratory Data Analysis in Empirical Research, pp. 104–109. Springer, Berlin-Heidelberg (2003)
17. Kryka, B.: Providing sustainable energy in Poland in comparison to the European Union in light of the seventh goal of the 2030 Agenda. Ekonomia i Środowisko **1**(68), 22–36 (2019)
18. Kryk, B.: Implementation of the concept of sustainable development in Poland in the context of goal 12 of the Agenda 2030. In: Conference Proceeding 5th International Conference SGEM 2018 SOCIAL SCIENCES & ARTS Scientifically. Economics and Tourism. Mod. Sci. **5**(1.4), 391–398 (2018)
19. Kumar, A., Sah, B., Singh, A.R., Deng, Y., He, X., Kumar, P., Bansal, R.C.: A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. Renew. Sustain. Energy Rev. **69**, 596–609 (2017). https://doi.org/10.1016/j.rser.2016.11.191
20. Mardani, A., Zavadskas, E.K., Khalifah, Z., Zakuan, N., Jusoh, A., Nor, K.M., Khoshnoudi, M.: A review of multi-criteria decision-making applications to solve energy management problems: two decades from 1995 to 2015. Renew. Sustain. Energy Rev. **71**, 216–256 (2017). https://doi.org/10.1016/j.rser.2016.12.053
21. Mastalerz-Kodzis, A., Pośpiech, E.: Zastosowanie wybranych elementów analizy fundamentalnej do wyznaczania portfeli optymalnych. Studia Ekonomiczne **146**, 68–78 (2013)
22. Radovanovic, M., Ivanivic, O.M., Teodorovic, N.: The review of the achieved degree of sustainable development in South Eastern Europe-the use of linear regression method. Renew. Sustain. Energy Rev. **15**(1), 766–772 (2011). https://doi.org/10.1016/j.rser.2010.07.064
23. Shen, K.-Y., Tzeng, G.-H.: Advances in multiple criteria decision making for sustainability: modeling and applications. Sustainability **10**, 1600 (2018). https://doi.org/10.3390/su10051600
24. Szopik-Depczyńska, K., Cheba, K., Bąk, I, Kiba-Janiak, M, Saniuk, S, Dembińska, I., Ioppolo, G.: The application of relative taxonomy to the study of disproportions in the area of sustainable development of the European Union. Land Use Policy 68 (2017). https://doi.org/10.1016/j.landusepol.2017.08.013
25. Tarczyński, W., Łuniewska, M.: Classic, fundamental and horizontally diversified portfolios—comparative analysis (in Polish). Acta Univ. Lodziensis. Folia Oeconomica **177**, 171–189 (2004)
26. Tarczyński, W.: A taxonomic measure of the attraction of investments in securities (in Polish). Przegląd Statystyczny **41**, 275–300 (1994)
27. United Nations: Addis Ababa Action Agenda of the Third International Conference on Financing for Development. Resolution adopted by the General Assembly A/RES/69/313 (2015)
28. United Nations: The Global Partnership for Development: Time to Deliver (MDG Gap Task Force Report). https://www.un.org/millenniumgoals/2011_Gap_Report/11-38394%20(E)%20MDG%20Gap%20Report%202011_WEB%20NEW.pdf (2011). Accessed 10 Apr 2019
29. United Nations: Transforming Our World: The 2030 Agenda for Sustainable Development. A/RES/70/1 (2015)
30. Walesiak, M.: Multivariate statistical analysis in marketing research (in Polish). Research Papers Wroclaw University of Economics 101(654), Wrocław (1993)

Chapter 24
Pilot Study as a Case of Applying an Experimental Approach to the Implementation of Tasks in the Public Sector

Beata Zofia Filipiak and Marek Dylewski

Abstract Experiments in the public sector are not widely used. However, as the practice shows, they are used to determine the expected utility of the proposed solutions or implemented changes, and the dominant method is pilotage. The study presents the theoretical approach to the experiment using pilotage, so as to show that the experiment gives the opportunity to indicate a "better" solution and provide justification for such a selection. In the chapter, the authors analysed the adopted assumptions of selected pilot implementations in the PSUs, the aim of which was to test and evaluate the solution to be used in practice as an expression of public participation. The analysis indicated that the visible use of piloting in the behaviour modelling of individuals included in the society is a good tool to maximise the usefulness of public tasks designed by LGUs.

Keywords Task of public sector · Public policies · Pilot program · Theory of expected utility

24.1 Introduction

The importance of economic experimentation in the practice of the public sector is still underestimated, which is caused by political conditions, but also by those resulting from the development of economic thought. The implementation of public tasks in public sector units (PSUs) is more and more often associated with the existing risk, and in particular with the risk of an optimal (effective) decision of public bodies on allocating public funds. The human behaviour in risk conditions is closely related

B. Z. Filipiak (✉)
Department of Finance, Faculty of Economics and Management, Institute of Finance, University of Szczecin, Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: beata.filipiak@usz.edu.pl

M. Dylewski
Faculty of Finance and Banking, Institute of Finance, WSB University in Poznań, Powstańców Wlkp. 5, 61-895 Poznań, Poland
e-mail: marek.dylewski@wsb.poznan.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_24

to the theory of expected utility formulated by von Neumann and Morgenstern [48]. They have hypothesised that individuals avoid risk and are guided by the rationality of choice, so they decide on options that maximise their expected value.

This is consistent with a rational decision-making model, i.e. determining the choices that are consistent and that maximise value within given limits [9]. In such situations, we are dealing with problems with a clear structure, known and easy to define, as well as with problems with an unclear structure and about which the information is ambiguous or incomplete [39].

In addition, it should be pointed out that the concept of economic rationality usually refers to decisions concerning choices, where apart from the decision-making process, there is also an element of creating alternatives that become an element of choice (the more we create them, the greater the chances of achieving decision-making success) [41]. Furthermore, given the Prospect Theory (TP) [24] and its normalisation, the Cumulated Prospect Theory (CPT) [25] should be emphasised that policy-makers are characterised by risk aversion, give a subjective weight values and probability and these weights determine differently depending on whether decisions are made in the face of losses or profits (public effects).

These problems require nonprogrammed decisions that craft novel solutions to meet the demands of the unique situation at hand. Managers, including those in the public sector, also differ in tendencies towards "systematic" and "intuitive" thinking during decision process. In systematic thinking, a person approaches problems in rational, step-by-step and analytical fashion. Managers who are intuitive can be expected to deal with many aspects of a problem or new task at once, jump quickly from one issue to another and consider "hunches" based on experience or spontaneous ideas [40].

When planning new tasks whose effects are not known (unstructured problems) and are associated with both high expenditure and the need for public acceptance, public bodies can use a pilot to:

- Determine whether the planned changes will lead to choices in which the public avoids risk and will be guided by the rationality of the choice of the planned public service in order to maximise the potential expected value. Which means that public decision-makers avoid the risk of unlikely losses and highly probable effects (according to the studies of: [4, 8, 16, 28, 32, 36]);
- Increase the likelihood of obtaining confirmation of the answer to the question about the assumed effects of changes in their economic policy. Which means eliminating the situation where decision-makers tend to risk highly likely losses and unlikely effects (according to research indicating the consequences of the choice made: [21, 31]);
- Initiate activities related to experimentation with new solutions or institutions [5, 14, 15, 33].

This approach assumes rationality in task design, but also an intuitive approach to expected results. For this reason, piloting and experimentation allow for verification of intuitive assumptions in terms of expected values.

A pilot study is a type of preliminary study and is intended to test and evaluate a solution to be used more widely and to a common extent in the practice of the PSUs. Generally speaking, the solution is tested by way of an experiment, in particular in regard to its assumed social usefulness (including individual), effectiveness, effects and information adequacy for the PSUs.

The paper will present a theoretical approach to the experiment with the use of piloting so as to confirm the thesis that it gives the opportunity to indicate a "better" solution from the point of view of allocation of public money resources (costs and effects) and to provide justification for such selection or to give up the initiative due to lower benefits than expected [17, 37].

24.2 Development of the Concept of Implementation of Public Tasks in Public Sector Units

The conditions under which PSUs pursue the objectives for which they have been established force them to enter into cooperative relations, but also into actions based on increased efficiency and savings of public funds. This approach is consistent with one of the principles of contemporary management of organisations, i.e. the principle of reducing the hierarchical distance between the place where the management problem has arisen and the place of its solution [35], as well as with the principle of obtaining the best results from given expenditures (Act of 2009, Article 44 [23]).

The literature on the subject indicates that PSU decision-making consists in choosing the way of achieving a public goal that we consider to be the best from the point of view of benefits and risk of choices [33]. The idea behind this approach is that in the decision-making process the measures that best lead to the desired goals are selected. Acting in accordance with the objectives-measures logic is the essence of Mr. Simon's concept of administrative rationality [42]. This approach overlaps with the concept of New Public Management (NPM).

Popular in recent decades, the NPM, derived from the theory of economics (e.g. theory of public choice and theory of transaction costs), is based on experience gained from management practices in the private sector. This approach focused on improving the performance of the public organisation and reducing costs, rather than on the effectiveness of the organisation in terms of its ability to meet social needs. In addition, the concept focuses on optimising the delivery of public services with the use of techniques from the private sector [1, 12].

On the wave of criticism of the NPM, but with regard to those features of the NPM that have a positive impact on the implementation of public policies or objectives, the concept of good governance has evolved. This particular approach is based on the pursuit of an effective, efficient and accountable process of managing public policies, which aims to strengthen the democratic system through, for instance, the participation of citizens. It is worth noting that on 25 July 2001, the European Parliament adopted a resolution to strengthen the democratic system through the participation

of citizens. The European Commission has adopted *The White Paper on European Governance* [10] public choice on the basis of co-decision, in particular the opinions of people, social organisations and public benefit organisations. Greater openness and accountability of all actors involved is proposed. Against this background, it is understood that good governance can achieve "(...) the institutional capacity to provide goods and services that are of direct interest to the public or their representatives in an effective, transparent, impartial and accountable manner, depending on the constraints of access to resources." [3].

This concept points to the basic principles of implementation of territorial self-government units' public objectives, which are: openness, participation, accountability, effectiveness and coherence [27, 29].

Searching for new solutions to increase effectiveness (of organisations, service provision or public funds expenditure) of the tasks performed, and in the final effect of the broadly understood public services, as well as to ensure consistency of undertaken actions, openness in the face of responsibility and with the use of participation—all this requires flexibility and often non-standard behaviour. On the one hand, partnership between the public and private sectors is applicable [13] and, on the other hand, the possibility of using information and telecommunication technologies is being sought both in the implementation and in improving the quality of the tasks performed [5]. The third direction is to use the results of pilot studies in the process of deciding how to carry out tasks with the involvement of public money resources [22].

24.3 Piloting as an Expression of Practical Application of the Expected Utility Theory in PSU Decisions

What plays an important role in the theory of decision-making are the theorems of rational decision-making, according to which people are guided by the principle of maximum benefit. This principle is related to maximising expected utility and was formulated by Nicholas Bernoulli in 1713. Numerous studies and experiments have shown that when making decisions, public authorities often violate the axioms of rational decision-making. The reasons for this state of affairs should be seen particularly in the impact of politics on the decision-making process, limited resources, especially the financial ones, objective changes in the socio-economic situation, or the lack of sufficient knowledge about the planned projects [20, 45, 46]. It should be stressed that if the preferences of an individual or the society are met by certain axioms, then their behaviour can be described as maximising its expected utility. Axioms are treated as conditions of rational behaviour, i.e. the normative principles of expected usability.

Kahneman and Tversky [26] in his study, considered criticism of the theory of expected usefulness, based on the results of conducted experiments found that the choice between risky variants leads to the choice of a variant with some effects, and

this behaviour is not consistent with the main principles theory of expected usability. Kahneman and Tversky noticed that people usually overstate the weight for a certain win compared to weights for probable results. This tendency is called the effect of certainty or the effect of certainty of Allais and causes risk aversion to appear in the election, where we have variants with some effects while the risk-taker's attitude appears in the elections related to losses [2, 34].

In modelling, the behaviour of individuals in a society which is driven by the goal of maximising usability (in public finance science it is understood as the personal sense of well-being of an individual being a member of a given community) the economists do not deny the fact that people care about their families, friends or communities. But public choice, similarly to the economic model of rational behaviour on which it is based, assumes that people are guided mainly by their own interests and, more importantly, that in the political process they behave no differently from people in the commercial goods market (Dobrowolski 2013) [18, 49]. This also has consequences in the decision-making process and in the perception of risk by public authorities (decision-makers).

It should be stressed that societies will always need public services realised as public tasks, however, the resources available to public entities to meet these needs are increasingly changing. Constitutional acts and subsequent normative acts, dedicated to specific types of these entities [6], impose a permanent obligation to provide services or provide public goods on the PSUs. The basis for providing such goods and services is a contractual relationship. This does not change the fact that the society views such contract from the perspective of the quality of public goods and services provided, and therefore from the point of view of meeting the needs reported by people.

The development of a new paradigm of management in public sector units (as indicated above), with emphasis on increasing the effectiveness of public spending, has also forced changes in the use of tools and instruments employed by public decision-makers. As the resources for the implementation of public tasks are limited, it is the decision-makers' perception of effects that enforces the often non-standard approaches. Moreover, the perception of risk becomes an important element because in a changing social and economic environment, with rapid advancement of technology, the people's needs are changing as well, which creates the risk to effective allocation of public funds.

Taking into account the fact that very often piloting is used for testing and launching social innovations, understood as the implementation of practices perceived as new in the public system [11, 44], not always is this risk (associated with the public authorities' decisions) taken into account [14]. When making their choices, public authorities must consider not only the risks (as the legislator requires an analysis of their impact and effects) but also the effects of their decisions on the society. This is why the experiment based on piloting is so important.

In its essence, piloting reduces decision-making risk and contributes to the rational use of public money, as it is carried out in order to test whether the solution proposed by public authorities will be carried out as intended and will allow to obtain specific data in order to make final decisions on the general accomplishment of the task at

hand. During a pilot study, one can examine everything from the degree of acceptance of a solution by the public to the quality of satisfying the expressed need by means of this solution, to the effects obtained in respect to the costs incurred to achieve them.

Being in line with the theory of expected utility, piloting not only takes into account the subjectivity of evaluation made by the PSUs but also gives the public a sense of shared decision-making and responsibility for the allocation of public money. Thus, one can say that it makes it possible to hold public authorities accountable for the decisions taken and to assess the results achieved efficiently, with due account of the citizens' participation. It should be remembered, however, that the concept of expected utility includes a subjective element, since a particular event may have a different expected utility for different investors. Thus, through the use of piloting, the PSU authorities may happen to erroneously assess the planned task (being a service or provision of a specific good) because after the task has been implemented in the whole PSU, the majority of the local community will not assess the project positively (for various reasons).

One of the factors affecting subjectivity in maximising the expected usability is the public approach to risk. The community may have a risk-neutral attitude, in which case it chooses to maximise the expected value. A formalised theory of making decisions in risk conditions, assuming that the public maximises the expected utility, was presented in 1944 by John von Neumann, a mathematician and Oskar Morgenstern, an economist (1958). In their theory, referred to as the theory of expected utility, they formulated axioms that must meet public preferences "in order for them to be considered comparable and to be described by means of indifference curves" [7]. According to this theory, these axioms, the fulfilment of which makes preferences rational, are as follows: comparability,[1] transitoriness,[2] independence,[3] stability[4] and continuity. Although, at first glance, they seem obvious and are met in all conditions by investors' preferences, for each of them there are practical examples where they cannot be met. The factors of particular importance are the public attitudes and a way in which public authorities make decisions.

[1]The Axiom of comparability means that a society unit can always determine its preferences for every public task.

[2]The Axiom of transitoriness means that a unit of society prefers A to task B and B to sentence C, and it also prefers sentence A to sentence C.

[3]The Axiom of independence (it was not literally formulated in the work of Neumann and Morgenstern) means that if the X and Y tasks are indifferent to the society unit, the choice should be indifferent between the implementation of the task:

- X with probability p and event V with probability $1 - p$;
- Y with probability p and event V with probability $1 - p$.

[4]The Axiom of stability means that the choice between tasks should depend only on the characteristics of these sentences and should not change when both sentences are subject to the same transformations.

Therefore, it seems justified to put limits on their impact in relation to the risks, by the means of pilot studies, which allows for mitigating the negative effects of the decisions taken. Firstly, the PSU will gain an important argument regarding the participation of citizens in the decision-making process. Secondly, the results of the pilot study will contribute to maximising the expected utility, as they will allow decision-makers to correctly estimate the probability of future events [11, 19, 51]. The third argument is the limited risk of failure to achieve a political effect. This means that through participation and positive results of the pilot study, when assessing the efficiency and effectiveness of allocation of public money, the PSUs gain an argument that all aspects of the decision-making process have been properly carried out, and that the limited effectiveness of resources involved in the task results from changes in social preferences (as demonstrated above when outlining the assumptions of the limited utility theory). It must always be assumed that a member of society, in their judgments and decisions about the choice of public goods and services, will demonstrate the attitude, according to which "with the inflow of new information, people can automatically update the probability of events occurring (as they draw conclusions based on the law of probability calculation)", which will have an effect on the perception of goods and services provided (i.e. on consumption and change of individual preferences) [9]. As a result, there is a mismatch between decisions and actions taken by the authorities the in order to meet the public needs, through the task at hand. There is also a mismatch to minimise the risk of suffering limited effects from the measures taken.

Piloting is associated with the problem shown in the literature of Kahneman and Tversky [26] and Mongin [34], where society will be based on this solution, which leads to the choice of a variant with some effects, but the risk-taker's attitude may be eliminated in advance in elections related to losses. Moreover, as demonstrated by Filipiak and Dylewski [15], this situation may discourage the public from making choices.

Despite this failure, it is thanks to piloting that it is possible to test those situations (however, they cannot all be eliminated) in which the change of information will cause a change in the decisions of the society members (individuals) about the goods and services provided. It can also be pointed out that piloting is a kind of experimental approach to the execution of public tasks, the results of which are public goods and services [15], because:

- the conditions for conducting a pilot study meet the requirements for the experimental approach (the attributes of an experiment) [47, 50];
- the pilot study fulfils the functions of the experimental approach [43, 49];
- it ensures the effect of a good study [30, 47].

24.4 Example of Practical Application

24.4.1 Scientific Approach

It is assumed that the pilot gives the possibility to the PSU to plan tasks in such a way as to limit the risk of incurring losses (too high expenditures). To this end, we have analysed:

- the adopted assumptions of selected pilot implementations in the PSUs, the aim of which was to test and evaluate the solution to be used in practice as an expression of public participation;
- socially useful (public) objectives, which were subject to public consultation as part of the pilot project; on the basis of their presentation, the society would evaluate the proposed solutions and make decisions;
- the evaluation of the objectives of pilot projects to improve the efficiency of the allocation of public resources and, ultimately, the provision of public services in the broad sense, as well as to ensure consistency of undertaken activities as well as their transparency to meet the requirements of accountability.

This analysis makes it possible to draw conclusions about the solution to the research problem: does the pilot provide an opportunity to identify a "better" solution and ultimately contribute to a better allocation of resources and to maximising the utility expected by the community at a predetermined level of risk acceptance?

Given that public authorities are free to formulate public policies and have decision-making autonomy, their legislation is local or territorial in nature and may vary from one type of PSU to another. Public authorities, by virtue of the size of the budget at their disposal, take into account the mandatory and optional measures that have to be implemented. Therefore, there are different projects (in terms of their scope, size and type of tasks) that are carried out in pilot system where the leading unit is a central government body, and other ones led by a local government unit.

For the purposes of this analysis, public information given in the Public Information Bulletin was used. For the purpose of implementing the adopted thesis and in order to solve the research problem, two nationwide pilots were selected for the study that had been initiated by central and self-governmental PSUs.

The study used the monographic method. The first example of piloting covered nine out of 12 cities associated in the Union of Polish Metropolises (it is an association of the largest Polish cities with poviat rights) and the city of Olsztyn, which is the largest city in the Warmińsko-Mazurskie Voivodeship. For the second example of a pilot project, the research sample was determined by a PSU at the central government level (Ministry of Family, Labour and Social Policy) basing on a task-based approach. The pilot covered three voivodeships (Mazowieckie, Dolnośląskie and Podkarpackie) and 3117 unemployed persons.[5] At the same time, the group of individuals taking part in the pilot was defined on the basis of the labour market analysis which had

[5]Information about the project Raport [38] and www.1 [52].

revealed that the groups for whom finding employment was the most difficult were the long-term unemployed, as well as women who did not take up employment after giving birth to a child, people who are older than 50 and people with disabilities.

24.4.2 Findings and Discussion

The paper will present the theoretical approach to the experiment with the use of piloting, so as to demonstrate the thesis that it gives the opportunity to indicate a "better" solution, and to provide justification for such selection or for not taking action due to benefits lower than expected.

The first example of a pilot, used in a form of an experiment as a method to verify the usability expected by communities in risk conditions, is participatory budgeting. The pilot study was initiated by local governments. Its choice was dictated by the vast scope of its implementation, as pilot participatory budgeting was launched in the largest Polish cities. After the trial period, in 2019 the participatory budgeting was introduced into the national law. Is it possible to speak of a successful experiment? Are the decision-makers in territorial self-government units able to ensure maximisation of the utility expected by their communities while operating in risk conditions? The participatory budgeting, as an expression of public participation, by way of a process of consulting potential projects (local government's own tasks) to be financed from public funds, in its principle assumes the rationality of choices in conditions of risk and of decisions made by particular individuals. However, the consultation framework, methods and scope of selection are prepared by the relevant PSU bodies. This is the most far-reaching and direct expression of citizens' participation in decision-making processes. Such a system of decision-making, as a form of public consultation, entails numerous unknown variables, due to difficulties in assessing the behaviour of an individual or the whole local community in terms of actual and direct participation in the decision-making process. Therefore, through piloting implemented as an experiment, one can obtain objective results, which will allow to obtain highly probable results that will provide the PSUs with information about both the costs and social benefits. Thus, the PSUs will or will not be able to confirm whether or not they are maximising the expected utility of the goods and services piloted by society.

The analysis covered ten cities with poviat rights (the largest cities in Poland), which decided to launch participatory budgeting. As an expression of acknowledgement of that form of decision-making in risk conditions, the PSUs determined the amounts of public funds that they were willing to allocate for the experiment. In determining the amount of public funds, the PSUs accepted it as a pool for participatory budgeting, in the scope of which direct decisions were taken by citizens in the form of a vote (Table 24.1). Respondents (society) are people over 18 years of age. In accordance with applicable law, PSUs can not limit the possibility of voting on participatory budgeting programs. Which means that each authorised person can take part in the research, and the entire local community is the research sample.

Table 24.1 Value of participatory budgets in selected cities with poviat rights in Poland in thou. PLN

	2013	2014	2015	2016	2017	2018	2019
Warszawa			26,237.2	51,215.6	58,588.9	61,419.9	64,784.2
Łódź		20,000	40,000	40,000	40,000	40,000	40,000
Wrocław	2000	20,000	20,000	25,000	25,000	25,000	25,000
Poznań	10,000	10,000	10,000	15,000	17,500	18,000	20,000
Gdańsk		9000	11,000	12,500	12,500	14,000	19,819
Lublin			10,000	15,000	15,000	15,000	15,000
Białystok		10,000	12,000	20,000	10,000	10,000	10,000
Kraków		4500	14,000	10,855	10,000	8000.1	9257.2
Szczecin		5000	5000	6000	7000	8000	9000
Olsztyn		2150	3300	3530	3730	3730	3730

Source Own study based on BIP data

The analysis, based on formal documents, of the assumptions of projects included in the pilot study in ten cities with poviat rights shows that:

- local communities opted for socially useful projects, characterised by maximisation of expected usability by the society (maximum number of people voting for a given project or projects);
- the risk of decision error was reduced, as the communities indicated a project that they accepted, which was an expression of a "better" choice and better allocation of public funds;
- the risk of failure to achieve a political effect was reduced, i.e. through public participation and positive results of public choices, the PSUs gained an argument in assessing the efficiency and effectiveness of public funds allocation.

The data presented in Table 24.1 show that the amounts allocated in the first year of the experiment were significantly lower than in subsequent years. In most cities, the amounts allocated in subsequent years increased regularly (in particular in Warsaw, Poznań and Szczecin) or once only (in Łódź, for example). The exceptions were two cities, Kraków and Białystok, where the amounts involved in the participation budgeting decreased. The analysis of formal documents, which were carried out by the authors, revealed that the first year was treated as a pilot, followed by an evaluation of the effects of implementing a new solution about deciding on the allocation of public funds. One more important conclusion can also be drawn from the presented data: the participation budgeting implemented under the conditions in which Polish local governments operate in Poland is not a method that will effectively replace the existing methods of PSU decision-making from indirect to direct decisions made with public participation.

It should be highlighted that a one-off or cyclical substantial increase in participatory budgeting amounts after the piloting period shows that this method makes it possible to identify a "better" solution to meet public needs (i.e. select projects

that better meet public needs) and, ultimately, the PSU bodies are more efficient in allocating resources and maximising expected public utility while accepting a certain level of risk. In the cities where there has been a regular increase in the amounts of the participatory budgeting, public participation has been gradually expanded (as a result of the positively assessed effects of piloting). However, in the cities where the process of raising participation has been stopped or frozen, it can be assumed that local community did not approve of this method of decision-making and thus did not achieve the maximum expected utility of the submitted projects. Moreover, the community, under the influence of inflowing new information and after an automatic analysis of the choices made, made decisions that changed the perception of the provided goods and services (individual consumption and preferences). In such a situation, some PSUs withdrew from expanding the funds available for participatory budgeting.

In general, the experimental participatory budgeting can be regarded as positive, because local communities, when faced with risks and environmental changes, achieve the expected utility of the projects adopted for implementation within the participatory budgeting. This can be clearly seen in the dynamics of the amounts regularly allocated to participatory budgeting. By using piloting as an experimental approach to assessing the expected utility, better allocation of resources is achieved and better assessment of expected utility by society at a certain level of risk is reached through public participation. The general assessment of the experiment carried out by the government of the Republic of Poland resulted in the introduction of the idea of participatory budgeting in Poland in 2019 as an obligatory process of involving citizens in local decision-making. However, only the minimum level of such participation was stated.

The second of the analysed examples of experiments in the PSU decision-making process was a pilot project "Partnership for Work", which was implemented as a state government project. The labour market problems and the high rate of unemployment in Poland prompted the search for solutions aimed at providing more effective assistance to the unemployed and at improving and developing public–private partnerships in the labour market. According to the assumptions of the programme, private entities based on partnership should, to a greater extent than before the pilot, engage in the model of activation of the unemployed, mainly those who are the furthest from the labour market.

The strategic objective of the pilot project "Partnership for Work" was to develop an optimal and supportive model of cooperation between voivodeship labour offices or poviat labour offices and recruitment agencies and other labour market partners in order to create an effective system of activation of the unemployed at the voivodeship (regional) level. The main inspiration for launching that pilot project was the need to test a method other than the existing schemes for coping with long-term unemployment as well as with the number of people at the risk of chronic unemployment.

The experiment, which was carried out in three regions, was based on the following assumptions:

- testing the model of commissioning measures to locate the unemployed on the labour market which was inspired by solutions based on cooperation of employment services with non-public entities in other countries and by solutions developed nationally;
- evaluation of the functionality and effectiveness of the proposed solutions concerning commissioning measures to locate the unemployed on the labour market, in particular the development of detailed rules and steps necessary for the systemic implementation of this solution on the labour market in Poland.

The implementation of the pilot project "Partnership for Work" started in 2012 in three voivodeships (regions): Mazowieckie, Podkarpackie and Dolnośląskie—on the basis of contracts for implementation and was completed in 2014. The activation activities for the participants of the pilot project were carried out for nine months.

The project assumed that the unemployed would take part in activation activities offered by employment agencies as a part of the tested method of commissioning activation measures, which were selected by the voivodeship self-government pursuant to the provisions of the Public Procurement Law. Moreover, the service provider (employment agency) selected in a given voivodeship activised the unemployed recommended by poviat employment offices engaged in the pilot project and designated by the Marshal of the voivodeship.

The participants in the pilot were defined on the basis of the analysis of the labour market, which showed that the greatest difficulty in finding employment was faced mainly by the long-term unemployed, women who did not take up employment after giving birth to a child, people over 50 years of age and people with disabilities. The long-term unemployed, women who did not take up employment after giving birth to a child and people over 50 years of age and people with disabilities they were respondents. Government assumptions of the program have set this respondents for the program. In the three voivodeships in total, 3117 unemployed persons were directed to the service provider under the pilot scheme. The amount of PLN 30 million was earmarked for the project.

The remuneration of the service provider depended on the employment effect achieved. The pilot project tested a payment model where remuneration was paid in four tranches due for specific activities undertaken by the service provider, i.e.:

- I tranche—receivable for the opening balance—20%.
- II tranche—receivable for the activation balance—30%.
- III tranche—Bonus I—25%.
- IV tranche—Bonus II—25%.

Full remuneration was paid if the service provider reached a certain level of employment effectiveness and retention rates. None of the service providers received tranche IV (Bonus II).

Reports on the evaluation of experimental activities under the new model of cooperation concerning employment activation in the region pointed to significant weaknesses and lack of implementation of the assumed project objectives. One of the main drawbacks was differing levels of competence of entities involved in the pilot

project, i.e. self-government voivodeships and poviats—as local government units independent of one another. On the basis of the experiment/pilot project, dysfunctions were identified which had not been taken into account when adopting the project objectives:

- due to the assumed main objective of the project, i.e. to obtain the employment effect, there was a negative effect of "buying employment";
- many persons recommended to the project actually were not interested in finding employment;
- the activation measures under the pilot project were too short in duration;
- defective communication and exchange of information within the bodies involved in the experiment;
- significant difficulties in developing substantive criteria for the selection of partners (employment agencies) in the public procurement process;
- insufficient role of poviat labour offices in monitoring, control and management of the project implementation, as entities closest to the problems of local labour markets;
- poor results in terms of quality and levels of wages offered to the unemployed;
- the contracting method of employment agencies favoured only large entities, that were not always interested in the social and local aspects of the project.

One of the positive sides of the experimental pilot project was the possibility to commission activation measures at the regional level and provides the unemployed with an opportunity to use the institutional support in a broader extent than before. The programme was not continued. According to the Ministry of Labour and Social Policy, the pilot project "Partnership for Work" was designed to help formulate legal regulations at the statutory (national) level, rather than to test this solution as a permanent one. The development of solutions in Poland was preceded by an analysis of solutions implemented by other European countries with regard to commissioning the activities as well as by piloting the model of commissioning the activities. However, taking into account the cost of the programme, the conclusion may be drawn that it was not a cheap experiment.

To recapitulate, the following effects were achieved as a result of the experiment:

- it has been demonstrated that the proposed solution, which was tested by way of piloting, does not allow the public to achieve the expected utility;
- the risk of error in decision-making was reduced, as the conclusions from the pilot study indicated the irrelevance and, in the future, the lack of public approval for the designed solutions which were the subject of the pilot study;
- the need to look for a "better" solution that would contribute to a better allocation of public funds was demonstrated;
- the risk of failure to achieve political effect was reduced.

24.5 Conclusions

The analysis has confirmed that the experimental approach is important for the adequate allocation of public funds and allows for the choice of a "better" solution that will contribute to obtaining the expected utility of the public goods provided.

The question: Does the pilot project provide an opportunity to identify a "better" solution and in the end contributes to a better allocation of resources and maximisation of utility expected by the community at a certain level of risk acceptance? has been positively verified. Taking into account the frequency of changeability of tasks in the PSUs (they show relative permanence regarding their own and commissioned tasks that are defined by law, while non-binding own tasks and commissioned tasks constitute a small percentage of tasks of low durability), the above two examples of piloting and their impact on the PSU tasks confirm that piloting as an element of the experiment has a positive effect on increased rationality of decision-making, i.e. it affects decisions that are consistent and maximise value within given limits. It allows to reduce problems of unclear structure, about which information is ambiguous or incomplete. Thus, the analysis additionally reinforces the theses defined by Robbins [39].

It should also be stressed that piloting is used to model the society members' behaviour which is driven by the desire to maximise utility [52]. This is an important aspect, as it allows, through the definition of axiomatic principles, both for modelling social behaviour and for estimating the maximised expected utility of the goods provided in the risk situation, as perceived by the public.

Although Kahneman and Tversky [26] in their study, showed that the choice between risky variants leads to the choice of a variant with some effects, and such behaviour is not consistent with the main principles of the theory of expected utility, it should be said that in the local government there are no variants of solutions for the tasks being implemented. The PSU policy-makers act differently (as demonstrated by the practice) than stock market players or managers of enterprises. On the PSU's, only one variant for which a piloting is applied is analysed. The affirmation is confirmed [2, 34], that risk aversion appears in the elections of the PSU authorities. What it means that by eliminating, the risk PSU authorities are guided by measurable effects. The piloting is to confirm decision-makers that the risk, which is most often manifested in the allocation of public funds accepted by the society, will be limited.

The pilot does not provide an opportunity to identify a "better" solution because the PSU's authority not building variant's of new tasks but the pilot ultimately contribute to a better allocation of resources by verifying the adopted assumptions of the pilot program. This maximises the tasks performed by the PSU, which means increasing the usability from the allocation of public funds.

The authors see the need to further explore the problem of measurement of utility expected by the community of pilot programs. This is a gap in practice and empirical research in Poland. In the world at PSU, only the effects of spending are measured. There is a lack of extensive research on the assessment of utility expected by the

community by the society with the use of pilot programs, taking into account the specificity of PSUs in the other countries.

References

1. Ackroyd, S., Kirkpatrick, I., Walker, R.: Public management reform in the UK and its consequences for professional organization: a comparative analysis. Public Adm. **85**(1), 9–26 (2007)
2. Allais, M.: Allais's paradox. In: The New Palgrave. A Dictionary of Economics. MacMillan (1987)
3. Attacking Poverty: World Bank, Washington (2000)
4. Brooks, P., Zank, H.: Loss averse behavior. J. Risk Uncertainty **31**(3), 301–325 (2005)
5. Bulbul, A.: The linkage between E-governance and good governance: an analysis on Bangladesh public administration. Humanit. Soc. Sci. **6**(4), 114–120 (2018)
6. Csink, L., Kurunczi, G., Varga, A.Z.: The role of legislation in pilots. In: Varga, A.Z. (ed.) Pilot Projects in Public Administration Management, pp. 25–35. Pázmány Péter Catholic University, Budapest (2010)
7. Czerwonka, M., Gorlewski, B.: Finanse behawioralne. Zachowania inwestorów i rynku, Oficyna Wydawnicza Szkoły Głównej Handlowej, Warszawa (2012)
8. Davies, G.B., Satchell, S.E.: The Behavioral Components of Risk Aaversion. J. Math. Psychol. **51**, 1–13 (2007)
9. Dobrowolski, K.: Teoria rynków efektywnych i model racjonalnego inwestora - od warunków ryzyka do warunków konfliktu. Contemp. Econ. **5**(1), 1–12 (2014)
10. European Governance—A White Paper: Official Journal of the European Communities, COM (2001) 428 final, (2001/C 287/01) (2001)
11. Fedorowicz, J., Gelinas Jr., U.J., Gogan, J.L., Williams, C.B.: Strategic alignment of participant motivations in e-government collaborations: the internet payment platform pilot. Gov. Inf. Quart. **26**, 51–59 (2009)
12. Ferlie, E., Hartley, J., Martin, S.: Changing public service organisations: current perspectives and future prospects. Br. J. Manag. **14**, S1–S87 (2003)
13. Filipiak, B.Z.: Funding of joint investment projects of public institutions with private entities. Acta Aerarii Publici **13**(2), 15–23 (2016)
14. Filipiak, B.Z., Dylewski, M.: Ryzyko w działalności jednostek samorządu terytorialnego w nowych uwarunkowaniach prawnych. Roczniki bezpieczeństwa 2012/2013, 28–40 (2013)
15. Filipiak, B.Z., Dylewski, M.: Participatory budgeting as example of behavioural impact of public policies. In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics, pp. 231–247. Springer, Cham (2018)
16. Fischburn, P., Kochenberger, G.: Two-piece von Neumann-Morgenstern utility functions. Decis. Sci. **10**, 503–518 (1979)
17. Fowler Jr., F.J.: The case for more split-sample experiments in developing survey instruments. In: Presser, S., Rothgeb, J.M., Couper, M.P., Lessler, J.T., Martin, E., Martin, J., Singer, E. (eds.) Methods for Testing and Evaluating Survey Questionnaires. Wiley, New York (2004)
18. Goodwin, N.R., Harris, J.M., Nelson, J., Weisskopf, T.E.: Economic behavior and rationality. Microeconomics in Context, pp. 145–157. M.E. Sharpe, New York (2014)
19. Grzeszkiewicz-Radulska, K.: Metody badań pilotażowych. Folia Sociol. **42**, 113–141 (2012)
20. Hanna, R., Daim, T.U.: Decision-making in the service sector—comparison of information technology acquisition between private and public institutes. Int. J. Innov. Technol. Manag. **4**(1), 41–58 (2007)
21. Hastie, R., Dawes, R.M.: Rational Choice in an Uncertain World, 2nd edn. Sage, Thousand Oaks, CA (2010)
22. Hills Jr., R.M.: Federalism and public choice. MPRA Paper 13625, pp. 1–54 (2009)

23. Internetowy System Aktów Prawnych. Sejm RP: Ustawa z dnia 27 sierpnia 2009 r. o finansach publicznych. Dz.U. nr 157, poz. 1240 z późn. zm. http://isap.sejm.gov.pl/DetailsServlet?id=WDU20091571240. Accessed 15 Dec 2018 (2009)
24. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica **47**(2), 263–291 (1979)
25. Kahneman, D., Tversky, A.: Advances in prospect theory: cumulative representation of uncertainty. J. Risk Uncertainty **5**(4), 297–323 (1992)
26. Kahneman, D., Tversky, A.: Choices, values, and frames *Notes*. In: Kahneman, D. (ed.) Thinking Fast and Slow. Farrar, Straus and Giroux, New York (2011)
27. Kampen, J.K.: Good governance at the local level: toward a global village or a city republic? Econ. Environ. Stud. **9**(1), 11–29 (2009)
28. Köbberling, V., Wakker, P.: An index of loss aversion. J. Econ. Theory **122**, 119–131 (2005)
29. Konkov, A.: Global Problems for Global Governance. valdaiclub.com, Moscow (2014)
30. Krawczyk, M.: Ekonomia Eksperymentalna. Wolters Kluwer Polska, Warszawa (2012)
31. Lau, R.R., Kleinberg, M.S., Ditonto, T.M.: Measuring voter decision strategies in political behavior and public opinion research. Public Opin. Quart. **82**, 325–350 (2018)
32. Levy, M., Levy, H.: Prospect theory: much ado about nothing? Manag. Sci. **48**(10), 1334–1349 (2002)
33. Marzuki, A.: Challenges in the public participation and the decision making. Process Sociologija i Prostor **53**(1), 21–39 (2015). https://doi.org/10.5673/sip.53.1
34. Mongin, F.: The Allais paradox: what it became, what it really was, what it now suggests to us. Econ. Philos. **30**(01), 1–37 (2019)
35. Nalepka, A., Kozina, A.: Podstawy badania struktury organizacyjnej. Akademia Ekonomiczna, Kraków (2007)
36. Payne, J.W., Laugham, D.J., Crum, R.: Further tests of aspiration level effects in risky choice behavior. Manag. Sci. **27**, 953–958 (1981)
37. Presser, S., Couper, M.P., Lessler, J.T., Martin, E., Martin, J., Rothgeb, J.M., Singer, E.: Methods for pretesting and evaluating survey questions. Public Opin. Quart. **68**(1), 109–130 (2004)
38. Raport: Przeprowadzenie monitoringu i ewaluacji projektu pilotażowego pn. "Partnerstwo dla Pracy", Zespół Sektora Publicznego IBC GROUP Central Europe Holding S.A., Warszawa. https://psz.praca.gov.pl/documents/10828/167868/RAPORT%203%20-%20pp%20Partnerstwo%20dla%20pracy.pdf/403a2ba4-66a0-43d6-8e4e-079b06d06899?t=1433499052000. Accessed 15 Dec 2018 (2014)
39. Robbins, S.P., DeCenzo, D.A.: Podstawy zarządzania. PWE, Warszawa (2002)
40. Schermerhorn Jr., J.R.: Management. Willey, New York (2005)
41. Secchi, E.: Extendable Rationality: Understanding Decision Making in Organizations (Organizational Change and Innovation). Springer, New York (2010)
42. Simon, H.A.: Administrative Behavior, 4th edn. The Free Press, New York (1997)
43. Smith, V.: Economics in the laboratory. J. Econ. Perspect. **8**(1), 113–131 (1994)
44. Sułek, A.: Eksperyment w badaniach ekonomicznych. PWE, Warszawa (1979)
45. Świrska, A.: Methods for improving the effectiveness of the financial management of local government units in Poland. Sci. Stud. Account. Finan. Prob. Perspect. **10**(1), 177–185 (2016). https://doi.org/10.15544/ssaf.2016.17
46. Szymaniec-Malicka, K.: Proces podejmowania decyzji w organizacjach publicznych – propozycja metody badawczej. Organizacja i Zarządzanie **74**, 437–447 (2014)
47. Varga, A.Z.: Legal limitations of public administrative pilot studies. In: Varga, A.Z. (ed.) Pilot Projects in Public Administration Management, pp. 15–24. Pázmány Péter Catholic University, Budapest (2013)
48. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, London (1953)
49. Vreugdenhil, H., Rault, P.K.: Pilot projects for evidence-based policy-making: three pilot projects in the Rhine basin. German Policy Stud. **6**(2), 115–151 (2010)
50. Wójcicki, S.: Zasady eksperymentu. Wydawnictwo Ministerstwa Obrony Narodowej, Warszawa (1994)

51. Woods, E.H.: Measuring the social impacts of local authority events: a pilot study for a civic pride scale. Int. J. Nonprofit Volunt. Sect. Mark. **11**(3), 165–179 (2006)
52. www.1: http://psz.praca.gov.pl/-/168475-projekt-pilotazowy-partnerstwo-dla-pracy. Accessed 15 Dec 2018

Chapter 25
What Really Helps Us Make a Choice? An Experimental Evaluation of AHP

Jacek Cypryjański, Aleksandra Grzesiuk and Kamil Bortko

Abstract In this study, we used a laboratory experiment to test how well AHP is coping with identifying our preferences and what persuades us to accept the AHP outcomes: the result itself or perhaps a "scientific" way of obtaining it. Participants were asked to give rankings at three stages of the experiment: one at the beginning, one after filling in the AHP matrices, and the final one after learning the AHP ranking. Half of the participants were presented with the actual and half with the manipulated AHP ranking. Using nonparametric statistical tests, we found evidence that among the participants of the experiment, the process of filling in the AHP matrices itself had an impact on the final decision. However, the reformulation of the final decisions under the influence of the AHP ranking took place only among the participants who were presented with manipulated results. In addition, the experiment showed that the AHP ranking was more distant from the final ranking than the two previous rankings given by the participants.

Keywords Decision analysis · Multi-criteria decision making · Analytic hierarchy process (AHP) · Preferences · Experimental economics

J. Cypryjański (✉)
Faculty of Economics and Management, University of Szczecin, ul. Mickiewicza 64, 71-101 Szczecin, Poland
e-mail: jacek.cypryjanski@usz.edu.pl

A. Grzesiuk
Faculty of Economics, West Pomeranian University of Technology, ul. Żołnierska 47, 71-210 Szczecin, Poland
e-mail: aleksandra.grzesiuk@gmail.com

K. Bortko
Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, ul. Żołnierska 49, 71-210 Szczecin, Poland
e-mail: kbortko@wi.zut.edu.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_25

25.1 Introduction

The application of multi-criteria decision-making (MCDM) methods is generally focused on supporting managerial decisions. The aim is to support the decision-making process in such a way as to bring managers closer to a rational choice. A slightly different way of applying MCDM methods is to use them to identify our preferences. Logic dictates us to believe that these methods allow us to better understand our preferences and eliminate heuristic strategies that we unconsciously use such as "I choose what I know." Unfortunately, we cannot be sure of that and, moreover, we cannot verify it directly because it is difficult to determine the "objective truth" about our preferences (if indeed there is one) to which we could compare the results. That is why different types of experiments are designed, in which the issue is analyzed in a variety of indirect ways [1–7].

In our research, we used a laboratory experiment to test how practical in identifying our preferences can be the analytic hierarchy process (AHP) and what persuades us to accept the AHP outcomes: the result itself or perhaps a "scientific" way of obtaining it. The decision problem that we tested was the choice of wallpaper for the computer screen among the five presented alternatives. Most of us set up our own wallpaper on a computer or smartphone screen, and the selection process itself can really absorb some of us. Counting on it, we gave up the monetary stimuli usually used in experimental economics.

The procedure of our experiment was based on an experiment conducted by Ishizaka et al. [2] in which participants were asked to give rankings at three stages of the experiment. The participants made their first ranking after seeing all five alternatives. They were then asked to carry out all the activities required to prepare the ranking according to the AHP method, but the AHP ranking itself was not presented to them. After this task, the participants were again asked to perform the ranking. Finally, the AHP ranking was presented to the participants and they were asked to make the third definitive ranking.

Such an experiment provides an opportunity to examine two aspects of the impact of AHP on the expression of our preferences: (1) If the B ranking is nearer to the final D ranking than the initial A ranking, it means that the process of filling in the AHP matrices itself brings the participants closer to the final D ranking; (2) if the final ranking D is nearer to the ranking C than B is to C, it means that the participants have reformulated their final decisions on the basis of the ranking made by the AHP method. To see if the reason for a possible reformulation of the final decisions under the influence of the AHP ranking is the ranking itself, or rather the belief in the "scientific nature" of the method, half of the participants were presented with the AHP ranking, which we changed in comparison to the actual results.

Using nonparametric statistical tests, we found evidence that among the participants of the experiment, the process of filling in the AHP matrices itself had an impact on the final decision. As for the second of the examined aspects, the reformulation of the final decisions under the influence of the AHP ranking took place only among the participants who were presented with manipulated results. In addition,

the experiment showed that the AHP ranking was more distant from the final ranking than the *A* and *B* rankings.

25.2 Literature Review

AHP has gained wide acceptance among researchers and practitioners and is probably the most widely used MCDM method [8–13]. The AHP method, introduced by Saaty [14], originally was applied to the marketing sector. It is a tool for dealing with complex decision making and is based on the pairwise comparison. It may help the decision-maker to set priorities and make a rational decision. By reducing complex decisions to a series of pairwise comparisons, and then synthesizing the results, the AHP helps to capture both subjective and objective aspects of a decision. The main advantages of AHP are that it allows hierarchical modeling of the problem, and the possibility to make verbal judgments and to confirm consistency [11]. The AHP method has also proven effective for dealing problems in various disciplinary domains. As can be seen from the review prepared by Ishizaka and Labib [11] and Emrouznejad and Marra [9], the usefulness of AHP for decision making has been shown in the healthcare sector, banks, supply chain management, education, operators evaluation, site selection, software evaluation, project selection, customer requirement rating, and many others. In recent studies, Krejči [15] and Krejči et al. [16] advocate the validity of the fuzzy extension of AHP. The mentioned authors showed that by neglecting the information about the uncertainty of intensity of preferences, an important part of knowledge about the decision-making problem can be lost and can cause the change in the ordering of alternatives.

AHP has been the subject of a number of laboratory experiments investigating various aspects of its use. Milet [17] in an experiment in which the participants compared the relative areas of five geometric shapes, measured the actual accuracy of four different methods of elicitation (verbal, numerical, and two graphical), and compared them with a simple direct estimation. The results showed that AHP, regardless of the elicitation method used, outperformed the direct estimation, while verbal and numerical pairwise comparisons were the most accurate, although at the same time the slowest and difficult to use. This laboriousness and time-consuming pairwise comparison used in AHP, unfortunately, means that with slightly more complex tasks, participants may experience fatigue that disrupts the results of experiments. This was observed by Nermend [5] in an experiment conducted with the use of neuroscience techniques (electroencephalography, galvanic skin response, and eye-tracking). In this experiment, fatigue among participants occurred on average after 45 pairwise comparisons, while, for example, the assessment of five variants according to five criteria requires as many as 60. Another interesting conclusion comes from an experiment carried out by Linares [4], which involved a car comparison by means of AHP. In the opinion of the participants of this experiment, the removal of inconsistencies from the AHP matrices did not improve the representation of their preferences.

Experiments conducted by Pöyhönen and Hämäläinen [6], Cypryjański and Grzesiuk [1], and Ishizaka and Siraj [3] focused on the comparison of MCDM methods. While the first two analyzed the causes of differences in results generated by particular methods, in the third incentive-based experiment, the three methods (AHP, SMART, and MACBETH) were evaluated for their usefulness. The results of the third experiment showed that both AHP, SMART and to a slightly lesser extent (although statistically insignificant) MACBETH helped participants in their decision-making. In addition, the questions asked at the end of the experiment showed that AHP (implemented in Expert Choice) and SMART (Right Choice), in contrast to MACBETH (M-MACBETH), were perceived by participants to be useful. AHP and SMART were also evaluated by participants as easy to use, but this was less the case for MACBETH.

As it was pointed out in the introduction, the inspiration for our study was a research paper written by Ishizaka et al. [2]. In this paper, experimental economics methods were used to test how well AHP works as a choice support system in a real decision problem. The decision problem, tested in their incentive-based experiment, was the problem of selecting a box of chocolates from five alternatives. The authors themselves selected the criteria according to which the participants evaluated alternatives using the AHP method (at the end of the experiment, a questionnaire was carried out in which the participants agreed with the selection of criteria made by the authors). The results of their experiment indicated that while AHP detected clear top and least priorities well, the three rankings given by the participants were closer to each other than they were to the AHP ranking. The authors found evidence that participants tended to follow the ranking provided by AHP. However, it was not observed that the process of filling in the AHP matrices itself brought the participants closer to the final decision.

25.3 Experimental Procedure

The experiment involved 64 participants—34 students and 30 academic and administrative staff members of the Faculty of Economics and Management at the University of Szczecin. The decision problem was the choice of wallpaper for the computer screen. We asked the participants to rank from the most to the least preferred five known images: “Wind from the Sea” by Andrew Wyeth (1947), “Impression, Sunrise” by Claude Monet (1872), “The Lovers of Venice” by Marc Chagall (1957), “Slaves stopping a horse” by Théodore Géricault (1817), and “The Bedroom” by Vincent van Gogh (1888). This decision problem was chosen on the basis of a discussion and a test experiment conducted with another group of students as capable of engaging the participants and getting them to give thoughtful answers. For the purposes of the experiment, a web application was developed. The application guided the participant through the experiment, asked him/her questions, collected answers, and performed all necessary calculations.

At the beginning, the participants were instructed on how the experiment will be conducted and on the principles of performing the evaluation by means of the AHP method. The experiment consisted of the following steps:

1. The participants were presented with paintings together with information about their titles, authors, and years of creation and then asked to give the first ranking (ranking A).
2. The participants were asked to choose between 3 and 5 criteria that they consider to be of importance when choosing the screen wallpaper among the presented paintings. Next, they were asked to make pairwise comparisons of the criteria they had chosen as well as, according to each criterion, pairwise comparisons of the paintings.
3. Again, the participants were asked to rank the paintings (ranking B).
4. The AHP ranking was calculated and presented to the participants. However, the participants were divided into two groups (random distribution). One group (32 participants) was presented with a real ranking (ranking C) and the second group (also 32 participants) was presented with manipulated results of AHP ranking (ranking C') in which we swapped the first position with the second and fourth with the fifth.
5. The participants were asked to give a final ranking of the paintings (ranking D).

At the end of the experiment, the participants were asked to answer one question: To what extent did the AHP ranking reflect your preferences? After completing the experiment, we calculated the distances between the individual rankings. To calculate them, we used the squared Euclidean distance:

$$XY = \sum_{i=1}^{5} (x_i - y_i)^2$$

where XY—distance between X and Y rankings, $i = 1, \ldots, 5$—numbers of assessed alternatives A, x_i—place of alternative i in X ranking, and y_i—place of alternative i in Y ranking. Table 25.1 shows an exemplary calculation of the distance between rankings X and Y.

In total, we have obtained six distances for each of the participants. Three of them represented the differences between the participants rankings (AB, AD, BD) and the

Table 25.1 Example of calculating the distance between ranking X: A2, A4, A1, A5, A3 and ranking Y: A2, A1, A3, A5, A4

Alternative	x_i	y_i	$(x_i - y_i)^2$
A1	3	2	1
A2	1	1	0
A3	5	3	4
A4	2	5	9
A5	4	4	0
XY distance			14

remaining three between the individual rankings and the ranking obtained by means of AHP (*AC*, *BC*, *CD*).

25.4 Results

The first issue we analyzed was to determine the size of the differences between the participants' rankings. To do this, for each participant, we calculated the sum of the distances between these rankings ($AB + BD + AD$, Fig. 25.1, axis $0Y$). The results showed that 31% of the participants in three consecutive rankings presented the same order of alternatives (total distance was equal to zero). Another 39% of the participants slightly changed their preferences during the experiment (the total distance was from 4 to 8). For the remaining 30% of the participants, the total distance between the rankings varied from 10 to 56, which may indicate that some of them were not particularly involved in the experiment or significantly changed their preferences during it. As for the changes in preferences during the experiment, as shown in Fig. 25.2, the initial *A* ranking differs most from the two subsequent

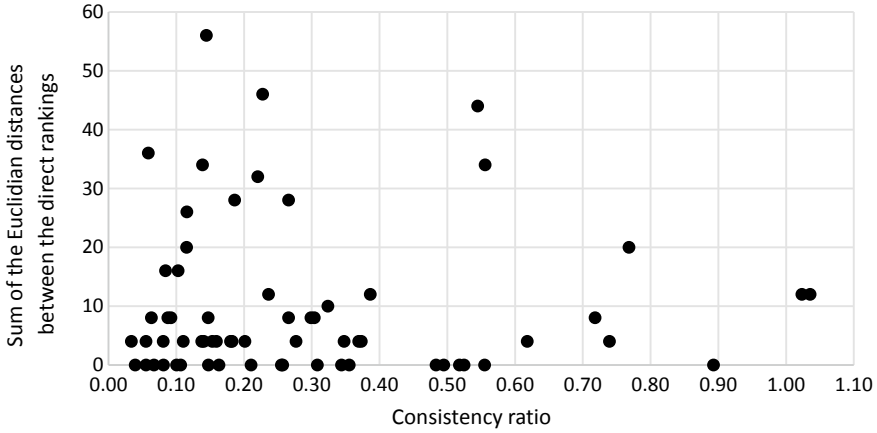


Fig. 25.1 No correlation between the consistency ratio and the variation across the rankings *A*, *B*, and *D*

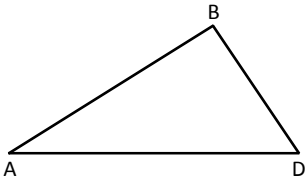


Fig. 25.2 Triangle visualizing average distances between direct rankings ($AB = 6.9$; $BD = 4.4$ and $AD = 8.3$)

ones, which means that the participants made the biggest changes after performing the AHP evaluation (selection of the criteria, pairwise comparison).

Then, we studied how the participants of the experiment were able to deal with the AHP method, particularly with the pairwise comparison. For this purpose, we calculated for each participant the value of the consistency ratio Cr [14]. It is commonly accepted in the literature that an individual is consistent in his or her pairwise comparison, if the consistency ratio for his or her assessments does not exceed 0.1. In our experiment, only 19% of participants had a consistency ratio $Cr < 0.1$. If we add to this group another 28% of those with Cr not exceeding 0.2, there is still more than half of the participants who did not cope well with the AHP method (Fig. 25.1, axis $0X$).

Although the participants in the Ishizaka et al. [2] experiment had slightly lower Cr values, both experiments show that it is quite difficult for individuals to make consistent pairwise comparisons. The results of our experiment also confirm another observation from the Ishizaka et al. [2] experiment—no correlation between the consistency ratio and the variation across the rankings A, B, and D (Pearson's correlation coefficient is -0.0138, Fig. 25.1).

Next, we have verified five research hypotheses based on Euclidean distance, which we adopted from the experiment by Ishizaka et al. [2]. These five hypotheses are presented below, along with the interpretation proposed by their authors.

Hypothesis 1 Euclidian distance $AD >$ Euclidian distance BD—the ranking B is nearer to the ranking D than A is to D.

Hypothesis 1 implies that the process of filling in the AHP matrices brings the participants closer to the final D ranking. This would mean that the process itself can help individuals improve their final ranking.

Hypothesis 2a Euclidian distance $AC >$ Euclidian distance CD—the ranking D is nearer to the ranking C than A is to C.

Hypothesis 2b Euclidian distance $BC >$ Euclidian distance CD—the ranking D is nearer to the ranking C than B is to C.

Hypotheses 2a and 2b imply that in order to build ranking D, participants take into account AHP advice (C ranking) and modify their previous direct rankings A and B.

Hypothesis 3a Euclidian distance $AD <$ Euclidian distance CD—the ranking A is nearer to the ranking D than C is to D.

Hypothesis 3b Euclidian distance $BD <$ Euclidian distance CD—the ranking B is nearer to the ranking D than C is to D.

Hypotheses 3a and 3b mean that the first two direct rankings (A and B respectively) better reflect the preferences of the participants than the C (AHP) ranking.

Adopting the same hypotheses allowed us to compare the results of both experiments. However, it should be stressed that the experiments were not identical: they are concerned with different decision problems; in our experiment, the participants

defined the criteria themselves; we did not use monetary stimuli, and in addition, half of the participants in our experiment were presented with a manipulated AHP ranking.

Before presenting the results of the verification of the hypotheses, two issues still need to be addressed. The first is to confirm that the manipulation of AHP rankings, which we performed on half of the participants, actually distanced their C (AHP) rankings from A and B rankings (from 5.8 to 8.2 and from 3.9 to 6.7, respectively). The second concerns the verification of hypotheses. Since half of the participants had manipulated AHP rankings, the individual hypotheses had to be verified differently. When verifying hypothesis 1, we took into account all the participants because this hypothesis did not directly refer to the C (AHP) ranking. In turn, hypotheses 2a and 2b were verified separately for participants who were presented with actual AHP results—rankings C and those whose results were manipulated—rankings C'. The last two hypotheses 3a and 3b were verified only in relation to participants who were presented with actual AHP results. The verification of these hypotheses for participants with manipulated results is not justified because the manipulated rankings C' presented to them were more distant than their actual rankings C and these, in turn, were not presented to them.

The hypotheses were verified using the sign test. The results of verification are presented in Table 25.2. The results show that among the participants of the experiment, the process of filling in the AHP matrices itself had an impact on the final decision (hypothesis 1) and that both direct rankings (A and B) better reflected the preferences of the participants than the AHP ranking (hypotheses 3a and 3b, respectively). As for hypotheses 2a and 2b, the results show that the reformulation of the

Table 25.2 Verification of the research hypotheses (for comparison, the results obtained in the experiments by Ishizaka et al. [2] are presented at the bottom of the table)

Hypothesis		N	True	Indeterminate	False	p-value (%)
H1	$AD > BD$	64	**25**	**32**	**7**	**0.1**
H2 a	$AC > CD$	32	8	17	7	–
	$AC' > C'D$	32	**13**	**16**	**3**	**1.1**
H2 b	$BC > CD$	32	8	19	5	–
	$BC' > C'D$	32	**12**	**18**	**2**	**0.6**
H3 a	$AD < CD$	32	**16**	**9**	**7**	**4.7**
H3 b	$BD < CD$	32	**17**	**9**	**6**	**1.7**
Ishizaka et al. [2]						
H1	$AD > BD$	19	3	16	0	–
H2 a	$AC > CD$	19	**7**	**11**	**1**	**3.5**
H2 b	$BC > CD$	19	**6**	**13**	**0**	**1.6**
H3 a	$AD < CD$	19	9	7	3	–
H3 b	$BD < CD$	19	**10**	**6**	**3**	**4.6**

Statistically, significant hypotheses (p-value < 5.0%) are marked bold

Table 25.3 Answers to the question "To what extent did the AHP ranking reflect your preferences?" according to the group of participants

Group of participants	Reflected	To some extent	Did not reflect
With actual AHP ranking	20	7	5
With manipulated AHP ranking	18	12	2

final decisions under the influence of the AHP ranking took place only among those participants of the experiment who were presented with manipulated results.

Finally, we analyzed the answers to the question "To what extent did the AHP ranking reflect your preferences?" given by the participants at the end of the experiment. Table 25.3 shows separately the responses given by the participants who were presented with the actual AHP ranking and those who were presented with the manipulated one. After performing Yate's correction, the chi-squared showed that the answers for the two groups of participants were not significantly different ($\chi^2 = 1.440; p = 0.258$).

25.5 Discussion and Conclusions

In discussing the results of our experiment, we will refer to the experiment from the paper by Ishizaka et al. [2, which further will be called in short the experiment-2011]. Such a comparison, having regard to the aforementioned differences between these two experiments, allows us to formulate a few additional remarks and conclusions.

First of all, the decision problem we chose did not arouse such concentration in the participants as was the case with the experiment-2011. It is clearly visible both in the higher sums of the distances between the direct rankings and the higher values of the consistency ratio (Fig. 25.1). However, taking into account the positive opinions of the participants expressed at the end of the experiment, it seems that the reason for these differences was the use of monetary stimuli in the experiment-2011. If this is the case, it is an indication of the effectiveness of experimental economics methods.

This lower degree of concentration of the participants in our experiment may be helpful in explaining why hypothesis 1, which assumes that the process of filling in the AHP matrices brings the participants closer to the final decision, was confirmed in our experiment but not in the experiment-2011. If we look at the average distances between the successive direct rankings *A*, *B,* and *D* (Fig. 25.2), we will notice that the main reason why the sums of these distances were higher in our experiment is the initial ranking *A*. This lends credence to scenario in which our participants (or at least some of them) were less concentrated at the beginning of the experiment and formulated their initial ranking quite rashly. It was only when they made a series of pairwise comparisons of both the criteria and the alternatives against each criterion that they started to think about their preferences in more detail. This suggests that in

some cases, the process of filling AHP matrices itself can have a significant impact on a better understanding of our preferences and, consequently, on our decisions.

Both experiments confirmed hypothesis 3b, according to which the direct ranking *B* better reflected the preferences of the participants than the AHP ranking. In addition, in our experiment, this was also the case for the initial ranking *A* (hypothesis 3a). It can be presumed that in both experiments the reason for the poor quality of AHP in reflecting preferences was the low percentage of participants (19% and 31%, respectively) who filled the AHP matrices consistently ($Cr < 0.1$). However, this conclusion contradicts the results of the experiment carried out by Linares [4], which showed that after removing the inconsistency, the participants' preferences were not better represented. In order to determine this issue, we undoubtedly need further experiments and new ideas for conducting them. On the other hand, both experiments showed that it is quite difficult for individuals to make consistent pairwise comparisons, regardless of their level of involvement.

Finally, let us address the question of what really prompts us to accept AHP results and reformulate our preferences. In trying to answer this question, we have obtained quite surprising results, since hypotheses 2a and 2b have only been confirmed for those participants who were presented with manipulated AHP rankings. Unfortunately, we were unable to find clear reasons why the participants who were presented with their actual AHP rankings did not include them in their final decisions, as was the case in the experiment-2011. However, this does not change the fact that participants with manipulated AHP rankings included them in their final decisions, although they were more distant from *A* and *B* rankings than their actual AHP rankings. Moreover, their assessment of the extent to which the AHP ranking reflected their preferences did not differ from the assessments of those to whom their actual AHP rankings were presented. Is there a halo effect due to the scientific nature of the AHP method? Or perhaps we are dealing here with an anchoring heuristics, as AHP ranking was the last thing the participants saw on the screen before making the final decision? Our experiment does not provide answers to these questions, but we hope it prompts a broader study of the issue.

References

1. Cypryjański, J., Grzesiuk, A.: Expressing our preferences with the use of AHP: the game is not worth the candle? In: Nermend, K., Łatuszyńska, M. (eds.) Problems, Methods and Tools in Experimental and Behavioral Economics. Springer Proceedings in Business and Economics, pp. 155–165. Springer, Cham (2018)
2. Ishizaka, A., Balkenborg, D., Kaplan, T.: Does AHP help us make a choice? An experimental evaluation. J. Oper. Res. Soc. **62**(10), 1801–1812 (2011)
3. Ishizaka, A., Siraj, S.: Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. Eur. J. Oper. Res. **264**(2), 462–471 (2018)
4. Linares, P.: Are inconsistent decisions better? An experiment with pairwise comparisons. Eur. J. Oper. Res. **193**(2), 492–498 (2009)

5. Nermend, K.: The implementation of cognitive neuroscience techniques for fatigue evaluation in participants of the decision-making process. In: Nermend, K., Łatuszyńska, M. (eds.) Neuroeconomic and Behavioral Aspects of Decision Making. Springer Proceedings in Business and Economics, pp. 329–339. Springer, Cham (2017)
6. Pöyhönen, M., Hämäläinen, R.P.: On the convergence of multiattribute weighting methods. Eur. J. Oper. Res. **129**(3), 569–585 (2001)
7. Ziemba, P., Wątróbski, J., Jankowski, J., Piwowarski, M.: Research on the properties of the AHP in the environment of inaccurate expert evaluations. In: Nermend, K., Łatuszyńska, M. (eds.) Selected Issues in Experimental Economics. Springer Proceedings in Business and Economics, pp. 227–243. Springer, Cham (2016)
8. Brunelli, M.: Introduction to the Analytic Hierarchy Process. Springer Briefs in Operations Research. Springer, Cham (2015)
9. Emrouznejad, A., Marra, M.: The state of the art development of AHP (1979–2017): a literature review with social network analysis. Int. J. Prod. Res. **55**(22), 6653–6675 (2017)
10. Gass, S.: Model world: the great debate—MAUT versus AHP. Interfaces **35**(4), 308–312 (2005)
11. Ishizaka, A., Labib, A.: Review of the main developments in the analytic hierarchy process. Expert Syst. Appl. **38**(11), 14336–14345 (2011)
12. Mu, E., Pereyra-Rojas, M.: (2017) Practical Decision Making. An Introduction to the Analytic Hierarchy Process (AHP) Using Super Decisions. Springer Briefs in Operations Research, vol. 2
13. Saaty, T., Vargas, L.: Models, Methods, Concepts & Applications of the Analytic Hierarchy Process, 2nd edn. Springer, New York (2012)
14. Saaty, T.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
15. Krejči, J.: Pairwise Comparison Matrices and Their Fuzzy Extention. Multi-criteria Decision Making with a New Fuzzy Approach. Studies in Fuzziness and Soft Computing. Springer (2018)
16. Krejči, J., Pavlačka, O., Talašova, J.: A fuzzy extension of analytic hierarchy process based on the constrained fuzzy arithmetic. Fuzzy Optim. Decis. Making **16**(1), 89–110 (2017)
17. Millet, I.: The effectiveness of alternative preference elicitation methods in the analytic hierarchy process. J. Multi-Criteria Decis. Anal. **6**, 41–51 (1997)

Chapter 26
Declared Preference Method—Research Experiment Concerning Air Quality

Małgorzata Burchard-Dziubińska, Elżbieta Antczak and Agnieszka Rzeńca

Abstract Externalities take place when an economic operator does not experience all costs (in case of negative externalities) or benefits (in case of positive externalities) involved in his activities, which are passed on to other economic operators. Specific challenge to economists is posed by negative technological externalities, which have also triggered experimental research studies using the declared (stated) preference method. The aim of the experiment was to find the most preferred solution geared towards the elimination of the source of air pollution in a village and to learn about respondents' willingness to pay (WTP) for the public good supplied as a result of such elimination. Out of the five proposed options, one of the most preferred by respondents was a *voluntary fundraising initiative to collect money for the construction of the neighbour's connection that would go on until they collect PLN 10 k but not longer than for a year.* By using the WTP method, we arrived at an unambiguous valuation of the good expressed in terms of money. The most preferred tax rate revealed as a result of the survey suggests that the valuation of the public good was underestimated. As a result, the situation described in the case study cannot be resolved successfully.

Keywords Public goods · Externalities · Declared preference method · Air pollution

26.1 Introduction

Today, quite rightly, attention is clearly being paid to the role of human behaviour as a significant source of environmental pollution. In Poland, we used to blame industry

M. Burchard-Dziubińska (✉) · E. Antczak · A. Rzeńca
Faculty of Economics and Sociology, University of Lodz, Lodz, Poland
e-mail: malgorzata.burchard@uni.lodz.pl

E. Antczak
e-mail: elzbieta.antczak@uni.lodz.pl

A. Rzeńca
e-mail: agnieszka.rzenca@uni.lodz.pl

K. Nermend and M. Łatuszyńska (eds.), *Experimental and Quantitative Methods in Contemporary Economics*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-30251-1_26

for air pollution and for the contamination of groundwater and soil. Currently, as a result of restrictive environmental policy equipped with legal and economic instruments, industry has lost its top position of the major polluter. Poor quality of the environment and, above all, air pollution is caused by individual human activities. Low emission linked with suburbanisation and many years of lagging behind when it comes to the standards of heat and electricity supply systems available to households are the main reasons why all users of the environment experience tangible negative effects. Due to air pollution, all people who breathe it become victims of the polluters while, looking from the opposite perspective, those who are working to eliminate air pollution provide benefits to other individuals who become covered by their protection activities. Air pollution and its protection are externalities with the range of impact covering all users of the environment. No one may be easily excluded from the circle of either victims or beneficiaries.

More precisely, externalities take place when an economic operator does not experience all costs (in case of negative externalities) or benefits (in case of positive externalities), which are passed on to other economic operators. The division into positive and negative externalities is normative by nature. Negative externalities manifest the presence of public anti-goods (public bads), such as, e.g., air pollution. On the other hand, positive externalities are sources of public goods, such as, e.g., protection of the atmosphere. Common features of public bads and goods are: non-rivalry nature of their consumption and inability to exclude anyone from their consumption. Assuming that overall benefits prevail over the disadvantages, we may say that the carrying out protection activities is a kind of public good linked with positive externalities. However, when trying to ensure adequate supply of protection efforts, we may face two serious economic problems: the *free rider problem*, i.e. using the good for free and the *prisoner's dilemma*, which in environmental context means that the lack of communication and ability to agree on a common strategy leads to a situation when all parties make a rational choice of not pursuing any protection activities although working in polluted environment deteriorates their individual performance and reduces the welfare.

Specific challenge to economists is posed by negative technological externalities. They have also triggered experimental research studies using the declared (stated) preference method. Usually, the declared preference method was used in connection with the polluter pays principle and the willingness to pay environmental costs, when the market does not generate adequate information about the issue. The proposed experiment concerns "pollution elimination at source" principle and civic support to an environmentally friendly investment project launched to produce common good, i.e., clean air.

The main objective of the undertaken research was to identify the tendency of environmental users, members of the local community to bear the costs of environmental protection investments and to determine the level of declared expenditures. This hypothetical case concerns members of a local community exposed to negative externalities of activities pursued by one of their neighbours. The undertaken research in the field of experimental economics is an attempt to search for effective solutions for environmental protection and to determine the dependence between the

effectiveness of undertaken actions and the declared willingness to pay for a good. An important aspect and background of these considerations were also the issue of the co-responsibility of citizens for the quality of the environment.

26.2 Public Goods and Issues of Demand and Supply: Literature Review

How to identify the size of output of a public good is a theoretical but also practical question. Problems arise when we want to measure preferences and have to deal with congestion. For public goods, equilibrium is achieved when the overall willingness to pay for the good equals the price for which the producer is willing to supply this good and different individuals are willing to pay different prices for a certain output of the good. These individual prices are referred to as Lindahl pricing (or Lindahl taxes). Effective valuation of a public good assumes that the sum of individual prices equals the marginal cost [14]. In literature, we may come across diverse ways of finding the demand curve for public good. According to Musgrave, the sum of individual prices equals the marginal cost [8]; Bowen [2] and Samuelson [15] argue that the sum of marginal rates of substitution equals the marginal cost while Buchanan [6] claims that the sum of marginal values equals marginal cost. Musgrave assumed that public and private needs are "individual wants." The principal difference between the two needs consists in the fact that "goods and services supplied to satisfy public needs must be consumed in equal amounts by all" [13]. There are two implications of such an approach. Firstly, if a voluntary solution was available, the aggregate demand curve for inseparable good would be the effect of vertical addition of the individual demands. The result may in a way be obtained from the Lindahl model and was identified against these categories by Bowen [2], and repeated by Samuelson [15]. Secondly, "because the same amount will be consumed by all, everybody knows they cannot be excluded from the benefits stemming from that." Thus, people are not forced to reveal their preferences by offering their bids in the market. The "exclusion principle," which is crucial for the exchange, may not be applied and the market mechanism does not work. To a consumer, it is always more attractive to hide or falsify her/his true preferences not to reveal her/his actual willingness to pay for public good. In accordance with Lindahl concept, anyone may be asked to specify her/his demand for public good at a hypothetical price or when faced with the distribution of tax burden involved in its supply. Prices for the same amount of a given public good may differ. It allows, at least theoretically, arriving at equilibrium. In practice, every consumer wishes to pass the costs of the supply of the good on to others by declaring demand lower than the actual one because (s)he assumes the good will be supplied to her/him anyway. Lindahl model does not eliminate the problem of a free rider. By being unable to exclude anyone from consumption, we may not find out the real value of the public good to individuals expressed in the price they are willing to pay. Investigating into preferences is a complex exercise but it does not

mean that economists are totally helpless. Literature provides a number of examples of mechanisms that force consumers to reveal their true preferences and prices linked with them. Vickrey–Clarke–Groves mechanism (VCG) is one of such examples. This auction mechanism helps in achieving a socially optimal solution through its strategy consisting in bidding for a realistic valuation. It is assumed that the cost of a project is given. The project should be launched if the sum of values declared by all citizens is higher than the cost or equal to it. In case of the VCG mechanism with the Clarke pivot rule, it means that a citizen pays a non-zero tax for the project if and only if it is pivotal, i.e. without her/his declaration the total value is less than the given cost and with this declaration, the total value is more than the cost. This mechanism is designed to discourage making unreliable declarations. The pivotal subject is charged with tax burden. Such tax is referred to as motivational [16]. An alternative solution is offered by Groves–Ledyard mechanism. In this case, the output of public good is decided by a central agency (authority) while expectations of individual subjects either increase or reduce it. Everybody sends their choices as to the amounts of the good without knowing declarations made by other participants. Tax paid by an individual depends on how much his expectations concerning the amount of the good in question deviate from the mean of other players' bids [4]. A real problem consists in the fact that as a result of the willingness to lower the declared prices, demand will be underestimated meaning the output of the public good will be insufficient or will not take place at all.

To learn if we are able to supply a given public good, it is worth conducting an experiment, which explores human behaviour in a near real-life environment. For that purpose, we may conduct a survey in laboratory conditions (on an isolated group). For questions about the valuation of a public good, it is justified to apply the conditional valuation method, in which we either want to learn about the willingness to pay (WTP) or the willingness to accept (WTA) [10, 11]. The willingness to pay is measured through the deployment of diverse techniques, among which the most popular include questionnaire studies or interviews. Account is taken of social and economic aspects of the examined population, such as, e.g.: types of households, domicile, education, age, gender and income per person in a household. It enables matching the answer to the question concerning the valuation with other features typical of the respondent. In the case of the WTP/WTA method, we want to arrive at a clear-cut valuation expressed in money, which helps in drawing the demand curve for the good covered by the survey. By using the declared preference method, we may also ask about the most desired ways, in which the public good in question could be provided and/or about the internalisation of negative externalities. It was applied in the experiment discussed in this chapter.

26.3 Methodology and Data

Consumer preferences can be measured based on historical observations of consumer behaviour in the market, i.e. on real decisions taken by consumers faced with a choice or based on data that describe consumer intentions. Because of these two sources of

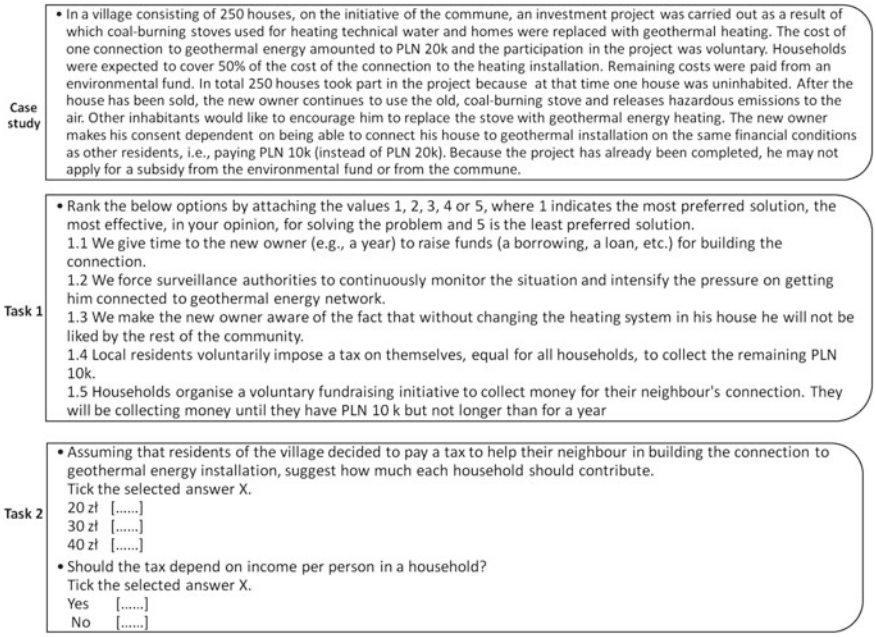


Fig. 26.1 Logical structure of the experiment

information about preferences, we may distinguish two methods, one of the analyses revealed preferences while the other one focuses on declared (stated) preferences (Table 26.1). Despite a wide spectrum of applications for both methods, measuring preferences of natural environment users, in which we analyse the environment and its quality, is a special case. The use of declared preferences method, its conditions

Table 26.1 Comparison between revealed and declared preference methods

Method	Revealed preference method	Declared preference method
Goal	Finding out about real market decisions	Finding out about declared consumer behaviour and attitudes (a hypothetical situation)
Time	Measurement based on historical observations of consumer behaviour in the market	Measurement made using methods that reveal consumer attitudes during the survey
Assumptions	Investigating into revealed preferences concerns real choices that have already been made The major obstacle is the fact that the method cannot be applied to hypothetical situations	The survey uses respondent individual experience as to their preferences vis-à-vis a set of options The limitation in this case lies in the assumption that respondents behave in a rational manner and maximise the utility

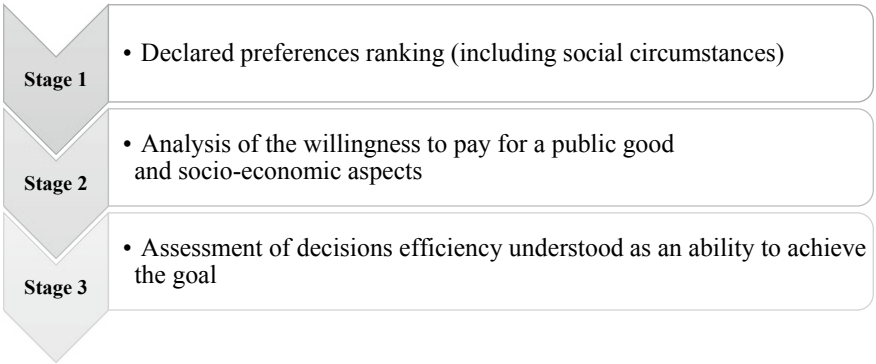


Fig. 26.2 Stages of experiment. *Source* Authors' own compilation

and limitations have been widely discussed in literature [1, 5] also in the context of their stability [3].

Based on a formalised questionnaire, the participants to the research experiment are expected to find a solution to the problem described in Fig. 26.1. Their choices made within the framework of hypothetical scenarios help us draw conclusions about their preferences and value of non-market goods. By making a decision, the participants allocate a particular degree of utility to each attribute. Thus, we assume that they select an option representing the maximum total utility to them. It is presumed that the questionnaire developed for the experiment covers all factors that respondents might consider when making a decision. As a result of the experiment, we will learn about preferred solutions and how they link with respondents' socio-economic attributes.

For the sake of reliability of the experiment, we validated the research tool by conducting a pilot survey. A survey on a group of 15 people helped us to fine-tune the experiment, to precisely set border conditions and to develop the final version of the questionnaire.

Finally, the experiment covered all first-year students of full-time undergraduate and graduate courses at Spatial Economics (Polish: Gospodarka Przestrzenna) at the Faculty of Economics and Sociology of the University of Lodz. We purposefully selected the students of this particular course as respondents. Over the time of their studies, these students will improve their knowledge and skills because the teaching content includes, inter alia, environmental economics, environmental protection, eco-innovation, local development policy, spatial development and legal framework for environmental protection. It means they will be able to validate results of the experiment in the future when they will have completed their education. The research study was anonymous and organised as a paper-based questionnaire study. Questionnaires were handed out directly to students who were asked to fill them out. We managed to collect back 90 questionnaires. The ultimate rate of return was 70% ($N = 86$ individuals). Some questionnaires were invalidated because they were filled out incorrectly or incompletely.

Collected research material helped us to process and analyse data in three stages (Fig. 26.2). All results of analyses were carried out using graphic and statistical tools available in the IBM SPSS Statistics software.

26.4 Results of the Survey

26.4.1 Declared Preference Ranking (Including Social Circumstances)

Students in a research sample represent small cities—up to 20 k residents (13%); average cities—between 20 k and 100 k residents (21%); large cities—a population over 100 k residents (24%) and rural areas (42%). Among respondents, 60% were women and 40% men.

According to respondents, an optimal solution to the problem in question would be a fundraising initiative through which local residents would voluntarily donate money to build a connection for their neighbour. This solution was preferred by most respondents, 34%, with 30% believing in its effectiveness (Option 1.5). About 30% students preferred to give time to the new owner to himself raise resources for building the connection (Option 1.1). This solution is seen as the second most effective to the problem. Another top-ranking way of helping the neighbour in funding the investment is self-taxation where all residents would pay the same amount (per household) to collect the remaining PLN 10 k (1.4). About 35% students decided it is an effective solution and 16% believed it is very effective. In turn, only 13% respondents declared they would force surveillance authorities to continuously monitor the situation and intensify pressure on building the connection (1.2). According to respondents making the new neighbour aware that without changing the source of heating, he will not be liked by the rest of the community was the least promising idea (1.3), 7% respondents assessed it as very effective (Fig. 26.3).

Statistically, significant differences have been observed ($p = 0.10$) in the number of men and women who decided that a voluntary fundraising initiative was the best solution to the problem (35% female students and 32% male students gave preference to this option). Moreover, 2% female and 12% male students considered the above solution little effective (Fig. 26.4). 27% female students and 35% male students would give time to the new owner to raise funds for the construction of the connection ($p = 0.72$). No significant differences were reported in the number of women (19%) and men (24%) who believed the above solution is ineffective. At the same time, the assessment of effectiveness of self-taxation option is significantly different for women and men ($p = 0.09$). Statistically, more women (54%) than men (27%) preferred this solution as a way of preventing air pollution. Besides, significantly more men (15%) than women believe self-taxation is ineffective with only 2% women decisively not supporting the initiative (Fig. 26.4).

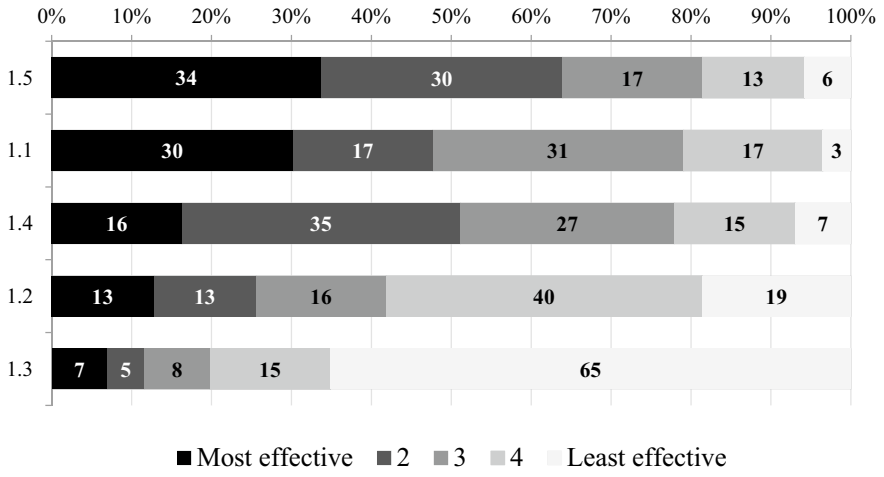


Fig. 26.3 Ranking of options in terms of their effectiveness. 1.1—We give time to the new owner to raise funds for the construction of the connection; 1.2—We force surveillance authorities to constantly monitor the situation and intensify pressure on the connection to geothermal installation; 1.3—We make the new owner aware that without changing the source of heating for his house, he will not be liked by other residents; 1.4—Residents of the village impose a tax on themselves, equal for all households, to collect the remaining PLN 10 k; 1.5—Households organise a voluntary fundraising initiative to collect money for their neighbour's collection. They will go on until they collect PLN 10 k but not longer than for a year; $N = 86$

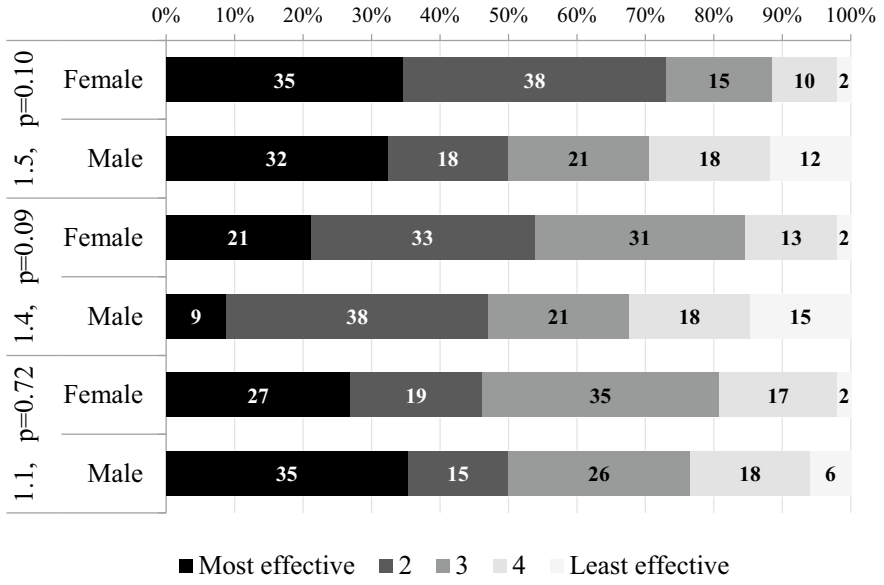


Fig. 26.4 Relationship between effectiveness assessments of the most preferred solutions and respondent gender. Female, $n = 52$, men, $n = 34$; we conducted the Chi-square tests of independence for significance level $\alpha = 0.10$ [7]

Results of the experiment show that residents of middle-sized cities assessed the solution consisting in giving time to the new owner to raise funds for the construction of the connection as significantly less effective ($p = 0.03$). Half of the students permanently domiciled in cities whose population ranges from 20 to 100 k inhabitants consider this option moderately effective and 28% believe it is not effective. Only 22% respondents claim such assistance could be effective in the struggle for air quality protection. On the other hand, clear majority of respondents from small and large cities as well as rural areas argue that the new owner should be given time to raise funds for the construction of the connection. Residents of medium-sized cities assessed the effectiveness of self-taxation, equal amount per household, to collect PLN 10 k significantly higher ($p = 0.10$). About 40% students from cities with population ranging between 20 and 100 k believe this is an effective solution and only 9% respondents permanently residing in small and large cities as well as in rural areas would prefer this solution (Fig. 26.5). No statistically significant differences

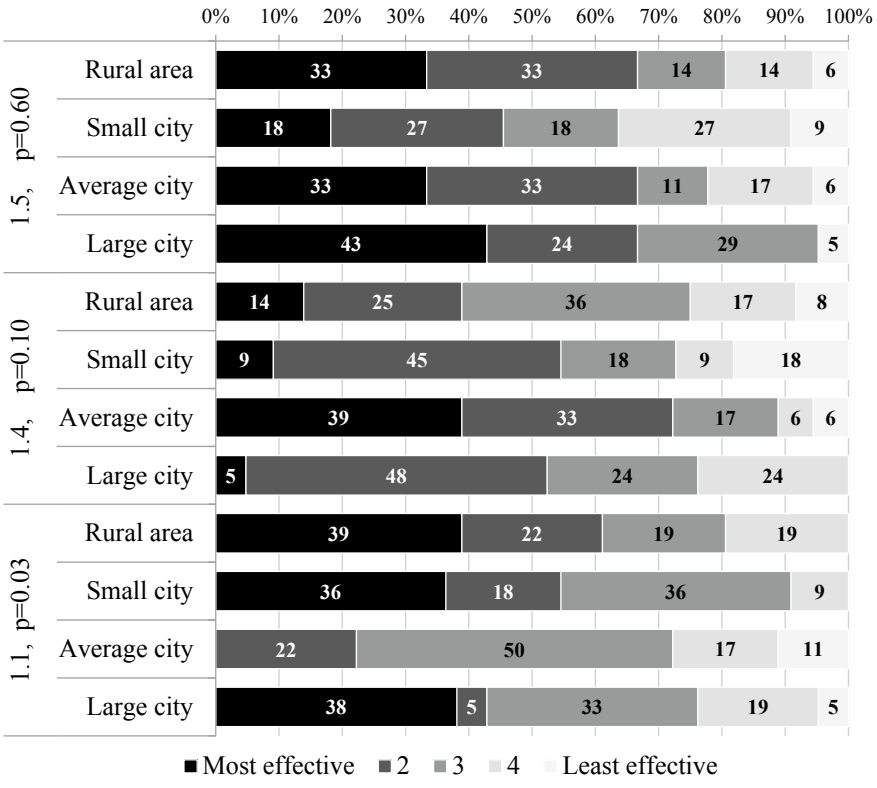


Fig. 26.5 Relationship between assessment of effectiveness of a solution and place of permanent residence. Rural area, $n = 36$; small city, $n = 22$; average city, $n = 18$; large city, $n = 21$. Detailed Fisher independence tests were carried out for the level of significance $\alpha = 0.10$, because over 20% cells have expected counts lower than five and the minimum expected count is smaller than one [9]

were reported between respondents' domicile and the assessment of effectiveness of the solution providing for a voluntary fundraising initiative to collect money for the neighbour's connection ($p = 0.60$). The smallest group of respondents from small cities, 18%, considered this form of assistance to the new owner highly effective. A clear majority, 44% students from large cities, preferred this very option. In turn, on average, 7% respondents permanently residing in locations indicated in the questionnaire assumed the solution is ineffective (Fig. 26.5).

26.4.2 Analysis of the Willingness to Pay for a Public Good and Socio-economic Aspects

Assuming that residents decided to impose a tax on themselves to financially support the construction of their neighbour's connection, the majority, as many as 41% respondents, were willing to contribute PLN 20. Slightly fewer, 37% respondents (out of 86), declared to pay PLN 30 per household. The rest of students (22%) opted for a tax amounting to PLN 40 (Fig. 26.6). Moreover, to the question, if self-taxation payment should depend on per capita income in a household, most respondents answered "yes" (66% out of 86 in the sample).

We also observed that significantly more male students (35% out of 34 respondents) than female students (13% out of 52 female respondents) are willing to pay PLN 40 ($p = 0.01$). In turn, clearly more women than men who were willing to pay the tax would like to pay PLN 30 (48% female students and 21% male students). Also, slightly more men than women declared the willingness to pay PLN 20 to subsidise the construction of their neighbour's connection (44% male students and 38% female students). Nevertheless, the difference in amounts amongst respondents is clearly the smallest (6% points) compared to the other two declared tax rates (Fig. 26.7).

Moreover, independently of the declared tax rate (20 PLN, 30 PLN or 40 PLN), a clear majority of respondents opted for the rate to be adjusted to per capita income ($p = 0.15$) (Fig. 26.8).

Tax rate declared by respondents is independent of their domicile ($p = 0.37$). Circa 40% students were willing to pay PLN 20 in the form of tax to help the neighbour

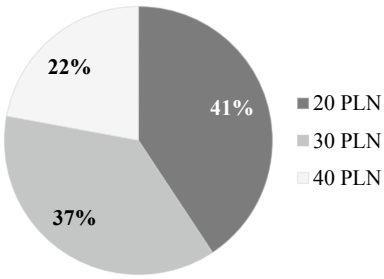


Fig. 26.6 Structure of declared amount of tax imposed by residents on their households

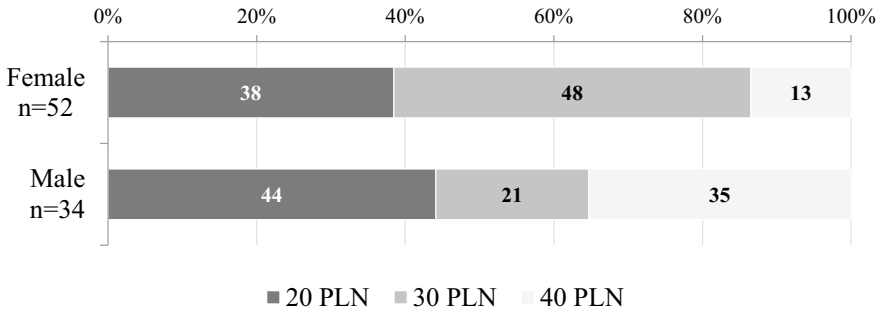


Fig. 26.7 Relationship between declared amount in self-taxation option and respondent gender. The Chi-square test of independence was carried out for the level of significance $\alpha = 0.10$

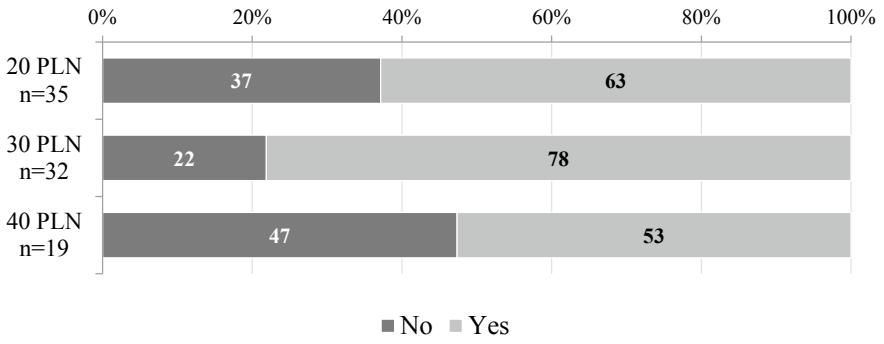


Fig. 26.8 Relationship between tax rate and its dependence on per capita income in a household. The Chi-square test of independence was carried out for the level of significance $\alpha = 0.10$

to construct the connection independently of their permanent place of residence. We also observed that a similar number of students from rural areas, small and large cities declared to pay the tax of PLN 30 (ca. 40% respondents). The majority, 39% students residing permanently in medium-sized cities, opted for tax rate PLN 40. Nevertheless, the latter amount was the least often declared. This is why despite differences between the place of residence and this category of tax, they were not statistically significant (at adopted level of significance) (Fig. 26.9).

26.4.3 Assessment of Decisions Efficiency Understood as an Ability to Achieve the Goal

Results of the experiment suggest significant differentiation in amounts declared in the self-taxation exercise intended to subsidise the construction of neighbour's connection stems from the assessment of the effectiveness of the tool ($p = 0.08$),

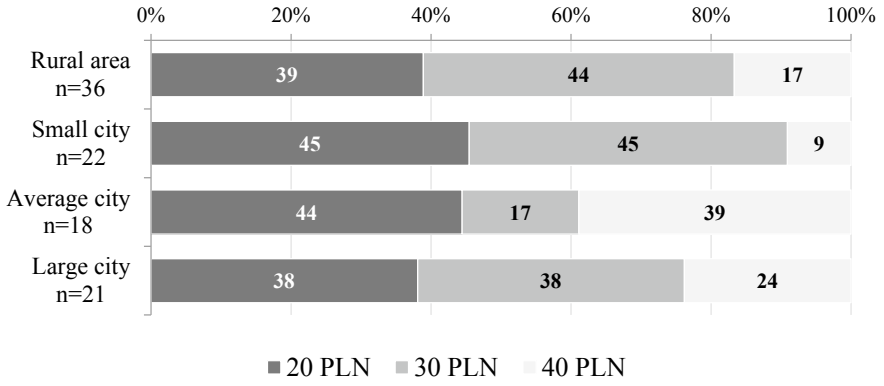


Fig. 26.9 Declared self-taxation rate and respondent domicile. We conducted detailed Fisher tests of independence for the significance level $\alpha = 0.10$, because over 20% cells have expected counts lower than five, and the minimum expected count is smaller than one

Fig. 26.10. Amongst respondents who declared the highest tax rate (40 PLN) as many as 47% believe, it is an effective tool and 32% consider it very effective (in total as many as 79% opt for this solution). At the same time, the group of respondents who declared the lowest tax rate (20 PLN) and believe this is an effective solution is significantly smaller, i.e. 43% (out of them only 9% claim this is the most effective solution and 34% believe it is effective). Similarly, 44% respondents who declared PLN 30 assessed this option as effective and very effective, and 16% out of this group believe self-taxation is the best solution to the problem (Fig. 26.10).

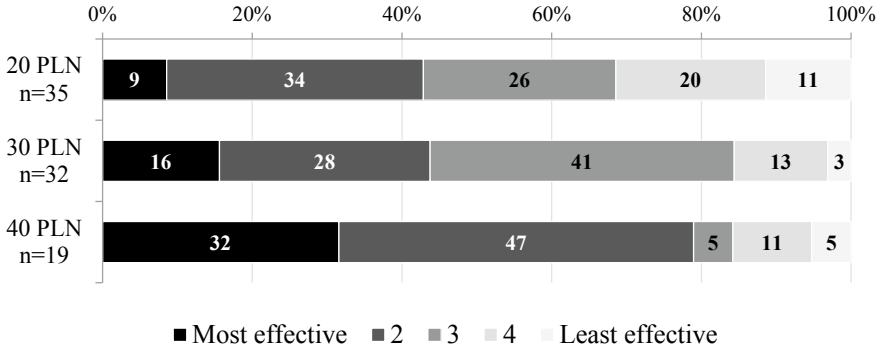


Fig. 26.10 Relationship between amounts declared for self-taxation and assessment of effectiveness

26.5 Discussion

Together with the increase in the tax rate, respondents' willingness to pay decreases (Fig. 26.6). However, the bigger the willingness to pay, the more likely the goal will be achieved (Fig. 26.10), which finally translates into increased willingness to pay more for clean air, i.e. demand for the good in question (Fig. 26.11). Thus, more effective methods of working out the common public good encourage respondents to pay higher rates (the value of β parameter is the highest in the function describing respondents' willingness to pay the highest declared rate, Fig. 26.11).

Considering assumptions adopted for the experiment and information collected from respondents, we made an attempt to extrapolate the results to the population (250 households sharing the cost of thermo-modernisation). By the same token, we estimated the subsidy offered to the construction of the neighbour's connection (Table 26.2). The willingness to pay a specific amount declared by the residents, their preferences and socio-economic circumstances would bring in PLN 7025. It means the investment was underestimated to the amount of PLN 2975 (10,000 PLN-7025 PLN).

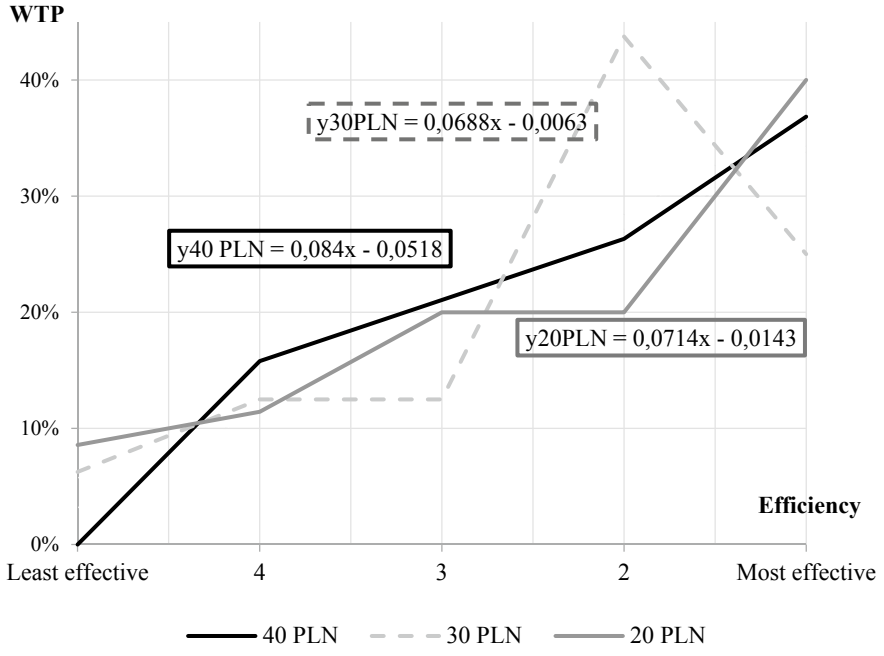


Fig. 26.11 Relationship between the effectiveness of proposed activities and declared willingness to pay for the good. *Note* The chart presents estimated equations of the linear trend models expressed in general terms as: $y = \alpha + \beta x$, where positive β represents a continuous increase in the value of analysed variable [12]

Table 26.2 Declared willingness to pay and financial resources collected for the connection

Demand (in number of households)	WTP in PLN	Unit value of the project in PLN (Demand * WTP)
103	20	2050
93	30	2775
55	40	2200
Total value of the project: 7025 PLN		

26.6 Conclusions

The experiment was designed to find the most preferred solution to the problem consisting in the elimination of the source of air pollution in a village and to learn about residents' willingness to pay for the public good supplied as a result of such elimination. The most preferred option was a *voluntary fundraising initiative to collect money for the construction of the neighbour's connection that would go on until they collect PLN 10 k but not longer than for a year* (1.5). Next was option 1.1 *We give time to the new owner (e.g., a year) to raise funds (a borrowing, a loan,* etc.*) for building the connection,* followed by options: 1.4 *Local residents voluntarily impose a tax on themselves, equal for all households, to collect the remaining PLN 10 k.;* 1.2 *We force surveillance authorities to continuously monitor the situation and intensify the pressure on getting him connected to geothermal energy network,* and 1.3 *We make the new owner aware of the fact that without changing the heating system in his house he will not be liked by the rest of the community* (Fig. 26.3).

As a next step, to arrive at more unambiguous value of the good expressed in terms of money, we used the WTP method. We need to bear in mind that if all households joined the self-taxation initiative, tax rate of PLN 40 would bring the lacking PLN 10 k and solve the problem. However, it turned out that the lowest rate, PLN 20, was the most preferred one (Fig. 26.6; Table 26.2), meaning the valuation of the public good was underestimated. It may mean that the project will not be accomplished in the voluntary taxation option.

References

1. Baker, R., Ruting, B.: Environmental policy analysis: a guide to non-market valuation. In: Productivity Commission Staff Working Paper, Canberra. Available via Media and Publications. https://www.pc.gov.au/research/supporting/non-market-valuation/non-market-valuation.pdf. Accessed 04 Apr 2019 (2014)
2. Bowen, H.R.: The interpretation of voting in the allocation of economic resources. Q. J. Econ. **58**, 27–48 (1943). https://doi.org/10.2307/1885754. Accessed 04 Jan 2019
3. Brouwer, R.: Constructed preference stability: a test–retest. J. Environ. Econ. Policy **1**, 70–84 (2012). https://doi.org/10.1080/21606544.2011.644922. Accessed 04 Jan 2019
4. Brown, C.V., Jackson, P.M.: Public Sector Economics. Blackwell, London (1992)

5. Brown, T.C.: Introduction to stated preference methods. In: Champ, P.A., Boyle, K.J., Brown, T.C. (eds.) A Primer on Nonmarket Valuation. The Economics of Non-market Goods and Resources, vol. 3, pp. 99–110. Springer, Dordrecht (2003)
6. Buchanan, J.: The demand and supply of public goods. In: Buchanan, J.M. (ed.) The Collected Work, vol. 5, p. 32. Liberty Fund, Indianapolis (1999)
7. Cochran, W.G.: The chi-square goodness-of-fit test. Ann. Math. Stat. **23**(3), 315–345 (1952)
8. Desmarais-Tremblay, M.: On the Definition of Public Goods. Assessing Richard A. Musgrave's contribution. Documents available via. de travail du Centre d'Economie de la Sorbonne. https://halshs.archives-ouvertes.fr/halshs-00951577/document. Accessed 04 Apr 2019 (2014)
9. Fisher, R.A.: Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh (1934)
10. Hanemann, W.M.: Valuing the environment through contingent valuation. J. Econ. Perspect. **8**(4), 19–43 (1994)
11. Loomis, J., Brown, T., Lucero, B., Peterson, G.: Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay. Land Econ. **72**(4), 450–461 (1996)
12. Maddala, G.S.: Introduction to Econometrics. Macmillan, New York (1992)
13. Musgrave, R.A.: A multiple theory of budget determination. Finanzarchiv **17**(3), 333–343 (1957)
14. Roberts, D.J.: The Lindahl solution for economies with public goods. J. Public Econ. **3**(1), 23–42 (1974)
15. Samuelson, P.A.: Diagrammatic exposition of a theory of public expenditure. Rev. Econ. Stat. **37**(4), 350–356 (1955)
16. Żylicz, T.: Ekonomia środowiska i zasobów naturalnych (Economics of the environment and natural resources) [in Polish]. PWE, Warszawa (2004)