STATISTIKA

STATISTICS AND ECONOMY JOURNAL

VOL. **100** (3) 2020

CZECH STATISTICAL OFFICE

CONTENTS

ANALYSES

About Statistika

The journal of Statistika has been published by the Czech Statistical Office since 1964. Its aim is to create a platform enabling national statistical and research institutions to present the progress and results of complex analyses in the economic, environmental, and social spheres. Its mission is to promote the official statistics as a tool supporting the decision making at the level of international organizations, central and local authorities, as well as businesses. We contribute to the world debate and efforts in strengthening the bridge between theory and practice of the official statistics. Statistika is professional double-blind peer reviewed open access journal included in the citation database of peer-reviewed literature **Scopus** (since 2015), in the **Web of Science** *Emerging Sources Citation Index* (since 2016), and also in other international databases of scientific journals. Since 2011, Statistika has been published quarterly in English only.

Publisher

The Czech Statistical Office is an official national statistical institution of the Czech Republic. The Office´s main goal, as the coordinator of the State Statistical Service, consists in the acquisition of data and the subsequent production of statistical information on social, economic, demographic, and environmental development of the state. Based on the data acquired, the Czech Statistical Office produces a reliable and consistent image of the current society and its developments satisfying various needs of potential users.

Contact us

Journal of Statistika | Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic
e-mail: statistika.journal@czso.cz | web: www.czso.cz/statistika_journal

What Can Be Learned on the Living Standard of Households from National Accounts?

Stanislava Hronová[1] | *University of Economics, Prague, Czech Republic*
Richard Hindls[2] | *University of Economics, Prague, Czech Republic*
Luboš Marek[2] | *University of Economics, Prague, Czech Republic*

Abstract

The living standard of households can be assessed with the aid of several indices. The standards of the System of National Accounts (SNA 1993 and ESA 1995) have introduced the concept of actual final consumption and social transfers in kind apart from the final consumption expenditure. However, the information which could be taken from these indices is utilized only to a very small extent. So, what can they actually tell us? Can they be used for assessing the living standards of households and their comparing on an international scale? The paper gives answers to these questions by means of comparing the data from the Czech Republic and France.

Keywords	***JEL code***
Actual final consumption, social transfers in kind, households, disposable income, living standard	*E21*

INTRODUCTION

The living standard in general, and the living standard of households, are comprehensive notions that are rather difficult to quantify; the reason for that stems from the use of different approaches to identify what the living standard actually is and which aspects (economic, social, environmental, etc.) should be included in it. Subsequently, a question is addressed of how to synthesise all those aspects under one index. Such a synthesis is hard or even impossible to carry out; hence similar attempts quite often result in simplified views and the indices such as the gross domestic product (hereinafter GDP) per capita or the national disposable income per capita are used. Both of these indices are based on the national accounts and are considered the basic aggregates whose relative (per capita) values are tools suitable

[1] Faculty of Informatics and Statistics, Department of Economic Statistics, University of Economics, Prague, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: hronova@vse.cz.
[2] Faculty of Informatics and Statistics, Department of Statistics and Probability, University of Economics, Prague, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: hindls@vse.cz.
[3] Faculty of Informatics and Statistics, Department of Statistics and Probability, University of Economics, Prague, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. E-mail: marek@vse.cz.

for comparing economic levels in different countries; but they cannot be looked upon as indices of households' living standard levels. The national accounts, and, in particular, the account of the household sector, contains additional indices that are much closer to the living standard concept than the two mentioned aggregates.

The GDP shows the outcome of the economic activities and does not reflect the living standard or the economic well-being; and neither do other indices on the national accounts. The primary goal of the national accounts is not to measure the living standard but to describe economic activities taking place in the given country and its relationships with abroad, using a system of interconnected indices. The national accounts have never been intended to express the living standard or well-being as its primary goal.

Objections frequently occur, stating that the GDP only expresses the outcome of economic activities, and these objections arise regularly.[4] Sometimes they are caused by efforts to reflect in the GDP damages on the environment, sometimes to express the population's satisfaction, social progress, economic performance of the given country or the well-being in general (conscious of the fact that it is not even clear what the well-being actually is). All such thoughts stem from misapprehension of the true purpose of the GDP – expressing the outcome of economic activities in a given territory. Hence it makes no sense to blame it for not providing the information about the living standard or even the well-being. Any efforts to adapt the GDP to reflect, say, the well-being, lead to artificial constructions that are far from the reality and have very little, or even nothing, in common with the mission the GDP has. Experts try hard to provide in the GDP an image of the economic activities' outcome that is as plausible as possible. The GDP is a basic aggregate whose design is based on a simple consideration: every productive activity generates an instance of value added, and its summary value is expressed by the GDP at the national-economy level. Because of its direct relationship with productive activities (income from production) and theoretical features (no duplicity occurs when summing up the value added amounts), the GDP is rightly viewed as a basic (even if not sole) aggregate on the national accounts. It is the root of all considerations aimed at generating, distributing and utilising income; other aggregate indices are derived from the GDP in compliance with the economic principles of the national accounts.

Due to its unambiguous definition and standardised estimation procedures, the GDP is suitable for international comparisons of a country's economic maturity. If this is the case, GDP per capita is used, expressed by the purchasing power parity; it is often presented as an indicator of the living standard. This view is undoubtedly simplified, but the point is that it depends on the definition of the living standard (similar to the importance of the definition of the well-being). If we relate the living standard or well-being of each inhabitant solely with the outcome of economic activities, this simplification may be acceptable. If a wider definition of the living standard or well-being is accepted, additional solutions become available – such as comparing the so-called actual consumption by households or selecting another suitable index, or a system of indices, directly concerning households.

[4] Let us, e.g., mention the Human Development Index – HDI of 1990, the Index of Sustainable Economic Welfare – ISEW of 1989, or the Genuine Progress Index – GPI of 1995. The Commission on the Measurement of Economic Performance and Social Progress (CMEPSP), generally referred to as the Stiglitz Commission, created by the French Government in the beginning of 2008 on initiative of the French President Nicolas Sarkozy is viewed as the most important activity in this direction. The work of this expert group in the area of economics and statistics should have resulted in suggestions for novel ways of quantifying the results of each individual country; these new ways should be more suitable (than the currently used indices) for expressing economic performance and social progress. The Commission consisted of 27 experts, mainly university economists from France, the USA, representatives of OECD, World Bank, and the United Nations, as well as eight Rapporteurs – French statisticians from INSEE, OECD, and OFCE. The Commission set out three areas in which it should find answers to the following questions: whether or not the GDP properly expresses the economic performance of a given country; whether (and how) social progress can be quantified; and whether mankind's development is sustainable from the economic, social and environmental viewpoints.

National disposable income (net or gross) is often presented as characteristic of the living standard or well-being that is more suitable than GDP because it is affected by redistribution processes. Nonetheless, the national disposable income is an outcome of the economic activities and the subsequent distribution and redistribution of income of all economic subjects, not only households (even though households' proportion in its value is substantial).[5] Distinguishing between the national disposable income and households' disposable income is therefore important because of not only the value, but – especially – of the content. The disposable income is the most important index for the household sector. It is an outcome of the distribution of the income coming from production, i.e., value added. Its most significant component is represented by labour income (wages and salaries); the additional ones include business income (operating surplus), and property income (or possibly the balance values of those components: in particular, interest and dividend values). The disposable income further includes social benefits, in particular old-age pensions, and other current transfers; its value is decreased by current taxes (income tax), social contributions, and other current transfers. In this respect, GDP and households' disposable income cannot, in general, be imposed upon one another. The GDP is an aggregate related to the national economy as a whole; but for the disposable income, while it is also monitored at the national-economy level, it does not make much sense to make use of the national value of the disposable income when measuring the population's living standard; households' disposable income should be used then. The latter can be related to the number of households, or even households' disposable income per consumption unit (equivalence scale).

The disposable income is understood, and designed on the national accounts, purely as an outcome of distribution. The fact that it does not reflect redistribution in kind is not detrimental because the national accounts also provide a concept of the adjusted disposable income; this is the disposable income adjusted to reflect the so-called social transfers in kind. In the case of households, the amount of the adjusted disposable income is higher than that of the disposable income by the value of the goods and services (education, health care, culture, etc.) paid in favour of households by the general government and the non-profit institutions serving households. The concept of adjusted disposable income was introduced by the United Nations national accounts standard SNA1993[6] and the EU standard ESA 1995.[7] The adjusted disposable income, and the related index of the actual final consumption, have, unfortunately, been tools rarely used in evaluating and comparing households' living standard levels. At the same time, those very indices better express the scope of households meeting their needs. The said indices came out of the system of national accounts; this fact enables us to evaluate and compare households' living standard levels in time and space. The present contribution is also aimed at comparative analysis based on the data taken from the household sector's account.

1 GENERAL RESOURCES FOR MEASURING THE LIVING STANDARD OF HOUSEHOLDS

The access route to quantification of the living standard is a proper definition of that notion. Hardly any other economic category is covered by such a wide spectrum of possible definitions: from strictly material, via social, political and other aspects, to purely subjective perception of the living standard as a certain degree of (not only material) well-being. All such definitions are based on qualitative approach; however, evaluations and comparisons require the notion to be quantified. And, unfortunately, we can only quantify what is quantifiable. In other words: if we are able to express the living standard in financial values, we are also able to measure and compare it. On the contrary, it is utterly impossible to quantify joy, happiness, and feeling of safety or satisfaction. Let us, therefore, focus on a definition of the living standard that we are able to quantify.

---

5 The EU average value equals $^2/_3$.

6 *System of National Accounts 1993.* New York: United Nations, IMF, OECD, Eurostat, World Bank, 1993.

7 *European System of Accounts – ESA 1995* (*Système Européen des Comptes – SEC 1995*). Luxembourg: Eurostat, 1996.

Keeping in mind that different aspects of the living standard and well-being cannot be synthesised, the broadest approach is that of the OECD's basic concept (cf. OECD, 2011). For individual member countries, levels of different aspects of households' quality of life are evaluated – both tangible (income, wealth, wages, and housing) and intangible (health care, safety, education, environment, etc.). In the area of expressing and comparing the living standard levels there is an indisputably interesting OECD "Better Life Index" project; it has been offering, since 2011, comparisons of living standard and well-being levels among the member states based on 11 areas in the material and quality-of-life conditions.[8] Another option offered by OECD in a number of research publications (cf. e.g., Van Zanden, 2014; or Exton and Shinwell, 2018; or Boarini, 2016) is synthesis of individual aspects with the aid of multi-dimensional analysis.

We can, alternatively, focus on quantifying households' living standard with respect to one selected aspect, i.e., restrict this wide notion to a single, but important index.

One of the possible restrictions of households' living standard and well-being notion is to focus on wealth.[9] Undoubtedly, it is important to monitor the status and evolution of wealth when considering measurements of well-being.[10] Statistically, it is not easy to estimate the status of households' tangible and intangible assets, even though we can rely on relatively trustworthy data sources, based on the definitions of individual assets. Nonetheless, Stiglitz (2009) contemplated to include the human and social capital in households' wealth (in addition to the components already contained on the national accounts). However, adding the concept of human and social capital to estimating households' wealth brings about new elements that are difficult to grasp statistically. If we extend the existing definitions by wealth components for which the estimate of the value is dubious, not only do we not approach the describable measurement of the living standard level, but also emphasise the doubts about the informative value of the data published by statisticians.

Accardo (2017) brought an interesting token into the discussion on the possibility to measure and compare the living standard levels on the basis of using data from the national accounts. His way to express the living standard level, quality of life and well-being goes via setting up an account of a "single" household (single in the sense of household categories identified by age, income, social circumstances, etc.). In fact, the standard summary account of the household sector does not have the necessary informative value from the viewpoint of expressing the quality of life; moreover, it does not contain such activities and aspects of households' life like housework, health, habits, free time, satisfaction, happiness, etc. To create such a "single household" account, we would have to combine different data sources and (again) resolve the problem of assigning financial valuation to activities and phenomena that do not have such financial "magnitudes". For example, we can estimate the amount of free time by analysis the daily timetables of selected people. But how should we actually evaluate the free time? On the basis of opportunity cost, i.e., net wages? What if the person in question has never worked? Is the value of free time, viewed in the sense of contributing to the well-being, the same of a person who is employed as a person who is unemployed, possibly for a long time? Can we, on an international scale, compare the data from the accounts set up for a "sole" household? We have mentioned just a few from among the questions asked by J. Accardo. His and similar works have, to a different extent, contributed to the bulk of the options we have at our disposal to describe certain socio-economic phenomena; but we must admit that scientist still encounter numerous obstacles, in many instances insurmountable.

Another way to go may be to restrict the notion of the living standard to the area of households' earnings/income (cf. OECD, 2011). Here we most often make use of the index of households' (adjusted) disposable income. The households' adjusted disposable income is a very good characteristic

[8] Cf. <*http://www.oecdbetterlifeindex.org*>.

[9] See Stiglitz, Sen, Fitoussi (2009) for one of such suggestions.

[10] Even though it remains unclear that well-being really is and whether the wealth measured by the wealth is or is not connected with well-being as such.

of the resources households can utilise for meeting their needs in the area market and non-market goods and services. It is a sum of disposable income (net or gross) and social transfers in kind that express the value of goods and services paid in favour of households by the general government and the non-profit institutions serving households. The adjusted disposable income represents the resources; and the index of actual final consumption expresses the value of the consumed goods and services. The latter is necessary (and, unfortunately, this fact often remains forgotten) to be monitored when evaluating and comparing households' living standard levels. As a matter of fact, that was the reason for including the concept of actual final consumption into the national accounts within the SNA 1993, or ESA 1995 standards. The national accounts provide us with long time series of data and their concepts and definitions are mutually comparable – that fact makes available to us new possibilities of evaluation and, in particular, of (international) comparisons of the living standard levels. The existence of long time series is an indisputable advantage here because the usual international comparisons (such as OECD, 2011) provide us with statistical views on one evaluated year for all the countries to be compared.

2 COMPARATIVE ANALYSIS OF HOUSEHOLDS' LIVING STANDARD LEVELS

As already stated above, our analysis will be based on the annual data of the national accounts, or rather the data on the account of the household sector. The period of our evaluation is from 2000 to 2018 and the countries to be compared are the Czech Republic and France. These two countries are different from each other not only by size, but also by their levels of economic development and the social policies applied by their governments. These circumstances are reflected not only in the different position held by the household sector within the national-economy framework (cf. Table 1), but also in different households' living standard levels.[11]

Table 1 Selected indices of the household sector position in the national economy (%)

	2000	2007	2009	2013	2018
France					
GVAh/GDP	17.2	17.3	17.1	16.3	15.5
GDIh/GNDI	60.1	61.3	63.4	61.3	60.3
AFCh/GDP	67.5	68.6	70.9	70.2	69.2
NWh/NWt	80.5	75.2	73.8	73.9	75.0
Czech Republic					
GVAh/GDP	23.2	18.7	19.3	17.5	16.4
GDIh/GNDI	57.2	54.8	59.5	57.8	55.5
AFCh/GDP	60.1	55.7	59.1	60.1	58.1
NWh/NWt	25.2	29.3	29.8	32.0	36.5

Note: GVAh – Gross Value Added of Households, GDP – Gross Domestic Product, GDIh – Gross Disposable Income of Households, GNDI – Gross National Disposable Income, AFCh – Actual Final Consumption of Households, NWh – Net Worth of Households, NWt – Net Worth of Total Economy.
Source: *<www.insee.fr>*, *<www.czso.cz>*

[11] Due to the impossibility to present the data in an integral time series, in addition to the beginning and ending years of the given period, the authors have decided to present the pre-crisis year 2007, crisis year 2009, and recovery year 2013.

The data shown in Table 1 indicates that the economic importance of households is approximately the same in both of the countries, if this importance is expressed as their proportions in the GDP. However, from the viewpoints of other selected indices we can see that French households are better off regarding the income distribution, redistribution, and use processes. The difference in proportions in the gross national disposable income amounts to about four percentage points, but the difference in the proportions of the actual final consumption values in the GDP gets to as high as 10 percentage points. In other words, French households' standing is better than that of the Czech households in the distribution and redistribution processes, especially in the in-kind form.

A principal difference can be seen with respect to households' proportion in the net worth. The very high proportion held by the French households is mainly implied by the excessive indebtedness of the French general government, which is pronounced in the low difference between the general government's financial assets and liabilities, a very low value of the net worth, and the low proportion in the corresponding national-economy value (cf. Hronová, Hindls, Marek, 2019). In the Czech Republic, general government's indebtedness is low; hence the situation is completely opposite.[12]

Consequently, the income distribution and redistribution processes reflected on the national accounts will help us identify the differences in households' outcomes from the economic activities, but also the results of the governments' social policies.

Table 2 Selected indices of the household sector_1 (%)

	2000	2007	2009	2013	2018
France					
AFC/FCE	130.0	131.2	132.8	133.5	133.5
STK/GDI	25.9	26.6	27.5	28.7	28.8
FCE/GDI	86.5	85.2	83.9	85.8	85.8
Wages/GDI	62.6	61.4	61.3	63.4	64.5
SB/GDI	30.9	31.3	33.1	35.4	35.5
Czech Republic					
AFC/FCE	120.5	122.6	123.4	123.3	124.1
STK/GDI	18.2	20.2	20.4	21.1	21.7
FCE/GDI	88.9	89.2	87.2	90.4	90.2
Wages/GDI	52.2	58.1	57.6	58.3	63.8
SB/GDI	22.8	24.1	24.8	25.5	23.7

Notes: AFC – Actual Final Consumption, FCE – Final Consumption Expenditure, STK – Social Transfers in Kind, GDI – Gross Disposable Income, Wages – Wages and Salaries, SB – Social Benefits.
Source: <*www.insee.fr*>, <*www.czso.cz*>

[12] The net worth of the French general government went down by 30.0% in the period under assessment; it amounted to a mere 2.0% of the national-economy net worth as of the end of 2018. The net worth of the Czech general government amounted to 40.0% of the national-economy value as of the end of 2018. This difference is implied by the substantially lower debt of the Czech general government (32.6% of GDP in 2018) than that of the French general government (98.4% of GDP in 2018).

The starting point for such considerations is, undoubtedly, the comparison of the actual final consumption and the expenses incurred on the final consumption. The difference between these two values consists of social transfers in kind, that is, the value of the goods and services paid in favour of households by the general government and the non-profit institutions serving households – such as education, health care, culture, sports, etc. The value of the social transfers in kind therefore covers goods and services consumed by households without paying for them, i.e., increasing households' living standard level regardless of their (in)ability to pay the corresponding expenses incurred on the final consumption.

The data shown in Table 2 clearly implies that, on a long-term basis, the Czech households incur a larger proportion of their disposable income on the final consumption. This difference will, inter alia, be reflected in the lower saving rate[13] of Czech households as compared with French ones (approximately 11% for the Czech households on a long-term basis, and approximately 14% for the French households). The most significant component that affects households' ability to satisfy their needs from their income, and consequently the most significant component of the disposable income, is undoubtedly represented by wages. The Czech Republic has always been characterised by a lower level of wages (and of labour productivity). However, the economic results of the Czech national economy, as well as the economic policy in the most recent years[14] have led to increasing the proportion of wages in households' gross disposable income (hereinafter GDI) by more than 10 percentage points, by which the value of this index got balanced between the Czech and French households in 2018. In a simplistic formulation, the initial conditions of the Czech and French *employee* households got balanced as regards the final consumption expenditure. The differences in the proportions of the final consumption expenditure with respect to the GDI are undoubtedly caused by the significantly lower purchasing power of old-age pensioners in the Czech Republic as compared with France; in fact, a typical pensioners' household spends its entire disposable income on consumption, mainly due to the low level of old-age pensions.[15]

The households' living standard does not exclusively depend on the ability to obtain the market products; it also follows from households' access to goods and services provided to them for free or almost for free by the general government and the non-profit institutions serving households. French households' actual final consumption is, on a long term basis, by one-third higher than the expenses incurred on final consumption; it is only by one-fourth for Czech households (cf. Table 2). In other words, French households cover three-quarters of their needs from their resources, while Czech households cover four-fifths. This is an important aspect that indicates a higher living standard and better living conditions of French households than those prevailing in the Czech Republic. Another consequence is a higher (by about eight percentage points) ratio of the social transfers in kind with respect to the GDI[16] for French households as compared with Czech ones (cf. Table 2).

The proportion of social benefits in the GDI is similar regarding a comparison between Czech and French households. The initial difference of eight percentage points was increased by one-half, that is, to 12 percentage points. Since the largest component of the social benefits is represented by old-age pensions, this growing disproportion can be assigned to the above-mentioned difference in the levels of the old-age pensions.[17]

---

13 Saving rate = Gross saving/Gross disposable income.

14 In particular, salaries were increased in public institutions with the consequent pressure on wage increases in the private sphere, as well as repeated increases of the minimum wages.

15 In 2018, the replacement rate was 60% in France but just 46% in the Czech Republic. Cf. <*https://www.oecd-ilibrary.org/finance-and-investment/gross-pension-replacement-rates/indicator*>.

16 Keeping in mind that the social transfers in kind are not included in the disposable income but are included in the adjusted disposable income. That is why we speak about a ratio, not a proportion.

17 A substantial proportion in social benefits corresponds to unemployment benefits, whose amount has been undergoing different evolution paths in the considered countries: the unemployment rate in France fluctuated between 8% and 10% during the entire period in question, while in the Czech Republic it was on a comparable level until 2015, when it started to fall down to its current value of 2%.

Table 3 Selected indices of the household sector_2 (%)

	2000	2007	2009	2013	2018
France					
SC/GDI	9.9	10.4	10.4	11.1	10.3
CT/GDI	15.1	14.1	13.9	16.3	17.5
(SC + CT)/GDI	25.0	24.5	24.3	27.4	27.8
SB/SC	314.0	301.3	316.7	318.6	344.1
SB/(SC + CT)	124.2	128.5	135.7	129.0	127.8
Czech Republic					
SC/GDI	10.1	12.6	11.2	11.9	12.6
CT/GDI	8.2	8.6	6.8	7.5	9.3
(SC + CT)/GDI	18.3	21.1	18.0	19.4	21.9
SB/SC	226.1	191.9	220.7	214.4	193.6
SB/(SC + CT)	124.8	114.1	137.2	131.4	108.2

Notes: SC – Social Contributions, GDI – Gross Disposable Income, CT – Current Taxes on Income, Wealth, etc., SB – Social Benefits.
Source: *<www.insee.fr>*, *<www.czso.cz>*

A higher proportion of social benefits in the GDI and a higher amount of unpaid consumption in the case of French households not only means a growing pressure on the general government's expenditure[18] and, consequently, debt (cf. above); it also brings a higher load on households regarding the (direct) taxes and social contributions paid by them.

The data shown in Table 3 enables us to take a "reverse view" on households' living standard. The situation is comparable in both of the countries regarding social contributions by households (related to the GDI): the proportion of social contributions by Czech households is by a mere one to two percentage points. However, French households' income-tax (current tax) load is, on a long term basis, substantially higher than that in the Czech Republic. We can include both of these components under a sole "umbrella" index of the so-called mandatory payments (here a sum of social contributions and current taxes paid by households), and compare the value of this umbrella index with that of the GDI. The result is again to the detriment of French households and the difference has, on a long term basis, been fluctuating around six percentage points. Let us compare the proportions in the GDI of the social benefits and social contributions in each of the countries in question. We can thus identify the balance between households' payments to the social system and their gain (in the form of social benefits) from the same system. Czech households pay by one to two percentage points (as related to the GDI) more on social contributions than French households (cf. Table 3) but the former get by eight to 12 percentage points (again, as related to the GDI, cf. Table 2) less in social benefits than the latter. This disproportion to the detriment of Czech households is even emphasised when we compare the volumes of received social benefits and paid social contributions (cf. Table 3).

[18] The expenditure of the French general government amounts to 52% to 57% of the GDP on a long-term basis; in the Czech Republic, this proportion fluctuates between 39% and 45% (cf. Hronová, Hindls, Marek, 2019).

However, this comparison is not entirely correct: let us realise that if there is a lack of resources from the paid social contributions[19] (of course paid by all subjects, not only households), the government may provide funding for social benefits from other resources, such as the taxes. Payments by households in the form of current taxes can thus, in connection with the received social benefits, be viewed on as a load analogous to the payments of social contributions. If we compare the volumes of the social benefits received by households[20] with those of the mandatory payments paid by households, this ratio is not so significantly better for French households; or rather, this ratio has been decreasing for Czech households after 2013, getting as low as 108.2% in 2018, which is by nearly 20 percentage points lower than the value valid for French households.

In the years of the economic growth that began in the Czech Republic in 2013, logically, the smaller value of the unemployment rate and the higher level of wages worsened the balance between the amounts Czech households pay to the social system and the benefits they receive from it. In any case, the increasing level of the living standard should bring about an increase in the social transfer amounts paid in favour of households by the general government and non-profit institutions serving households. Unfortunately, Czech households' position in this respect is worse than that of French households.

After formulating this partial conclusion, we must get back to the fact that the largest proportion in the social benefits is represented by old-age pensions. The economic growth and low unemployment rate in the Czech Republic have led to a faster growth in the volume of the social contributions paid by households (with a growth rate of 180.4% in current prices in the period in question) than in the volume of social benefits received by households (with a rate of 133.5%);[21] that is, the value of the SB/SC, or SB/(SC + CT), index went down. At the same time it means that people not participating in the labour process, in particular old-age pensioners, are hurt by the overall economic prosperity. Old-age pensioners' households are those that make use, to the greatest extent, of the healthcare services (including payments for medicines) and other social services whose values are included in the amount of the social transfers in kind. It is therefore clear that the insufficient scope of the services provided for free, or nearly for free, and the volume of the social benefits are, in addition to the expenses incurred on final consumption, important attributes of households' living standard – and these very attributes may be decisive in the perception of "which households are better off". Nonetheless, the larger proportion of households' unpaid consumption and the higher proportion of the social benefits (whether related to the GDI or to the paid social contributions) generates pressure on the general government's expenditure.[22] This is the case of France, where the government's social policy on the one hand ensures a higher level of households' living standard, on the other hand it decreases the living standard of the society as a whole due to the problematic sustainability of public finances.

---

[19] This is the case of France, where – since 2009 – the volume of the social benefits paid by the general government has been higher than the amount of the social contributions received by the general government. This balance has been opposite in the Czech Republic on a long-term basis; the volume of the social contributions received by the general government is by about one-third higher than the volume of the social benefits paid by the general government.

[20] The authors are aware of a certain margin of error here: households receive social benefits not only from the general government (within the framework of the mandatory social insurance) but also from financial corporations (within the framework of the optional social insurance) and from employers (as an equivalent of employers' imputed social contributions). However, such additional benefits are by far less important.

[21] In France, the amount of social contributions paid by households went up by 67.3%, while the amount of the social benefits received by households by 83.4%.

[22] Partly to the expenditure incurred by non-profit institutions serving households; however their proportion in the social transfers in kind in favour of households and in the non-market output is not very significant in this context. The proportion in the social transfers in kind of non-profit institutions serving households amounts to about 5% in the Czech Republic and about 10–12% in France; their proportion in the non-market output is again about 5% in the Czech Republic and about 10% in France.

CONCLUSIONS

Evaluating and comparing households' living standard levels is a recurring and politically sensitive topic. In order to quantify this problem, we must first define the content of the notion and then find indices that will enable us to evaluate and compare them. OECD is active in the evaluations and international comparisons; on the basis of the individual indices, OECD compares the countries to be evaluated. This approach is static; the living standard levels are evaluated in particular years but the results can hardly be used for comparative analysis in time.

The national accounts, or rather the account of the household sector, provide a tool that is very suitable to evaluate and compare the living standard levels in time and space. The possibility to compare in space is based on the comparable methodologies of national accounts (ESA 2010,[23] or SNA 2008,[24] the temporal comparisons stem from not only the long and methodologically comparable time series, but – in particular – the use of relative indices. In addition to the indices usually used for the evaluation of particular components of the living standard, such as the disposable income and adjusted disposable income, the household account provides other indices, whose values bear information about households expenditure (incurred on final consumption expenditure),[25] as well as what households consumed regardless of who pays for it (actual final consumption). The difference between the actual final consumption and the final consumption expenditure consists of the social transfers in kind.

When comparing the data for the household sectors in the Czech Republic and France, we revealed the fact that the area of households' unpaid-for consumption (in the form of transfers in kind) is an important aspect by which households' living standard levels differ from each other in the two countries we consider here. Czech households obtain a relatively lower proportion of the unpaid-for final consumption; and the ratio of the social benefits as related to the disposable income is also smaller in the Czech Republic than in France. On the other hand, Czech households' load by income tax is smaller; and the relative value of their payments on social contributions is about the same as that of French households. The economic growth in the Czech Republic after 2013 has been accompanied by a worsening ratio of the received social benefits with respect to the mandatory payments by households; this fact indicates, inter alia, a significant lag in the growth of social benefits (mostly consisting of old-age pensions) behind the that of wages and other types of income. Having in mind the low level of social benefits (in particular, old-age pensions), we clearly see that the living standard level can be improved by increasing the unpaid-for final consumption. On the other hand it turns out that the general government's accommodating social policy (manifested in the large volume of the non-market output and social transfers in kind) in favour of households leads to problems with sustainability of public finances, from which root the social tension grows, being detrimental to the well-being of the society as a whole (even though we still cannot be sure what the well-being in fact is); this is also the case of France.

It is difficult to say where the level of households' living standard is actually higher. Hence we should not resort to a more-or-less intuitive comparisons between easily grasped and popularising indices. In each such comparison, different indices have different meanings; only when we evaluate them comprehensively, with a wide use of the data from the national accounts, can we find a qualified answer to the question where the living standard level is higher (or lower). That is why our analysis (on an example of comparing the Czech Republic and France) has been aimed at illustrating the possibilities offered by the system of the national accounts not only in comparing the living standard levels on the basis of the data taken from the household sector's account but also in the context of other data that can be found on the national accounts.

[23] *European System of Accounts – ESA 2010 (Système Européen des Comptes – SEC 2010).* Luxembourg: Eurostat, 2013.
[24] *System of National Accounts 2008.* New York: United Nations, IMF, OECD, Eurostat, World Bank, 2009.
[25] The authors are aware of a certain margin of error here: the final consumption expenditure also includes items consumed by households but paid by someone else, in particular, imputed rents and self-supplied agricultural production.

ACKNOWLEDGEMENTS
This article was written thanks to support from the Institutional Support to Long-Term Conceptual Development of Research Organisation, the Faculty of Informatics and Statistics of the University of Economics, Prague.

References

ACCARDO, J. An Appreciation of Living Standards, Well-Being and Quality of Life [online]. In: *Texts of the 16th Conference of the Association de Comptabilité Nationale*, Paris: INSEE, 2017. <https://insee.fr/fr/information/2834065>.

BOARINI, R. et al. Multi-dimensional Living Standards: A Welfare Measure Based on Preferences [online]. *OECD Statistics Working Papers*, No. 2016/05, Paris: OECD Publishing. <https://doi.org/10.1787/5jlpq7qvxc6f-en>.

CLARK, A. AND SENIK, C. Will GDP growth increase happiness in developing countries? *Paris School of Economics Working Paper*, 2010, No. 2000/43.

EXTON, C. AND SHINWELL, M. Policy use of well-being metrics: Describing countries' experiences [online]. OECD *Statistics Working Papers*, No. 2018/07, Paris: OECD Publishing. <https://doi.org/10.1787/d98eb8ed-en>.

GCEE. *Monitoring economic performance, quality of life and sustainability.* Wiessbaden: German Council of Economic Experts, 2010.

HALLIWELL, J. F. Life Satisfaction and the Quality of Development. *NBER Working Paper* 14507, National Bureau of Economic Research, 2008.

HINDLS, R. AND HRONOVÁ, S. Unemployment and inflation rates in the years of economic turbulences or do the Phillips curve relationships still hold? In: *8th Annual Conference of the EuroMed Academy of Business*, Verona: EuroMed Press, 2015, pp. 861–870.

HRONOVÁ, S., HINDLS, R., MAREK, L. Economic Behaviour of the General Government and Sustainability of Public Finances – Comparative Analysis of the Czech Republic and Selected EU Countries [online]. *Statistika*, 2019, 99(4), pp. 350–368. <https://www.czso.cz/documents/10180/88506452/32019719q4_350_hronova_analyses.pdf/059f497e-dbf6-4ca1-a93b-3af73d369c34?version=1.0>.

INSEE. Deux ans après le rapport Stiglitz-Sen-Fitoussi: quelles mesures du bien-être et de la soutenabilité? *Contributions de l´INSEE*, Paris, 2011.

LLENA-NOZAL, A., MARTIN, N., MURTIN, F. The economy of well-being: Creating opportunities for people's well-being and economic growth [online]. *OECD Statistics Working Papers*, No. 2019/02, Paris: OECD Publishing. <https://doi.org/10.1787/498e9bc7-en>.

OCDE. *Comment va la vie? Mesurer la bien-être"*. Paris: Editions OCDE, 2011.

OECD. *Compendium of OECD Well-Being Indicators.* Paris: OECD, 2011.

STIGLITZ, J. E., SEN, A., FITOUSSI, J.-P. *Report by the commission on the Measurement of Economic Performance and Social Progress* [online]. <https://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>.

VAN ZANDEN, J. et al. eds. *How Was Life? Global Well-being since 1820* [online]. Paris: OECD Publishing, 2014. <https://doi.org/10.1787/9789264214262-en>.

Can the Business Tendency Survey Predict the Economic Indicators in Czechia?

Veronika Ptáčková[1] | *University of Economics, Prague, Czech Republic*
Jakub Fischer | *University of Economics, Prague, Czech Republic*

Abstract

In uncertain times, it is crucial to have some statistics, which can help with the prediction of future development in the national economy. Business Tendency Survey is one of the most essential and favourite tools for predictions in economic statistics. The article aims to determine which confidence indicators help predict the Czech economic development and which base of the confidence indicators is the best for making predictions. Using the Granger causality test, we prove the Business Tendency Survey indicators are beneficial for predictions. The economic sentiment indicator and the confidence indicator for industry predict the gross value added better than the gross domestic product. The long-term average is a slightly better base than the base indices. The predictions are most accurate in the horizon of two quarters. Individual composite indicator for the industry well predicts both the industrial production index (for the next month) and the gross value added in the industry (for the next quarter).

Keywords

Business Tendency Survey, prediction ability, gross domestic product, business cycle, base indices

JEL code

C10, C22

INTRODUCTION

The development of the economy in the European Union is a frequently discussed issue. The Business Tendency Survey is one of the standard data sources used. The European Commission coordinates its methodology. Thanks to this, we obtain harmonised qualitative data across all European countries. The weighting scheme (during the construction of the indicators) and the base are disputable parts of the calculation. The European Commission uses the long-term average (over the period 1990–2018). For countries where the European Commission has missing values, the period starts at the earliest available data point onwards to 2018. Every January, the European Commission recalculates the indicators. They have to add the values from the ending year (i.e. in January 2020, the European Commission

[1] University of Economics, Prague, Faculty of Informatics and Statistics, W. Churchill Sq. 4, 130 67 Prague, Czech Republic. Also Czech Statistical Office, Na padesátém 81, 100 82 Prague 10, Czech Republic. Corresponding author: e-mail: veronika.ptackova@vse.cz, phone: (+420)224095435.

is going to review the European Sentiment Indicator so that it will be based on a simple average stretching from 1990 to 2019). However, some countries have a different base. We had a simple average of 2005 in the Czech Republic by the end of 2019. From the beginning of 2020, the Czech Statistical Office uses the long-term average. In Italy, they use the simple average of 2010, while the simple average of 2015 is used in Germany and Slovakia (Ptáčková, Štěpánek, Hanzal, 2019; Czech Statistical Office, 2020).

The Business Tendency Survey has two forms of publishing the results: (i) base indices, and (ii) balance, respectively. This paper has two aims. The first is to verify which economic indicators are best predicted by business cycle surveys. We will compare predictions of gross domestic product, gross value added, the industrial production index and gross value added in industry. The second objective will be the base of the BTS indicators. Does the base affect the results? If so, which base is best for predictions in the Czech economy?

We divided the article into three chapters. The literature review is the first one. The second one consists of methodology and data, while the last part allows readers to become informed about the most relevant results of the calculation. After that, we discuss the recommendation of the base used.

1 LITERATURE REVIEW

The Business Tendency Survey (BTS) is a popular tool for predictions of the economy. The question is posed as to what this survey forecasts. The respondent should evaluate the current situation and changes in the ensuing three (or six) months (for example, the development of the national economy, financial situation or employment). Businessmen or consumers choose (typically) from three options: "decrease", "stable" or "growth". The Business Tendency Survey is a rich and reliable source for quick estimates of the results of the national economy - both as a whole and in individual sectors. Unfortunately, they are rarely used in practice (Marek, Hronová, Hindls, 2019).

1.1 Predictions of the economic indicators

The Business Tendency Survey is an essential and favourite tool for the predictions of the economic variables – Gross Domestic Product (GDP), Gross Value Added (GVA) or business cycle. Modugno, Soybilgen, Yazgan (2016) created a model re-estimation. The model has good results during a global economic crisis, but it worsens after the crisis.

Sorić and Marković (2010) discuss the fact that indicators calculated from the Business Tendency Survey responses are leading indicators for macroeconomics variables. However, they have to predict the selected variables. Balances which analysts estimate from Business Tendency Survey are an essential tool for the forecasting of and support for policy decisions (Cesaroni and Iezzi, 2015). Hainisch and Scheufele (2017) warned that data revisions could affect the quality of the survey data and the forecasting.

Bergström (1995) found the relationship between the industrial production growth rate and the Swedish indicators from the Business Tendency Survey. Österholm (2013) confirmed this conclusion. The survey data help with the Swedish short-term forecast of GDP growth. The author informs that it has essential informational value. Martinsen, Ravazzolo and Wulfsberg (2014) predicted GDP growth for the selected regions and sectors. They used a factor model approach for these predictions. Zeman, Lojka and Obst (2016) analysed the possibilities of using confidence indicators for the development of the q-o-q GVA changes by using Vector Autoregressive Models and co-integration analysis. In Czech data, they discovered no relationship between the questions and confidence indicator of the industries (except for the manufacturing industry where they identified significant short-term relationship).

Bruno and Lupi (2004) focused on the business cycle turning points and the prediction ability of these indicators. According to Cesaroni and Iezzi (2015), analysts should use an accurate indicator calculated from the selected question (e.g. the number of employees, the company's expectation about the economic development or investments) for short-term predictions. But the indicators help with the policy decisions and improve the forecasts, too. Lehmann and Weyh (2016) analysed only employment expectations. They confirmed that these results help with employment growth forecasting. Abberger (2005) calculates that responses to capacity assessment are essential for the evaluation of investment.

Białowolski (2014) changes the approach of the construction confidence indicator. He warns that this concept should be tested for a particular country. But they confirmed that we could use indicators for forecasting the main macroeconomic variables. Marek, Hronová and Hindls (2019) have shown that views on developments in industry and construction with a time lag assist with the predictions of the quarterly gross value added of the national economy. The confidence indicator in trade and services was not significant. The authors argue that the respondents in the mentioned sectors have a different view of the assessment and prediction of their business development. At the same time, they remind of the dominance of manufacturing industries in the Czech economy.

1.2 Relationship of industrial production index and gross value added in industry

The monthly industrial production index (IPI) is the first approximation of the development of gross value added in the industry as the IPI for the last month of the given quarter is released around 20 days before the estimate of gross value added for individual CZ-NACE sections. There are several methodological differences between IPI and gross value added.[2] Based on the empirical analysis of the coherence of indicators, Fischer et al. (2014) note that the IPI is much closer to the development of industrial enterprises' sales at the level of individual CZ-NACE subsections than the gross value added in these subsections. However, at the level of the whole industry (sections B+C+D) the IPI is strongly correlated with the development of revenues from industrial activity, total production (gross output) of industry and the gross value added of the industry as well (Fischer et al., 2014).

2 DATA AND METHODOLOGY

2.1 Business Tendency Survey

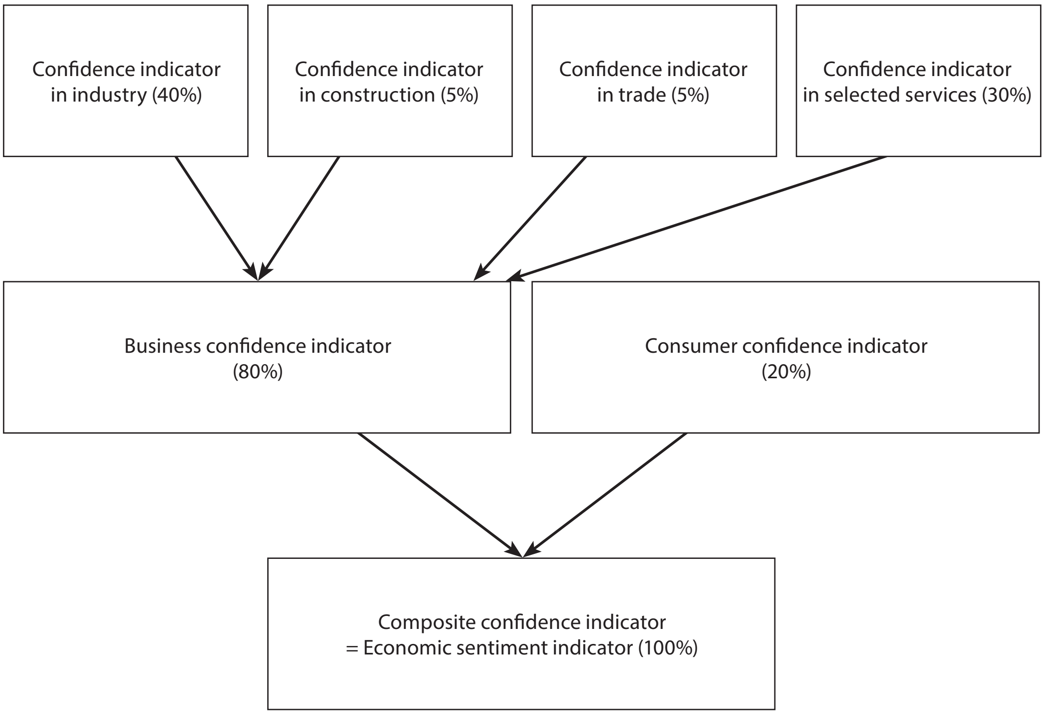
The Business Tendency Survey defines confidence indicators, which are useful for the economic sentiment indicator construction (see Figure 1).

- The composite confidence indicator (economic sentiment indicator) is calculated as the weighted average of seasonally adjusted confidence indicators in the industry, trade, construction and in the selected service and consumer confidence indicator.
- The business confidence indicator is calculated as the weighted average of seasonally adjusted confidence indicators in industry, construction, trade and selected services:
  - The confidence indicator in the industry is calculated as the simple average of seasonally adjusted balances of three questions (the assessment of total demand, assessment of stocks of final production – with an inverted sign – and the expected development of production activity).
  - The confidence indicator for construction is calculated as the simple average of the assessment of total demand and the expected development of employment.
  - The confidence indicator in trade is calculated as the simple average of the assessment of the economic situation, the evaluation of stocks (with an inverted sign) and the expected development of the economic situation.
  - The confidence indicator in selected services is calculated as the simple average of the assessment of the economic situation, the assessment of demand and expected demand.
- The consumer confidence indicator is calculated as the simple average of the expected financial situation of consumers, the expected total economic situation, expected total unemployment (with an inverted sign) and savings expected in 12 months (Czech Statistical Office, 2019).

[2] See ECB (2016).

Figure 1 summarizes previous information about the Economic sentiment indicator construction. One can see the weights of the individual confidence indicator in brackets. The European Commission defines mentioned weighting scheme.

Figure 1 Weighing system for the construction of composite confidence indicator (in %)



Source: Czech Statistical Office (2019)

2.2 Granger causality test

We can use the Granger causality test to determine whether a concrete time series is beneficial for the forecasting of another one. The main idea of this test is: if the variable *X* is causal for variable *Y* in the Granger way, we can say that the information from variable *X* is essential for accurate predictions of variable *Y* (Wang, 2019).

The test uses two stationary time series (X_t and Y_t). The time series Y_t is correlated with lagged values of X_t. The test helps with the understanding of how the change in the one time series impact the changes in the second time series – it does not matter which one is causation and which one is a consequence (Hušek, 2007). The null hypothesis means that variable *X* does not influence variable *Y* in the Granger way. Formula (1) describes the restricted regression model:

$$Y_t = \propto_0 + \sum_{i=1}^{m} \propto_i Y_{t-i} + \mu_t \,, \tag{1}$$

where Y_t and Y_{t-i} are two continuous observations of the time-series in times t and t-i, respectively, α_0 and α_i are coefficients of the toughness of the dependency between Y_t and Y_{t-i} and μ_t is a non-systematic (white-noise) component of the model.

The next equation is different than the first one. There are lagged values of the time series X used as an explanatory variable. This Formula (2) is an unrestricted regression model:

$$Y_t = \propto_0 + \sum_{i=1}^{m} \propto_i Y_{t-i} + \sum_{j=1}^{m} \beta_j X_{t-j} + \mu_t , \tag{2}$$

where X_t and Y_t are values of the time series X and Y in the time t, the α_0 is a linear constant (we have to estimate this value), μ_t is the white noise sequence, α_i and β_j are coefficients of the variables X_t and Y_t and m is the number of lagged terms (Wang, 2019). Granger causality test uses the past values of both time series – X_t and Y_t. It is important to mention again that we use the known (past) values of the economic indicators and differently lagged values of the Business Tendency Survey indicators.

The test compares the unrestricted model and restricted model. The analysis is focused on data from January 2003 until December 2018 for all sectors (industry, construction, trade, and selected services, respectively) and consumers in the Czech Republic. All the computations were performed using the R language and environment for statistical computing and graphics (R Core Team, 2014).

We consider the following economic indicators:

- economic sentiment indicator (qualitative data, BTS, base indices = base in 2005, 2010, 2015, long-term average),
- business confidence indicator (qualitative data, BTS, base indices = base in 2005, 2010, 2015, long-term average),
- the individual confidence indicator in the industry (qualitative data, BTS, base indices = base in 2005, 2010, 2015, long-term average),
- gross domestic product (quantitative data, national accounts),
- gross value added (quantitative data, national accounts),
- industrial production index (quantitative data, short-term statistics, base indices = base in 2005, 2010),
- gross value added in the industry (quantitative data, national accounts, base indices = base in 2005, 2010).

Economic sentiment indicator, the business confidence indicator, individual confidence indicators in the industry and the IPI are published monthly. That is why we recalculate these values quarterly (for comparison with the gross domestic product and gross value added). We calculate the simple average of the three-monthly values of the mentioned economic indicators. We have to adjust data (seasonally) using the X-13 Census ARIMA procedure. After that, we extract a linear trend from data. It means that all time-series were de-trended thanks to the decomposition (we get data devoid of a trend after this step). Before modelling, it is necessary to remove the seasonality (Granger, 1979). We test the stationarity of all-time series using the augmented Dickey-Fuller test. This test is based on the analysis of the first-order autoregressive process AR(1) (Box and Jenkins, 1970).

This paper follows the paper from Ptáčková, Štěpánek and Hanzal (2019). They concluded that we should use the base indices within the base given as a simple average of 2005, 2010 or 2015 years, respectively for the economic indicators prediction (in terms of achieving the tightest relations between the obtained and predicted values of the economic indicators). As in the mentioned article, the decisions will be made using only the p-values and F-statistics. We will focus only on the one-way relationship: it means that the results from the Business Tendency Survey describe/predict the values of the economic indicators (GDP, GVA, IPP).

3 RESULTS

Firstly, we focus on the relationship between economic sentiment indicator in time t and gross domestic product, gross value added, respectively, in time $t + 2$. It means that the information from the BTS can help to predict the development of the economic indicators in the following six months. In Table 1, ESI_BI is the economic sentiment indicator in the base indices form (where the base is the simple average of the monthly results in 2005, 2010, or 2015, respectively). ESI_LTA is economic sentiment indicator in the base indices form (where the base is the simple average of the monthly results from January 2003 to December 2017).

In Table 1, ESI shows significant results for both economic indicators: gross value added and gross domestic product. But we have a better relationship between ESI and gross value added in terms of achieving the tightest relations between the obtained and predicted values of the economic indicators.

Table 1 Granger causality tests – ESI (t), GDP, GVA ($t + 2$)

	ESI_BI (2005, 2010, 2015)		**ESI_LTA (2003–2017)**	
	***p*-value**	**F-test**	***p*-value**	**F-test**
GDP (BI = 2005)	0.0008	12.4980	0.0003	14.5380
GDP (BI = 2010)	0.0008	12.4840	0.0003	14.5920
GVA (BI = 2005)	0.0000	31.2120	0.0000	35.7380
GVA (BI = 2010)	0.0000	31.1390	0.0000	35.5930

Source: Own construction

Secondly, we analyse the prediction ability of the business confidence indicator in time t and the gross domestic product or gross value added in time $t + 1$ (in the next three months). In this case, we do not have significant results for any combination (see Table 2).

Table 2 Granger causality tests – BCI (t), GDP, GVA ($t + 1$)

	BCI_BI (2005)		**BCI_BI (2010)**		**BCI_BI (2015)**		**BCI_LTA (2003–2017)**	
	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**
GDP (BI = 2005)	0.3015	1.0840	0.4457	0.5895	0.3884	0.7549	0.3809	0.7793
GDP (BI = 2010)	0.2532	1.3312	0.3837	0.7703	0.3312	0.9599	0.3244	0.9876
GVA (BI = 2005)	0.3375	0.9350	0.6704	0.1830	0.5365	0.3866	0.5189	0.4210
GVA (BI = 2010)	0.3596	0.8527	0.6973	0.1528	0.5625	0.3392	0.5447	0.3711

Source: Own construction

After that, we continue in the calculation, and we focus on the next period: gross domestic product and gross value added in time $t + 2$ (in the next six months). We have better results for the long-term average. Still, we should consider the following: the business confidence indicator helps with the predictions of the key economic indicators in the following six months. We obtain similar results for this when we analyse the partial confidence indicator in the industry instead of the business confidence indicator.

Table 3 Granger causality tests – BCI (*t*), GDP, GVA (*t* + 2)

	BCI_BI (2005)		BCI_BI (2010)		BCI_BI (2015)		BCI_LTA (2003–2017)	
	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**
GDP (BI = 2005)	0.0015	11.1520	0.0050	13.5870	0.0007	12.8400	0.0007	12.7230
GDP (BI = 2010)	0.0016	10.9760	0.0005	13.5940	0.0007	12.7740	0.0008	12.6480
GVA (BI = 2005)	0.0000	26.9770	0.0000	34.9110	0.0000	32.3280	0.0000	31.9390
GVA (BI = 2010)	0.0000	26.8170	0.0000	34.2760	0.0000	31.8790	0.0000	31.5140

Source: Own construction

In the next step, we try to find the best base for the individual confidence indicator in the industry (CI_BI) and industrial production index (IPI) in the time *t* + 1 (next month), …, *t* + 6 (next six months). It is essential to mention that we analyse monthly data.

Table 4 Granger causality tests – CI_BI (*t*), IPI (*t* + 1, …, *t* + 6)

	CI_BI (2005, 2010, 2015)		CI_LTA (2003–2017)	
	***p*-value**	**F-test**	***p*-value**	**F-test**
IPI (t + 1)	0.0029	9.1179	0.0029	9.1179
IPI (t + 2)	0.0050	8.0605	0.0050	8.0605
IPI (t + 3)	0.0160	5.9068	0.0160	5.9068
IPI (t + 4)	0.0649	3.4485	0.0649	3.4485
IPI (t + 5)	0.0508	3.8670	0.0508	3.8670
IPI (t + 6)	0.6281	0.2354	0.6281	0.2354

Source: Own construction

In Table 4, we have significant results (in terms of getting the tightest relations between the obtained and predicted values of the economic indicators) for these combinations: individual confidence indicator in the industry (based in 2005, 2010 and 2015) and IPI (*t* + 1), IPI (*t* + 2), individual confidence indicator in the industry (LTA) and IPI (*t* + 1), IPI (*t* + 2). The results are equal when we use different bases or long-term average for the confidence indicator.

Finally, we calculate the Granger causality tests for the individual confidence indicator in the industry (based in 2005, 2010 and 2015) and the gross value added in the industry. Table 5 describes the results of the Granger causality test. We can say that we have essential results for the combined gross value added in the industry in the time *t* + 1 (the number of the GVA in the industry is in a million Czech crowns from the national accounts – constant prices) and confidence indicator in the industry. Again, it does not matter which base we use.

Table 5 Granger causality tests – CI_BI (*t*), GVA in the industry (*t* + 1, *t* + 2, *t* + 3)

	CI_BI (2005)		CI_BI (2010)		CI_BI (2015)		CI_LTA (2003–2017)	
	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**	***p*-value**	**F-test**
GVA (t + 1)	0.0014	11.2740	0.0014	11.2740	0.0014	11.2740	0.0003	14.5400
GVA (t + 2)	0.7669	0.0887	0.7669	0.0887	0.7669	0.0887	0.7808	0.0782
GVA (t + 3)	0.7611	0.0933	0.7611	0.0933	0.7611	0.0933	0.8086	0.0592

Source: Own construction

DISCUSSION AND CONCLUSION

The Business Tendency Survey is an important data source indicating the current and future development of the national economy. A lot of analysts confirm that it works. Österholm (2013) used BTS results for the short-term forecast of GDP growth. Marek, Hronová and Hindls (2019) confirm that these results help with the predictions in the industry and construction sector.

In this paper, we focus on finding the best base for the predictions (in terms of getting the tightest relations between the obtained and predicted values of the economic indicators) of the Czech economy. By the Granger causality test, the economic sentiment indicator has significant results for the predictions of the gross domestic product and gross value added in the next six months. The business confidence indicator has the same predictive results. In both cases (economic sentiment indicator and business confidence indicator), we obtain higher *F*-values for the gross value added than for the gross domestic product. The individual confidence indicator in the industry well predicts the monthly industrial production index. The results are significant for the following month and the two months predictions, regardless of the base we used.

In conclusion, we proved that we could use the Business Tendency Survey for the predictions of economic indicators. When we talk about the gross domestic product, we can recommend using the long-term average as the base, but there is no significant difference comparing to the base indices. In this analysis, we see that the BTS is better for GVA predictions compared to GDP. Also, we can use the confidence indicator in the industry for the gross value added in the industry predictions.

In the next research, we can improve the weighting system for getting better results of the economic sentiment indicator. Furthermore, we plan to use the results from the survey-on-survey, which is currently provided in co-operation with the Czech Statistical Office.

ACKNOWLEDGEMENT

This paper has been prepared under the support of a project of the University of Economics, Prague – Internal Grant Agency, project No. F4/14/2018 "Business Tendency Survey in the Czech Republic and Possibilities of increasing its predictive ability" together with the Institutional Support for Long-Term and Conceptual Development of Research and Science at the Faculty of Informatics and Statistics, University of Economics, Prague.

References

ABBERGER, K. *The Use of Qualitative Business Tendency: Surveys for Forecasting Business Investment in Germany No. 13.* Ifo Working Paper, 2005.

BERGSTRÖM, R. The relationship between manufacturing production and different business survey series in Sweden 1968–1992. *International Journal of Forecasting*, 1995, 11(3), pp. 379–393.

BIALOWOLSKI, P. Concepts of confidence in tendency survey research: an assessment with multi-group confirmatory factor analysis. *Social indicators research*, 2015, 123(1), pp. 281–302.

BOX, G. E. AND JENKINS, G. E. *Time series analysis: Forecasting and control.* San Francisco: Holden-Day, 1970.

BRUNO, G. AND LUPI, C. Forecasting industrial production and the early detection of turning points. *Empirical economics*, 2004, 29(3), pp. 647–671.

CESARONI, T. AND IEZZI, S. The predictive content of business survey indicators: evidence from SIGE. *Journal of Business Cycle Research*, 2017, 13(1), pp. 75–104.

CZECH STATISTICAL OFFICE. *Business cycle surveys – Methodology* [online]. Prague: CZSO. [cit. 16.10.2019]. <https://www.czso.cz/csu/czso/business_cycle_surveys>.

ECB. A closer look at differences between industrial gross value added and industrial production. *Economic Bulletin*, Issue 1/2016, Box 2. [cit. 30.1.2020]. <https://www.ecb.europa.eu/pub/pdf/ecbu/eb201601.en.pdf>.

FISCHER, J. et al. *Manažerské shrnutí výsledků metodického auditu statistiky ČSÚ* [online]. 2014. [cit. 16.11.2019]. <https://www.czso.cz/documents/10180/39785642/audit_6.pdf/fa8a9583-c69e-4a56-b84a-517cf43354dc?version=1.1>.

GRANGER, C. W. Seasonality: causation, interpretation, and implications. In: *Seasonal analysis of economic time series*, NBER, 1978, pp. 33–56.

HEINISCH, K. AND SCHEUFELE, R. Should forecasters use real-time data to evaluate leading indicator models for GDP prediction? German evidence. *German Economic Review*, 2018.

HLOUŠEK, M. *Stationarity* [online]. 2019. [cit. 30.5.2019]. <http://www3.econ.muni.cz/~hlousek/ teaching/stationarity.pdf>.

HUŠEK, R. *Ekonometrická analýza.* Oeconomica, 2007.

LEHMANN, R. AND WEYH, A. Forecasting employment in Europe: Are survey results helpful? *Journal of Business Cycle Research*, 2016, 12(1), pp. 81–117.

LOJKA, J., OBST, J., ZEMAN, J. Assessment of the Ability of the Business and Consumer Surveys to Predict the Gross Value Added of the Czech Republic [online]. *Statistika: Statistics and Economy Journal*, 2016, 96(4), pp. 19–36.

MAREK, L., HRONOVÁ, S., HINDLS, R. Možnosti odhadů krátkodobých makroekonomických agregátů na základě výsledků konjunkturních průzkumů (Possibilities of Estimations of Short-term Macroeconomic Aggregates Based on Business Surveys Results). *Politická ekonomie*, 2019(4), pp. 347–370.

MARTINSEN, K., RAVAZZOLO, F., WULFSBERG, F. Forecasting macroeconomic variables using disaggregate survey data. *International Journal of Forecasting*, 2014, 30(1), pp. 65–77.

MODUGNO, M., SOYBILGEN, B., YAZGAN, E. Nowcasting Turkish GDP and news decomposition. *International Journal of Forecasting*, 2016, 32(4), pp. 1369–1384.

ÖSTERHOLM, P. Survey data and short-term forecasts of Swedish GDP growth. *Applied Economics Letters*, 2014, 21(2), pp. 135–139.

PTÁČKOVÁ, V., ŠTĚPÁNEK, L., HANZAL, V. Is one time-series in the business tendency survey able to predict another one? Granger causality between time series. In: *Applications of Mathematics and Statistics in Economics (AMSE 2019)*, Paris: Atlantis Press, 2019, pp. 39–48. ISBN 978-94-6252-804-8

R CORE TEAM. *R: A language and environment for statistical computing* [online]. Vienna, Austria: R Foundation for Statistical Computing, 2017. <http://www.R-project.org>.

SORIĆ, P. AND MARKOVIĆ, M. Predicting downturn: Are tendency surveys a good estimator of retail activity in Croatia. *Ekonomski pregled*, 2010, 61(9–10), pp. 559–575.

WANG, X. A Granger causality test of the causal relationship between the number of editorial board members and the scientific output of universities in the field of chemistry. *Current Science (00113891)*, 2019, 116(1).

The Impact of Migration on Well-Being in a Remittances Dependent Economy

Azizbek Tokhirov[1] | *Charles University, Prague, Czech Republic*

Abstract

The aim of this study is to examine the impact of migration on migrant-sending communities. The particular attention is drawn to the well-being of households residing in Central Asia, where migration is a crucial issue due to constant and high outflow of labor force. More specifically, in an attempt to determine the extent to which international migration influences subjective well-being of household members left behind, the research uses panel data collected in Tajikistan from 2007 to 2011. The results indicate that, on average, there is a positive relationship between migration of a family member and improvements in satisfaction with life as-a-whole and current financial situation of those who are staying behind. The positive effects are even more pronounced when the sample is restricted to migrant-sending households that receive remittances. The further split-sample analysis also documents that the impact of migration appears to be heterogeneous across different economic and geographic contexts.

Keywords

International migration, remittances, subjective well-being, life satisfaction, financial satisfaction

JEL code

F22, F24, I31

INTRODUCTION

Internal market distortions all over the world force many individuals to migrate in search of better life conditions and to escape different types of deprivation. The statistics show that international migration stock of nearly 272 million people accounts for 3.5% of the world population (United Nations, 2019). Due to constraints to move freely between countries in the form of strict immigration policies and transportation costs, international migrants in most of the cases not only leave their communities but also family members behind (Démurger and Wang, 2016). Migrant transfers of money and goods, commonly referred as remittances, therefore, have become one of the main sources of household income in increasing number of developing countries. In many cases, remittances are not determined after the process of migration, but the prospect of remitting might significantly affect the decision to migrate. Hence, it is not surprising that the volume of remittances to developing countries had significantly increased over the last years: from USD 228.6 billion in 2006 to USD 529 billion in 2018 (World Bank, 2016; World Bank, 2019). At the household level, the share of remittances might even reach 50% of the income

[1] Department of Social Geography and Regional Development, Faculty of Science, Charles University, Albertov 6, 128 43 Prague 2, Czech Republic. E-mail: azizbek.tokhirov@natur.cuni.cz.

(Duval and Wolff, 2010). Moreover, migrant-sending communities receive not only financial remittances, but there are also social changes brought about by migrants, usually defined in rather intangible terms (e.g. ideas, behaviors, values) and referred as social remittances (Bailey et al., 2018; Cingolani and Vietti, 2019). Given the considerable number of migrants, there is an important question regarding the well-being of those staying behind: whether cross-border transfers can compensate emotional and financial losses connected with the absence of household members?

The answer for this query cannot be found in the literature because there is no theoretical consensus on the direction of the development impacts of migration on migrant-sending communities. Historically, classical (Ravenstien, 1885) and neoclassical (Harris and Todaro, 1970) paradigms emphasized the positive contribution of migration to economic development of source communities through relocation of excessive capital and labor. This point of view is in line with the generally accepted theory of labor migration proposed by Stark and Bloom (1985), where migration is regarded as a joint utility maximizing decision between migrants and other family members undertaken with the aim of diversifying risks perpetual to domestic environment. On the contrary, there is an alternative school of thought which emphasizes that negative effects of migration would prevail over positive ones, and that "migration leads to economic dependency and stunted development in migrant-sending societies" (Bohra-Mishra, 2013, p. 173).

Along with theoretical studies, a significant number of empirical papers were dedicated to examination of how migration in general and migrants' remittances in particular affect the well-being of those who are left behind. Though, these researches tend to explore changes in the well-being of migrant-sending families indirectly based on economic indicators at the individual or household levels. Among them, we can mention, for example, consumption and investment (Adams and Cuecuecha, 2010), education (Gyimah-Brempong and Asiedu, 2015), fertility decision-making (Šimková and Langhamrová, 2015), health (Lu, 2013), housing (Strielkowski and Weyskrabová, 2013), income (De and Ratha, 2012), labor supply (Justino and Shemyakina, 2012) or poverty (Esquivel and Huerta-Pineda, 2007). The standard approach to infer welfare from observed behavior is not a sole empirical option. Another, less explored approach is related to personal judgments of people (Syrovátka, 2007). Well-being indicators in this case are usually derived from answers of individuals to general and specific questions about their life satisfaction or happiness levels. However, the area of direct evaluation of well-being by members of migrant-sending households be they children, seniors or spouses is relatively new and consequently, less investigated (Nguyen et al., 2007).

The variety of seemingly contradictory theoretical and empirical predictions suggests that a dilemma over the relationship between migration and well-being will not be solved in the foreseeable future. Moreover, the available research is mostly based on the data from Latin American, African and Southeastern Asian countries with relatively few empirical studies concentrating on the impact of migration on households in Central Asia, one of the vulnerable regions as identified by Collier (2008), which lies on the cross-roads of active migratory movements. This study aims to address indeterminacy in the literature by analyzing the extent to which migration of a family member affects subjective well-being of those staying behind by applying the latest data from nationally-representative surveys conducted all over Tajikistan.

The choice of the country is not arbitrary. Due to severe consequences of the Soviet Union collapse and prolonged transitory period, many Tajik households chose to migrate as a coping strategy. For many years, Tajikistan was the world leader in terms of dependence on migrant financial transfers measured by a considerable margin of personal remittances in the country's national income (Danzer and Ivaschenko, 2010). The issue of persistent labor emigration had also been acknowledged at the governmental level of Tajikistan by the establishment of the migration service agency (International Labour Organization, 2011). As it can be seen from Figure 1, over the last years, the value of received personal remittances constantly accounted for approximately 30% of Tajikistan's gross domestic product (GDP), with the lowest observed value being far higher than the unweighted global average value.

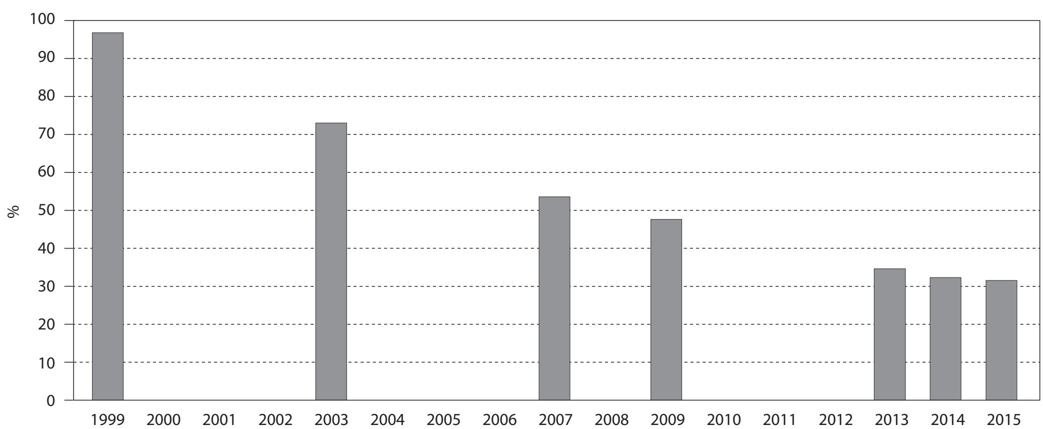
The overall increasing trend is explained by local economic deficiencies pushing people to become migrants (Clément, 2011), while abrupt downturns in 2009 and 2014 can be attributed to external economic crises (Danzer and Ivaschenko, 2010; Petrović et al., 2017).

Figure 1 Personal remittances received (% of GDP) by Tajikistan

Source: WDI (2019)

Despite high remittance inflows, the country is characterized by considerable number of households experiencing financial difficulties. Based on the latest available estimations, more than 30% of the population was below the national poverty line (Figure 2). This economic situation is nothing new, Tajikistan was also among the poorest states of the former Soviet Union (Clément, 2011).

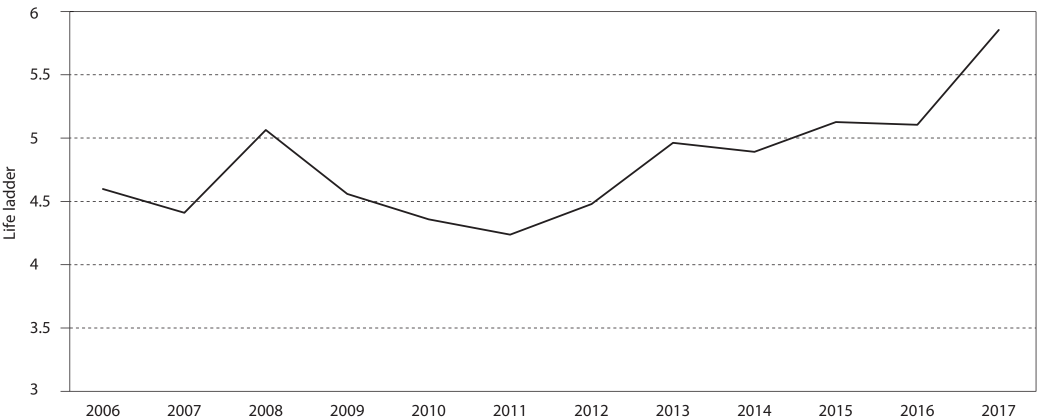
Figure 2 Poverty headcount ratio at national poverty lines (% of population) in Tajikistan



Source: Trading Economics (2015)

With respect to subjective well-being measures, Tajikistan's country-average happiness score (Figure 3) was ranked only 74th out of 156 countries in the 2016–2018 Ranking of Happiness; however, when the increase of 0.764 in the index for the period between 2005–2008 and 2016–2018 is considered the country was among the top 18 nations (Helliwell et al., 2019).

Figure 3 Happiness score of Tajikistan



Source: Sustainable Development Solutions Network (2019)

Given interconnectivity between the previous three indicators, they can be viewed at the same time; it should be noted that because the data are derived from different sources, the coverage is not identical. Firstly, we can say that the percentage of Tajikistan's population living in poverty decreased by three times from 1999 to 2015. Over this period, the country's dependence on remittances and happiness score noticeably increased. The data also suggest that the positive changes in the poverty rates were steady, while the latter indicators fluctuated noticeably.

Taking into consideration economic situation in Tajikistan, it is important to obtain a comprehensive overview of household migration experience. Despite the importance of international migration for the economy of Tajikistan, only relatively few studies have attempted to investigate the well-being of migrant-sending households residing in this country. The studies mainly concentrated on "objective" well-being measures. Therefore, the aim of this research is to explore additional ways through which migration could possibly affect households in Tajikistan, so that to contribute to the understanding of the impact of migration on regions with high outflows of people. In this context, evaluative well-being is chosen to be the outcome variable. Based on Graham and Nikolova (2015), it is assumed that life evaluations are more appropriate measure of choice and opportunities than other categories of subjective well-being (hedonic or eudaimonic); and might represent actual capabilities and means that individuals have, allowing policymakers to better target poor and socially deprived people. This type of well-being is less considerably affected by short-term fluctuations and expected to capture economic consequences of migration. Accordingly, research conclusions can contribute to the existing economic knowledge of whether migration can promote development or not.

1 DATA DESCRIPTION

The study draws on the Tajikistan Living Standards Survey (TLSS) conducted under auspices of the World Bank in 2007 and in 2009, and the Tajikistan Household Panel Survey (THPS) collected by the Institute for East- and Southeast European Studies in 2011. The representative data were collected from urban and rural areas of the country's each administrative region. Initially, 4 860 households were randomly selected to participate in the survey connected with measuring the quality of life in Tajikistan (Gang et al., 2018). After 2 years, the survey organizers, motivated by the same purpose, re-interviewed a random subsample of 1 503 households within the 2007 TLSS (Danzer and Ivaschenko, 2010). In 2011, another large-scale questionnaire was distributed to 1 503 households, most of them being from the 2007 TLSS and the 2009 TLSS, to investigate the migration patterns in Tajikistan (Danzer et al., 2013).

Ideally, the analysis should be based on the random sub-sample of 1 503 households. There are several practical issues that reduce the actual number of households. Firstly, the information about only 1 458 households was updated in 2011. In addition to the panel attrition, variables describing various household characteristics have missing values across surveys. As a result, the sample with complete information about households tracked over each survey wave comprises 1 283 data points. The number of households is comparable to the recent research in the similar settings by Gang et al. (2018), where the analysis is based on the balanced panel of 1 257 households.

The study mainly concentrates on a single binary regressor, which is equal to 1 if a household had at least one current member residing abroad when the survey data were collected. Table 1 summarizes information about the number of households with and without migrants in every wave of the surveys. The values suggest that from the total number of observations, 726 of them are identified as being a migrant-sending household. It should be noted that the survey probability to observe households with migrants was much higher in 2011, whereas the difference in the number of migrant-sending households between the years 2007 and 2009 is less significant. When we explicitly consider whether migrant-sending households received cash or in-kind transfers over the course of the year, we can see that in the context of Tajikistan, migration is mostly motivated by remittances. The correlation between “migration” and “remittances” is positive and strong (0.9).

Table 1 Distribution of sampled households by decision to migrate and send remittances

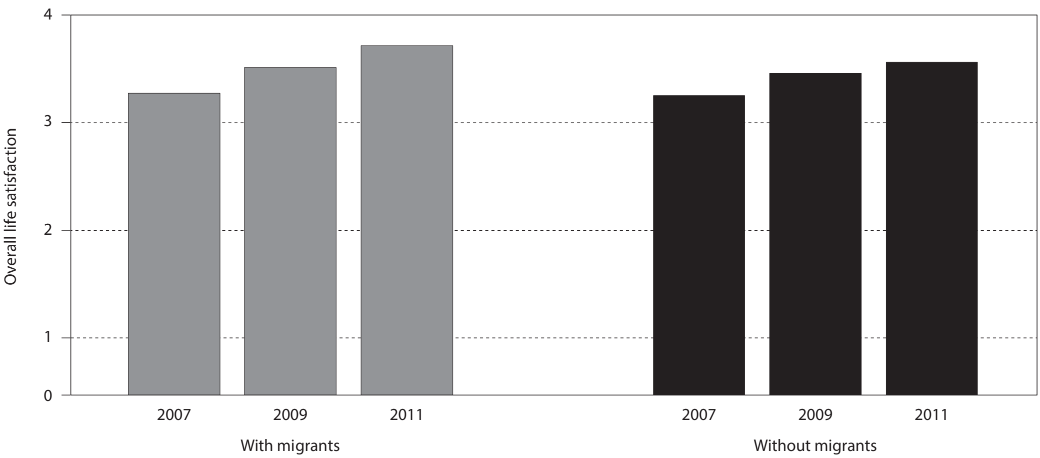
	Without migrants 2007	With migrants 2007	Remittance-receiving
Without migrants 2009	966	145	
With migrants 2009	113	59	151
Without migrants 2011	836	97	
With migrants 2011	243	107	305
Remittance-receiving		179	

Source: TLSS (2007), TLSS (2009), THPS (2011)

For the outcome variables, the study considers the survey answers of the most informed household member to the two following Likert scale questions: (1) Overall how satisfied are you with your life? and (2) How satisfied are you with your current financial situation? Following the new economics of labor migration, it is assumed that migration is a joint household decision and its impacts affect all household members. Therefore, we can consider the opinion of household head as a proxy for overall household well-being. Although, the wording of questions remained the same across surveys, the division of answers into categories was changed from 4 to 5 between the waves of the survey. With the aim of enabling a comparison between periods, two ordinal variables with 4 categories are constructed by merging interim categories; particularly, the answers are categorized ascendingly as: not at all satisfied, less than satisfied, rather satisfied or satisfied.

Given the specifications of treatment and outcome variables, now we can consider their joint progression over five years: Figure 4 illustrates the changes in household satisfaction with overall life conditions, while Figure 5 shows the evolution of household satisfaction with current financial situation.

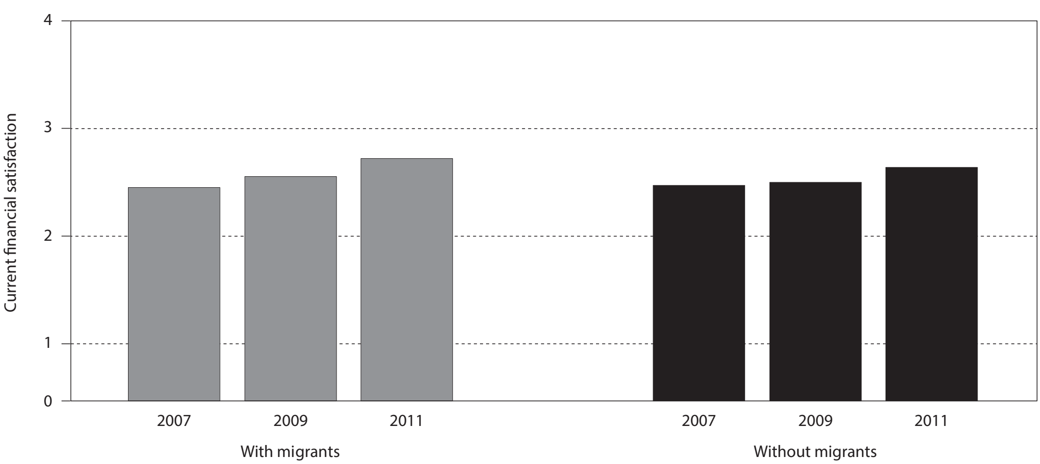
Figure 4 Distribution of sampled Tajik households by overall well-being



Source: TLSS (2007), TLSS (2009) and THPS (2011)

Based on the graphs, we can see that both subjective well-being measures experienced a growth during the period under observation. The average value of overall life satisfaction increased from 3.21 to 3.46 and then to 3.64, while the average values of satisfaction with current financial situation were 2.45, 2.53 and 2.72 in 2007, 2009 and 2011 respectively. In addition to the time-related changes, the indicators also diverge with respect to migration decision. When the indicators are disaggregated, in almost all periods, migrant-sending households had higher satisfaction levels than households without migrants.

Figure 5 Distribution of sampled Tajik households by financial well-being



Source: TLSS (2007), TLSS (2009) and THPS (2011)

To allow comparability with the previous literature, in addition to migration, the study estimates household subjective well-being as a function of several covariates. To avoid the possibility of endogeneity, the chosen variables should affect the outcome variables; and not be systematically influenced by the treatment status. Mostly, individual and regional household characteristics satisfy these requirements and are proposed by the literature on the typical changes associated with migrant-sending households; given the recommendations; information related to household head is not considered due to possibility of endogeneity (Démurger and Wang, 2016). Table 2 presents a summary of variables selected to be used for the analysis.

Table 2 Descriptive statistics

	Mean	St dev	Mean	St dev	Mean	St dev
	2007		2009		2011	
Household size	6.37	2.83	6.78	3	6.42	3.14
Household composition						
Share of children below the age of 6	0.10	0.13	0.107	0.13	0.111	0.14
Share of children aged between 6 and 15	0.234	0.19	0.21	0.19	0.195	0.19
Share of adults aged between 16 and 65	0.61	0.24	0.631	0.23	0.635	0.23
Share of elderly over the age of 65	0.056	0.15	0.052	0.15	0.059	0.16
Share of female adults	0.33	0.18	0.33	0.17	0.346	0.18
Household education level						
Average education of household members	2.56	0.88	2.62	0.89	2.68	0.95
Number of household members with tertiary education	0.27	0.45	0.29	0.45	0.29	0.45
Household location						
Urban	0.34	0.47	0.34	0.47	0.33	0.47
Districts of Republican Subordination	0.207	0.41	0.207	0.41	0.207	0.41
Dushanbe	0.162	0.37	0.162	0.37	0.162	0.37
Gorno-Badakhshan Autonomous	0.102	0.3	0.102	0.3	0.102	0.3
Khatlon	0.264	0.44	0.264	0.44	0.264	0.44
Sughd	0.265	0.44	0.265	0.44	0.265	0.44
Migration network						
Proportion of households with migrants in primary sampling unit	0.15	0.15	0.13	0.13	0.27	0.24

Source: TLSS (2007), TLSS (2009), THPS (2011)

Overall, characteristics of households are comparable across surveys with expected time-induced changes. The average household size slightly increased from 2007 to 2011 with a small upsurge in 2009. As for the household composition: the share of adults increased resulting from the decline of the share of teenagers, while the share of children and elderly stayed almost at the same percentage. The household educational level experienced a growth based on the increase in the average years of schooling and number of people with higher education. Given the previously mentioned increase in the number of migrant-sending households in 2011, it is not surprising to observe a notable increase in the relative size of the existing migration network of this period. In relation to the areal distribution, nearly twice as many households were residing in rural area than in urban agglomeration. Finally, we can notice that there were no considerable changes with respect to the aggregate regional distribution of households, which might signify that main household units are spatially immobile.

2 EMPIRICAL SPECIFICATION

This section briefly reviews several regression techniques which are applied to determine the well-being effects of migration. The first option is to disregard the fact that selected subjective well-being measures are ordinal and treat them as they are continuous. In this case, we can assume that there is a linear relationship between the variable representing subjective well-being and selected household characteristics, and evaluate the respective model with the ordinary least squares (OLS) estimator based on pooled data.

Appropriate statistical treatment of ordinal data would be to consider non-linear models. In this case, we should start by introducing a latent measure (Y^*) of the indirect utility for a group of representative agents living together (i) in period (t):

$$Y_{it}^{*}=x_{it}^{*}\beta + \varepsilon_{it}, \tag{1}$$

where x represents a set of observable characteristics, including a "migration" variable (M), which define household well-being, with a vector of coefficients (β) and an error term (ε).

However, we can only observe y with k discrete categories and individual intercept terms (μ) specified as follows:

$$y_{it} = k \;\; if\, \mu_k \leq Y^{*} \leq \mu_{k+1}, \quad k = 1, 2, ..., K. \tag{2}$$

Depending on the assumptions regarding the distribution of the error term, the probability of observing outcome k can be estimated by different methods, such as ordinal logit or probit regression.

It should be noted that we cannot include all household characteristics into Formula (1). Therefore, the previously discussed regression estimates might be biased due to the presence of unobserved heterogeneity. To solve this issue, we can use the panel structure of data and re-define ε_{it} by explicitly introducing unobserved household traits, which might be affecting household well-being, to the model:

$$Y_{it}^{*} = x_{it}^{*}\beta + \alpha_i + u_{it}. \tag{3}$$

If we assume that selected covariates are not correlated with unobserved household characteristics, we can estimate Formula (3) with a Random Effects (RE) model. Conversely, if we think that $E(\alpha_i \mid x_{it}^{*}, M_{it}) \neq 0$, we should consider a Fixed Effects (FE) model. Given the structural differences between models, by applying both of them we can verify robustness of the findings. For the former case, the study considers a random effects ordered probit model (described by Alsakka and Ap Gwilym, 2010), while for the latter case, the study considers a fixed effects ordered logit model ('blow-up and cluster' (BUC) estimator introduced by Baetschmann et al., 2015); both estimated by a conditional maximum likelihood method.

Yet, when we are examining the well-being, characteristics of households with respect to their intention to migrate may be a reason for the possible divergences. In other words, we should also take the possibility of selection bias into account. McKenzie et al. (2006) tested the major methods to address endogeneity in migration studies by conducting a natural experiment in New Zealand. Based on their results that instrumental variable (IV) regressions with valid instruments provide the most accurate results, this method is chosen for the current setup with migration network as an instrumental variable. The study considers a community-level measure of migration network proposed by Justino and Shemyakina (2012), which is a proportion of households with migrants in a primary sampling unit. Given the nature of the Tajik immigration, the decision to migrate depends significantly on external economic and political situation, and the information about current and historical migration density rates should ideally capture this constantly changing environment. The composite instrumental variable is expected to comply with both exogeneity and relevance assumptions because migrant networks should positively affect the probability to migrate without directly affecting household well-being. The foremost argumentation is that social ties tend to direct future migrants providing them instructions on the ways of possible migration destinations (Démurger and Wang, 2016); while, clustering households based on the respective primary sampling units should tackle possible well-being changes induced by the instrument. Due to specificity of the variable (i.e. it is generated by aggregating migration network variable of each period) and to restore original distribution of the outcome variables (5 categories), a one-period version of the estimation based on the THPS 2011 is applied. Particularly, using a two-stage extended ordered probit regression, we can attempt to uncover exogenous variation of the treatment variable with additional probit model for migration decision (M):

$$M_i = 1(z_i\beta + u_i > 0), \tag{4}$$

where a vector z contains variables from Formula (2) and other strictly exogenous covariates which determine the treatment variable, and u_i – unobserved errors, which are multivariate normal with mean 0.

3 BASELINE RESULTS

Columns 1–4 of Table 3 report the panel analysis for the determinants of overall life satisfaction of households over the whole period under consideration. The overall number of observations are same for the pooled OLS and RE ordered probit regressions, while additional artificial observations are created for the FE ordered logit regression. The results are derived based on the clustered standard errors at primary sampling units, which can be considered as a common practice for the studies based on Living Standard Measurement Surveys (Kan and Aytimur, 2018). Due to methodological considerations, the BUC estimator requires clustering at the household level.

Table 3 Impact of migration on overall life satisfaction

	OLS	RE	BUC		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Relatives abroad	0.114 (3.07)***	0.179 (3.02)***	0.283 (2.07)**		0.932 (4.45)***	
Receive remittances				0.376 (2.56)**		0.998 (4.4)***

Table 3 (continuation)

	OLS	RE	BUC		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Needed medical assistance	–0.071 (3.38)***	–0.104 (3.58)***	–0.15 (2.53)**	–0.151 (2.54)**	–0.158 (3.08)***	–0.158 (3.06)***
Employment status	0.041 (3.08)***	0.06 (3.09)***	0.104 (2.31)**	0.105 (2.33)**	0.079 (2.65)***	0.083 (2.78)***
Experience financial difficulties	–0.267 (7.84)***	–0.359 (7.76)***	–0.498 (5.38)***	–0.497 (5.37)***	–0.338 (3.88)***	–0.338 (3.84)***
Household size	0.053 (2.63)***	0.078 (2.69)***	0.116 (1.18)	0.128 (1.3)	0.085 (1.81)*	0.088 (1.83)*
Household size^2	–0.002 (2.38)**	–0.003 (2.45)**	–0.002 (0.54)	–0.002 (0.6)	–0.005 (2.53)**	–0.005 (2.52)**
Children (<6)	0.003 (0.15)	–0.001 (0.03)	–0.03 (0.32)	–0.038 (0.41)	–0.036 (0.72)	–0.038 (0.75)
Children (6–15)	–0.001 (0.07)	–0.002 (0.06)	–0.208 (2.33)**	–0.215 (2.43)**	0.056 (1.2)	0.054 (1.13)
Elderly (>65)	–0.083 (3.32)***	–0.131 (3.64)***	–0.239 (1.37)	–0.238 (1.37)	–0.136 (1.99)**	–0.13 (1.9)*
Female adults	–0.074 (3.41)***	–0.11 (3.47)***	–0.077 (0.79)	–0.088 (0.9)	–0.056 (1.17)	–0.06 (1.23)
Average education	0.04 (1.55)	0.077 (2)**	0.132 (1.4)	0.135 (1.44)	0.143 (2.22)**	0.143 (2.2)**
Tertiary education	0.053 (1.32)	0.095 (1.5)	0.049 (0.3)	0.044 (0.27)	0.132 (1.34)	0.132 (1.33)
First-stage: Migration network					4.976 (13.11)	4.972 (13.23)
Urban/rural	Yes	Yes	No	No	Yes	Yes
Regions	Yes	Yes	No	No	Yes	Yes

Table 3 (continuation)

	OLS	RE	BUC		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Survey waves	Yes	Yes	Yes	Yes	No	No
Observations	3 849	3 849	4 548	4 548	1 283	1 283
Res. cross corr.					–0.506 (4.03)***	–0.499 (3.6)***
R^2	0.09					
Log likelihood		–3 748	–1 480	–1 478	–1 953	–1 909

Note: T- and z-scores are based on clustered standard errors and their absolute values are displayed in parentheses. * Significance at 10%, ** 5%, and *** 1% level.
Source: Author's own calculations

As it can be seen, the difference between conditional average well-being patterns of migrant-sending households and households without migrants is statistically significant within all specifications, which might suggest that there are structural changes caused by migration. As for direction of the relationship, the overall change is positive: migrant-sending households are expected to experience higher levels of overall life satisfaction than their non-migrant counterparts. As in Dickerson et al. (2014), the obtained findings indicate that linear and non-linear models yield comparable results. Therefore, we can consider the coefficients of the linear model because they can be interpreted as marginal effects without additional concerns about the underlying latent variables. Given the interval between 1 and 4 in which the outcome variable is bounded, the coefficient of 0.11 is somewhat small in terms of magnitude. However, in relative terms, the impact of migration is more pronounced than the effect of domestic employment status of household members. Moreover, once we exclude from the analysis migrant-sending households that do not receive remittances, the impact of migration increases nearly by 25%.

When the estimated coefficients of the control variables are considered, their values are in line with economic reasoning as well as previous empirical studies. The burden of health and financial issues has significant and adverse effects on household well-being. Conversely, overall life satisfaction is expected to improve as the number of working household members increases. The correlation between household educational level and overall evaluative well-being is also positive. The impact of household composition is negative with respect to the reference category, which is a share of children below the age of 6. The estimates capturing the effect of household size indicate that a relationship between household size and overall life satisfaction follows a U-shaped pattern.

The further regression results for the impact of migration on the alternative measure of subjective well-being of remaining household members are presented in columns 1–4 of Table 4. As anticipated, households that send out migrants have higher probability to be satisfied with current financial situation. The respective positive sign of the coefficient for migration variable is preserved across all panel regressions. Although, we cannot compare the regression coefficients, it should be stated that the changes brought about by migration is less prominent in the case of current financial satisfaction than overall life satisfaction. The signs and magnitudes of control variables are comparable between Table 3 and 4, as well as the tendency that the impact of remittances is greater than the gross effects of migration.

Table 4 Impact of migration on satisfaction with current financial situation

	OLS	RE	BUC		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Relatives abroad	0.092 (2.95)***	0.147 (2.85)***	0.2 (1.66)*		0.84 (2.89)***	
Receive remittances				0.287 (2.25)**		0.89 (2.97)***
Needed medical assistance	−0.099 (5.4)***	−0.163 (5.21)***	−0.23 (3.84)***	−0.228 (3.83)***	−0.215 (4)***	−0.214 (3.97)***
Employment status	0.034 (3.15)***	0.053 (2.9)***	0.048 (1.26)	0.051 (1.33)	0.063 (2.13)**	0.065 (2.2)**
Experience financial difficulties	−0.277 (10.6)***	−0.444 (9.94)***	−0.577 (6.52)***	−0.576 (6.51)***	−0.406 (4.4)***	−0.406 (3.84)***
Household size	0.052 (3.19)***	0.082 (2.99)***	0.035 (0.4)	0.044 (0.52)	0.075 (1.61)	0.076 (1.61)
Household size^2	−0.002 (2.65)***	−0.003 (2.39)**	−0.000 (0.00)	−0.0001 (0.04)	−0.004 (2.05)**	−0.004 (2.05)**
Children (<6)	−0.009 (0.47)	−0.018 (0.53)	−0.044 (0.52)	−0.051 (0.61)	−0.032 (0.62)	−0.032 (0.62)
Children (6–15)	0.011 (0.64)	0.013 (0.45)	−0.044 (0.54)	0.051 (0.62)	0.034 (0.73)	0.032 (0.69)
Elderly (>65)	−0.05 (2.02)**	−0.079 (1.93)*	0.109 (0.69)	0.114 (0.73)	−0.084 (1.23)	−0.08 (1.16)
Female adults	−0.053 (2.97)***	−0.082 (2.73)***	−0.001 (0.01)	−0.011 (0.13)	−0.068 (1.4)	−0.069 (1.41)
Average education	0.08 (3.77)***	0.124 (3.44)***	0.05 (0.6)	0.052 (0.64)	0.101 (1.54)	0.1 (1.51)
Tertiary education	0.113 (3.56)***	0.184 (3.5)***	0.109 (0.78)	0.104 (0.74)	0.179 (2.1)**	0.178 (2.09)**

Table 4						(continuation)
	OLS	RE	BUC		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
First-stage: Migration network					5.001 (13.3)***	5 (13.6)
Urban/rural	Yes	Yes	No	No	Yes	Yes
Regions	Yes	Yes	No	No	Yes	Yes
Survey waves	Yes	Yes	Yes	Yes	No	No
Observations	3 849	3 849	4 011	4 011	1 283	1 283
Res. cross corr.					–0.44 (2.5)***	–0.44 (2.4)***
R^2	0.1					
Log likelihood		–4 025	–1 367	–1 365	–1 978	–1 935

Note: T- and z-scores are based on clustered standard errors and their absolute values are displayed in parentheses. * significance at 10%, ** 5%, and *** 1% level.
Source: Author's own calculations

Robustness tests are performed in columns 5–6 of Tables 3 and 4 by explicitly accounting for the endogeneity of migration and remittance decision. The results of the first stage binary probit estimations are comparable with the previous literature (Démurger and Wang, 2016) and can signalize that the vector of instruments can properly explain household migration decision. Particularly because there is a strong and positive relationship between the size of migration network and probability to send migrants and subsequently, receive remittances. The estimated correlations between the errors from the first and second stage equations are negative and significantly different from zero, possibly indicating that the choice of sending migrants is indeed endogenous and unobserved factors affecting the selection are negatively associated with the higher levels of subjective well-being.

When the predicted values from the first stage regression are used, the migration and remittances variables retain their sign values and relative within-regression magnitudes. The estimates are comparable not only with respect to the variables of interest, but we can also find similarities in the way how confounding variables affect outcome variables. However, there is a noticeable difference in the cross-regression magnitudes. More specifically, the cross-sectional regressions yield much higher magnitude of estimates than the panel regressions due to the increase of outcome variable categories.

4 HETEROGENEITY ANALYSIS

In addition to migration, several exogenous factors might affect well-being patterns of households. In the case of satisfaction with current financial situation, household wealth might be an important element (Démurger and Wang, 2016). Although, all external surroundings might be relevant for overall life satisfaction, economic opportunities available for households might not be the same in different parts of the country (Robinson and Guenther, 2007). Therefore, for the sake of investigating heterogeneity in the relationship between migration and subjective well-being, we should adjust the sample to different contexts, specifically according to the regional location and income of a household unit. Precisely, the heterogeneity analysis draws on the 2011 THPS with the initial distribution of households from the 2007 TLSS. This choice stems from the considerations to minimize the possibility of correlation between changes in household characteristics and out-migration. To capture all aspects of migration experience, the study considers only "migration" variable for the heterogeneity analysis.

4.1 Regional decomposition

The effects of migration are positive in both urban and rural settings in the case of overall life satisfaction (Table 5); the statistical significance of the estimates are higher in comparison to the case when the whole sample is considered. The increase of statistical significance can be explained by the decrease in the exogenous variation between variables. Separate calculations also demonstrate that households with migrants in urban settings are likely to be more satisfied with life as-a-whole than rural households. Conversely, the impact of migration on satisfaction with current financial satisfaction is only statistically significant (and positive) for rural households. The estimations indicate indeterminacy in the way how migration affects urban households.

Table 5 Heterogeneous effects of migration: location

	Overall life satisfaction	Current financial satisfaction
Rural	0.862 (3.33)***	0.902 (3.43)***
Urban	1.26 (4.69)***	0.122 (0.17)

Note: Z-scores are based on clustered standard errors and their absolute values are displayed in parentheses. *** significance at 1% level. Individual controls are included in all regressions.
Source: Author's own calculations

4.2 Income-based disaggregation

Table 6 provides additional insights into the impact of migration on subjective well-being measures across "pre-migration" household income categories. The measure of monetary deprivation is represented by a poverty headcount index derived from the expenditures-based poverty line of buying 2 250 calories and affording a certain amount (36%) of non-food items (TLSS, 2007). The well-being changes (measured

Table 6 Heterogeneous effects of migration: income

	Overall life satisfaction	Current financial satisfaction
Poor	0.839 (3.61)***	0.75 (2.62)***
Non-poor	1.1 (5)***	0.939 (3.11)***

Note: Z-scores are based on clustered standard errors and their absolute values are displayed in parentheses. *** significance at 1% level. Individual controls are included in all regressions.
Source: Author's own calculations

by both indicators) caused by migration is positive and statistically significant for poor and non-poor households. Though, the magnitude of regression coefficients tends to increase with household income. When we compare the coefficients across regressions, the difference in the impact of migration between poor and non-poor households is more significant in the case of overall life satisfaction than current financial satisfaction.

DISCUSSION AND CONCLUSION

Since the early 1990s, labor migration has made notable contributions to addressing financial vulnerability of households in transition economies. The question posed by this study was whether the access to international migration can also promote other aspects of well-being of those who are left behind. The research to date on the topic has produced mixed and debatable results. On this occasion, in an attempt to link well-being and migration, the study starts with investigating prospective transmission mechanisms and develops an economic model of life evaluations based on linear and nonlinear estimations. After controlling for self-selection of households and endogeneity of migration decision, this research finds a particularly clear evidence of the positive effects of international migration on source countries.

The results suggest that there is a strong positive relationship between subjective well-being and international migration, with the main mechanism being financial remittances. More specifically, the findings demonstrate that having family members abroad and subsequent receipt of remittances, on average, is expected to increase the probability of being satisfied with life as-a-whole and with current financial situation at the household level. The migration-induced changes are more pronounced in the case of overall life satisfaction than current financial satisfaction which might imply that the impact of out-migration is structurally different with respect to financial and non-financial measures of subjective well-being.

The further analysis showed that specific characteristics of households may also be a source of heterogeneity in the well-being effects of migration. The impact is heterogeneous in terms of household location and income level. When we consider satisfaction with life as-a-whole, richer and urban households benefit more from migration in comparison to their respective counterparts. Conversely, rural migrant-sending households are expected to be more satisfied with their current financial situation, while the impact of migration is not observed for urban households; as for the divergence in financial well-being levels between households from different income groups, the change introduced by migration is positive and relatively homogeneous. Based on the split-sample analysis, we can firstly hypothesize that not only financial remittances affect households but also social remittances. This might contribute to resolving the economic quest of why economic agents might be emotionally indifferent to further economic improvements after reaching a certain income threshold (Stevenson and Wolfers, 2013). Secondly, it is documented that Tajik rural households encounter extra costs to receive international remittances due to limited access to transport infrastructures and financial intermediaries (Clément, 2011). At the same time, according to Robinson and Guenther (2007), rural areas of Tajikistan are more prone to natural hazards and households engage in migration to diversify income. Therefore, it is plausible that rural households with migrants act more economically responsibly and achieve greater financial security.

The derived results are highly relevant for the economic setting of Tajikistan. Extensive labor emigration as a response to poverty and lack of employment has become a routine occurrence for the people living in Tajikistan. The government's migration policy and the institutional agenda for applying migration policies have been a complicated matter. However, given the occurrence that the country experiences considerable outflows of people and unable to ensure sufficient funding for social provision, the contribution of migration to refining living conditions can be significant and should draw attention of policymakers. More particularly, when the migration situation of Tajikistan and the research findings are considered together, there are several policy implications for the bodies operating in the field of or affected by migration and remittances.

Firstly, remittances improve the financial well-being of Tajik households and might provide them opportunity to exit a poverty trap. This is especially actual for rural households, given their limited access to labor opportunities. However, we cannot conclude that international migration from Tajikistan is purely "pro-poor" because richer households have higher returns to migration. When we also consider the positive effects of migration on overall subjective well-being, it can be hypothesized that migration might be construed not only in terms of wealth-expanding economic activity but also as an important factor contributing to life satisfaction. As far as we are concerned with development policies, the differences in the well-being patterns of households with and without migrants should be taken explicitly into account in the process of policy formulation. Since policymakers are concerned with a tradeoff between provision of social assistance to population and maintaining a balanced budget, this practice may facilitate better targeting of households in need. Consequently, monetary injections by municipal or non-governmental organizations can be used more efficiently. Finally, if migration and remittances are encouraged under a certain policy, it should be noted that their impact is relatively susceptible to observed and unobserved household characteristics. In this regard, the more effective approach might be to concentrate on small-scale policy reforms rather than introducing migration-related initiatives at the national level.

References

ADAMS JR., R. H. AND CUECUECHA, A. Remittances, household expenditure and investment in Guatemala. *World Development*, 2010, 38(11), pp. 1626–1641.

ALSAKKA, R. AND AP GWILYM, O. A random effects ordered probit model for rating migrations. *Finance Research Letters*, 2010, 7(3), pp. 140–147.

BAETSCHMANN, G., STAUB, K. E., WINKELMANN, R. Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A*, 2015, 178(3), pp. 685–703.

BAILEY, A. J., DRBOHLAV, D., SALUKVADZE, J. Migration and pastoral power through life course: Evidence from Georgia. *Geoforum*, 2018, 91, pp. 97–107.

BOHRA-MISHRA, P. Labor migration and investments by remaining households in rural Nepal. *Journal of Population Research*, 2013, 30(2), pp. 171–192.

CINGOLANI, P. AND VIETTI, F. Social remittances and local development in the Republic of Moldova. A critical analysis of migrants as agents of change. *Southeast European and Black Sea Studies*, 2019. DOI: 10.1080/14683857.2019.1657685.

CLÉMENT, M. Remittances and expenditure patterns in Tajikistan: a propensity score matching analysis. *Asian Development Review*, 2011, 28(2), pp. 58–87.

COLLIER, P. *The bottom billion: Why the poorest countries are failing and what can be done about it.* 1st Ed. New York, NY: Oxford University Press, 2008.

DANZER, A. M., DIETZ, B., GATSKOVA, K. *Tajikistan household panel survey: Migration, remittances and the labor market survey report.* Regensburg: Institute for East- and Southeast European Studies, 2013.

DANZER, A. M. AND IVASCHENKO, O. Migration patterns in a remittances dependent economy: Evidence from Tajikistan during the global financial crisis. *Migration Letters*, 2010, 7(2), pp. 190–202.

DE, P. K. AND RATHA, D. Impact of remittances on household income, asset and human capital: Evidence from Sri Lanka. *Migration and Development*, 2012, 2(2), pp. 163–179.

DÉMURGER, S. AND WANG, X. Remittances and expenditure patterns of the left behinds in rural China. *China Economic Review*, 2016, 37, pp. 177–190.

DICKERSON, A., HOLE, A. R., MUNFORD, L. A. The relationship between well-being and commuting revisited: Does the choice of methodology matter? *Regional Science and Urban Economics*, 2014, 49, pp. 321–329.

DUVAL, L. AND WOLFF, F-C. Remittances matter: Longitudinal evidence from Albania. *Post-Communist Economies*, 2010, 22(1), pp. 73–97.

ESQUIVEL, G. AND HUERTA-PINEDA, A. Remittances and poverty in Mexico: a propensity score matching approach. *Integration & Trade Journal*, 2007, 11(27), pp. 45–71.

GANG, I. N., GATSKOVA, K., LANDON-LANE, J., YUN, M.-S. Vulnerability to poverty: Tajikistan during and after the global financial crisis. *Social Indicators Research*, 2017, 138(4), pp. 925–951.

GRAHAM, C. AND NIKOLOVA, M. Bentham or Aristotle in the development process? An empirical investigation of capabilities and subjective well-being. *World Development*, 2015, 68, pp. 163–179.
GYIMAH-BREMPONG, K. AND ASIEDU, E. Remittances and investment in education: Evidence from Ghana. *Journal of International Trade & Economic Development*, 2015, 24(2), pp. 173–200.
HARRIS, J. R. AND TODARO, M. P. Migration, unemployment and development: a two-sector analysis. *American Economic Review*, 1970, 60(1), pp. 126–142.
HELLIWELL, J. F., HAIFANG, H., WANG, S. Changing world happiness. In: HELLIWELL, J. F., LAYARD, R., SACHS, J. D., eds. *World happiness report 2019*, New York: Sustainable Development Solutions Network, 2019.
INTERNATIONAL LABOUR ORGANIZATION. *Tajikistan sets up migration service* [online]. Moscow: International Labour Organization, 2011. [cit. 23.3.2020]. <https://www.ilo.org/moscow/news/WCMS_429179/lang--en/index.htm>.
JUSTINO, P. AND SHEMYAKINA, O. N. Remittances and labor supply in post-conflict Tajikistan. *IZA Journal of Labor & Development*, 2012, 1(8). DOI: 10.1186/2193-9020-1-8
KAN, S. AND AYTIMUR, R. E. Labor force participation of women left behind in Tajikistan. *Oxford Development Studies*, 2018, 47(1), pp. 1–28.
LU, Y. Household migration, remittances and their impact on health in Indonesia. *International Migration*, 2013, 51(1), pp. 202–215.
MCKENZIE, D., STILLMAN, S., GIBSON, J. How important is selection? Experimental vs. non-experimental measures of the income gains from migration. *Journal of the European Economic Association*, 2006, 8(4), pp. 913–945.
NGUYEN, L., BRENDA, S. A., TOYOTA, M. Migration and the well-being of the 'left behind' in Asia. *Asian Population Studies*, 2007, 2(1), pp. 37–44.
PETROVIĆ, M., HARRAP, G., KARDIKULOV, J. *"From Russia to Tajikistan": Changing the way money moves* [online]. United Nations Development Programme, 2017. [cit. 23.3.2020]. <https://www.eurasia.undp.org/content/rbec/en/home/blog/2017/7/25/-From-Russia-to-Tajikistan-changing-the-way-money-moves.html>.
RAVENSTEIN, E. G. The laws of migration. *Journal of the Statistical Society*, 1885, 48(2), pp. 167–235.
ROBINSON, S. AND GUENTHER. T. Rural livelihoods in three mountainous regions of Tajikistan. *Post-Communist Economies*, 2007, 19(3), pp. 359–378.
ŠIMKOVÁ, M. AND LANGHAMROVÁ, J. Remittances and their impact on the Czech economy. *Prague Economic Papers*, 2015, 24(5), pp. 562–580.
STARK, O. AND BLOOM, D. E. The new economics of labor migration. *American Economic Review*, 1985, 75(2), pp. 173–178.
STEVENSON, B. AND WOLFERS, J. Subjective well-being and income: Is there any evidence of satiation? *American Economic Review: Papers & Proceedings*, 2013, 103(3), pp. 598–604.
STRIELKOWSKI, W. AND WEYSKRABOVÁ, B. Ukrainian labour migration and remittances in the Czech Republic. *Tijdschrift voor Economische en Sociale Geografie*, 2014, 105(1), pp. 30–45.
SYROVÁTKA, M. Aggregate measures of welfare based on personal consumption. In: ŠAUER, P., eds. *Environmental economics, policy and international relations,* Conference proceedings, Prague: Litomyšl Seminar Publishing, 2007.
TODARO, M. P. A model of labor migration and urban unemployment in less-developed countries. *American Economic Review*, 1969, 59(1), pp. 138–148.
TOKHIROV, A. Remittances and subjective well-being of the left behinds in Tajikistan. *Economics Bulletin*, 2018, 38(4), pp. 1735–1747.
UNITED NATIONS. *Total international migrant stock* [online]. New York: United Nations, 2019. [cit. 23.3.2020]. <https://www.un.org/en/development/desa/population/migration/data/estimates2/estimates19.asp>.
WORLD BANK. *Migration and remittances factbook 2016.* 3rd Ed. Washington, DC: World Bank Group, 2016.
WORLD BANK. *Record high remittances sent globally in 2018* [online]. Washington, DC: World Bank Group, 2019. [cit. 23.3.2020]. <https://www.worldbank.org/en/news/press-release/2019/04/08/record-high-remittances-sent-globally-in-2018>.

Equivalence of Fault Trees and Stochastic Petri Nets in Reliability Modeling

Ondřej Vozár[1] | *University of Economics, Prague, Czech Republic*

Abstract

Modeling of reliability of the complex systems (machines, large networks, human body) is an important area of recent research. There are two main approaches applied: i) fault trees, ii) Petri nets. For the probabilistic study of a system is vital to know its minimal cut/minimal path sets. Both for fault trees and Petri Nets it is an NP-hard problem. Liu and Chiou (1997) described the equivalence of both representations for a given system. Furthermore, they found a top-down matrix algorithm to find critical cuts and minimal paths of the Petri net of the system. They claim without proof that their algorithm is more efficient than the ones for fault trees. We present both representations of a system. The algorithm is illustrated on a simple example of a three-masted vessel and a more complex "three-motor" system by Vesely et al. (1981).

Keywords	JEL code
Reliability, time to failure, fault tree, stochastic Petri net, exponential distribution	*C10, C18, C60*

INTRODUCTION

The demand for a more precise estimation of the reliability of complex systems has steadily increased because of both legal regulation and the growing complexity of real industrial systems. Such a complex system is for example a power plant, airplane, machine, reactor, large computer or transportation network, a human body. Applications of the methods range from energetics, engineering, transportation, computer science to safety studies, and medicine.

For reliability analysis of complex systems the following standard methodologies are used: i) fault trees (for the state of the art of this approach see Limnios, 2007), ii) Petri Nets (see monograph of Bause and Kritzinger, 2002).

Both methodologies have been further developed mostly assuming exponentially distributed time to failure. This assumption is unrealistic because it implies that parts of the system do not age.

In both approaches a system is represented as a tree in the language of the graph theory. If the fault tree method is applied, it is vital to know minimal cuts and minimal paths sets of a fault tree. However, finding all minimal cuts and minimal path sets of a fault tree is an NP-hard problem (Rosenthal,

[1] Faculty of Informatics and Statistics, Department of Statistics and Probability, University of Economics, Prague, Nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic. Also Czech Statistical Office, Na padesátém 81, 100 82 Prague 10, Czech Republic. E-mail: vozo01@vse.cz.

1975). On the other hand, a fault tree representation is suitable to derive the probability distribution of the reliability of a system and its estimators (Limnios, 2007).

Liu and Chiou (1997) established a one-to-one relation between a fault tree and a Petri Net of a system. They also proposed a top-down matrix algorithm to find all minimal cuts and minimal paths of the system. They claim without any proof this algorithm is more computationally efficient than algorithms for fault trees. To our best knowledge no complexity analysis or computational studies have been carried out. From the theoretical point of view Petri Nets is not a framework suitable to derive the probability distribution of reliability of a system and its estimators (Bause and Kitzinger, 2002).

If claim of Liu and Chiou (1997) is true, the following strategy for derivation of distribution of reliability functions and its estimators in general setting (gamma, log-logistic distribution of time to failure) would be efficient. First, transformation a fault tree of a system to a Petri Net is carried out to find minimal cut sets and minimal path sets. Then, all derivations are done in the fault tree setting.

The paper is organized as follows. The first part introduces elements of reliability theory. Secondly, concepts of fault trees and Petri Net are introduced. Thirdly, top-down matrix algorithm of Liu and Chiou (1997) to find all minimal cut sets and minimal paths sets is presented. Then, all the methods are illustrated in the model example of a three-masted vessel (Kubelka, 2016). Finally, both methods are applied to a more complex case of the "three-motor" system (Vesely et al., 1981) to assess Petri Nets based method to find minimal cut/minimal path sets.

1 ELEMENTS OF RELIABILITY THEORY

Systems are classified by their complexity to *Single-Component Systems* and *Multi-Component Systems*. The second classification distinguishes *Repairable* and *Non-Repairable Systems*. We focus on Non-Repairable Systems in this article only.

1.1 Single Component Systems

Let X be a continuous random variable representing time to failure of the system with cumulative distribution function $F(t) = P(X \leq t)$ and its density function $f(t)$.

Survival function (reliability) *is the complementary function to cumulative distribution function*:

$$R(t) = 1 - F(t) = P(X > t) = \int_t^\infty f(x)\, dx. \tag{1}$$

Note, that we have $R(0) = 1$ and $R(\infty) = 0$.

Hazard rate (instantaneous failure rate) at time t is defined as:

$$\lim_{\Delta t \to 0+} = \frac{P(t < X \leq t + \Delta t \,||\, X > t)}{\Delta t} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)}. \tag{2}$$

Mean time to failure (MTTF) is defined as the mean of the time to failure, e.g. $E(X)$.

The exponential distribution is the most used for modeling time failure in reliability theory. It gives a system without memory (a system is not aging) or a Markovian system, e.g. for fixed $x > 0$, $t > 0$ we have:

$$P(X > x + t \,||\, X > x) = P(X > t). \tag{3}$$

For fixed time $t > 0$ and parameter $\lambda > 0$ we get:

$$f(t) = \lambda e^{-\lambda t},\ F(t) = 1 - e^{-\lambda t},\ R(t) = e^{-\lambda t},\ h(t) = \lambda,\ MTTF = E(X) = 1/\lambda. \tag{4}$$

The assumption on exponentially distributed time to failure will be used in the section Application.

1.2 Multiple Components Systems

Let us consider a binary system with n components: $C = \{1, 2, \dots, n\}$. For each component i we define a binary variable x_i (0: the component is in good state, 1: the component is down).

Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \{0,1\}^n$ be the vector jointly describing the states of the components. We define a structural function $\boldsymbol{\varphi}(\boldsymbol{x})$ with values {0,1} as:

$\boldsymbol{\varphi}(\boldsymbol{x}) = 1$, if the system is in a good state, $\boldsymbol{\varphi}(\boldsymbol{x}) = 0$, if the system is down.

A system is said to be series if its good functioning depends on the functioning of all its components. If at least one component fails, then the system also fails. In fault tree setting is modeled by gate OR. The structural function of the *series system* is given by:

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \prod_{i=1}^{n} x_i \,. \tag{5}$$

A system is said to be parallel, if its good functioning is assured by functioning of least one of its components. Only if all components fail, then the system also fails. In fault tree setting is modeled by gate AND. The structural function of the *parallel system* is given by:

$$\boldsymbol{\varphi}(\boldsymbol{x}) = 1 - \prod_{i=1}^{n}(1 - x_i). \tag{6}$$

1.3 Coherent Systems

In reliability theory most of the techniques are limited only to *coherent systems*. A coherent system has these properties:

- it consists only of parallel and series systems (e.g. gates AND and gates OR),
- it has no redundant component (i.e. its states do not affect the state of the system),
- it does not contain a component and its negation simultaneously,
- it contains neither loops nor circuits in its graph representation.

For state assessment of a system following concepts are defined (Limnios, 2007):

- *path:* a subset of components whose simultaneous good functioning assures good functioning of the system regardless of the functioning of the other components,
- *minimal path:* a path which does not contain another path,
- *cut set:* a subset of components whose simultaneous failure leads to the system failure regardless of the failure of the other components,
- *minimal cut set:* a cut set that does not contain another cut set.

Set of the minimal paths is denoted as:

$$C = \{C_1, C_2, \dots, C_c\}.$$

Set of minimal cuts is denoted as:

$$K = \{K_1, K_2, \dots, K_k\}.$$

By minimal paths set or minimal cuts set its structural function is simplified as:

$$\varphi(\boldsymbol{x}) = 1 - \prod_{j=1}^{c} \left(1 - \prod_{i \in C_j} x_i\right), \tag{7}$$

$$\varphi(\boldsymbol{x}) = \prod_{j=1}^{k} \left(1 - \prod_{i \in K_j} (1 - x_i)\right). \tag{8}$$

1.4 Probabilistic Study of Coherent Systems

Let's consider a coherent system $S = (C, \varphi)$ of order $n \geq 1$ (number of its components). Let X_i be a Bernoulli random variable with parameter p_i, which describes the state of the component i ($i = 1, 2, \ldots, n$) with values $x_i \in \{0,1\}$. Then *reliability of the system* $R(\mathbf{p})$, where $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ is the probability, that the system is in a good state. We can express reliability also by minimal cut sets or minimal path sets as:

$$R(\mathbf{p}) = P(C_1 \cup C_2 \cup \cdots C_c) = 1 - P(K_1 \cup K_2 \cup \cdots K_k). \tag{9}$$

Different bounds for reliability function was established (see Limnios, 2007) using knowledge of minimal paths set and minimal cuts set.

Minimal sets bounds are established as follows. A lower bound is derived through minimal cut sets and an upper bound is derived through minimal paths sets:

$$\prod_{j=1}^{k} \left[1 - \prod_{i \in K_j} (1 - p_i)\right] \leq R(\mathbf{p}) \leq 1 - \prod_{j=1}^{c} \left(1 - \prod_{i \in C_j} p_i\right). \tag{10}$$

If the minimal cuts are 2-by-2 disjoint sets ($K_i \cap K_j$ is empty, if $i \neq j$), then upper bound equals to $R(\mathbf{p})$. The same rule applies to minimal paths and the lower bound of $R(\mathbf{p})$.

Trivial bounds are based on the observation that the reliability of a coherent system lies between the reliability of the series system and reliability of the parallel system:

$$\prod_{i=1}^{n} \mathbf{p}_i \leq \mathbf{R}(\mathbf{p}) \leq 1 - \prod_{i=1}^{n} (1 - \mathbf{p}_i). \tag{11}$$

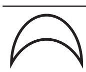
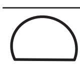
2 FAULT TREES AND PETRI NETS

From now all the methods will be illustrated on the example of a three-masted vessel (Kubelka, 2016). The vessel is maneuverable, if all its three components work (see Figure 1):

- the keel is not broken,
- the helm is maneuverable,
- at least one of the three masts are not broken,
- each mast is not broken, if both sail and spare are not broken.

2.1 Fault Trees

Table 1 Fundamental operators of fault trees

Graphic Symbol	Name	Meaning
	OR	The output is generated if at least one of the inputs exists
	AND	The output is generated if all the inputs exist

Source: Limnios (2007)

The basic notions and symbols of fault trees used in this paper are summarized in Tables 1 and 2.

Note that in this form fault trees representation of a system enables only static analysis of the reliability of the system. It is necessary to extend this setting by adding a time variable. However, this generalization is not the goal of the paper.

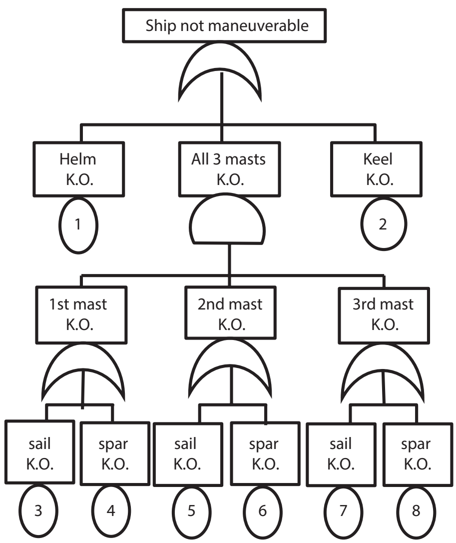
Table 2 Events of fault trees

Graphic Symbol	Meaning
Rectangle	Top or intermediate event (the system is down)
Circle	Basic event
Triangle	Transfer (fault tree is developed further)

Source: Limnios (2007)

Then the fault tree of the three-masted vessel is then constructed as – see Figure 1.

Figure 1 Fault tree of the three-masted vessel



Source: Kubelka (2016)

For such a simple system is easy to find all minimal cuts and paths manually. It has 10 minimal cuts: $K_1 = \{1\}$, $K_2 = \{2\}$, $K_3 =\{3,5,7\}$, $K_4 = \{3,5,8\}$, $K_5 = \{3,6,7\}$, $K_6 = \{3,6,8\}$, $K_7 = \{4,5,7\}$, $K_8 = \{4,5,8\}$, $K_9 = \{4,6,7\}$, $K_{10} = \{4,6,8\}$. For example minimal cut $K1$ means that the vessel is broken because the keel of the vessel is broken.

It has 3 minimal paths: $C_1 = \{1,2,3,4\}$, $C_2 = \{1,2,5,6\}$, $C_3 = \{1,2,7,8\}$. Minimal path C_1 means that the vessel is maneuverable, because of the keel, the helm, and the first mast (i.e. both its spar and sail are not broken) of the vessel work well.

Note that finding all minimal cuts of a coherent system is an NP-hard problem (Rosenthal, 1975). It means that the computational complexity of any algorithm grows exponentially with the number of components of a system. Even for a system with moderate size (a system with more than 20 components) is quite computationally demanding.

2.2 Petri Nets

Petri Nets (Petri, 1962) was designed to study information systems in computer science. They have been further developed and applied in many areas, also in modeling reliability of complex systems (see monograph of Bause and Kritzinger, 2002, among others). The basic notions and graphic symbols of Petri Nets are summarized in Table 3.

Table 3 Symbols and notions of Petri Nets

Graphic Symbol	Notion	Meaning
Circle	Place	Objects, components of a system
Dot	Token	Specific value, state of the object, component
Rectangle	Transition	Activities changing state or value of the object
Arrow	Arc	Connection of places and transitions

Source: Own construction by Bause and Kritzinger (2002)

For reliability modeling, Place-Transition Petri Nets and Stochastic Petri Net are used.

A Place-Transition Petri Net (Bause and Kritzinger, 2002) is defined as 5-tuple $PN = (P, T, I^{+}, I^{-}, M_0)$, where:

- $P = \{p_1, p_2, \ldots, p_m\}$ is a finite and non-empty set of places,
- $T = \{t_1, t_2, \ldots, t_n\}$ is a finite and non-empty set of transitions,
- $P \cap T = \emptyset$
- $I^{+}, I^{-}: P \times T \to N_0$ are oriented incidence functions (arcs),
- $M_0: P \longrightarrow N_0$ is a vector of the initial state of the systems.

Note that places represent for example a server or a hardware/software component of a system or a module of a software system. Transitions represent relations of different components of a system (i.e. transactions between servers or software components).

Place-Transition Petri Nets enable an only static analysis of the coherent system. To study the reliability of the system in time domain Place-Transition Petri Nets were generalized to Stochastic Petri Nets.

Stochastic Petri Net (Natkin, 1980; Molloy, 1981) with continuous time $SPN = (PN, \Lambda)$ is defined as Place-Transition Petri Net $PN = (P, T, I^{+}, I^{-}, M_0))$ equipped by a parameter set $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

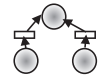
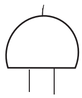
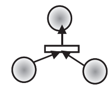
The role of a parameter set Λ is as follows:

- the parameter λ_i serves for modeling activation time of transition t_i,
- the transition t_i can be then activated only if there is a token in the corresponding place p_i,
- transition times T_i of a transition t_i are usually modeled as independent exponential random variables $T_i \sim Exp(\lambda_i)$.

2.3 Equivalence between fault tree and Place-Transition Petri Net

Liu and Chiou (1997) found in their seminal paper one-to-one relationship between Place-Transition Petri Net and corresponding fault tree.

Table 4 Fundamental operators of fault trees and its Petri Nets representation

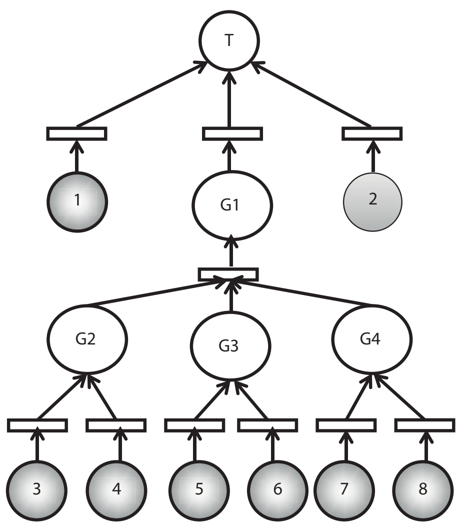
Fundamental Operator		Structural Function φ	Petri Net
Graphic Symbol	**Name**		
	OR	$\boldsymbol{\varphi}(\boldsymbol{x}) = 1 - (1 - x_1)(1 - x_2)$	
	AND	$\boldsymbol{\varphi}(\boldsymbol{x}) = x_1 x_2$	

Source: Own construction by Liu and Chiou (1997)

Both top and intermediate events and basic events are modeled as places in Petri Nets setting. This is why these events in the fault tree setting resemble a hardware/software component of a system.

Fault tree of the three-masted vessel (see Figure 1) is transformed into Stochastics Petri Net – see Figure 2.

Figure 2 Stochastic Petri Net of the three-masted vessel



Note: T: top event, G: gate (fundamental operators AND, OR), a place with token: .
Source: Own construction

Liu and Chiou (1997) also developed a recursive top-down matrix algorithm to find both minimal cut sets and minimal path sets simultaneously.

This method proceeds as follows:

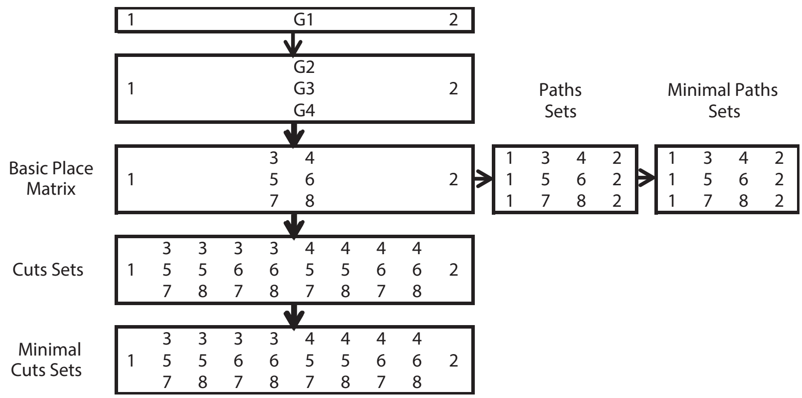
- write down the numbers of places horizontally if the output place is connected by multi-arcs to transitions,
- write down the numbers of places vertically if the output place is connected by an arc to a common transition,

- when all places are replaced by places representing basic events, a matrix is created. If there is a common entry located between rows or columns, it is also the entry present in each row or column. The column vectors of the matrix contain cuts sets, the row vectors then paths sets,
- finally select the minimal cuts sets and minimal path sets.

Liu and Chiou (1997) claim without any proof that this algorithm is much simpler than the corresponding ones for fault trees. By my best knowledge there have not been done any computational studies of algorithms to find minimal cut/path sets. Also, the efficient implementation of the algorithm is still an open problem. If the conjecture of the authors is true, then it would be efficient to transform the fault tree to Petri Net and find the minimal cuts/path sets. Anyway, to find minimal cut sets for a Petri Net is also an NP-hard problem. It follows from a one-to-one relation between a fault tree and a Petri Net and NP-hardness of minimal cut sets problem for a fault tree.

The schematic description of the method for the three-masted vessel is given in Figure 3.

Figure 3 Method for finding minimal cuts/paths sets by Liu and Chiou (1997)



Source: Own construction

The algorithm browses the Petri Net from its top places to down. The places from the current level are written down in one row of a matrix for a series system. On the other hand, the places from a parallel system are written down in one column of a matrix. Basic event (a place in a Petri Net) is represented by a number, intermediate event (logical gate of a Petri Net) is represented by symbols like G1, G2. It symbolizes another level of the Petri Net to be browsed in the next step. The algorithm stops, if we reached the lowest level of a Petri Net, i.e. it browsed all logical gates of a system. The all minimal cut/path sets of the three-masted vessel are found to derive its reliability function.

3 APPLICATIONS

Methods presented above are applied to the example of a three-masted vessel (Kubelka, 2016) and a more complex "three motors system" (Vesely et al., 1981). Time to failure of the components of both systems is assumed to be independent exponentially distributed random variables as well.

3.1 Reliability Function of the Three-Masted Vessel

Let us consider that time to failure all components of the three-masted vessel (Kubelka, 2016) are independent exponentially distributed with parameters given in Table 5.

Table 5 Time to failure distribution of the components

i	Component	MTTF (hours)	Ri(t)
1	Keel	2 000	$R_1(t) = e^{-t/2\,000}$
2	Helm	1 000	$R_2(t) = e^{-t/1\,000}$
3	Sail of the 1st mast	200	$R_3(t) = e^{-t/200}$
4	Spar of the 1st mast	500	$R_4(t) = e^{-t/500}$
5	Sail of the 2nd mast	200	$R_5(t) = e^{-t/200}$
6	Spar of the 2nd mast	500	$R_6(t) = e^{-t/500}$
7	Sail of the 3rd mast	200	$R_7(t) = e^{-t/200}$
8	Spar of the 3rd mast	500	$R_8(t) = e^{-t/500}$

Source: Own construction

The reliability functions of the keel and the helm are known. To derive the reliability function of the vessel the survival function of the three masts must be derived first.

The survival function of one mast is a product of survival functions of its sail and spar. It gives for $t > 0$:

$$R_m(t) = e - \frac{t}{200e} - \frac{t}{500} = e - \frac{7t}{1\,000}. \tag{12}$$

The system of all three masts is a parallel system of the masts, therefore its survival function is for $t > 0$:

$$R_{3m}(t) = 1 - (1 - e - \frac{7t}{1\,000}). \tag{13}$$

The three-masted vessel is a series system consisting of the keel, the helm, and three masts. Therefore, the survival function of the vessel is the product of survival functions of these three components, i.e. for $t > 0$:

$$R_{vessel}(t) = e - \frac{t}{2\,000}\, e - \frac{t}{1\,000}\left(1 - (1 - e - \frac{7t}{1\,000})^3\right) = e - \frac{3t}{2\,000}\left(1 - (1 - e - \frac{7t}{1\,000})^3\right). \tag{14}$$

The reliability of the vessel is strongly affected by the reliability of the system of three masts. Therefore, safety measures should focus mainly on the reliability of the sails and spars (see Table 6).

Table 6 Reliability function of the vessel

Time (hours)	Reliability function of			
	Keel	Helm	3 masts	Total
100	0.9512	0.9048	0.8724	0.7509
200	0.9048	0.8187	0.5724	0.4240

Table 6 (continuation)

Time (hours)	Reliability function of			
	Keel	Helm	3 masts	Total
300	0.8607	0.7408	0.3242	0.2067
400	0.8187	0.6703	0.1716	0.0942
500	0.7788	0.6065	0.0879	0.0415
600	0.7408	0.5488	0.0443	0.0180
700	0.7047	0.4966	0.0222	0.0078
800	0.6703	0.4493	0.0111	0.0033
900	0.6376	0.4066	0.0055	0.0014
1 000	0.6065	0.3679	0.0027	0.0006

Source: Own construction

Trivial bounds of the reliability function provide an extremely poor approximation of the reliability function (see Table 7). Minimal cuts set bound provide quite tight intervals for reliability function. Neither minimal cuts set, nor minimal paths sets are 2-by-2 disjoint, therefore neither lower bound nor upper bound equals the reliability function.

Table 7 Minimal sets bounds and trivial bounds of reliability function of the vessel

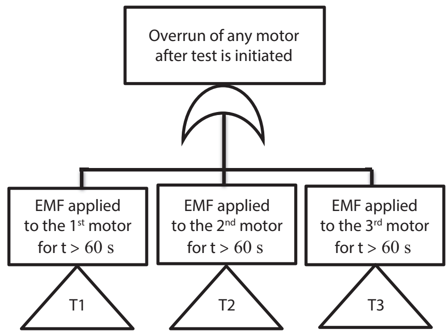
Time (hours)	Lower bound		Reliability function R(t)	Upper bound	
	Trivial	Minimal cuts sets		Minimal path sets	Trivial
100	0.10539922	0.70945428	0.75090010	0.81227589	0.99999832
200	0.01110900	0.28226174	0.42401196	0.45402751	0.99984388
300	0.00117088	0.07071146	0.20673107	0.21643080	0.99844529
400	0.00012341	0.01321632	0.09415493	0.09681565	0.99354886
500	0.00001301	0.00207403	0.04151348	0.04218520	0.98299794
600	0.00000137	0.00029443	0.01801734	0.01817896	0.96576308
700	0.00000014	0.00003957	0.00775945	0.00779717	0.94201188
800	0.00000002	0.00000517	0.00332898	0.00333761	0.91268718
900	0.00000000	0.00000067	0.00142551	0.00142745	0.87906209
1 000	0.00000000	0.00000009	0.00060985	0.00061028	0.84244015

Source: Own construction

3.2 Three-Motor System

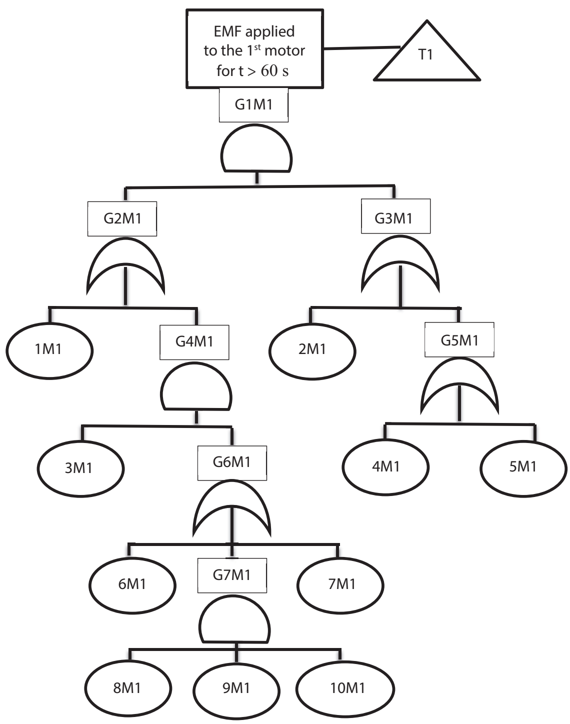
The three-motor system by Vesely et al. (1981) is often used as a benchmark for the assessment of algorithms in the field of reliability. It models a real-life control system of three motors. To shut down the system, it impresses a 60-second signal test. After 60 seconds, it is supposed to shut down all three

Figure 4 Fault tree for "three-motor" system



Source: Own construction, by Vesely et al. (1981)

Figure 5 Fault tree for the 1st motor in "three-motor" system

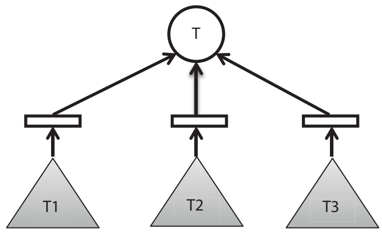


Source: Own construction by Vesely et al. (1981)

motors. The control then reveals failure of the system if the electromagnetic force (EMF) is applied to any of the three motors for more than 60 seconds after the signal test started.

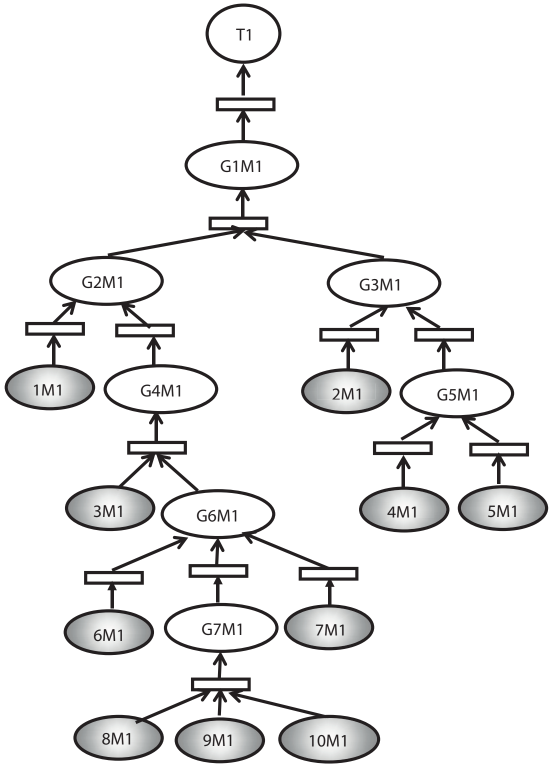
The motors are in series wiring to the control system (see Figure 4 for its fault tree and Figure 6 for its Petri Net). All three motors are the same (see Figure 5 for its fault tree and Figure 7 for its Petri Net). We refer readers to Vesely et al. (1981), p. 116 for the technical details of the motor components.

Figure 6 Place-Transition Petri Net for "three-motor" system



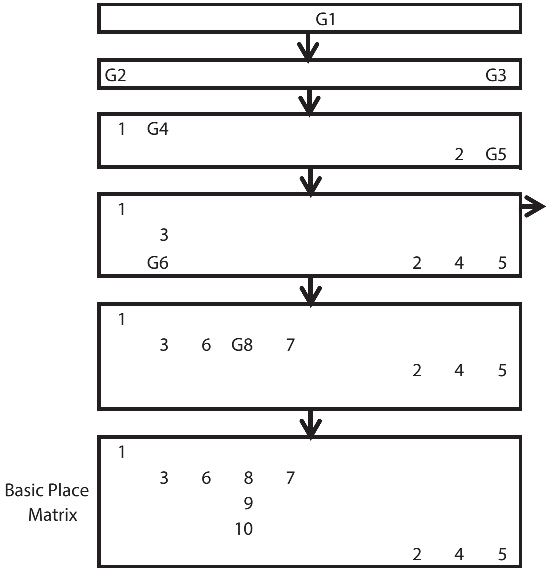
Source: Own construction by Vesely et al. (1981)

Figure 7 Place-transition Petri Net for the 1st motor in "three-motor" system



Source: Own construction by Vesely et al. (1981)

Figure 8 Basic Place Matrix of the 1st motor in "three-motor" system for finding minimal cut/path sets by method Liu and Chiou (1997)

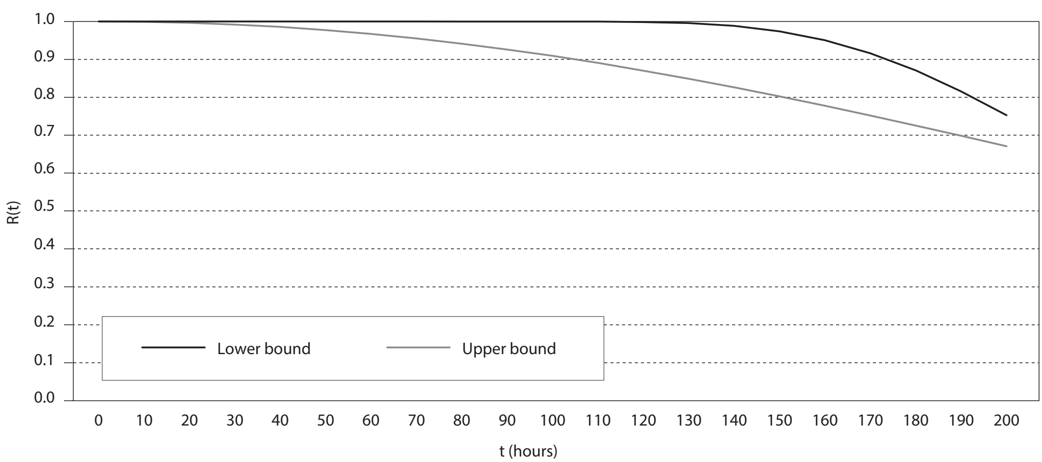


Source: Own construction

For simplicity, only a schematic description of the method of Liu and Chiou (1997) to find the Basic Place Matrix of the first motor is presented below (see Figure 8).

The first motor has 12 minimal cuts: $K_1 = \{1,2\}$, $K_2 = \{1,4\}$, $K_3 = \{1,5\}$, $K_4 = \{2,3,6\}$, $K_5 = \{3,4,6\}$, $K_6 = \{3,5,6\}$, $K_7 = \{2,3,7\}$, $K_8 = \{3,4,7\}$, $K_9 = \{3,5,7\}$, $K_{10} = \{2,3,8,9,10\}$, $K_{11} = \{3,4,8,9,10\}$, $K_{12} = \{3,5,8,9,10\}$. It has 4 minimal paths can: $C_1 = \{1,2,3,4,5\}$, $C_2 = \{1,2,4,5,6,7,8\}$, $C_3 = \{1,2,4,5,6,7,9\}$,$C4 = \{1,2,4,5,6,7,10\}$. We skip subscript of the first motor $1M$ to keep notation simple.

Figure 9 Bounds of the survival function of "three-motor" system



Source: Own construction

The system is a series system of three identical motors with $n = 30$ components. It means that the minimal cut sets of the system are a union of the minimal cut sets of the three identical motors. The minimal path sets of the system are the Cartesian product of the minimal path sets of the three identical motors. Therefore "three-motor" system has $3 \times 12 = 36$ minimal cuts and $4^3 = 64$ minimal paths. In such a complex case reliability function cannot be derived analytically.

Trivial bounds of the survival function (Formula (11)) usually fail for the systems with a high number of components. The bounds by Formula (10) often work quite well. For the system with independent, identically exponentially distributed times to failure (with MTTF = 1 000 hours, i.e. $\lambda = 1/(1\ 000)$ of the component's bounds are very tight (see Figure 9). After time t = 250 hours lower and upper bounds are almost identical.

CONCLUSION

Firstly, we reviewed standard approaches to modeling the reliability of complex systems – *fault trees* and *Petri Nets*. One-to-one relation between a fault tree and Petri Net of Liu and Chiou (1997) and their recursive top-down algorithm to find minimal cut sets and minimal paths sets were presented.

The algorithm was illustrated on a simple model example of the three-masted vessel and a more complex "three-motor" system of Vesely et al. (1981). The bounds of the reliability functions were established. Trivial bounds fail in the real complex systems. The minimal sets bounds seems to be a good approximation.

The two examples showed, that the recursive top-down matrix algorithm by Liu and Chiou (1997) is computationally demanding. It seems that the transformation of the fault tree to Petri Net to find minimal cuts and minimal paths sets more efficiently is not effective.

In future work we focus on comparison, computational complexity assessment, and computational study of minimal cut/minimal path sets algorithms for both settings.

ACKNOWLEDGMENT

This study was supported by the grant IGS F4/20/2020 (*Metody pro výpočet rizika, spolehlivosti a pohotovosti složitých systémů*), which has been provided by the Internal Grant Agency of the University of Economics in Prague. The author would like to thank two unknown referees for their valuable and constructive comments that helped considerably to improve the contents of the paper.

References

BAUSE, F. AND KRITZINGER, P. S. *Stochastic Petri Nets. An Introduction to Theory.* Braunschweig/Wiesbaden: Friedr. Wieweg & Sohn Verlagsgesellschaft GmBh, 2002. ISBN 3-525-15535-3

KUBELKA, V. *Use of random processes for calculation of risk, reliability and availability of complex industrial systems.* Master thesis, Prague: Faculty of Mathematics and Physics, Charles University, Department of Probability and Mathematical Statistics, 2016.

LIMNIOS, N. *Fault trees.* Newport Beach, CA: ISTE Ltd., 2007. ISBN 978-1-905209-30-9

LIU, T. S. AND CHIOU, S. B. The application of Petri Nets to failure analysis. *Reliability Engineering and System Safety*, 1997, 57, pp. 12–142.

MOLLOY, M. K. *On the Integration of Delay and Throughput Measures in Distributed Processing Models.* PhD thesis, University of California, 1981.

NATKIN. S. *Les Reseaux de Petri Stochastiques et leur Application a l'Evaluation des Systemes Informatiques.* PhD thesis, Paris: CNAM, 1980.

PETRI, C. A. *Kommunikation mit Automaten.* PhD thesis, Bonn: Universitaet Bonn, 1962.

ROSENTHAL, A. A computer scientist looks at reliability computations. In: *Reliability and Fault Tree Analysis*, SIAM, Philadelphia, 1975, pp. 133–152.

VESELY, W. E. et al. *Fault tree handbook.* Washington, D.C.: U.S. Nuclear Regulatory Commission, 1981. NUREG-0492

Using Decision Trees to Improve Variable Selection for Building Composite Indicators

Adrian Oțoiu[1] | *Bucharest University of Economic Studies, Bucharest, Romania*
Emilia Țițan | *Bucharest University of Economic Studies, Bucharest, Romania*

Abstract

The established variable selection methods for building composite indicators have strong limitations with respect to the results obtained. Some of them focus on getting an index structure with a high alpha reliability and/or a high percentage of the total data variance explained. These methods are likely to omit variables with strong explanatory power, and lead to an unsatisfactory classification of countries. Decision trees can also be used in selecting variables that are the most relevant for building composite indicators. An example of variable selection for building a composite indicator, which compares results using Cronbach's coefficient alpha, factor analysis, and decision trees, shows that the latter method yields comparable, or better results. Using cluster analysis on the selected variables, we show that the decision tree variable shortlist has better discrimination power than those obtained with the other methods, even in the presence of outliers and missing values.

Keywords

Composite indicators, decision trees, Cronbach's Alpha, factor analysis, indicator methodology, entrepreneurship indicators

JEL code

C43, I31, L26

INTRODUCTION

The interest in designing and publishing composite indicators is not new. However, in the past few years, there was a strong increase in the number of composite indicators. According to The Economist (2014a), their number increased from just under 20 before 1994 to about 100 in 2000–2004, and to over 150 active indicators for the period 2010–2014. They cover just about any dimension of human activity, from human development and social progress to the state of the environment, education performance, ease of doing business, democracy, corruption, entrepreneurship, etc. There is an index for any social issue or public policy (The Economist, 2014b).

While the aim of these indicators is a laudable one, to aggregate several measures relevant to one domain in a simple, easy-to-understand index by governments, think-tanks and campaigners (The Economist, 2014b), it is very often the case that they are found wanting. In some cases they are misleading, leading to rankings that defy common sense and are disproved by empirical evidence

[1] Bucharest University of Economic Studies, Piata Romana 6, Bucharest, Romania. Corresponding author: e-mail: otoiu.adrian@gmail.com.

(Otoiu et al., 2014). In other cases, there are significant overlaps between the aims and contents of several indicators, e.g. Global Entrepreneurship Monitor, Global Competitiveness Index, and several other indicator frameworks (Pekka et al., 2013), Global Gender Gap, Gender Inequality Index, Index of Women's Power and Glass Ceiling Index for gender inequality (The Economist, 2014a), Human Development Index, Legatum Prosperity Index and Social Progress Index for human development and social progress (Otoiu et al., 2014), Climate Analysis Indicators Tool (CAIT) and Climate Change Performance Index (CCPI) for climate change (Bandura, 2008).

The information provided by the composite indicators is of great, and sometimes crucial, potential value. Nowadays, however, there is a great danger of misusing this information, and hiding essential evidence behind the global public attention/debate generated following (new) releases of composite indicators. In the age of artificial intelligence, this translates into a new level of accountability for the information provided by composite indicators with respect to their ability to provide an accurate picture of the state of affairs of a particular domain, as a whole and for specific issues they address, as big data analytics and the use and linkages between different datasets is now the main generator of knowledge (Rometty, 2018).

In spite of the popularity, widespread use and interest in composite indicators, it does not appear that competition among them has been effective in weeding out the good indices from the bad ones. Rather, it seems that some indices manage to survive despite having dodgy methodologies, which are in some cases not disclosed, due to a good management of celebrity endorsements and media coverage that creates headlines (The Economist, 2014b).

However, the key to having an index that is a meaningful measure of a domain is the use of a sound methodology that ensures it is an effective multidimensional gauge of that domain. This consists, in the first stages, in identifying variables relevant to the domain that represents the main focus of the index, which is in many cases followed by making a selection from them based on their statistical properties. Then, appropriate methods of weighting and aggregation would ensure that composite indexes and/or sub-indexes are built in a reliable way, that yield measures relevant to the multidimensional concepts that they are supposed to quantify.

While there is sufficient knowledge and expertise available in choosing the variables for constructing models, there are few techniques singled out as established methods for variable selection. Reference manuals and handbooks point of to both ensuring relevance, timeliness, availability and trustworthiness of the source variables (Hsu et al., 2013), while the main statistical methods used for selecting the component variables are mainly principal component analysis/factor analysis, Cronbach alpha coefficient, and cluster analysis (OECD, 2008). In other cases, correlation analysis was used to remove candidate variables that were similar in content and had similar correlation patterns with other variables (Otoiu et al., 2014).

Considering the fact that all methods used in selecting component variables for composite indicators are based on multivariate analysis, by taking into account their relative contribution in explaining their overall variation, this paper proposes a novel method for selecting variables using the decision tree method. In order to test our approach, we make a comparison between it, the Cronbach's alpha measure, and exploratory factor analysis for selecting variables used to construct two human progress indexes, the Human Development Index 2014 and the Legatum Prosperity Index 2014. We used then cluster analysis to assess the extent to which selected variables enable us to obtain a better cluster structure in terms of the ability to explain the total variability of the data and of the quality of clusters obtained, assessed using silhouette widths (Rousseeuw, 1987).

1 LITERATURE REVIEW

Selecting input variables is done primarily on the basis of their relevance to the construction of the index. This is based on a thorough knowledge of the field and the availability of the data. According to OECD (2008) the first step in constructing a composite indicator is "developing a theoretical framework" by defining the concept that will be measured by the indicator, determining the sub-groups corresponding to it in the case of conceptual multidimensionality (e.g. well-being), and identifying the selection criteria and methodology for component variables. The next step, which constitutes the focus of this paper, consists of selecting the variables that make up the index, or sub-indexes when there are subgroups, based on their strengths and their quality (OECD, 2008). Quality is often addressed in the methodological papers of several indexes, which shows that variables are selected based on their reliability, availability and timeliness (Porter and Stern, 2014; Legatum Institute, 2012).

OECD (2008) establishes that "an index is above all the sum of its parts" which should be kept in mind when building and assessing the outcomes of composite indicators. With respect to the specific properties that composite indicators should have, Paun (1983), cited by Pele (2008) establishes that "A composite indicator sensitive and anti-catastrophic is compensatory". This means that: 1) the improvement recorded in one component should be reflected in the improvement in the overall indicator Pele (2008), 2) it is not possible that a small change in one component effects a large change in the composite indicator. Both properties mean that a big change of one component is not accompanied by an opposite modification of another, so that the resulting values of the composite indicator will equal those obtained in the absence of this change (Pele, 2008).

Michalos et al. (2011) propose the following properties for an acceptable indicator: 1) relevant to the concerns of our main target audiences, 2) easy to understand, 3) reliable, valid, and sensitive to changes, 4) politically unbiased, 5) timely, easy to obtain, and periodically updated, 6) comparable across jurisdictions and groups, 7) objective or subjective, 8) positive or negative, 9) obtained through an open, transparent, and democratic consultative review process.

There are essentially two types of indicators, ones that are based on a relatively large number of variables e.g. Legatum Prosperity Index, Social Progress Index, and others that are based on only a few variables, such as the Human Development Index. While there are differences in the methodological approaches, most composite indicators that are reliable and enjoy a certain degree of prestige are based on a well-documented and scientifically sound methodology. A part of this methodology covers how component variables are selected. According to OECD (2008), the main statistical methods of selection are the following: principal component analysis/factor analysis, Cronbach alpha coefficient, and cluster analysis, which are widely used in practice to establish the shortlist of variables. A recent review paper (Gan et al., 2017) points out that the main analytical techniques used for constructing 96 sustainability indicators are principal component analysis/factor analysis and regression analysis, which both account for about 59% of all analytical techniques used. Of the total (analytical, opinion based, and equal weights) techniques used in weighting component variables of composite indicators and sub-indicators (weighting being a step in constructing aggregate indicators that follows variable selection), 11.46% employ principal component analysis/factor analysis, and another 6.25% employ regression method based on the use of a dependent variable relevant to the 'target', that is to the composite indicator or a sub-indicator (Gan et al., 2017).

Recently, some approaches have expanded and challenged the variable selection methods. One issue is the tradeoff between the choice of either a few variables, or a wide range of variables to capture the latent aspects of multidimensional concepts for which no specific variables to describe it are available (Foa and Tanner, 2012). Another one is the use of a "target measure" to help select the relevant variables for building a composite indicator. In this respect, the work of Abberger et al. (2018) shows that, even if established methods such as principal components and correlation analysis are used for variable selection, there is still the need to establish some reference variables that will guide construction and revision of composite indicators.

In this context, decision trees represent a predictive data mining technique whose primary use is to predict the outcome of one target variable based on the evolution of several explanatory variables.

Its results recommend it for the use of variable selection as the algorithm has the ability to discriminate between variables that have a significant explanatory power of the evolution of the target (dependent) variable. While there are other similar techniques which model the evolution of the dependent variables based on one or more independent variables (e.g. linear regression), decision trees are deemed to have the following advantages (Enachescu, 2009): ease of implementation and interpretation, robustness with respect to outliers and missing values, variable selection done taking into account interactions between variables.

Decision trees are also used in modelling the relationship between one target variable and several explanatory variables. In several fields, e.g. credit risk modelling and marketing, they are considered to be established techniques, which yield results comparable to other methods. Results by Hand and Henley (1997) show that decision tree methods yield comparable results to linear regression and logistic regression. Also, Siddiqui (2005) identifies decision trees as one of the key classification techniques used in statistical-based customer segmentation for credit scoring, and lists them as one of the methods used in building scorecards, along with logistic regression and neural network techniques. Customer segmentation using decision trees is popular in the telecom industry, used in predicting customer value (Weiss, 2005). Other applications of decision trees are found in medical sciences, in assessing the relative importance of variables identified as risk factors for major depressive disorder (Song and Lu, 2015) or for other diseases.

2 A BRIEF OVERVIEW OF THE VARIABLE SELECTION METHODS

The OECD handbook for constructing composite indicators (OECD, 2008) mentions three methods for selecting variables: factor analysis/principal components analysis (FA/PCA), scale reliability and cluster analysis.

FA/PCA emerges as a method of choice for selecting variables at it enables researchers to see the relationships between variables (OECD, 2008). Essentially, for variable selection, the exploratory factor analysis methods are used, which entail representing the observed variables as a function of parameters computed for unobserved factors (Preacher et al., 2013).

As Cooper (1983) points out, factor analysis comprises a set of multivariate statistical techniques, among which principal components analysis (PCA). While similar to FA in terms of results and use, PCA does not assume any relationship between the underlying structure of the variables (Cooper, 1983), and the components extracted (the equivalent of factors in FA) are, by design, orthogonal to each other. Due to this consideration, we have chosen FA as it may be better suited for exploring the relationships between variables based on their communalities (Cooper, 1983).

The scale reliability method measures the internal consistency of several variables used in building a single indicator (OECD, 2008). Its use for variable selection ensures that selected variables measure a single dimensional item (Nardo et al., 2005). The consistency is measured by the Cronbach coefficient alpha, reported in two forms, an unstandardized and a standardized form. Following the recommendation of Falk and Savalei (2011), we will use the standardized form, due to the fact that later stages of composite index construction will certainly require the use of a normalization technique that makes possible the aggregation of variables with different units of measurement into a single aggregate measure.

The third method, cluster analysis, is mostly a descriptive tool used to group countries (OECD, 2008) with the purpose to give some insights of the overall structure of the variables involved (OECD, 2008). It does not give a direct assessment of the contribution of each variable, and, in our opinion, its value mostly lies with checking whether results of variable selection done with other methods match the clusters obtained for shortlisted variables, and are compatible with index scores (Otoiu et al., 2014). Given these considerations we will use cluster analysis to compare the results obtained using other variable selection methods in terms of the quality of the cluster structures obtained.

3 USING DECISION TREES TO SELECT VARIABLES

Decision trees is an important technique used in data mining/knowledge discovery. It consists of classification of a target variable into a set of classes, based on the values of explanatory variables (Rokach and Maimon, 2014). This is performed using an algorithm known as recursive partitioning (Izenman, 2008), which is a step-by-step process by which a node is split, or not, into child nodes (Izenman, 2008) by asking a sequence of Boolean questions of type: is a value of a variable, on which the split is done, lower than a threshold value, or not? The process starts with a first, or root, node, and ends when an optimal solution is found based on criteria specific to the algorithm used (Izenman, 2008).

Among the competing algorithms used for building decision trees, we have chosen one of the most widely used, Classification and Regression Tree (CART), developed by Breiman et al. (1984) due to the fact that is one of the most commonly used, yielding results that are clear and easy to interpret.

Many software packages compute variable importance, which shows which variables control the classification process (Izenman, 2008). This is achieved through calculating the sum of the improvement scores for each explanatory variable, for all the nodes where it acts as a primary or surrogate splitting variable (Gebre-Sellasie et al., 2011; Thierneau and Atkinson, 2019). In the rpart package, which contains an implementation of CART in R, and will be used for variable selection later in this paper, variable importance is expressed as a percentage for each variable selected for tree construction by scaling the improvement sums for all variables to 100, and discarding those with proportions smaller than 1 (Thierneau and Atkinson, 2019).

The use of decision trees for variable selection is not as straightforward as the use of factor analysis and cluster analysis, as these techniques do not require the existence of a target variable and thus are ideal candidates for establishing the structure of the new variable, the resulting composite indicator or sub-index. The apparent difficulty of using this technique, given by the fact that it is a "supervised learning method" which requires the existence of a target variable (Bishop, 2007), can be overcome by the use of a target variable whose explanatory power is well-known, trusted, and relevant with respect to the goals and concept behind the developed composite indicator. This practice is used by the Legatum Institute (2013) in developing its Legatum Prosperity Index based on linear regressions of potential input variables on GDP per capita and overall self-reported life satisfaction, both considered to be the most reliable actual dimensions of human well-being. A similar approach is used by the Abberger et al. (2018) for improving the Swiss composite leading economic indicator by selecting variables based on their relationship with a reference, or target, variable.

4 CASE STUDY: USING DECISION TREES IN VARIABLE SELECTION

The feasibility and use of decision trees in selecting variables for constructing or revising a composite indicator will be shown in an application which attempts to select input variables for building a composite indicator of well-being. Data comes from Otoiu et al. (2014), where a validation of three of the most popular indexes of well-being and human progress is performed, namely the Human Development Index (HDI), Happy Planet Index (HPI), and Legatum Prosperity Index (LPI). Validation is based on a list of candidate variables deemed relevant for defining the multidimensional concept of well-being: 1) growth of CO^2 emissions (CO2GRW), 2) CO^2 emissions per capita (CO2PerC), 3) forest area as a share of total land area (Forest), 4) Gross National Income (GNI) per capita (GNIPC), 5) greenhouse gas emissions per capita (GHGPC), 6) Gini coefficient for income (GINI), 7) life expectancy at birth (LifeExpB), 8) natural resource depletion (NResDep), 9) mean years of schooling (YRSSch), 10) labor force participation rate for men (PRM), 11) labor force participation rate for women (PRF), 12) total labor force participation rate (PRT), 13) overall life satisfaction (SATISF), 14) share of fossil fuels in fuel consumption (ShareFF), 15) share of renewables in resource consumption (ShareRen), 16) unemployment rate (UE),17) urban pollution (Upoll), 18) well-being (WellB), computed as the arithmetic mean of individual responses to the Ladder of Life question in the Gallup World Poll.

In order to prove the validity and relative performance of decision trees used as a variable selection method, we will compare its results with those obtained using Cronbach's coefficient alpha and FA method. For the target variable used in the decision tree method, we chose GNIPC and WellB, as the best available proxies for the subjective and, respectively, objective well-being. These are the two major sides of the concept of well-being, described both in terms of material progress and an improvement in the living standards and conditions that will enable individuals to reach their goals given the opportunities available within a country at a certain time. A detailed explanation of the concept of well-being, which explains its bivalent nature, can be found in Otoiu et al. (2014), and the rationale for this approach can be found in the construction of the LPI (Legatum Institute, 2013).

All estimations are done using the R package **rpart** for decision trees, the **factanal** function for FA and the **reliability** function Cronbach's coefficient alpha, the latter two implemented in the RCommander graphical user interface. Results will be checked with the cluster analysis method Partition Around Medoids, implemented in the R "cluster" package. Compared to other clustering methods, this one produces graphs and diagnostics that enable an easy assessment of the quality of the cluster structure, both as a whole and for individual clusters.

Table 1 Results of decision trees estimation

Regression Tree Statistics

Target variable	CP	nsplit	rel error	xerror	xstd
GNIPC	0.597278	0	1	1.00662	0.145317
	0.153842	1	0.40272	0.56682	0.134489
	0.041019	2	0.24888	0.41913	0.111481
	0.04	3	0.20786	0.39637	0.088197
WellB	0.703880	0	1	1.00771	0.085633
	0.157171	1	0.29612	0.33884	0.033161
	0.048922	2	0.13895	0.17359	0.019016
		3	0.09003	0.12174	0.016763

Variable importance (percent)

Target variable	Explanatory Variables							
	CO2PerC	LifeExpB	YRSSch	PRM	WellB	SATISF	UE	ShareRen
GNIPC	24	19	15	13	13	12	4	1

Target variable	Explanatory Variables								
	SATISF	GNIPC	LifeExpB	CO2PerC	YRSSch	PRF	PRT	Forest	Upoll
WellB	35	19	18	13	8	3	2	1	1

Note: CP – complexity parameter, nsplit – number of splits, rel error – relative misclassification error, xerror – cross-validated error, xstd cross-validated standard deviation.
Source: Authors' calculation based on Otoiu et al. (2014) data

The results of the decision trees built using the two target variables are presented in Table 1. They indicate that decision tree structures obtained are fairly robust, with a goodness of fit (R^2) of about 79% for the GNIPC variable and, respectively, 91% for the WellB variable, computed from the misclassification errors for the last split (rel. error).

The resulting tree structures themselves were not particularly useful as they were too restrictive, with only three variables used in computing the solution for GNI per capita (CO2PerC, LifeExpB and UE) and only one for Well-being (SATISF, which is a strong proxy for well-being). However, 13 out of 20 variables reported as having a significant relative importance were retained and used to build a cluster structure that would discriminate countries based on their inferred level of well-being.

Variable selection done using the Cronbach's Alpha coefficients on the same variables has yielded the results shown in Table 2.

Table 2 Variable selection results using the scale reliability measure

Variable name	Alpha	Std.Alpha	R (item, total)
CO2PerC	0.0019	0.9277	0.7392
GNIPC	0.6876	0.9005	0.8226
LifeExpB	0.0016	0.9077	0.6737
SATISF	0.0026	0.8972	0.7285
WellB	0.0026	0.8992	0.7169
YRSSch	0.0024	0.9129	0.6522

Note: R (item, total) represents the correlation between one variable and the average behaviour of all variables.
Source: Authors' calculation based on Otoiu et al. (2014) data

Selection was done by deleting items for which R item scores, which compute the correlation between each item and the sum of the other items (Fox, 2012), is small. The final selection has a much lower number of variables, but the standardized alpha of 0.922 corresponding to the selected variables shown in Table 2, indicates a strong variable structure that can be used for calculating a composite indicator, fact confirmed by item scores well above 0.5, showing that each retained element has a significant contribution to the construction of the composite index.

Table 3 Factor loadings and diagnostic measures of factor analysis

	Factor loadings			
Variable	**Factor1**	**Factor2**	**Factor3**	**Factor4**
CO2PerC	0.269	0.196	0.927	0.157
Forest	0.106	0.282	–0.365	n/a
GNIPC	0.544	0.551	0.453	n/a
LifeExpB	0.555	0.548	0.221	n/a

Table 3 (continuation)

Variable	Factor1	Factor2	Factor3	Factor4
Factor loadings				
SATISF	0.948	0.236	0.2	n/a
ShareFF	0.134	0.162	0.266	0.938
Upoll	–0.151	–0.552	0.112	n/a
WellB	0.934	0.263	0.21	n/a
Diagnostic measures				
SS loadings	2.576	1.756	1.34	1.121
Proportion of variance explained	0.286	0.195	0.149	0.125
Cumulative variance explained	0.286	0.481	0.63	0.755

Source: Authors' calculation based on Otoiu et al. (2014) data

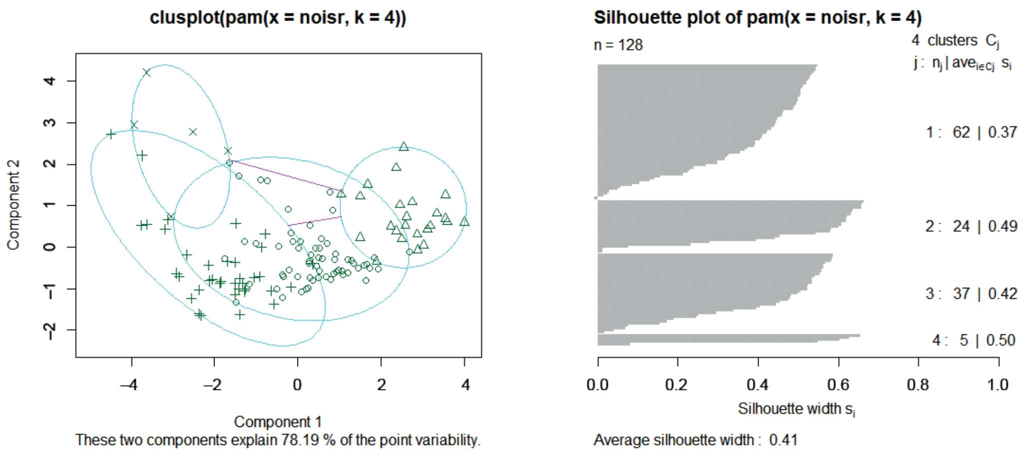
In achieving the factor solution presented in Table 3, we had to give up the GINI variable for which there was a high incidence of missing values. Results retain the variables for which factor loadings were higher than 0.5, as per the recommendation of the OECD composite indicators manual (OECD, 2008), and exclude the Forest variable which does not comply with this requirement. The solution presented still includes this variable as, by using it, we have obtained the best factor structure, which explains 75.5% of the total variability of the data set, and includes variables with significant factor loadings.

Further, in order to validate the results obtained using the three variable selection approaches, and show that they can be used to design a composite indicator that can classify countries based on their level of well-being, we use cluster analysis to compare the three sets of selected candidate variables in order to assess how well their total variability can be captured by cluster structures that classify countries into different groups based on their well-being.

The validation procedure is similar to the first part of the one done by Otoiu et al. (2014). An optimal cluster structure is obtained from selected input variables, with a high percentage of data variability being explained by the cluster structure (above 60%), silhouette plots averaging over 60% for the entire structure, and values of 50% or more, together with no misclassification for individual clusters (Otoiu et al., 2014). The cluster structures were obtained by eliminating observations with missing values. We considered this to be the right approach given the nature of the data and the exploratory aim of our approach. Using imputation techniques could have biased the results due to the fact that we deal with different countries whose unique characteristics may not be properly inferred through using a purely quantitative method. Due to this fact, in presenting our results, we have not compared the percentage of data variability explained by the cluster structure, as it would favor structures with fewer variables.

The optimal cluster structure obtained with Cronbach's Alpha coefficients, presented in Figure 1, shows that the selected variables, CO2PerC, LifeExB, SATISF, YRSSch, GNIPC, and WellB, manage to explain about 78% of the overall data variability. However, the cluster structure is fairly weak as average silhouette widths for the total cluster structure, and for three out of four individual clusters, is below 50%. Moreover, cluster membership is sizably unbalanced, with cluster 1 grouping almost half of the data, and cluster 4 having only 5 observations. The results are weak if the degree of misclassification is considered, shown

Figure 1 Optimal cluster structure with scale reliability variables

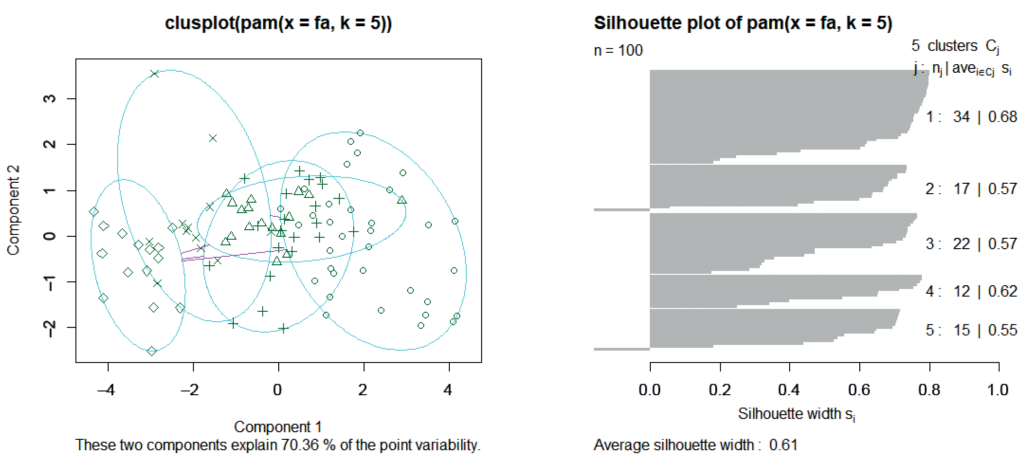


Source: Authors' calculation

as the part of the silhouette plots that extend below 0. As only marginal misclassification is observed, we conclude that, indeed, this structure is weak and cannot be used to achieve a satisfactory grouping of countries based on the selected variables.

For the factor analysis technique results, presented in Figure 2, show a balanced cluster structure, with an average silhouette width of 0.61, cluster individual widths above 0.55, and some sizable misclassification for cluster 2 and 5. Indeed, clustering based on 8 variables, CO2PerC, LifeExB, SATISF, YRSSch, GNIPC, WellB, Upoll, and ShareFF is able to achieve a fair discrimination of countries that would show their different levels of well-being in an appropriate way.

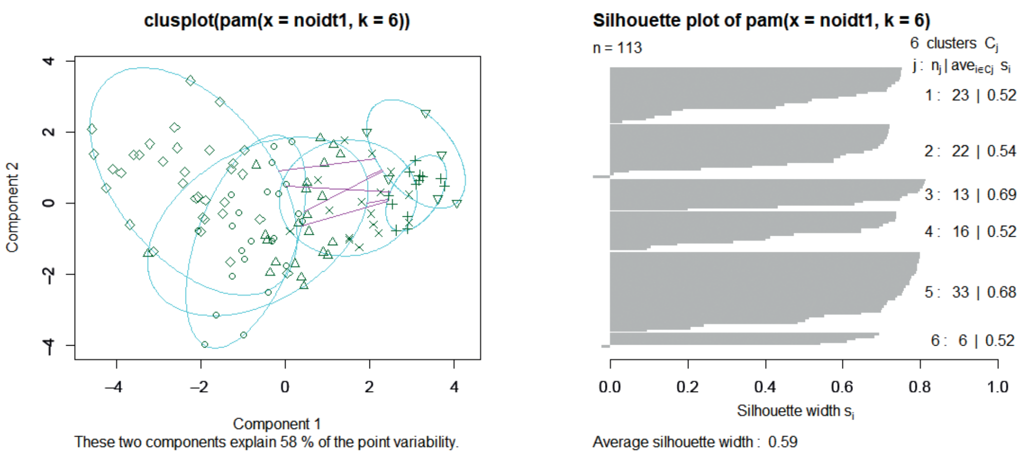
Figure 2 Optimal cluster structure with factor analysis variables



Source: Authors' calculation

The optimal cluster structure obtained for the variables selected with the decision tree algorithm, presented in Figure 3, has a slightly better performance than results obtained using factor analysis. The percentage of the overall variability explained is lower due to inclusion of 11 variables, CO2PerC, LifeExB, SATISF, YRSSch, GNIPC, WellB, Upoll, PRT, UE, Forest, and ShareRen. Albeit the average silhouette width is slightly worse (0.59 vs. 0.61) than for the structure obtained from variables selected using factor analysis, we see a very small of misclassification for cluster 2 and 6. Furthermore, all cluster-specific silhouette widths are above 50%, with rather high values of 68% or over for two clusters (3 and 5). Finally, for clusters 1 and 4, with the worst silhouette widths, no misclassification was observed. A comparison done with the six-cluster solution using variables selected with the FA method, presented in Appendix 1, shows a clear superiority of the solution using decision trees, with higher average silhouette width, strong misclassifications for 4 out of 6 clusters, and one cluster silhouette width below 0.5.

Figure 3 Optimal cluster structure with decision trees variables



Source: Authors' calculation

In sum, the decision tree method proved to be the most effective way of selecting candidate variables. The cluster structure obtained is able to discriminate six country groups with minor misclassification, obtaining clusters with a rather balanced number of elements. While some overall diagnostic measures may indicate that the cluster structure obtained using factor analysis for variable selection is better than the structure obtained using the variables selected with the rpart algorithm, interpretations should consider the decrease of average silhouette width which occurs when more variables are used, and the higher degree of misclassification observed for the former structure.

DISCUSSION AND CONCLUSIONS

Using decision trees to select variables for building composite indicators is a valid alternative to the established methods of variable selection. Its advantages lie in the fact that decision trees can work with virtually any type of variable, and that they are not very sensitive to outliers and missing values. Due to their features, variable selection is likely to be more complex and take into account the relevant features of the data set to a greater extent than when the other established methods are used.

The selection of relevant variables for defining well-being obtained with the decision tree algorithm, implemented with the rpart package, is better than the one which used Cronbach's alpha coefficients, and close to the one which employed factor analysis. The cluster structure obtained was markedly better, providing a higher number of country groups with little or no misclassification rate.

Several elements need to be taken into account when considering the use of decision trees as a variable selection method for composite indicators. The most important is the existence of variables which are representative proxies of the multidimensional concept that is analyzed. In our example, GNI per capita and Life satisfaction measures were used as the best available measures of well-being.

In the absence of an established proxy, building a naïve index with equal weights assigned to standardized/normalized candidate input variables, may provide an initial solution for the target variable to be used.

In some cases, decision trees can be extremely useful for designing parsimonious indexes, which employ only a few variables. A particular situation is relevant to the case of entrepreneurship indicators, when there are significant overlaps between variables (Pekka et al., 2013), and using decision trees can yield a better indicator structure that can clearly describe the multifaceted dimensionality of entrepreneurship.

Another example is HDI, which is calculated from 4 variables: Life expectancy at birth, Mean Years of Schooling, Expected Years of schooling, and GNI per capita. While this feature was not explored in this paper, further research may be able to assess the strengths of this method, and improve some of the composite indicators which use a large number of variables.

Another issue worth exploring is employing some rules for input variable selection using decision trees. It may be worth researching whether all variables selected as important are to be included in the development of a composite indicator, or a variable importance threshold should be established that would keep variables with large importance scores, and discard those with low importance, e.g. Urban pollution or Forest area in our case study.

ACKNOWLEDGEMENT

This paper was co-financed from the Human Capital Operational Program 2014–2020, project number: POCU/380/6/13/125245 No. 36482/23.5.2019 "Excellence in interdisciplinary PhD and post-PhD research, career alternatives through entrepreneurial initiative (EXCIA)", coordinator: The Bucharest University of Economic Studies.

References

ABBERGER, K., GRAFF, M., SILIVERSTOVS, B., STURM, J-E. Using rule-based updating procedures to improve the performance of composite indicators. *Economic Modelling*, 2018, 68, January, pp. 127–144.

BANDURA, R. A Survey of Composite Indices Measuring Country Performance: 2008 Update. *Studies Working Paper 2008–02*, United Nations Development Programme, Office of Development.

BISHOP, C. M. *Pattern recognition and machine learning.* New York: Springer-Verlag, 2006.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A., STONE, C. J. *Classification and Regression Trees.* Wadsworth, Belmont, 1984.

COOPER, J. C. B. Factor Analysis: An Overview. *The American Statistician*, 1983, 37, 2, pp.141–147.

GAN, X. et al. When to use what: Methods for weighting and aggregating sustainability indicators. *Ecological Indicators*, 2017, 81, pp. 491–502.

GEBRE-SELASSIE, G., VAN BORTEL, W., LEGESSE, W., YEWHALAW D. Malaria and Water Resource Development. In: HUNTER, W. 3rd Ed. *Recent Advances and Issues in Environmental Science*, Oakville: Apple Academic Press, 2011.

ENĂCHESCU, D. *Data Mining: metode și aplicații. Editura Academiei Române*, 2009.

FALK, C. F. AND SAVALEI, V. The Relationship Between Unstandardized and Standardized Alpha, True Reliability, and the Underlying Measurement Model. *Journal of Personality Assessment*, 2011, 93, 5, pp. 445–453.

FOA, R. AND TANNER, J. C. Methodology of the Indices of Social Development. *ISD Working Paper Series 2012–04*.

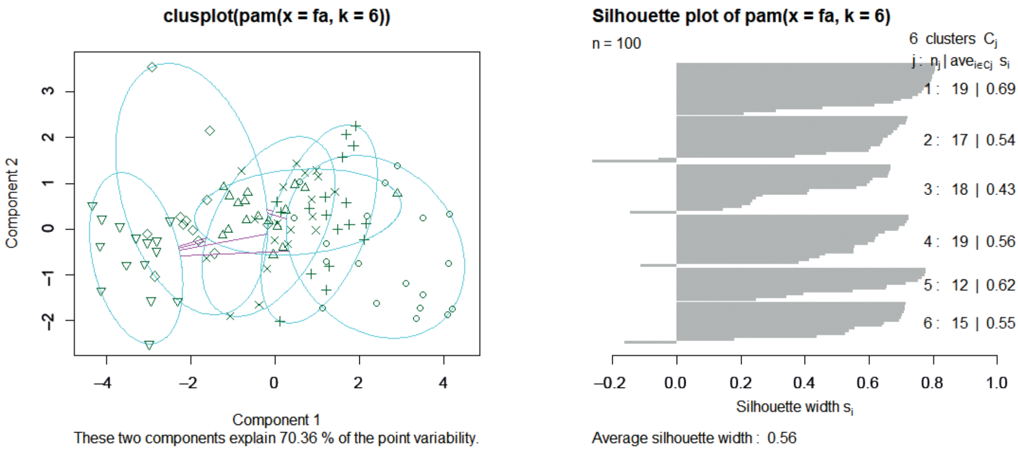
FOX, J. *Documentation for package 'Rcmdr' version 1.8–4* [online]. 2008. [cit. 5.11.2015]. <http://artax.karlin.mff.cuni.cz/r-help/library/Rcmdr/html/00Index.html>.

IZENMAN, A. J. *Modern Multivariate Statistical Techniques.* New York: Springer, 2008.

HAND, D. J. AND HENLEY, W. E. Statistical Classification Methods in Consumer Credit Scoring: a Review. *Journal of the Royal Statistical Society Series A*, 1997, 160(3), pp. 523–541.
HSU, A., JOHNSON, L. A., LLOYD, A. *Measuring Progress: A Practical Guide from the Developers of the Environmental Performance Index (EPI).* New Haven: Yale Center for Environmental Law & Policy, 2018.
LEGATUM INSTITUTE. *The 2012 Legatum Prosperity Index.* United Kingdom: The Legatum Institute, 2012.
LEGATUM INSTITUTE. *The Legatum Prosperity Index 2013. Methodology And Technical Appendix* [online]. United Kingdom: Legatum Institute, 2013. [cit. 2.11.2015]. <http://media.prosperity.com/2013/pdf/publications/Methodology_2013_FinalWEB.pdf>.
MICHALOS, A. C., SMALE, B., LABONTÉ, R., MUHARJARINE, N. T*he Canadian Index of Wellbeing. Technical Report 1.0* [online]. 2011. [cit. 6.2.2013]. <http://medcontent.metapress.com/index/A65RM03P4874243N.pdf>.
NARDO, M., SAISANA, M., SALTELLI, A., TARANTOLA, S. Tools for Composite Indicators Building. Joint Research Centre. *European Commission report EUR 21682 EN*, 2005.
OECD. H*andbook on Constructing Composite Indicators: Methodology and User Guide.* Paris: OECD, 2008.
OTOIU, A, TITAN, E, DUMITRESCU, R. Are the variables used in building composite indicators of well-being relevant? Validating composite indexes of well-being. *Ecological Indicators*, 2014, 46, pp. 575–585.
PAUN, G. An Impossibility Theorem for Indicators Aggregation. *Fuzzy Sets and Systems*, 1985, 9(2), pp. 205–210.
PEKKA, S., ZOLTAN, J. A., WUEBKER, R. Exploring country-level institutional arrangements on the rate and type of entrepreneurial activity. *Journal of Business Venturing*, 2013, 28(1), pp. 176–193.
PELE, D. T. About the Impossibility Theorem for Indicators Aggregation. *Journal of Applied Quantitative Methods*, 2009, 4(1), pp. 82–87.
PORTER, M. E. AND STERN, S. *Social Progress Index 2014* [online]. Social Progress imperative, 2014. [cit. 1.2.2018]. <https://www2.deloitte.com/content/dam/Deloitte/cr/Documents/public-sector/2014-Social-Progress-IndexRepIMP.pdf>.
ROUSSEEUW, P. J. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics*, 1987, 20, pp. 53–65.
PREACHER, K. J., ZHANG, G., KIM, C., MELS, G. Choosing the optimal number of factors in exploratory factor analysis: a model selection perspective. *Multivariate Behavioral Research*, 2013, 48(1), pp. 28–56.
ROKACH, L. AND MAIMON, O. *Data mining with decision trees: theory and applications.* 2nd Ed. Singapore: World Scientific Publishing Co. Pte. Ltd., 2014.
ROMETTY, G. *We need a new era of data responsibility* [online]. World Economic Forum Annual Meeting, 2018. [cit. 2.2.2019]. <https://www.weforum.org/agenda/2018/01/new-era-data-responsibility>.
SIDDIQI, N. *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring.* New York: Wiley and Sons, 2005.
SONG, Y. AND LU, Y. Decision tree methods: applications for classification and prediction. S*hanghai Archives of Psychiatry*, 2015, 27(2), pp. 130–135.
THE ECONOMIST. *Ranking the rankings.* November 8th 2014 print edition, 2014a.
THE ECONOMIST. *How to lie with indices.* November 8th 2014 print edition, 2014b.
THERNEAU, T. M. AND ATKINSON, E. J. *An Introduction to Recursive Partitioning Using the rpart Routine* [online]. 2019. [cit. 18.1.2020]. <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>.

ANNEX

Figure A1 The 6-cluster solution for variables selected with factor analysis



Source: Authors' calculation based on Otoiu et al. (2014) data

Bank Specific, Industry Specific and Macroeconomic Determinants of Bank Efficiency in Euro Area

Ehsan Rajabi[1] | *Agricultural Planning, Economic and Rural Development Research Institute (APERDRI), Tehran, Iran*
Reza Sherafatian-Jahromi[2] | *Universiti Putra Malaysia, Serdang, Selangor, Malaysia*

Abstract

This study analyses the cost and profit efficiency of the banking sector in all 17-Euro area Member States during the period from 1999 until 2012. The two-stage approach, the generalized method of moment (GMM) regression model is used to regress the efficiency level obtained from the first stage on factors that could influence the efficiency score. Therefore, the efficiency score measures that derived from the DEA estimations are used as the dependent variable and then regressed upon environmental variables. The result suggests that the cost and profit efficiency of 126 listed bank is found to be on average negatively related to population density, banking activity, loan management activity, and profitability while economic condition, financial deeping rate, and bank network extension have a positive influence on cost and profit efficiency. Overall, our results demonstrate that environmental variables contribute significantly to the difference in efficiency scores between the Member States.

Keywords

Efficiency, Euro area, generalized method of moment (GMM), environmental variables, banking characteristics

JEL code

G21, D61, L16

INTRODUCTION

The Eurozone, which is composed of the seventeen European countries that have joined together to form a common monetary union, represents more than 15% of global Gross Domestic Production (GDP) and 20% of world exports (European Commission, 2015). The European Central Bank (ECB) is the sixth of the seven institutions of the European Union (EU) as listed in the Treaty on European Union (TEU). It is the central bank for the Eurozone, one of the world's most important central banks, to administrate the monetary policy of the 17 EU Member States that constitute the Eurozone. In 1957, the European

[1] Assistant Professor, Department of Agriculture and Food Policy, Agriculture Planning, Economic and Rural Development Research Institute (APERDRI), Ministry of Jihad-e Agriculture, Tehran, Iran. Corresponding author: e-mail: rajabi.ehsan63@gmail.com.

[2] Department of Economics, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia. E-mail: rezasherafatian@yahoo.com.

Union was founded. From 1957 to 1999, economic integration has progressed as follows by Treaty of Rome (1957), the snake (1970s), European Monetary System (1979), the Single European Act (1986), Maastricht Treaty (1991), and Stability & Growth Pact (1997) and, finally, in 1999 the Euro is launched among 11 members of EU. Greece, Slovenia, Cyprus and Malta, Slovakia and Estonia have been joined Euro area from 2001 to 2011.

The economic impacts of this zone are resounding worldwide nations. Thus, both the success and failures of the European Central Bank policies will affect not only country members of the Eurozone, but also the global economy in general. The formation of the ECB and its currency, the Euro, has presented greater success to the euro Member States. Its long-term efficiency, productivity, and stability will base on the efficacy of the ECB policies in addressing some critical obstacles to its success.

The European banking markets, especially banks, in their function as financial intermediaries, contribute to economic activity in a number of ways. In this case, the banking systems in developing countries have undertook major reforms in order to create effective banking institutions with a high level of soundness and capable of facilitating economic growth (Andries, Apetri, Cocris, 2012). Furthermore, banking sector has undergone significant transformation worldwide in its operational environment during the last two decades. Both environmental and internal factors have influenced its structure, efficiency, and performance in banking industry, and effective banking system is better able to withstand negative shocks and contribute to the stability of the financial sector (Brissimis, Delis, Tsionas, 2010). Therefore, the efficiency of the banking system becomes an important issue to the academic world, the banking system decision-makers, and regulatory bodies.

Moreover, the banking sector is considered by any economy as a key sector for the smooth operating of its domestic economic system, so the development of the new banking status is an issue of major concern. When competition increases between banks or between banks and other financial institutions, internally and internationally, bankers and policy makers are to determine whether banks are managed efficiently and productively, and, if not, to take remedial action.

In other respect, analysing the efficiency level of banks is of interest from a policy perspective because if banks are becoming more efficient then, one would expect better performance, enhanced profitability, greater amounts of funds channelled through the system, competitive prices, and service quality for consumers, as well as increased safety and soundness (Casu, Girardone, Molyneux, 2004). In addition, looking at efficiency differences across countries may help to identify the achievement or perhaps of policy notations or, additionally, may highlight a variety of strategies carried out through banking systems. The information acquired about the evaluation of the bank's performance is known to improve its overall efficiency of operations and, in turn, it may improve its competitive frontier.

The banking industry exposes a multitude of new developments and challenges. Deregulation, liberalization, information technology, and the entry of new types of competitors have contributed to internationalization of the existing capital markets and to the developments of new markets of sophisticated financial instruments. The banking scenery changes even more radically in Europe, where the introduction of the single currency (Euro) in 1999 has removed institutional obstacles for banks to operate in other EU countries.

This changing banking environment concentrates on competitive conditions in the Eurozone and on the viability of its – until recently – sheltered banks. After all, for long, domestic banks have enjoyed comparative advantages on the domestic markets for bonds and equity in the field of underwriting and trading activities based on the existence of national currencies. Nowadays, especially after the constitution of the Euro, the efficiency of the euro area banks will become more and more crucial in the light of the current and expected increase in competition. Less efficient banks run the risk to be driven from the market. Efficiency can also be a decisive element in the game of mergers and take-overs, where inefficient banks are an easy and sought-after target.

Furthermore, in order to response a major concern of investors and policy makers how is the health performance of EU-17's banks, this paper seeks to discover the level and spread of bank efficiency in the Eurozone. In particular, it targets on differences across countries explained by environmental condition like various macroeconomic conditions, various banking specification, bank specific characteristics, and other sources over time. Therefore, this study attempts to determine the influencing factors of bank efficiency level in the Eurozone.

Our study contributes to the literature as follows. First, the literature is a treasure of country studies on efficiency in the banking industry. Studies on international comparison of efficiency are rare. Actually, such an international comparison is really a heroic attempt, as the differences in banking behaviour and economic and institutional conditions (in terms of institutions, supervisory rules, government interference, customer preferences and level of development) between countries are huge. As activities of banks diverge strongly and as part of these activities are truly challenging, it is actually for just one country quite a task to capture bank behaviour by one model. Moreover, international comparisons are easily distorted by national differences in macroeconomic condition, banking sector specification categories, as shown in this study. In other words, it can easily be confirmed that single-country studies are entirely unsuitable for international comparisons and provide misleading results.

Moreover, our study contributes to the literature by providing estimations of both cost and profit efficiency of banks based on nonparametric frontier analysis for all the euro area Member States; also, it compares efficiencies scores derived from Member States. Although, most of the studies of banking efficiency in the European countries are based on one-year analyses, and one year is not sufficient to observe the efficiency level. Hence, this study will evaluate the efficiency level for fourteen years. In addition, the sample of this study will include most significant banks (85 percent of total banking assets in the euro area) that operate in all seventeen Member States.

The rest of this paper is structured as follows: The literatures are reviewed in section two, followed by model specification in the third section. The sources of data are represented in section four. In section five, the empirical results and discussion are presented, followed by the conclusion and policy implications in the last section.

1 LITERATURES REVIEW

The literature on bank efficiency has a long tradition and has accumulated a significant amount of researches with various results, scopes, and methodologies. As an example, we can refer to Berger and Humphrey (1997), and Berger et al. (1999). In the United States, Berger and Humphrey (1991) published a report based on bank performance. In this report, it was stated that banks could improve their performance in terms of efficiency and productivity by using economies of scale, or economies of scope. Moreover, some studies looked at the conceptual background of the banks (Lovell, 1993) and the risks associated with their performance (e.g. Paradi and Zhu, 2013; Casu, Ferrari, Zhao, 2013; Berg, Førsund, Jansen, 1992; Berger and DeYoung, 1997; McAllister and McManus, 1993; Mester, 1996). It is also notable that the common point among these studies is that they ranked foreign banks less efficient compared to the local ones (DeYoung and Nolle, 1996; Edward Chang, Hasan, Hunter, 1998; Hasan and Hunter, 1996; Mahajan, Rangan, Zardkoohi, 1996; Peek, Rosengren, Kasirye, 1999).

One thing that remains the same though is the number of the cross-country comparative researches. They mostly concentrated on the European market. Studies show that changes in the sample size, production specifications, and assessment techniques lead to different results (Bos and Schmiedel, 2007). Notwithstanding their gap, some experimental results are valuable. Some of the studies done in the EU were in alliance with those done in the U.S.; both possessed cost efficiency levels of 70 to 80 percent, and the profit was around 50 to 60 percent. However, some countries were more efficient; for instance

Spain, and Belgium, showed greater degrees of efficiency compared to Germany, Austria, and the United Kingdom (Pastor, Perez, Quesada, 1997).

Some studies tried to do cross-country studies. For instance, Sheldon (1999) conducted a study on 1 783 banks in the EU using data from 1993 to 1997. By applying the Data Envelopment Analysis, Sheldon realized that usually larger banks showed better performance compared to smaller ones. It was also shown that specialized banks as well as retail banks were also more profit and cost efficient. On the other hand, smaller banks and wholesale banks showed lower efficiency levels. The efficiency on average was quite low at less than 46%. The profit efficiency of the banks was slightly better at 65%. In his study, Sheldon showed that banks from Denmark, Sweden, and France had the higher profit levels and the ones in Portugal, Greece, Spain, and U.K had the lowest profits.

Being concerned with a similar notion, Altunbas et al. (2001) concentrates on the German banking market in the period between the years 1989 to 1996. They discern between public savings banks, private commercial banks, and mutual cooperative banks. As a result of this comparison between the private commercial banks and others banks including the public ones they realized that private banks were more inefficient in terms of cost and profit.

In another study done in Germany, Bos et al. (2005) looked at the accounting factors that lead to differences in bank efficiency. It was also shown that banks from different regions have different performance. The size of the banks, their type, as well as their geographic origin was also found to be influential factors in the efficiency of the banks. However, they realized a huge difference between the results they obtained by the method employed for controlling the heterogeneity. That is how they arrive at the benchmarking paradox and elaborate on it: "we take part in a benchmarking activity to evaluate and measure the differences in performances; however, to do so, a general benchmark has to be assumed".

It can be said that such paradox is quite clear in the case of cross-country studies. The reason is that in such studies a common efficient frontier is often used as the basis of comparison. This can lead to wrong results as some countries have access to technologies that are absent in others. However, it is not possible to compare the efficiency outcomes across borders in cases where the frontier has been employed in every country in the sample and every banking institution's performance is weighed up against the bank which has the best-practice within that country. Recent studies tried to stay away from the bias estimation in cross-country bank efficiency comparisons by combining country-specific environmental variable (Barth, Lin, Seade, Song, 2013; Chaffai, Dietsch, Vivas, 2001; Dietsch and Vivas, 2000; Vivas, Pastor, Hasan, 2001; Vivas, Pastor, Pastor, 2002).

As an example, we can refer to Dietsch and Vivas (2000) who were more focused on the theory that the employment of a frontier benchmark can lead to misleading result. The reason for these misleading efficiency outcomes of companies from different nations is that these approaches often fail to control the cross-country economic conditions, demographic, and regulatory differences which are beyond a company's control. Therefore, using common frontier can lead to results that indicate a bank to be inefficient while it is regarded as efficient in comparison to the average performance of the banks operating within the national market.

In addition, in another study Vivas et al. (2001) make a test where they replicate every banking market's performance if the average banks chose to function in any other nation. That is how they realized that it would be good for some banks to operate in another country since their performance in another country was quite higher and better in comparison to their performance in their original country.

At the end, Bos and Kolari (2005) make a better quality comparison by weighing up small and large independent US and European banks. As for the none-common frontier sets of rules and regulations, they keep the profit and cost frontier the same in both the US and Europe and operate on that basis. Although, they found evidence in favour of a single profit frontier, they declined the single cost frontier.

As explained above, the initiatives cannot solve the problem of cross-border efficiency comparison of banks with different level and type of technologies in different nations. This study tries to use the single-frontier method to increase the level of the published researches in the literature, to forecast the common efficiency variables across the countries and to consider their specified environment circumstances to account for the main technologies in the EU financial industry.

In view of the literature discussed, we can say that empirical studies in banking efficiency have been conducted extensively (for example USA, Germany, and Spain) with financial variable and monetary policy but that few studies have been done to investigate banking efficiency in European countries, especially for all the euro area countries with these financial integration and monetary policy variables. Therefore, more empirical work is needed on the banking efficiency in the euro area Member States.

2 MODEL SPECIFICATION

To investigate the influence of different factors of environmental condition two-step procedure was employed; the data envelopment analysis (DEA), and GMM estimators. Therefore, in following section, we select inputs, outputs, and DEA approach.

There is a continuous discussion in the banking literature concerning the most appropriate interpretation of outputs and inputs. According to Bergendahl (1998): "There have been almost as many assumptions of inputs and outputs as there have been applications of DEA" (p. 235). Berger and Humphrey (1997) identified two major approaches including production approach and intermediation approach for selecting inputs and outputs; production approach and intermediation approaches. According to Berger and Humphrey (1997), neither intermediation approach nor production approaches are perfect because they cannot perform financial institutions' dual role as being financial intermediaries and provider of document or transactions processing services.

In some studies, earning assets is used as an output. This is consistent with asset approach proposed by Sealey and Lindley (1977) while deposits are considered as an additional output, by other researchers, which is related to an approach known as value-added approach. Recently, a different modified version of the intermediation approach is adopted (see Avkiran, 1999; Chu and Lim, 1998; Das and Ghosh, 2006; Drake, Hall, Simper, 2006; Sturm and Williams, 2004). This is known as profit oriented or operation approach that considers revenue components such as interest income, non-interest income as outputs and cost components such as personal expenses, interest expenses as inputs. According to Drake et al. (2006), based on input oriented DEA relative efficiency analysis, the more efficient units will be better at maximizing profit through minimizing different costs incurred in making different revenue flow. They also indicated that, this approach could be more suitable for taking the variety of strategic responses by financial companies in confronting dynamic changes in environmental and competitive situations.

To estimate cost and profit efficiency score, the present study focused on intermediation approach to construct the DEA frontier. Under the intermediation approach (following Berger and Humphrey, 1992), we assume deposits (X_1): demand, savings, and time deposits, Labor (X_2): staff of bank together with management expertise required for providing bank services ,physical capital (X_3): offices, branches, and computer hardware as inputs and loans (Y_1): is total amount of loans concerning each banking firm, investment (Y_2) total securities, equity investments and other investments as outputs.

Price of borrowed funds (w_1) was used as interest expenses over the sum of deposits price of labor (w_2) calculated by personnel expenses to the employees' number as the unit price of labor. Price of physical capital (w_3) was measured by non-interest expenses over fixed assets. Price of loan (p_1) was calculated by interest income on loans over total loan. Price of investment (p_2) was measured by total non-interest operating income plus other interest income over other earning assets.

As already recorded, if economic objective functions are reasonable and if reliable price information is available, however, DEA can be used to identify cost efficiency (Cooper et al., 2000). Since we assume indeed that banks minimize cost in the euro area, we consider in this study input-oriented efficiency with variable return to scale. The minimum cost is obtained by solving the DEA linear programing problem.

$$min\sum_{i=1}^{n} w_{io}x_i, \tag{1}$$

$$\text{subject to: }\sum_{i=1}^{n} x_{ij}\lambda_j \leq x_i \quad (i = 1, 2, \dots , n), \tag{2}$$

$$\sum_{i=1}^{n} y_{rj}\lambda_j \geq y_{r0} \quad (r = 1, 2, \dots , m), \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{4}$$

$$\lambda_j \geq 0 \quad (j = 1, 2, \dots , N), \tag{5}$$

where $j = 1, \dots , N$ are the number of banks, $i = 1, \dots , n$ are input volumes used by bank j, $r = 1, 2, \dots , m$ measures the volume of output r and w_{io} is the unit cost of the input i of bank DMU_o which is the benchmark projection that can be different from one bank to another. Although, the objective is to choose the x_i and λ_j values to minimize the total cost of satisfying the output constraints. The w_{io} in the objective represent unit costs. The minimization problem is calculated for each bank of and each year in the sample, thus identifying for each a benchmark combination of inputs and cost.

Every DEA model assumes a returns-to-scale characteristics that is represented by the ranges of the sum of the intensity vector λ, i.e., $L \leq \lambda_1 + \lambda_2 + \dots + \lambda_n \leq U$. Here, we compute variable returns to scale and use $L = U = 1$. We consider convex hull representation. Our model allows substitutions in inputs. Based on an optimal solution (x^*, λ^*) of the above problem, the cost efficiency of DMU_o is defined as:

$$CE_o = \frac{C_{min}}{C_o} = \frac{\sum_{i=1}^{n} w_{io}x_i^*}{\sum_{i=1}^{n} w_{io}x_{io}}, \tag{6}$$

where CE_o is the ratio of minimum cost to observed cost for the *oth* firm. Clearly, this approach implies that all observed input-cost combinations are measured with no error. Outliers may be classified as very efficient simply because data error.

Similar to cost efficiency, the profit efficiency (PE) can be estimated by solving the following linear programming problem n times; each time for a different bank in the sample. Therefore, the profit-maximization problem of a multiple-output, multiple-input firm facing input and output prices w and p, respectively, can be formulated as the following DEA problem:

$$\pi = py^* - wx^* = max\sum_{r=1}^{m} p_r y_r - \sum_{i=1}^{n} w_i x_i, \tag{7}$$

$$\text{subject to: }\sum_{j=1}^{N} x_{ij}\lambda_j \leq x_i \quad (i = 1, 2, \dots , n), \tag{8}$$

$$\sum_{j=1}^{N} y_{rj}\lambda_j \geq y_r \quad (r = 1, 2, \dots , m), \tag{9}$$

$$\sum_{j=1}^{N} \lambda_j = 1, \tag{10}$$

$$\lambda_j \geq 0 \quad (j = 1, 2, \dots , N). \tag{11}$$

The profit-maximizing input and output quantities $x_i^*(i = 1, 2, \dots , n)$ and y_r $(r = 1, 2, \dots , m)$ are obtained along with the other decision variables $\lambda_j^*(j = 1, 2, \dots , N)$ at the optimal solution of this problem. The maximization problem is calculated for each bank of the sample, thus identifying for each a benchmark combination of inputs and revenue. Based on an optimal solution (x^*, y^*, λ^*) of the above problem, return to the maximum profit attainable for bank *o_th* given the production technology and the input prices w_i it faces. The profit efficiency of DMU_o is defined as the ratio between the observed profits and the maximum profits as follows:

$$PE_o = \frac{\pi_o}{\pi_{max}} = \frac{\sum_{r=1}^{m} p_{ro} y_{ro} - \sum_{i=1}^{n} w_{io} x_{io}}{\sum_{r=1}^{m} p_r y_r^* - \sum_{i=1}^{n} w_{io} x_i^*}, \tag{12}$$

where $PE_o \leq 1$ and equals unity when the bank operates on the estimated frontier and is deemed profit efficient.

The efficiency scores generated from the DEA program were used as independent variables in a panel regression model to explain bank performance, the dependent variable (Lehmann, Warning, Weigand, 2004) .Using panel regression, a non-parametric method and multivariate analysis may assist in understanding and validating behavioural relationship in the banking sector (Jalan, 2002). For second step, a linear regression model is estimated to be in following form:

$$y_{it} - y_{it-1} = (1 - \alpha)y_{it-1} + \beta(\mathrm{L})X_{it} + \eta_i + \varepsilon_{it}, \tag{13}$$

where y_{it} represents the efficiency score of bank i at time t, X represents the set of explanatory variable (environmental variable), η_i can be described as an unobserved specific effect of the country and ε_{it} can be described as an error term.

In order to control for cross-country differences in the environment that banks operate, there are some approaches that can be utilized for incorporating environmental variables (modifying for the environment) in DEA applications. The term of "environmental variable" is usually utilized for describing factors which could have an impact on firm's efficiency but be outside of the manager that covers banking sector, macroeconomics countries specification (Coelli, Rao, Donnell, Battese, 2005). Finally, model one is developed to investigate relationship of cost and profit efficiency with environmental variables.

$$\begin{aligned} &\textit{Bank efficiency} \\ &\quad = f\,(\textit{Lag of Bank efficiency} + \textit{Population density} \\ &\quad + \textit{Economic condition} + \textit{Financial deepening rate} + \textit{Banking activity} \\ &\quad + \textit{Bank network extension} + \textit{Profitability} \\ &\quad + \textit{Loan management activity}). \end{aligned} \tag{14}$$

Formula (14) is extended for the purpose of reflecting the variables. The baseline regression model is formulated as below:

$$\begin{aligned} EF_{ijt} &= \alpha + \lambda EF_{ijt-1} + \beta_1 POPD_{ijt} + \beta_2 \text{In}(GDP)_{ijt} + \beta_3 FDEEP_{ijt} \\ &+ \beta_4 CLAIMS_{ijt} + \beta_5 BRANCH_{ijt} + \beta_6 ROE_{ijt} + \beta_7 LOANTA_{ijt} + \eta_j + \mu_{ijt}\,, \end{aligned} \tag{15}$$

$i = 1, \dots , 126, t = 1, \dots , 14, j = 1, \dots 17.$

In the banking literature, the significance of specifying the environmental variables for the purpose of avoiding bias in models has been identified (see, for example Dietsch and Vivas, 2000; Vivas et al., 2001; Vivas et al., 2002). Based on this first model, across country efficiency differences are ascribed to managerial inside the commercial banks while the difference in efficiency can be explained through different environment and economic regulatory across countries (see Berger and Mester, 1997). Three groups of environmental variables (bank specific and countries, banking industry factors) was employed for all 126 selected bank from all 17-Euro area Member States including Spain, Austria, Cyprus, Slovenia, Belgium, Portugal, Estonia, the Netherlands, Finland, Malta, France, Luxembourg, Germany, Italy, Ireland, and Greece from 1999 to 2012.

Based on above-mentioned studies, three groups of variables that are supposed to be connected with changes in efficiency across banks are formed. Country level variables are the first group that explains macroeconomics situations such as population density, economic condition, and financial deepening ratio. According to Yildirim and Philippatos (2007), demand of banking services' supply is positively affected by favourable economic conditions, which brings about an improvement in bank efficiency as well.

The second group includes variables that describe the structure of banking industry for each Member States (activity in the banking activity, bank network extension). According to Kasman and Yildirim (2006), the overall banking network that is measured through the market size of banking and the banking activity is measured by bank claim to private sector over the GDP encourage higher efficiency. Therefore, we employ those variables as explanatory variables in two-step approach.

Last group consists of variables describing bank specific characteristic in each bank. The bank specific variables are: *LOANTA* is a total loan of bank to total assets ratio and can be described as a measure of loan management activity; *ROE* can be described as pre-tax profit divided by equity. These variables have been utilized in previous studies to show the bank-specific characteristics that influence efficiency (Allen and Rai, 1996; Fries and Taci, 2005; Isik and Hassan, 2003; Pasiouras, 2008).

3 DATA

The current study explored the extent of a possible relationship between the efficiency of bank and environmental variables. For the current study, data was collected from various sources, "Bankscope" database of BVD-IBCA, Eurostat, World Bank, ECB during the period of 1999–2012. For a sample, unconsolidated accounting data of EU-17 Member States including Spain, Austria, Cyprus, Slovenia, Belgium, Portugal, Estonia, the Netherlands, Finland, Malta, France, Luxembourg, Germany, Italy, Ireland, and Greece was used. This sample is adequate for investigating the evolution in bank efficiency of EU Member States because main EU countries are included in it.

The sample of banks have been selected from about total 6,000 credit institutions (entire population) in the euro area which is covered by almost 85 percent of total banking assets in the euro area. From this list of credit institutions, 126 active credit institutions was selected (6 from Austria, 6 from Belgium, 3 from Cyprus, 3 from Estonia, 3 from Finland, 11 from France, 25 from Germany, 4 from Greece, 5 from Ireland, 15 from Italy, 6 from Luxembourg, 3 from Malta, 6 from the Netherlands, 4 from Portugal, 3 from Slovakia, 4 from Slovenia, 18 from Spain) based on the European Central Bank (ECB) decision which are prepared to deal with tasks of new banking supervision as a part of a single supervisory mechanism. These credit institutions are directly supervised by ECB based on one of significance criteria (including size, economic importance, cross-border activates, direct public financial assistance). They will work closely with the national well qualified authorities with the aim to supervise all other credit institutions.

Almost 85 percent of total banking assets in the area of Euro represented by these 126 credit institutions that will be identified in accord with significant criterion. Significance criteria include total value of assets exceeding €30 billion, significance for the specific country or the EU economy as a whole and significance cross-border activities by ratio of its cross-border assets/liabilities in more than one other participating Member State to its total assets/liabilities to be more than 20% and direct public financial assistance by requested or received funding from the European Stability Mechanism or the European Financial Stability Facility.

4 RESULTS

4.1 Summary Statistics

Based on practical consideration, to estimate cost and profit efficiency score, the present study focuses on the intermediation approach to construct the DEA frontiers. Under the intermediation approach (following Berger and Humphrey, 1992), we assume deposits (X_1): deposits & short term funding, labor (X_2): number of employees, physical capital (X_3): fixed assets as inputs and loans (Y_1): total loan, investment (Y_2): other earning assets as outputs. Although, we consider price of borrowed funds ($w1$): interest expenses over deposit, price of labour (w_2): personnel expenses over total labor, price of physical capital (w_3): non-interest expenses over fixed assets as input price and price of loan (P_1): interest income on loans over total loan, price of investment (P_2): total non-interest operating income plus other interest incomes over other earning assets as output price. Table 1 shows summarized descriptive statistics for inputs, outputs and their prices that were employed for the estimation of the cost and profit efficiency by the DEA approach.

Table 1 Descriptive summary of inputs, outputs, and prices for estimating efficiency

Variable (symbol)	Obs	Mean	Std. Dev.	Min	Max
Deposits (X_1)	1 142	124 000	224 000	30.1	1 740 000
No. of employees (X_2)	1 142	16 243.01	30 997.18	5	193 349
Fixed assets (X_3)	1 142	1 510	3 190	0.11	24 600
Total loan (Y_1)	1 142	104 000	178 000	7.3	1 030 000
Other earning assets (Y_2)	1 142	120 000	277 000	1.1	2 450 000
Price of fund (W_1)	1 142	0.078	0.412	0	9.598
Price of labor (W_2)	1 142	123 598.5	468 316.9	9 703.714	9 588 510
Price of physical capital (W_3)	1 142	10.010	87.977	–1.205	2 304
Price of loan (P_1)	1 142	0.191	1.466	0.0001	37.634
Price of investment (P_2)	1 142	0.087	1 171 443	–0.143	1.168

Note: The unit of input and output variables is US $ Million.
Source: Own construction

4.2 Estimation of Efficiency by DEA

The DEA models' results are summarized in Tables 2–4. Table 2 reports the results of the models run using the entire 126-listed bank dataset for 17-Euro area Member States from 1999 to 2012, while Table 2 provides cost efficiency score with the breakdown of the results by country and year. Table 3 reports the results of the profit efficiency models by country and year.

From Table 2, it is observed that the three lowest averages of cost efficiency of sample are reported for Slovakia, Greece, and Portugal banking sector which stood at 11.12%, 11.23%, and 16.79%, respectively. Conversely, for three highest averages of cost efficiency, the Netherlands, France, and Germany bank cost efficiency stood at 70.95%, 56.69%, and 48.03%, respectively. For whole the euro area, cost efficiency ranged for a low of 30.62% in 2005 to a high of 43.52% in 2002 and totally, average of cost efficiency is 35.84% from 1999 to 2012. The average cost efficiency level is 35.84%, meaning that on average they are 65.12% inefficient in costs. Inefficiency levels show by how much costs can fall if inputs were used as efficiently as the unit located on the best practice frontier was.

The trend of the dispersion of cost efficiency values as measured by its standard deviation remains indistinct at best. The standard deviation of the scores was of 30.84 for the Euro area while the minimum standard deviation of cost efficiency sore remained at 4.32 for Slovakia, and maximum standard deviation score stood at 56.69 for France.

It is observed that over the sample period, the Netherlands banking sector reported higher cost efficiency for all years except for 2003 that Luxembourg banking sector stood at 61.15%, and it was varying from a low of 63.2% in 2007 to high of 100% in 1999, 2000, and 2001. A plausible reason could be assigned to the fact that during the period under study, the Netherlands banking sector was relatively efficient in its basic function as an intermediary between savers and borrowers. It should be mentioned here that decreasing of cost and profit efficiency of the Netherland banking system is because of the SNS bank's and Nederlandsche Bank's NV (De) performance during 2005 to 2012. The SNS bank has a negative profit value (loss) and Nederlandsche bank NV has high level of deposit and low value of loan which is made a low level of efficiency. Although, lower cost efficiency was recorded for Slovakia's banking sector across the sample period except for 2004, 2006, and 2007 that Greece's banks were inefficient bank in the sample and it was varying from a low of 7.16% in 2012 to high of 16.27% in 2002. In constant, the empirical results indicated a large asymmetry between countries regarding their profit efficiency level. Although, the evaluation of profit efficiency in each country shows that there doesn't seem to be a clear trend in general, efficiency scores had been decreasing from starting to ending years in most of the countries in sample.

The results of the profit efficiency estimation are presented in Table 3. From Table 3, it could be observed that the Netherlands, Germany, and France are three highest average profit efficiency members across the euro area during the period under study with 63.95%, 44.58%, and 37.99% respectively. It means; these most efficient banking sectors did not use extra excess input resources in comparison with an inefficient banking sector, while the inefficient banking sector used more extra excess resources to produce the same amount of outputs as the efficient country banking sector.

In comparison, Portugal, Slovakia, and Greece had the lowest average profit efficiency members across the euro area with 2.56%, 3.29%, and 4.52%, respectively. It means these county banking sectors were operated in the most inefficient manner, all other conditions being equal. For the euro area, profit efficiency ranged from a low of 20.69% in 2010 to a high of 38.79% in 2000 and, finally, average of profit efficiency was 24.55 % from 1999 to 2012. The average profit efficiency level is 24.55%, meaning that on average, they were 75.45% inefficient in profit. Inefficiency levels show by how much the profit (costs) can increase (decrease) if inputs were used as efficient as the unit located on the best practice frontier. The trend of the dispersion of profit efficiency scores as measured by its standard deviation remains indistinct at best.

It is observed that over the sample period, the Netherlands banking sector had the highest profit efficiency for all years except 2003 that Luxembourg banking sector stood at 67.14%, and it is varying from a low of 48.26% in 2008 to high of 100% in 1999, 2000, 2001, and 2002. Although, lower profit efficiency was recorded for Estonia's banking sector across the sample period except for 2004, 2011, and 2012

Table 2 Summary of cost efficiency of banks across the euro area member states estimated by DEA form 1999 to 2012

Country	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Mean	Std.Dev/Obs
Austria	39.73	29.72	30.19	36.71	14.81	21.50	13.74	16.36	20.49	20.98	19.19	15.9	17.47	15.75	20.49	10.76/69
Belgium	28.06	23.69	28.46	29.55	26.39	24.05	41.56	37.92	37.22	36.74	37.20	34.30	27.84	33.22	34.46	25.52/50
Cyprus	17.42	19.23	21.25	23.16	21.70	23.01	12.76	17.58	22.29	22.57	20.84	14.61	11.62	11.55	18.54	9.73/42
Estonia	13.10	15.72	17.09	24.52	24.59	25.68	10.38	17.58	26.47	31.91	28.22	16.43	41.92	42.25	25.21	22.70/30
Finland	.	.	.	.	.	24.43	13.16	19.59	22.52	24.27	21.00	22.31	44.80	45.11	26.51	23.24/25
France	.	.	.		.	82.11	58.87	59.14	59.01	59.50	58.69	46.50	49.09	57.70	56.69	35.65/86
Germany	67.10	61.43	64.09	67.70	53.83	49.38	34.40	49.28	46.26	47.60	49.20	43.28	43.66	41.19	48.03	33.65/254
Greece	.	.	.	.	.	8.28	6.05	8.04	12.05	17.31	15.29	11.91	11.61	10.56	11.23	3.89/36
Ireland	.	.	.	.	.	25.69	45.99	52.79	53.39	48.92	45.19	40.27	40.00	36.72	44.03	31.73/43
Italy	.	.	.	.	.	8.50	23.17	25.16	31.15	35.99	31.90	23.25	26.72	25.47	27.77	24.62/119
Luxembourg	34.37	47.32	73.08	57.14	61.15	75.10	28.58	24.38	25.19	17.45	13.02	10.43	11.17	10.82	28.65	30.60/65
Malta	10.92	13.11	14.96	20.14	19.23	22.30	40.57	45.3	47.47	46.78	40.80	42.68	17.72	27.74	31.33	31.22/36
Netherlands	100	100	100	80.58	35.33	78.58	68.60	63.26	63.62	68.43	68.57	72.14	73.09	70.92	70.95	35.34/52
Portugal	.	.	.	.	.	10.07	8.59	10.45	15.79	23.33	21.28	20.83	19.37	16.33	16.79	6.23/33
Slovakia	7.56	7.73	12.82	16.27	14.53	17.67	5.36	9.34	13.54	14.11	13.38	7.24	6.67	7.16	11.12	4.32/40
Slovenia	18.48	18.59	17.46	26.64	25.08	25.37	20.27	37.37	36.32	30.83	30.51	13.05	12.52	13.49	23.65	18.03/48
Spain	.	.		.	.	31.90	31.07	40.96	37.73	40.87	37.60	31.93	39.10	33.02	36.05	27.44/114
Euro Area	35.91	35.43	40.77	43.53	35.40	38.08	30.62	36.87	37.72	39.30	37.73	32.25	34.12	33.37	35.84	30.84/1 142

Notes: Efficiency score is in percentage; Obs: number of observations; Std.Dev: standard deviation.
Source: Own construction

Table 3 Summary of profit efficiency of banks across the euro area member states estimated by DEA form 1999 to 2012

Country	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Mean	Std.Dev/Obs
Austria	8.03	8.22	8.64	10.38	6.88	9.02	8.97	6.02	4.92	3.75	4.41	4.20	3.99	3.26	6.06	4.08/69
Belgium	100	100	100	30.04	28.11	31.21	29.49	27.54	13.59	23.32	23.12	23.02	5.18	20.38	25.79	35.69/50
Cyprus	5.84	7.07	9.08	8.68	6.43	36.25	3.20	2.93	3.87	1.47	1.28	1.42	1.60	1.68	6.49	15.15/42
Estonia	1.01	1.66	2.10	2.34	2.44	3.85	1.93	2.07	1.61	0.06	0.41	0.060	34.14	34.04	8.19	24.98/30
Finland	.	.	.	.	.	6.10	10.60	14.34	10.30	6.22	11.73	18.13	34.20	34.61	17.06	28.02/25
France	.	.	.	.	.	73.57	60.10	44.44	31.30	34.17	47.84	45.34	42.86	45.68	44.58	41.19/86
Germany	85.51	85.12	69.24	56.25	47.96	42.53	31.17	45.82	36.77	29.79	29.19	29.15	25.45	26.68	37.99	43.40/254
Greece	.	.	.	.	.	6.16	7.75	7.12	5.02	2.38	3.61	4.03	2.55	2.06	4.52	3.06/36
Ireland	.	.	.	.	.	15.62	32.34	34.19	31.07	25.60	31.10	26.24	22.11	29.57	28.7	35.82/43
Italy	.	.	.	.	.	11.13	19.51	15.41	15.08	14.64	15.51	15.30	17.00	15.72	15.92	31.74/119
Luxembourg	35.39	66.74	75.50	71.68	67.14	68.22	42.07	36.02	35.18	17.30	17.53	1.79	19.08	3.28	33.04	45.05/65
Malta	7.23	8.38	8.96	8.85	6.25	3.23	35.21	34.98	35.52	34.17	34.02	33.98	34.38	34.53	25.43	40.51/36
Netherlands	100	100	100	100	1.86	77.00	62.26	61.93	49.40	48.26	61.06	56.09	67.56	68.27	63.95	45.80/52
Portugal	.	.	.	.	.	2.39	3.72	4.42	3.62	1.07	2.14	2.05	1.53	2.00	2.56	1.40/33
Slovakia	4.95	7.99	5.77	5.42	3.68	3.23	2.78	2.72	3.00	1.52	1.59	1.34	2.12	2.03	3.29	1.98/40
Slovenia	4.09	4.00	4.87	5.75	3.89	3.35	4.07	27.18	27.00	2.67	1.53	1.77	1.30	1.22	7.02	19.70/48
Spain	.	.	.	.	.	14.23	11.53	29.69	18.48	12.56	15.47	13.57	10.68	16.59	15.28	25.98/114
Euro Area	38.23	38.79	37.60	32.44	25.70	28.96	25.08	29.46	23.64	19.03	22.07	20.71	20.69	22.11	24.55	37.27/1 142

Notes: Efficiency score is in percentage; Obs: number of observations; Std.Dev: standard deviation.
Source: Own construction

that Slovenia's banks were inefficient banks in the sample, and it was varying from a low of 0.06% in 2010 to high of 2.34% in 2002.

To be useful for regulatory policy purposes, efficiency measures should be stable over the time. For this purpose, standard deviation was used for dispersion of profit efficiency. The standard deviation of the scores was of 37.27 for the euro area while the minimum standard deviation of profit efficiency sore remained at 1.40 for Portugal, and maximum standard deviation score stood at 45.80 for the Netherlands.

As a result, from both Tables 2 and 3, it can be mentioned that average cost efficiency which estimates for the banks over the sample period was higher than profit efficiency, and the Netherland banking sector was more efficient than others except for 2003. This result suggested that the banks' performance on the cost side was not matched by their ability to generate enough revenues. On the other hand, we found high levels of efficiency in costs and lower levels in profits, verifying the importance of inefficiencies on the revenue side of banking activity. Furthermore, we found the cost efficiency rankings were more stable over the time than profit efficiency according to standard deviation.

In addition, it is observed that over the sample period, both cost and efficiency measures display significant variation and did not archive sustained efficiency gain across the euro area banking sector. This result could be due to changing competitiveness and dynamic environments and new technologies in banking markets. In this case relative changes of efficiency rankings over the time are reasonable. Therefore, in next section, we try to investigate the influence of dynamic environment (at country, market, and bank level) on cost and profit efficiency.

4.3 Bank Efficiency and Environmental Conditions

For implementation of environmental variables (modifying for the environment) in the dynamic efficiency model which is used to avoid bias in efficiency models, two-step procedure was employed. At the first stage, analysis efficiency score obtained in Eurozone during the year 1999 and 2012 by the data envelopment analysis (DEA) was completed. Then, in the second step, GMM estimation was used to regress the measured efficiency on environmental variables. Therefore, in following section, we will present descriptive statistics, correlation matrix of explanatory variables and, finally, baseline regression results of GMM estimator.

In Table 4, the descriptive statistics is presented to analyse and measure the environmental variable into three categories of country level, industry of banking sector and bank specific characteristics. The table shows the mean, standard deviation, minimum, maximum, and number of observations. Mean represents the average value; standard deviation shows the deviation of value from the mean, minimum reports the smallest value of variable, maximum shows the biggest value of variable and number of observations calculated by number of time to cross section ($t \times n$).

The table shows that average cost efficiency is 0.358 with a standard deviation of 0.308 while profit efficiency mean is 0.246 with a standard deviation of 0.373. The population density (POPD) average in the euro area Member States is 207.711 with a standard deviation of 216.805, the minimum and maximum values are 17.16 and 1 302.1, respectively. The economic condition (LGDP) average is 26.842 and standard deviation is 1.843. The ratio of financial deepening (FDEEP) mean is 4.415 and standard deviation is 5.410. The banking activity average (CLAIMS) and bank network extension (BRANCH) average are approximately 1.259 and 85.701, respectively. The profitability (ROE) mean is 0.09 and shows standard deviation value 1.961 with the minimum and maximum values of 17.16 and 1 302.1, respectively, while the loan management activity (LOANTA) varies between 0.001 to 0.96 with a standard deviation of 0.210 and mean of 0.521. Number of observation is 1 124 for all variables that shows that data set is an unbalance panel.

Having estimated the cost and profit efficiency levels for each of the banks in the sample, it is of interest to analyse the factors that may explain the differences in efficiency among the banking

Table 4 Descriptive statistics of country, industry, and bank characteristics variables

Variable (symbol)	Obs	Mean	Std. Dev.	Min	Max
Cost efficiency (*CE*)	1 142	0.358	0.308	0	1
Profit efficiency (*PE*)	1 142	0.246	0.373	0	1
Country level variables					
Population density (*POPD*)	1 142	207.711	216.805	17.165	1
Economic condition (*LGDP*)	1 142	26.842	1.843	302.113	28.754
Financial deepening rate (*FDEEP*)	1 142	4.415	5.410	0.562	31.945
Industry of banking sector					
Banking activity (*CLAIMS*)	1 142	1.259	0.481	0.224	3.021
Bank network extension (*BRANCH*)	1 142	85.701	62.546	4.279	363.924
Bank specific characteristics					
Profitability (*ROE*)	1 142	0.094	1.961	–21.386	53.044
Loan management activity (*LOANTA*)	1 142	0.521	0.210	0.001	0.963

Notes: The independent variables, population density (POPD) calculated as people per sq. km of land area; economic condition (LGDP) calculated as the natural log of the GDP (constant 2005 US$); financial deepening rate (FDEEP) calculated as total bank asset (domestic and foreign bank) to GDP current price; banking activity (CLAIMS) calculated as domestic credit (bank claim) to private sector over the GDP; bank network extension (BRANCH) calculated as number of domestic and foreign branches of credit institutions (per 1 000 km^2); profitability (ROE) calculated as total pre-tax profit over total equity; loan management activity (LOANTA) calculated as total loan over total assets.
Source: Own construction

systems of the Eurozone. The baseline regression results focusing on the relationship between bank cost efficiency and explanatory variables (environmental variables) are presented in Table 5. Several diagnostic tests are made to show that results are warranted. The model performs reasonably well with most of the variables remaining stable across the various regressions tested. We report results for both difference and system GMM estimator for both one-step and two-step version. A lot of applied work using GMM estimator has rather focused on results for the two-step estimator than on one-step estimator. This is partly because in two-step estimation, the standard covariance matrix is robust to panel-specific autocorrelation and heteroscedasticity, but the standard errors are downward biased. Thereupon, in this study we have focused on the two-step estimator which suggested very modest efficiency gain than the one-step version, nevertheless, the one-step estimator was reported in all GMM estimation tables. System panel GMM requires more assumptions (which are employed to generate consistent and efficient parameter) than the first difference panel GMM, but if the assumptions hold, it will achieve a greater efficiency. Therefore, system panel GMM estimations are focused while first difference panel GMM is presented in all tables.

Although, for controlling huge number of instruments which is a real danger of over fitting the endogenous variables when the time period is long, GMM has instructed to use only maximum one lag depth for the endogenous variables as instruments. Furthermore, number of instruments is less than cross-sectional observations (banks) when the rule of thumb is to keep the number of instruments

Table 5 Baseline analysis for effect of country, industry, and bank characteristics on cost efficiency (controlling endogeneity)

Regressors	GMM-DIF One-step	GMM-DIF Two-step	GMM-SYS One-step	GMM-SYS Two-step	GMM-SYS♣ One-step	GMM-SYS♣ Two-step
Initial of cost efficiency (L1)	0.3024***	0.3001***	0.3664***	0.3659***	0.3659***	0.3646***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population density	–0.0024*	–0.0024***	–0.0006***	–0.0006***	–0.0006***	–0.0006***
	(0.091)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Economic condition[1]	0.1411	0.1394***	0.0102***	0.0105***	0.0106***	0.01061***
	(0.243)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Financial deepening rate	0.0039	0.0046***	0.0111***	0.0111***	0.0111***	0.0112***
	(0.577)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Banking activity	–0.0370	–0.0373***	–0.0746***	–0.0744***	–0.0706***	–0.0677***
	(0.349)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Bank network extension	0.0052***	0.0053***	0.0025***	0.0025***	0.0025***	0.0025***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Profitability	–0.0022	–0.0022***	–0.0011	–0.0012***	–0.0012	–0.0015***
	(0.317)	(0.000)	(0.622)	(0.000)	(0.600)	(0.000)
Loan management activity	–0.0810	0.0797***	–0.1759***	–0.1735***	–0.1800***	–0.1772***
	(0.405)	(0.000)	(0.004)	(0.000)	(0.004)	(0.000)
Sargan test (*p*-value)[2]	0.0001	0.1895	0.0001	0.4902	0.0001	0.4660
Serial correlation test						
AR(1) (*p*-value)[3]	0.0001	0.0006	-	0.0006	-	0.0006
AR(2) (*p*-value)	0.3227	0.2410	-	0.3232	-	0.3090
Wald test for joint significance (*p*-value)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
No. of instruments	96	96	120	120	121	121
Cross-sectional observations	123	123	126	126	126	126

Notes: The independent variables, population density (POPD) calculated as people per sq. km of land area; economic condition (LGDP) calculated as the natural log of the GDP (constant 2005 US$); financial deepening rate (FDEEP) calculated as total bank asset (domestic and foreign bank) to GDP current price; banking activity (CLAIMS) calculated as domestic credit (bank claim) to private sector over the GDP; bank network extension (BRANCH) calculated as number of domestic and foreign branches of credit institutions (per 1 000 km^2); profitability (ROE) calculated as total pre-tax profit over total equity; loan management activity (LOANTA) calculated as total loan over total assets.
♣ The regressions also include time trend variable for the different time periods that are not reported. [1] In the regression, this variable is included as log (variable). [2] The null hypothesis is that model and overidentifying conditions are correct specified. [3] The null hypothesis is that there is no serial correlation in the first-differenced disturbances. Values in parenthesis are *p*-value. ***, **, * indicates significance at 1%, 5% and 10% levels respectively.
Source: Own construction

less than or equal to the number of groups. It should be mentioned that in the baseline estimation, endogeneity problem of economic condition variable (GDP constant 2005 USD) is controlled by the variables which are instrumented with GMM-style instruments, *i.e.* lagged values of the variables in levels. For all the GMM estimation models discussed in the following subsections, the Sargent test (under Sargent thought)[3] for overidentifying restriction, and the Arrelano-bond tests (AR(2)) show that at the 5% significance level, our instruments are appropriately orthogonal to the error and no second-order serial correlation is detected, respectively (see Baum et al., 2010).[4]

The first two columns of Table 5 report the results for GMM-DIF and next two columns report GMM-SYS, respectively. Using the first-differenced GMM estimator in this panel, the coefficient on the lagged dependent variable is only 0.3001, suggesting implausibly low returns to scale. Using the system GMM estimator that exploits the moment conditions, the coefficient on the lagged dependent variable is 0.3659, which is important to speed up the adjustment of error to reach long run equilibrium.

The coefficients of all environmental variables are significant at least at 1% level system panel GMM in two-step version. Hence, macroeconomic, banking industry condition, and bank specific characteristic factors play an important role in determining cost efficiency. In the first set of variables, to capture the potential for retail banking services and its correlation with bank cost efficiency, *population density* variable has a negative sign, indicating that higher population and deposit densities contribute to higher banking costs. One reason can be found in the characteristics of banking competition. In particular, if banks compete by opening more branches for strategic reasons, this creates excessive bank operating costs. The sign of the *real GDP* is positive which shows that whenever development level of the economy is higher; the operating and financial costs are lower. Therefore, countries with a higher real income have a banking system that operates in a mature environment resulting in more competitive interest rates and profit margins. The sign of the coefficient of the *financial deepening ratio* variable is also positive, which suggests that higher level of financial intermediation of banking sector contributes to lower banking costs (*i.e.*, increase in cost efficiency). Therefore, on average a larger volume of financial intermediation through the banking sector can be associated with somewhat higher efficiency levels. This result is in line with the results of Dietsch and Vivas (2000).

A second set of variables was used to capture the banking market structure in specific countries. The sign of the coefficient of *banking activity* variable is negative, which suggests higher amounts of loans per unit of GDP will increase banking costs. The *bank network extension* index is positive, which implies that banks have higher cost efficiency in highly extended bank markets. Therefore, a larger number of banks should indicate a more competitive environment and banking service accessibility, indicating development of banking service could be connected with higher cost efficiency.

The third set of variables representing individual bank characteristics consists of the following variables. The coefficient on *profitability* has an expected negative sign, indicating that higher amounts of profit per unit of equity increase banking costs (*i.e.*, decrease in cost efficiency). In essence, the empirical findings seem to suggest that in case of the euro area banking system, the more profitable banks may not necessarily those which exhibit higher cost efficiency level. The coefficient on *loan management activity* is negative, which is indicating that whenever the loan per unit of asset rate is higher; the banking costs are higher. On the other hand, banks with a higher ratio of loans to assets are found to be less cost efficient.

3 The Sargent test is the most common diagnostic utilized in GMM estimation to evaluation the suitability of the model. A rejection of the null hypothesis implies that the instruments are not satisfying the orthogonality condition required for their employment (Baum et al., 2007).

4 Baum et al. (2010) points out that in a dynamic panel data context, the first order serial correlation could be expected, but the second-order serial correlation should not be detected if the instruments are appropriately uncorrelated with the error term.

The results of first-difference panel GMM in two-step version estimation shows that coefficients of all environmental variables are significant at least at 1% level as like as the system panel GMM with the same sign but different amount of coefficients. Moreover, system panel GMM regression which includes time trend variable was estimated (column 5 and 6), and results are consistent with two-step system GMM regarding sing and level of significance, but the amount of coefficients are different. Although, adjustment speed of lag dependent variable (0.3646) is significant and positive but less amount than two-step system GMM.

The first two columns of Table 6 report the results for first difference panel GMM and next two columns report system panel GMM, respectively. Using the first-differenced GMM estimator in this panel, the coefficient on the lagged dependent variable is only 0.3564, suggesting implausibly low returns to scale. Using the system GMM estimator which exploits the moment conditions, the coefficient on the lagged dependent variable is low (0.4233) and statistically significant. These results suggest that the profit efficiency of the previous year (L1) is significantly and positively related to the efficiency of the current year in both models.

It could be argued that the efficiency of the previous year may represent a certain level of accumulated knowledge and technological endowment that may help banks to generate higher outputs with their inputs by adapting relatively quick to the changes brought by the environmental conditions.

The results of both specification tests, that is AR(2) for testing the serial correlation and the Sargent test for testing the validity of instrument adopted are also valid. As shown in Table 10, the p-values for the AR(2) and Sargent tests are higher than 0.10, that is, statistically insignificant at the ten percent significance level. This implies that the empirical model has been correctly specified because there is no serial correlation (autocorrelation) in the transformed residuals, and the instruments (moments conditions) used in the models are valid.

In general, the influence of the environmental variables is in line with our expectations. All of the coefficients on the environmental variables in estimation of the profit efficiency model are significant at the 1% confidence level for first difference and system panel GMM in two-step version. It appears that macroeconomic, banking industry condition and bank specific characteristic factors in the banking sector play a more important role for both profit efficiency and cost efficiency. Therefore, these findings confirm our belief that the environmental variables are an important factor in explaining differences in euro area banking efficiency.

As can be seen in Table 6, in the two-step system GMM estimation, the first group which is called *country level variables* (including population density, economic condition and financial deepening ratio) is positively and statistically significant in influencing the banks' profit efficiency except population density sign. Contrary to expectations, the coefficient of the population density variable has a negative sign. Higher density contributes to an increase in banking costs instead of the expected decrease in costs, causing lower profit efficiency. Two reasons can be found in the characteristics of non-price banking competition. In particular, banks create an excessive bank operating cost when they compete by extending branches for the strategic reasons. Therefore, banks should suffer more cost where real estate is more expensive and salaries are higher for opening branches in large when supplying a given level of services. Real GDP is positively linked with bank profit efficiency. Specifically, a one percentage point increases in the real GDP leads to a contemporaneous increase in banks' profit efficiency by 0.0105 percentage point. This factor is expected to affect the demand and supply of deposits and loans in numerous ways. In particular, the countries with higher real income have a banking system that operates in a mature environment resulting in more competitive interest rates and profit margins. The effect of financial deepening ratio, measured by total bank asset to GDP in a given country, is also statistically significant and positively influence the banks' profit efficiency, indicating that on average a larger volume of financial intermediation through the banking sector can be associated with somewhat higher profit efficiency levels.

Table 6 Baseline analysis for effect of country, industry, and bank characteristics on profit efficiency (controlling endogeneity)

Regressors	GMM-DIF One-step	GMM-DIF Two-step	GMM-SYS One-step	GMM-SYS Two-step	GMM-SYS♣ One-step	GMM-SYS♣ Two-step
Initial of cost efficiency (L1)	0.3565***	0.3564***	0.4226***	0.4233***	0.4182***	0.4175***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population density	–0.0022	–0.0021***	–0.0006***	–0.0006***	–0.0006***	–0.0005***
	(0.250)	(0.000)	(0.005)	(0.000)	(0.005)	(0.000)
Economic condition[1]	–0.5147***	–0.5146***	0.0062**	0.0063***	0.0066**	0.0069***
	(0.001)	(0.000)	(0.018)	(0.000)	(0.015)	(0.000)
Financial deepening rate	0.0185*	0.0184***	0.0090**	0.0088***	0.0093**	0.0086***
	(0.057)	(0.000)	(0.040)	(0.000)	(0.036)	(0.000)
Banking activity	–0.0630	–0.0630***	–0.1075***	–0.1068***	–0.1048***	–0.1041***
	(0.248)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Bank network extension	0.0026**	0.0027***	0.0026***	0.0026***	0.0026***	0.0025***
	(0.038)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Profitability	–0.0064**	–0.0063***	–0.0057*	–0.0056***	–0.0058*	–0.0057***
	(0.037)	(0.000)	(0.070)	(0.000)	(0.068)	(0.000)
Loan management activity	0.1322	0.1326***	–0.0782	–0.0800***	–0.0850	–0.0891***
	(0.320)	(0.000)	(0.418)	(0.000)	(0.380)	(0.000)
Sargan test (*p*-value)[2]	0.0072	0.3278	0.0001	00.4942	0.0001	0.5090
Serial correlation test						
AR(1) (*p*-value)[3]	0.0001	0.0004	-	0.0003	-	0.0003
AR(2) (*p*-value)	0.5034	0.7047	-	0.4173	-	0.6676
Wald test for joint significance (*p*-value)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
No. of instruments	96	96	120	120	121	121
Cross-sectional observations	123	123	126	126	126	126

Notes: The independent variables, population density (POPD) calculated as people per sq. km of land area; economic condition (LGDP) calculated as the natural log of the GDP (constant 2005 US$); financial deepening rate (FDEEP) calculated as total bank asset (domestic and foreign bank) to GDP current price; banking activity (CLAIMS) calculated as domestic credit (bank claim) to private sector over the GDP; bank network extension (BRANCH) calculated as number of domestic and foreign branches of credit institutions (per 1 000 km^2); profitability (ROE) calculated as total pre-tax profit over total equity; loan management activity (LOANTA) calculated as total loan over total assets.

♣ The regressions also include time trend variable for the different time periods that are not reported. [1] In the regression, this variable is included as log(variable). [2] The null hypothesis is that model and overidentifying conditions are correct specified. [3] The null hypothesis is that there is no serial correlation in the first-differenced disturbances. Values in parenthesis are *p*-value. ***, **, * indicates significance at 1%, 5% and 10% levels respectively.

Source: Own construction

The second group consists of variables, describing the *structure of the banking industry* in specific countries. The sign of the coefficient of loans per unit of GDP ratio or so-called *banking activity* is negative. In essence, the empirical findings seem to suggest that in case of the euro area banking system, the more profit efficient banks may not necessarily be the one which performs in county with high loan per unit of GDP level. The accessibility of banking services for the customers, measured by dividing number of domestic and foreign branches of credit institutions by the number of 1 000 square kilometres, has a positive sign. A higher banking density may favour bank efficiency by making the access to banking products easier for customers. Thus, the higher density of bank branches, the higher is the banking profit efficiency.

The third group consists of variables, describing the bank specific characteristics in specific banks. *Profitability* prove to have a negative influence on profit efficiency, indicating that higher amounts of profit per unit of equity decrease banking profit (*i.e.*, decrease in profit efficiency). The results imply that more profitable banks may not necessarily be those which exhibit higher profit efficiency level. The *loan management activity* also seems to be negatively connected with profit efficiency, indicating that whenever loan per unit of asset rate is higher, the banking profits are lower. The empirical findings seem to suggest that banks with higher ratio of loans to assets are found to be less profit efficient.

The results of first difference panel GMM estimation in the two-step version shows that coefficients of all environmental variables are significant at least at 1% level as like as a system panel GMM with same sign in most cases, but the amount of coefficients is different. The economics condition and loan management activity variables has a different sign from system panel GMM. The sign of real GDP is negative and loan per unit of asset rate is positive which is not consistent with the two-step system GMM estimation. Moreover, the system panel GMM regressions includes estimated time trend variable (column 5 and 6) and the results are consistent with the two-step system GMM regarding the sign and level of significance, but coefficients are different. Although, the adjustment speed of lag dependent variable (0.4175) is significant and positive but less amount than the two-step system GMM.

4.4 Robustness Test

In order to check for the robustness of the results, we have estimated both the cost and profit model with similar regression models but endogeneity problem of economics condition variable has not controlled the model estimation with GMM estimator.

The empirical findings are presented in Tables 7 and 8 seem to suggest a perspicuous relationship between the level of environmental variables and bank efficiency in the euro area countries. Following Arellano and Bond (1991), three additional conditions should be satisfied to avoid model misspecification: a significant AR(1) serial correlation, lack of AR(2) serial correlation and a high Sargan test statistics.

Using the first-differenced GMM estimator in this panel, the coefficients on the lagged cost and profit efficiency variables are only 0.3634 and 0.4145, respectively, suggesting implausibly low returns to scale. Using the system GMM estimator which exploits the moment conditions, the coefficients on the lagged cost and profit efficiency variables are low (0.3871 and 0.4396) and statistically significant. These results suggest that cost and profit efficiency of the previous year (L1) is significantly and positively related to the efficiency of the current year in both models.

It can be observed from both tables that, all in all, the results remain qualitatively similar in terms of directions and significance level. Similar to baseline regression model, the population density, banking activity, profitability, and loan management activity variables have negative association with cost and profit efficiency in system GMM. Although, it can be observed from Tables 7 and 8

Table 7 Robustness analysis for effect of country, industry, and bank characteristics on cost efficiency (dynamic panel estimation)

Regressors	GMM-DIF One-step	GMM-DIF Two-step	GMM-SYS One-step	GMM-SYS Two-step	GMM-SYS♣ One-step	GMM-SYS♣ Two-step
Initial of cost efficiency (L1)	0.3622***	0.3634***	0.3874***	0.3871***	0.3912***	0.3897***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population density	–0.0043***	–0.0039***	–0.0010***	–0.0010***	–0.0010***	–0.0010***
	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Economic condition[1]	0.0130	0.0212	0.0086***	0.0085***	0.0083***	0.0077***
	(0.913)	(0.120)	(0.001)	(0.000)	(0.003)	(0.000)
Financial deepening rate	0.0144**	0.0139***	0.0166***	0.0164***	0.0165***	0.0165***
	(0.046)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Banking activity	–0.0510	–0.0559***	–0.0757**	–0.0777***	–0.0822**	–0.0844***
	(0.247)	(0.000)	(2.04)	(0.000)	(0.032)	(0.000)
Bank network extension	0.0039***	0.0040***	0.0040***	0.0040***	0.0040***	0.0041***
	(0.000)	(0.000)	(6.58)	(0.000)	(0.000)	(0.000)
Profitability	–0.0023	–0.0023***	–0.0021	–0.0013	–0.0021	–0.0016*
	(0.317)	(0.000)	(0.88)	(0.257)	(0.382)	(0.086)
Loan management activity	0.1399	0.1463***	–0.2401***	–0.2378***	–0.2379***	–0.2338***
	(0.167)	(0.000)	(3.44)	(0.000)	(0.001)	(0.000)
Sargan test (*p*-value)[2]	0.0001	0.1236	0.0001	0.087	0.0001	0.076
Serial correlation test						
AR(1) (*p*-value)[3]	0.0001	0.0006	-	0.0005	-	0.0006
AR(2) (*p*-value)	0.351	0.2665	-	0.3015	-	0.3137
Wald test for joint significance (*p*-value)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
No. of instruments	85	85	97	97	98	98
Cross-sectional observations	123	123	126	126	126	126

Notes: The independent variables, population density (POPD) calculated as people per sq. km of land area; economic condition (LGDP) calculated as the natural log of the GDP (constant 2005 US$); financial deepening rate (FDEEP) calculated as total bank asset (domestic and foreign bank) to GDP current price; banking activity (CLAIMS) calculated as domestic credit (bank claim) to private sector over the GDP; bank network extension (BRANCH) calculated as number of domestic and foreign branches of credit institutions (per 1 000 km^2); profitability (ROE) calculated as total pre-tax profit over total equity; loan management activity (LOANTA) calculated as total loan over total assets.

♣ The regressions also include time trend variable for the different time periods that are not reported. [1] In the regression, this variable is included as log (variable). [2] The null hypothesis is that model and overidentifying conditions are correct specified. [3] The null hypothesis is that there is no serial correlation in the first-differenced disturbances. Values in parenthesis are *p*-value. ***, **, * indicates significance at 1%, 5% and 10% levels respectively.

Source: Own construction

Table 8 Robustness analysis for effect of country, industry, and bank characteristics on profit efficiency (dynamic panel estimation)

Regressors	GMM-DIF One-step	GMM-DIF Two-step	GMM-SYS One-step	GMM-SYS Two-step	GMM-SYS♣ One-step	GMM-SYS♣ Two-step
Initial of cost efficiency (L1)	0.4149***	0.4145***	0.4413***	0.4396***	0.4426***	0.4402***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population density	0.0003	0.0003***	–0.0016***	–0.0016***	–0.0017***	–0.0017***
	(0.873)	(0.007)	(0.001)	(0.000)	(0.001)	(0.000)
Economic condition[1]	–0.4316***	–0.4029***	0.0075	0.0071***	0.0068	0.0067***
	(0.010)	(0.000)	(0.155)	(0.000)	(0.197)	(0.000)
Financial deepening rate	0.0182*	0.0173***	0.0074	0.0074***	0.0070	0.0069***
	(0.053)	(0.000)	(0.134)	(0.000)	(0.160)	(0.000)
Banking activity	0.0040	–0.0032	–0.0019	–0.0015	–0.0082	–0.0069***
	(0.947)	(0.419)	(0.968)	(0.432)	(0.865)	(0.007)
Bank network extension	0.0046***	0.0044***	0.0036***	0.0036***	0.0036***	0.0037***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Profitability	–0.0074**	–0.0077***	–0.0078**	–0.0073***	–0.0077**	–0.0079***
	(0.021)	(0.000)	(0.018)	(0.000)	(0.019)	(0.000)
Loan management activity	0.0325	0.0304***	–0.0205	–0.0197***	–0.0245	–0.0245***
	(0.813)	(0.000)	(0.867)	(0.000)	(0.842)	(0.000)
Sargan test (*p*-value)[2]	0.0001	0.5254	0.0001	0.2423	0.0001	0.2222
Serial correlation test						
AR(1) (*p*-value)[3]	0.0001	0.0003	-	0.0003	-	0.0003
AR(2) (*p*-value)	0.3227	0.6558	-	0.6567	-	0.7110
Wald test for joint significance (*p*-value)	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
No. of instruments	85	85	97	97	98	98
Cross-sectional observations	123	123	126	126	126	126

Notes: The independent variables, population density (POPD) calculated as people per sq. km of land area; economic condition (LGDP) calculated as the natural log of the GDP (constant 2005 US$); financial deepening rate (FDEEP) calculated as total bank asset (domestic and foreign bank) to GDP current price; banking activity (CLAIMS) calculated as domestic credit (bank claim) to private sector over the GDP; bank network extension (BRANCH) calculated as number of domestic and foreign branches of credit institutions (per 1 000 km^2); profitability (ROE) calculated as total pre-tax profit over total equity; loan management activity (LOANTA) calculated as total loan over total assets.
♣ The regressions also include time trend variable for the different time periods that are not reported. [1] In the regression, this variable is included as log (variable). [2] The null hypothesis is that model and overidentifying conditions are correct specified. [3] The null hypothesis is that there is no serial correlation in the first-differenced disturbances. Values in parenthesis are *p*-value. ***, **, * indicates significance at 1%, 5% and 10% levels respectively.
Source: Own construction

that the impact of economic conditions, financial deepening rate and bank network extension are positive on the cost and profit efficiency of banks operating in the euro area countries which is consistence with baseline estimation.

CONCLUSION AND POLICY IMPLICATION

The purpose of the current descriptive and quantitative correlational study was to examine the relationship between environmental variable and level of bank cost and profit efficiency in the euro area. The two-step quantitative research design was employed to accomplish the purpose of the current study: Data Envelopment Analysis (DEA) and panel regression analysis. At the first stage, we estimated the cost and profit efficiency level of entire 126 listed bank dataset for 17-euro area Member States by using the nonparametric DEA approach to investigate whether the cost and profit efficiency of the euro area banking system improved between 1999 and 2012, and to compare the efficiency scores of the financial sectors of the euro area Member States. At the second stage, we regressed the efficiency level obtained from the first stage on factors that could influence the efficiency of banks by using a GMM regression model the period of study. We choose an unbalanced panel rather than a balanced panel, to take banks gone into bankrupt or those being absorbed into account. Indeed, the use of a balanced panel may overestimate cost efficiency as it ignores these banks, which may be less efficient on average.

The average cost efficiency of the banks in the sample has been estimated to be 35.84%, indicating that banks could reduce their costs by 65.12% on average. Reassuringly, evidence on increasing average cost efficiency over the time is found. The study also detected differences in average cost efficiency of banks among countries. The highest average bank efficiency scores were achieved by the three states, the Netherlands, France, and Germany with value of 70.95%, 56.69%, and 48.03%, respectively. The lowest average efficiency scores were obtained for Slovakia, Greece and Portugal, where bank efficiency on average amounted to only 11.12%, 11.23%, and 16.79%, respectively.

Concerning profit efficiency, the average efficiency level is 24.55%, meaning that on average they are 75.45% inefficient in profit. Inefficiency levels show by how much profit (costs) can increase (decrease) if inputs where used as efficiently as the unit located on the best practice frontier. It could be observed that the Netherlands, Germany and France are three highest average profit efficiency members across the euro area during the period of study with 63.95%, 44.58%, and 37.99%, respectively. In comparison, Portugal, Slovakia, and Greece have the lowest average profit efficiency members across the euro area with 2.56%, 3.29%, and 4.52%, respectively. Finally, the empirical results indicate a large asymmetry between countries regarding their profit efficiency level. Although, the evolution of profit efficiency in each country shows that there does not seem to be a clear trend in general. The efficiency scores have been decreasing from starting to ending years in most of the countries in the sample.

This study also investigates the influence the environmental conditions on the cost and profit efficiency of banks operating in the banking industries of the seventeen members from the euro area over the period 1999–2012. In particular, the specific environmental conditions of each country play an important role in the definition and specification of the common frontier of different countries. We formed three groups of variables that are assumed to be associated with changes in efficiency across banks. The first group includes country level variables explaining macroeconomic conditions (population density, economic condition, and financial deepening rate). The second group consists of variables describing the industry of the banking sector in specific countries (banking activity and bank network extension). Finally, variables in the third group describe bank specific characteristics that could determine differences in efficiency levels achieved (profitability and loan management activity). Overall, our results demonstrate that environmental variables contribute significantly to the difference in efficiency scores between the Member States.

As for the first set of variables, a higher population density contributes to lower cost and profit efficiency, most likely due to an increase in bank operating costs. GDP (constant 2005 US$) proved to have a positive influence on cost and profit efficiency, while the effect of financial deepening seems to be positive, indicating that banks in financially more developed markets on average operate at somewhat lower inefficiency levels. Results indicated that higher banking activity, measured by the domestic credit to private sector over the GDP, results in lower average cost and profit efficiency levels. Other banking industry characteristics, namely bank network extension, measured by the number of banks operating in a given country is found to have a significant and positive influence on cost and profit efficiency. Furthermore, the cost and profit efficiency of a given bank was found to be on average negatively related to its banking activity and loan management activity which measured by loan over total assets. The bank profitability indicator, ROE, as expected, turned out to be negatively related to bank efficiency.

For policy implication, bank regulators and management in Slovakia, Greece, and Portugal (as the most inefficient banks), under the condition of a market economy and facing a fiercely competitive banking market, should focus on how to improve management, innovate technology and enhance the quality of employees. The result of this would contribute to improvement of their cost and profit efficiency and their competitive power in the euro area banking market, rather than to enlarging the scale of production and increasing outputs by increasing inputs. In addition, the least inefficient banks could be closed and their resources transferred to the relatively efficient banks, the result of this would be that the euro area banking industry could also improve the cost and profit efficiency and produce more outputs with less resources.

References

ALLEN, L. AND RAI, A. Operational efficiency in banking: An international comparison. *Journal of Banking & Finance*, 1996, 20(4), pp. 655–672.

ALTUNBAS, Y., EVANS, L., MOLYNEUX, P. Bank ownership and efficiency. *Journal of Money, Credit and Banking*, 2001, 33(4), pp. 926–954.

ANDRIES, A. M., APETRI, A. N., COCRIS, V. The impact of the banking system reform on banks performance. *African Journal of Business Management*, 2012, 6(6), pp. 2278–2284.

AVKIRAN, N. K. The evidence on efficiency gains: the role of mergers and the benefits to the public. *Journal of Banking & Finance*, 1999, 23(7), pp. 991–1013.

BARTH, J. R., LIN, C., MA, Y., SEADE, J., SONG, F. M. Do bank regulation, supervision and monitoring enhance or impede bank efficiency? *Journal of Banking & Finance*, 2013, 37(8), pp. 2879–2892.

BERG, S. A., FØRSUND, F. R., JANSEN, E. S. Malmquist indices of productivity growth during the deregulation of norwegian banking, 1980–89. *The Scandinavian Journal of Economics*, 1992, 94, pp. 211–28.

BERGENDAHL, G. DEA and benchmarks–an application to nordic banks. *Annals of Operations Research*, 1998, 82, pp. 233–250.

BERGER, A. N., AND HUMPHREY, D. B. Efficiency of financial institutions: International survey and directions for future research. *European Journal of Operational Research*, 1997, 98(2), pp. 175–212.

BERGER, A. N. AND DEYOUNG, R. Problem loans and cost efficiency in commercial banks. *Journal of Banking & Finance*, 1997, 21(6), pp. 849–870.

BERGER, A. N., AND HUMPHREY, D. B. The dominance of inefficiencies over scale and product mix economies in banking. *Journal of Monetary Economics*, 1991, 28(1), pp. 117–148.

BERGER, A. N. AND MESTER, L. J. Inside the black box: What explains differences in the efficiencies of financial institutions? *Journal of Banking & Finance*, 1997, 21(7), pp. 895–947.

BOS, J. W. AND KOLARI, J. W. Large bank efficiency in europe and the united states: Are there economic motivations for geographic expansion in financial services? *The Journal of Business*, 2005, 78(4), pp. 1555–1592.

BOS, J. W. B. AND SCHMIEDEL, H. Is there a single frontier in a single european banking market? *Journal of Banking & Finance*, 2007, 31(7), pp. 2081–2102. doi:10.1016/j.jbankfin.2006.12.004

BRISSIMIS, S. N., DELIS, M. D., TSIONAS, E. G. Technical and allocative efficiency in european banking. *European Journal of Operational Research*, 2010, 204(1), pp. 153–163. doi:10.1016/j.ejor.2009.09.034

CASU, B., FERRARI, A., ZHAO, T. Regulatory reform and productivity change in Indian banking. *Review of Economics and Statistics*, 2013, 95(3), pp. 1066–1077.
CASU, B., GIRARDONE, C., MOLYNEUX, P. Productivity change in european banking: a comparison of parametric and non-parametric approaches. *Journal of Banking & Finance*, 2004, 28(10), pp. 2521–2540.
CHAFFAI, M. E., DIETSCH, M., LOZANO-VIVAS, A. Technological and environmental differences in the european banking industries. *Journal of Financial Services Research*, 2001, 19(2–3), pp. 147–162.
COELLI, T. J., RAO, D. S. P., O'DONNELL, C. J., BATTESE, G. E. *An introduction to efficiency and productivity analysis.* Springer Science & Business Media, 2005.
COELLI, T. J. A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation. *CEPA working paper 96/97*, Department of Econometrics, University of New England, 1996.
DAS, A. AND GHOSH, S. Financial deregulation and efficiency: an empirical analysis of indian banks during the post reform period. *Review of Financial Economics*, 2006, 15(3), pp. 193–221.
DEYOUNG, R. AND NOLLE, D. E. Foreign-owned banks in the united states: Earning market share or buying it? *Journal of Money, Credit and Banking*, 1996, 28(4), pp. 622–636.
DIETSCH, M. AND LOZANO-VIVAS, A. How the environment determines banking efficiency: a comparison between french and spanish industries. *Journal of Banking & Finance*, 2000, 24(6), pp. 985–1004.
DRAKE, L., HALL, M. J., SIMPER, R. The impact of macroeconomic and regulatory factors on bank efficiency: a non-parametric analysis of hong Kong's banking system. *Journal of Banking & Finance*, 2006, 30(5), pp. 1443–1466.
EDWARD CHANG, C., HASAN, I., HUNTER, W. C. Efficiency of multinational banks: an empirical investigation. *Applied Financial Economics*, 1998, 8(6), pp. 689–696.
FRIES, S. AND TACI, A. Cost efficiency of banks in transition: Evidence from 289 banks in 15 post-communist countries. *Journal of Banking & Finance*, 2005, 29(1), pp. 55–81. doi:10.1016/j.jbankfin.2004.06.016
HASAN, I. AND HUNTER, W. C. Efficiency of japanese multinational banks in the united states. *Research in Finance*, 1996, 14, pp. 157–174.
ISIK, I. AND HASSAN, M. K. Efficiency, ownership and market structure, corporate control and governance in the turkish banking industry. *Journal of Business Finance & Accounting*, 2003, 30(9–10), pp. 1363–1421.
KASMAN, A. AND YILDIRIM, C. Cost and profit efficiencies in transition banking: the case of new EU members. *Applied Economics*, 2006, 38(9), pp. 1079–1090.
LEHMANN, E., WARNING, S., WEIGAND, J. Governance structures, multidimensional efficiency and firm profitability. *Journal of Management and Governance*, 2004, 8(3), pp. 279–304.
LOVELL, C. A. K. Production frontiers and productive efficiency. *The Measurement of Productive Efficiency: Techniques and Applications*, 1993, 20(2), pp. 3–67.
LOZANO-VIVAS, A., PASTOR, J. T., HASAN, I. European bank performance beyond country borders: What really matters? *European Finance Review*, 2001, 5(1–2), pp. 141–165.
LOZANO-VIVAS, A., PASTOR, J. T., PASTOR, J. M. An efficiency comparison of european banking systems operating under different environmental conditions. *Journal of Productivity Analysis*, 2002, 18(1), pp. 59–77.
MAHAJAN, A., RANGAN, N., ZARDKOOHI, A. Cost structures in multinational and domestic banking. *Journal of Banking & Finance*, 1996, 20(2), pp. 283–306.
MCALLISTER, P. H. AND MCMANUS, D. Resolving the scale efficiency puzzle in banking. *Journal of Banking & Finance*, 1993, 17(2), pp. 389–405.
MESTER, L. J. A study of bank efficiency taking into account risk-preferences. *Journal of Banking & Finance*, 1996, 20(6), pp. 1025–1045.
PARADI, J. C. AND ZHU, H. A survey on bank branch efficiency and performance research with data envelopment analysis. *Omega*, 2013, 41(1), pp. 61–79.
PASIOURAS, F. Estimating the technical and scale efficiency of greek commercial banks: the impact of credit risk, off-balance sheet activities, and international operations. *Research in International Business and Finance*, 2008, 22(3), pp. 301–318.
PASTOR, J., PEREZ, F., QUESADA, J. Efficiency analysis in banking firms: an international comparison. *European Journal of Operational Research*, 1997, 98(2), pp. 395–407.
PEEK, J., ROSENGREN, E. S., KASIRYE, F. The poor performance of foreign bank subsidiaries: Were the problems acquired or created? *Journal of Banking & Finance*, 1999, 23(2), pp. 579–604.
SEALEY, C. W. AND LINDLEY, J. T. Inputs, outputs, and a theory of production and cost at depository financial institutions. *The Journal of Finance*, 1977, 32(4), pp. 1251–1266.
SHELDON, G. Costs, competitiveness and the changing structure of european banking. *Fondation Banque De France Pour La Recherche Working Paper*, Paris, France, 1999.
STURM, J. AND WILLIAMS, B. Foreign bank entry, deregulation and bank efficiency: Lessons from the australian experience. *Journal of Banking & Finance*, 2004, 28(7), pp. 1775–1799.
Trade for all: Towards a more responsible trade and investment policy [online]. European Commission, 2015. [cit. 26.4.2019]. <http://trade.ec.europa.eu/doclib/docs/2015/october/tradoc_153846.pdf>.
YILDIRIM, H. S. AND PHILIPPATOS, G. C. Efficiency of banks: Recent evidence from the transition economies of Europe, 1993–2000. *European Journal of Finance*, 2007, 13(2), pp. 123–143.

Modeling of Trend in Sales of New Passenger Vehicles in Poland Using the Arima-X-12 and Tramo-Seats Methods

Wojciech Lewicki[1] | *West Pomeranian University of Technology Szczecin, Szczecin, Poland*
Aleksandra A. Olejarz-Wahba[2] | *University of Warmia and Mazury in Olsztyn, Olsztyn, Poland*

Abstract

The purpose of this article is to present development trends in sales of new passenger vehicles in Poland. The authors of the article suggest that this way of presenting sales results is incorrect. When decomposing the time series of the number of passenger car sales per seasonal component, the trend and accidental fluctuations and irregular fluctuations, it was proved that the analyzed phenomenon is more complex. In addition, it has been proved that the sale of new cars is a component of sales of passenger cars produced this year and sales of vehicles produced in the previous year, in which sales are characterized by different seasonality.

The monthly data from January 2013 to November 2018 were used to decompose the time series of sales of new passenger vehicles in Poland. The data from the Central Register of Vehicles and Drivers was made available. The ARIMA-X-12 and TRAMO-SEATS methods were used to decompose the analyzed time series.

Keywords

Time series, seasonality, vehicle sales, trend, TRAMO-SEATS, ARIMA-X-12, outliers

JEL code

C15, D24, O52, R40

INTRODUCTION

The available literature on the subject emphasizes that in the last decades in the field of motorization a very significant technological progress (Nolan, 2010) has been recorded. This process concerned both the global automotive industry and the countries of Eastern Europe. Almost all stages of motor vehicles production at that time introduced new, cheaper production technologies, which led to the current state, in which the car has became the goods available to almost every consumer (Kudłak et al., 2017).

Observations of market reality made by the authors of the article clearly show that in the trade press, as well as in the scientific literature, the sale of passenger vehicles is analyzed as the ratio of current sales

[1] West Pomeranian University of Technology Szczecin, Faculty of Economics, 43 Żołnierska St., 71-210 Szczecin, Poland. E-mail: Wojciech.Lewicki@zut.edu.pl.

[2] University of Warmia and Mazury in Olsztyn, Faculty of Economics, 4 Michała Oczapowskiego St., 10-719 Olsztyn, Poland. E-mail: wahba@wp.pl.

to the result from the previous month and analogous to the month from the previous year. As the sale of new passenger vehicles, the number of new vehicles registered for the first time is presented (Prieto et al., 2013; Nolan, 2010: George et al., 2009; Klimkowska, 2012; Sołtysiak, 2015; Sryjakiewicz, 2015; Michalak and Merło, 2015; Mutrynowski, 2015).

Moreover, as the in-depth analysis of the available literature has shown, no such analyzes exist. The authors of the article suggest that this way of presenting sales results is incorrect. When decomposing the time series of the number of passenger car sales per seasonal component, the trend and accidental fluctuations and irregular fluctuations, it was proved that the analyzed phenomenon was more complex. In addition, it was proved that the sale of new cars was a component of sales of passenger cars produced this year and sales of vehicles produced in the previous year. It was demonstrated that these sales were characterized by varying seasonality. Therefore, presenting the considerations regarding the development trends of sales of new passenger vehicles in Poland seems to be the appropriate marking of the research problem, covering both sales trends and the selection of analytical tools and methods for analyzing and forecasting the sales volume of automotive products (and studies referring to this market, which makes the subject even more interesting and worth considering, in order to diagnose the correctness of steps taken in this matter (Urban, 2000; Grudkowska and Paśnicka, 2007).

The article has been divided into individual sections. Section 1 discusses and justifies the selection of ARIMA-X-12 and TRAMO-SEATS research methods. Section 2 draws attention to the essence of research issues, indicating at the same time the irregularities in the current method of analysis of vehicle sales in Poland. Section 3 presents the first results of research, analyzing seasonal trends and fluctuations. The conclusions and next steps are included in the summaries in the last part of this article.

The addressees of the presented research results are selected entities from the automotive and insurance market, including dealers of new vehicles, authorized service stations, leasing partners and insurance companies. In addition, the presented research can also be used by public administration bodies to forecast the future number of vehicles on the road.

1 METHODS

As the literature on the subject indicates, describing the time series, the trend of the studied phenomenon is identified, as well as the strength and trend of seasonal and accidental factors. Isolation of individual components can be done using mechanical or analytical methods. Mechanical methods (moving average methods and smoothing methods) smooth time series by removing seasonal fluctuations and outliers. When examining the development trends of economic phenomena, the trend and the occurrence of seasonal fluctuations are most often made (Fischer, 1995). Analytical methods are most frequently used to evaluate these components, which result in econometric models describing the evolution of the studied phenomenon over time. Irregular effects, outliers are often eliminated as those that distort the general level of the phenomenon (Dagum, 1980). These effects may, however, be a source of more accurate knowledge about the formation of the phenomenon. Analyzing the sales of new passenger vehicles in Poland, the authors used the following methods: TRAMO-SEATS and ARIMA-X-12.

The TRAMO-SEATS procedure and the ARIMA-X-12 procedure are methods that combine analytical and mechanical analysis of time series (Sax et al., 2018) These methods allow the analysis of non-stationary time series, unambiguous identification of the trend, seasonal and accidental fluctuations and allow to determine the nature of these fluctuations (Cabrero, 2000). The ARIMA-X-12 procedure was developed and is used by the United States Census Bureau. The TRAMO-SEATS procedure was developed by A. Maravell and V. Gomez in 1996. In both procedures, the ARIMA model is estimated, with the TRAMO-SEATS procedure matching the ARIMA model to each of the time series components when the ARIMA-X-12 procedure estimates such a model only for the original series. The smoothing process is also different. In the first method, the test results determine the selection of the seasonal adjustment filters. In contrast, in the long-term method, ad hoc filters are used (Lian et al., 2018).

The TRAMO-SEATS and ARIMA-X-12 procedures identify seasonal effects when performing the Fredman test, the Kruskal-Wallis test (is one of the most popular alternatives to the one-way ANOVA variance analysis), the mobile seasonality test, the seasonally traceable test and the complex seasonality test (Dagum et al., 2016). Thanks to this, it is possible to identify seasonality and mobile seasonality, based on the evolution of the seasonal pattern in time. Testing is based on the isolated seasonal component (Baron et al., 2018).

The Friedman test, created by the American economist Milton Friedman, compares average levels over several periods. The null hypothesis of this test is: *"k different samples come from the same population"* or: *"k different samples come from several populations of equal average."* The studied series do not show seasonal fluctuations when there are no grounds to reject the null hypothesis (Friedman, 1940).

The Kruskal-Wallis test, like the Friedman test, examines whether the samples come from the same population. The test hypotheses are also consistent with the Friedman test hypotheses.

The mobile seasonality test is a two-factor analysis of variance of the SI component. The variance S^2 used in this test is divided into season-related S_b^2 variability, i.e. months or quarters, year-related S_m^2 variation, and intra-group variability S_s^2. The null hypothesis of this test is that the seasonal pattern does not evolve over time.

The identifiable seasonality test verifies the null hypothesis that there is identifiable seasonality in the analyzed time series (Kruskal and Wallis, 1952).

The seasonality composite test is a procedure that verifies the hypothesis of the occurrence of seasonality in time series. This test is a combination of the Friedman test, Kruskal-Wallis test, mobile seasonality test and traceable seasonality. The result of a composite seasonality test is one of the following: there is identifiable seasonality, there is no seasonality, there is a low probability of identifiable seasonality.

As a result, it is possible to identify seasonal and mobile seasonality, consisting in the evolution of the seasonal pattern over time. Testing is based on a separate seasonal component (Enders, 2010). In this case, the various types of outliers analyzed concerned the number of newly registered passenger vehicles produced both in the current year and the previous year.

The calendar effects identified in the seasonal adjustment process are effect business days and the effect of movable holidays. The effect of working days is the effect of different number of working days in particular periods on the phenomenon studied. It is recognized that the activity of business entities is more intense on business days (Maravall et al., 2014). The effect of movable festivals concerns in particular the world of Easter and Corpus Christi and concerns the variable economic activity in the period of the Christmas celebrations. Christmas holidays (Muirhead, 1986) are not included in holiday effects.

The irregular component, which is extracted through the TRAMO-SEATS and ARIMA-X-12 procedures, are unusual and irregular effects, such as random or unpredictable factors. Atypical observations are impulses that may have different character (Muirhead, 1986), Additive Outliers (one-off, significant variations from the expected value of the studied phenomenon, which does not affect the value in subsequent periods), Level Shift (permanent change in the level of the variable), Temporary Change (temporary change of the variable level and return to the initial level, usually in accordance with the exponential or linear function) and Innovation Outliers (innovative impulses caused, for example, by the use of new production technology, resulting in a change of the entire process generating data, including changing the form of the trend) (Cabrero, 2000).

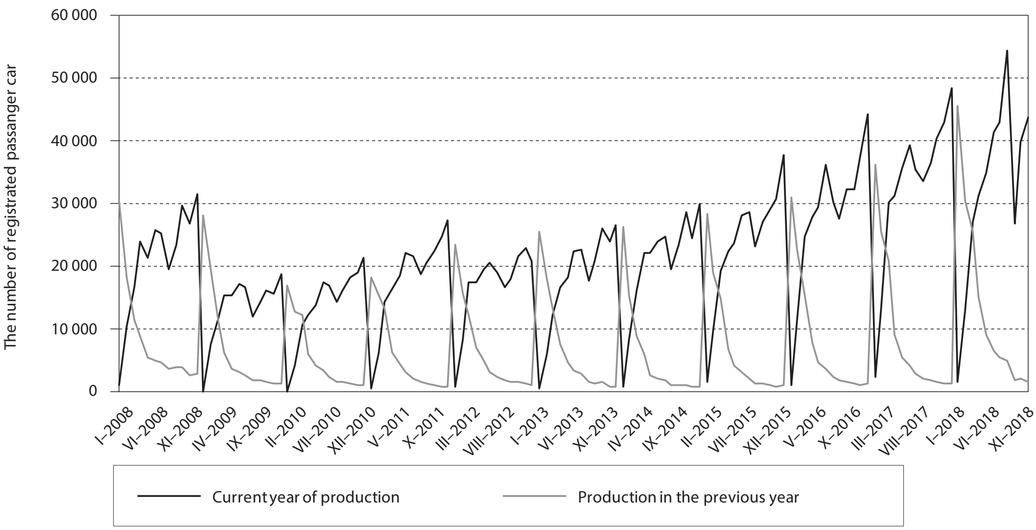
Comparing the TRAMO SEATS procedure and the ARIMA-X-12 procedure, it can be concluded that in the first of them there were more comprehensive possibilities of automatic seasonal adjustments methods. This procedure, like the ARIMA-X-12, provides decomposition of the multiplicative and additive time series and comprehensive model identification. However, the procedure for detecting outliers is different.

TRAMO automatically detects individual types of atypical observations in the form of transient changes (Ghysels and Osborn, 2010). Other types of outliers, i.e. long-lasting (LS) and one-time (AO), are detected as in ARIMA X-12. As a result, a non-stardard observation may be classified differently depending on the procedure (Atuk and Ural, 2002). It does not change the fact that these methods are commonly used by almost all statistical offices in the world (Sax and Eddelbuettel, 2018). Calculations were made using the package DEMETRA + (Grudkowska, 2011).

2 DATA

Previous analyzes of sales of new passenger vehicles in Poland, appearing in industry press, general internet websites, as well as in scientific literature, usually refer to the number of newly registered passenger vehicles in Poland in total and broken down into the most popular car brands. An example is the reports of the Polish Association of the Automotive Industry, in which data from the Central Register of Vehicles and Drivers is presented. Analyzing such data, the authors of this article consider it incorrect. According to the principles of statistics, one should analyze collections of a uniform nature. The set of the number of newly registered vehicles is not a homogeneous set in terms of the tested feature, because under the name "newly registered" there are vehicles registered for the first time and from the current production as well as from the production of the previous year. This approach is considered incorrect by the authors, pointing to the need to analyze these data divided into two subsets: newly registered passenger vehicles from current production and newly registered passenger cars produced in the previous year. Such data were obtained directly from the Central Register of Vehicles and Drivers, which records all vehicles moving around Poland, as well as records issued documents entitling to drive vehicles. The data obtained are monthly data from the period from January 2008 to November 2018. The study did not include December 2018 due to the fact that the vehicles registration in the database systems has changed.

Figure 1 Sales of new passenger cars in the current and previous year production in Poland in 2008–2018



Source: Own study based on CEPIK data

At the beginning of the surveyed period, in January 2008, the total number of newly registered vehicles in Poland was 32 000, with 95% of that number being sales of vehicles from production in 2007.

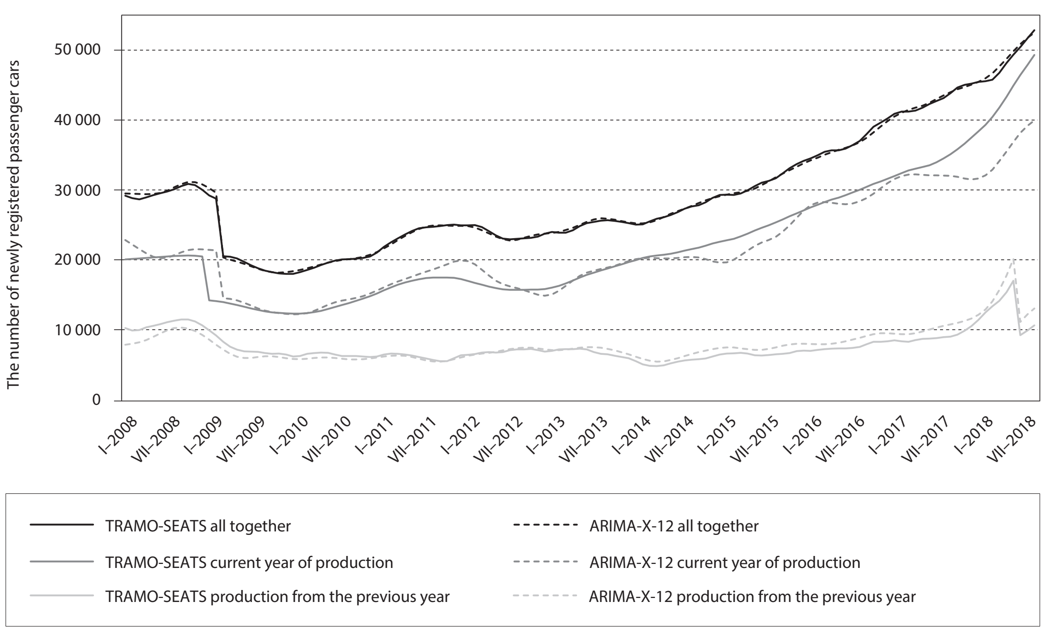
In December 2008, total sales amounted to approximately 35 000 vehicles, of which nearly 95% were sales of vehicles manufactured in the current year. In the following years, the situation was analogous. The share of newly registered vehicles from current production was small at the beginning of the year and then increased, reaching the maximum at the end of the year. The number of newly registered vehicles from production from the previous year was the reverse. The share of these vehicles represented a large share of the total number of newly registered passenger vehicles at the beginning of the year and decreased, reaching a minimum at the end of the year. It is noticeable, therefore, that the occurring fluctuations may indicate the occurrence of seasonality, different for cars from current production and from the previous year. In addition to seasonal fluctuations, an upward trend is visible (see Figure 1).

3 RESULTS AND DISCUSSION

3.1 Trend

The use of the TRAMO-SEATS and ARIMA-X-12 procedures allowed for the decomposition of the time series of the number of newly registered passenger vehicles in Poland. Firstly, the seasoned trend component was obtained (see Figure 2). As can be seen in the chart below, the number of newly registered passenger vehicles in total, vehicles registered and produced in the same year as well as vehicles manufactured in the previous year, were characterized by an increasing trend. However, a significant drop is noticeable at the beginning of 2009. Such a sharp drop in sales was a nipple of reaching the Polish global crisis. This decrease is relatively larger in the case of sales of new vehicles manufactured in the current year than in the case of vehicles from the previous year. The decrease translated into the overall result of sales of new passenger vehicles. However, in the case of the number of newly registered passenger vehicles from production from the previous year, a sudden increase is noticeable, followed by a decrease at the end of 2018. This was influenced by the introduction

Figure 2 The trend of the number of newly registered passenger vehicles in Poland in 2008–2018



Source: Own study based on CEPIK data

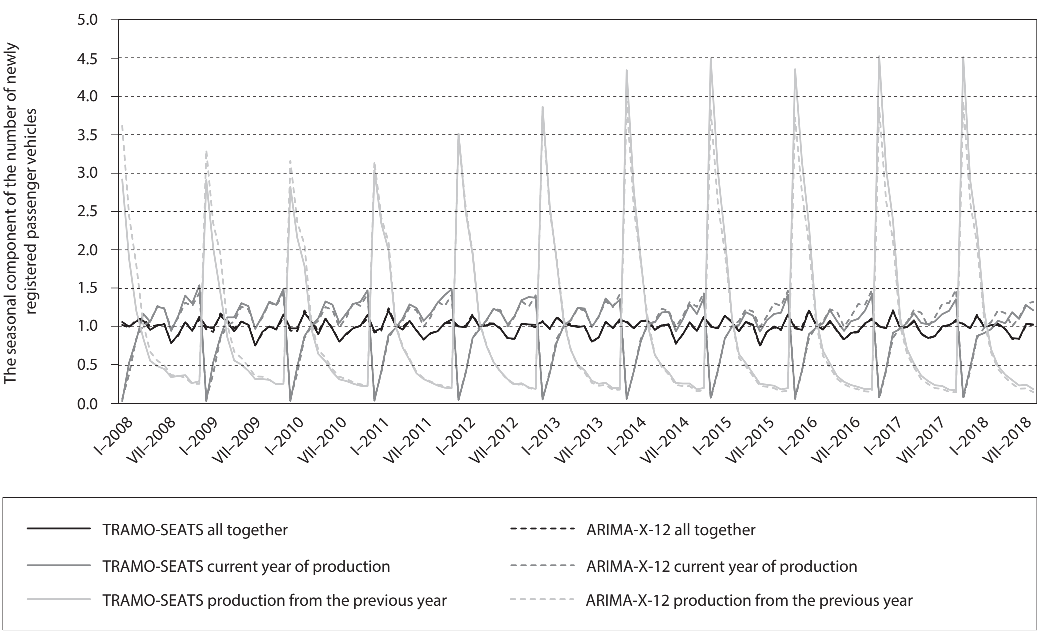
of the "worldwide harmonized light vehicle test procedure", i.e. procedures for fuel consumption, carbon dioxide and exhaust emissions standards.

The results obtained by using both procedures TRAMO-SEATS and ARIMA-X-12 were very similar in the case of the total number of newly registered passenger vehicles and in the case of new passenger vehicle registrations from the previous year production. Differences between the trend estimates occurred in the number of newly registered vehicles from current production.

3.2 Seasonality

Many of the effects is not only the trend and random fluctuations, but also a certain seasonality. It was assumed that the seasonal fluctuations are independent of the level assumed by the time series. The isolated seasonal component of the number of newly registered vehicles in Poland shows variations from the trend. Values below 1.0 indicate a decrease in the number of newly registered vehicles in relation to the trend. Values above 1.0 mean an increase in the number of newly registered vehicles. As can be seen (Figure 3), the higher values of variations from the trend were characterized by sales of passenger vehicles from current production. In addition, the decline in sales of passenger vehicles from current production was matched by the increase in sales of vehicles manufactured in the previous year. This indicates the occurrence of a different seasonality.

Figure 3 The seasonal component of the number of newly registered passenger vehicles in Poland in 2008–2018



Source: Own study based on CEPIK data

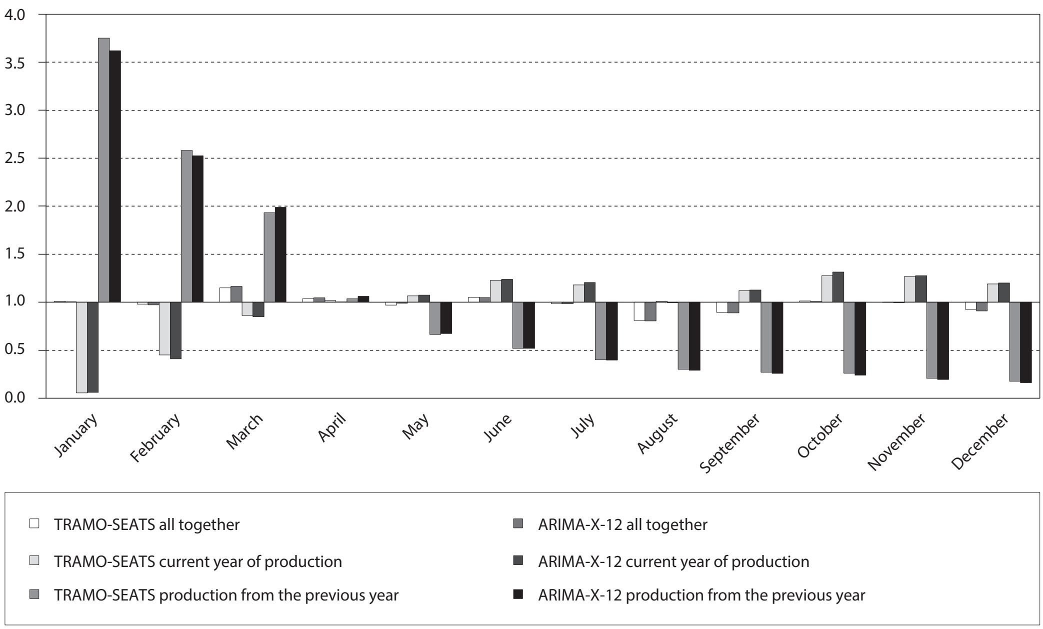
Occurrence of seasonality was confirmed by statistical tests (see Table 1). The results of both the Friedman test and the Kruskal-Wallis test indicate statistically significant differences between the average levels of the number of newly registered vehicles in individual months. This applies to both the general number of newly registered passenger vehicles as well as the sale of vehicles manufactured in the current year and produced in the previous year. The results of the seasonality test indicate, on the other hand, the evolution of the seasonal pattern over time, in the case of the total number of newly registered passenger vehicles – both for the value of the seasonal component extracted by the TRAMO-SEATS and ARIMA-X-12 procedures. Changes to the seasonal pattern were not found for newly registered vehicles manufactured in the previous year. As for the sale of passenger vehicles from the current year, different results were obtained using TRAMO-SEATS and ARIMA-X-12 procedures.

Table 1 Results of statistical tests of seasonality of newly registered passenger vehicles in Poland in 2008–2018

T – TRAMO-SEATS *A* – ARIMA-X-12		Friedman Test	Kruskal-Wallis Test	Evolutive Seasonality Test
		F-statistic		
Newly registered vehicles in total	T	43.4023***	107.2579***	0.2755
	A	101.4514***	62.3837***	0.1383
Newly registered vehicles from current production	T	144.5800***	113.5749***	4.2080***
	A	154.3249***	122.5364***	1.5453
Newly registered vehicles from the previous year's production	T	1 028.9032***	124.9036***	3.6385***
	A	996.4687***	121.4024***	5.7494***

Note: Statistical significance at the level *** –0.001, ** – 0.05, * – 0.10.
Source: Own study based on CEPIK data

Figure 4 The average deviation of the seasonal component of newly registered passenger cars in Poland in the years 2008–2018



Source: Own study based on CEPIK data

Analyzing the average monthly variations of the seasonal component (see Figure 4), it is possible to confirm the previously observed regularity, that increases in the registration of new vehicles produced in the current year correspond to the decrease in the number of passenger vehicle registrations produced in the previous year. Seasonal variations of new vehicle registrations were the biggest for vehicles manufactured in the previous year. Their sales increased the most (in relation to the trend) at the beginning of the year, when sales of yearbooks continue (first quarter) and then decreased until the end of the year. In the case of sales of new vehicles manufactured in the current year, sales increases were recorded in the second half of the year, and declines in the first quarter. However, variations from the trend in this case were much smaller. Seasonal fluctuations in the sales of vehicles from current production and production from the previous year translated into the seasonality of sales of new vehicles in total, characterized by an increase in sales in March and a drop in sales in particular in the months of August, September and December.

3.3 Outliers

The development of the studied phenomenon in time is not only a trend and seasonal fluctuations, but also outliers will result from random events, legislation, etc. TRAMO-SEATS and ARIMA-X-12 procedures allow for seasoning time series and identifying outliers along with the determination of their character. The largest amount of outliers was characterized by the sale of new vehicles manufactured in the current year. Using the TRAMO-SEATS procedure, 8 outliers were distinguished, their character and the moment of occurrence being different. Only for September 2018 the results of both procedures indicated a temporary change. It is difficult to determine which of the procedures distinguished outliers better. At present, it is difficult to clearly determine which events influenced individual outliers (Were these changes legislation also changes related to exhaust gas reduction standards?).

The number of newly registered passenger vehicles manufactured in the previous year was characterized by two unusual observations – in the case of the TRAMO-SEATS procedure and one – in the case of the ARIMA-X-12 procedure. The results of both procedures indicated a change in the trend level in September 2018, which was already signaled in (see Figure 2). In July and August 2018 there was a significant increase in the number of newly registered vehicles, after which a decline was noted. The effect of the increase in sales was caused by the announcement of changes in legislation regulating the possibility of deducting in full the purchase costs (leasing) of vehicles by entrepreneurs. This caused that entrepreneurs wishing to take advantage of the existing possibility of deducting the full leasing installment for the purchased vehicle, accelerated their purchasing decisions.

Outliers identified for the number of newly registered passenger cars from current production and produced in the previous year, have not translated into identifying the same outliers for the total number of newly registered passenger vehicles. For this time series, a change in the trend level

Table 2 Outlier values of the number of newly registered passenger vehicles in Poland in 2008–2018

Date / Method *T* – TRAMO-SEATS *A* – ARIMA-X-12		Total number of vehicles	Production vehicles from the year under analysis	Production vehicles from the previous year
2008-01	T	-	AO	-
	A	-	-	-
2008-02	T	-	-	-
	A	-	LS	-

Table 2 (continuation)

Date / Method *T* **– TRAMO-SEATS** *A* **– ARIMA-X-12**		**Total number of vehicles**	**Production vehicles from the year under analysis**	**Production vehicles from the previous year**
2009-01	T	-	LS	-
	A	-	AO	-
2009-03	T	LS	-	-
	A	LS	LS	-
2010-01	T	-	-	-
	A	-	AO	-
2010-02	T	-	AO	-
	A	-	-	-
2011-01	T	-	AO	-
	A	-	-	-
2012-01	T	-	TC	-
	A	-	-	-
2014-04	T	-	-	AO
	A	-	-	-
2015-01	T	-	-	-
	A	-	TC	-
2016-01	T	-	AO	-
	A	-	-	-
2017-01	T	-	-	-
	A	-	AO	-
2018-01	T	-	TC	-
	A	-	-	-
2018-08	T	AO	-	-
	A	AO	AO	-
2018-09	T	TC	TC	LS
	A	TC	TC	LS

Note: LS – Level Shift, TC – Temporary Change, AO – Additive Outliers.
Source: Own elaboration based on CEPIK data

in March 2009 was distinguished, a one-off shock in August 2018 and a temporary change in September 2018. These values were identified using both seasonal methods procedures: TRAMO-SEATS and ARIMA-X-12. The values, which were identified for 2018, reflect the reaction of buyers to the announcement of changes in legislation limiting the possibility of including in the costs of business the full leasing installment for the purchase of a vehicle.

3.4 Calendar Effects

An extraordinary advantage of using the procedures TRAMO-SEATS and ARIMA-X-12 is the ability to identify the effects of the calendars. These effects include the Easter Effect, which in the case of the number of newly registered vehicles in Poland has not been identified. Analysis made by the authors, however, allowed to the Week Days Effect, which is also a calendar effect. It is interesting because the effect, although it concerns days of the week, is identifiable on the basis of monthly data. No daily data is needed to identify this effect, as both procedures are based on calendars with working days, fixed and movable holidays. Comparing the number of working days and counting individual days of the week in each of the analyzed months, a decision is made about the occurrence of the tested effect. In the case of the total number of newly registered vehicles, a statistically significant increase in sales on Tuesdays (assuming a significance level of 0.05) and on Fridays and statistically significantly reduced

Table 3 Week Days effect of the number of newly registered passenger vehicles in Poland in 2008–2018

Day of the week / Method ***T*** **– TRAMO-SEATS** ***A*** **– ARIMA-X-12**		**Total number of vehicles**	**Vehicles from production from the previous year**
		Parameter – Value	
Monday	T	–0.0193	0.0180
	A	–0.0193	0.0043
Tuesday	T	0.0285**	0.0159*
	A	0.0301**	0.0270*
Wednesday	T	0.0169	0.0087
	A	0.0147	0.0116
Thursday	T	0.0050	0.0115
	A	0.0050	0.0080
Friday	T	0.0233**	0.0121
	A	0.0232**	0.0117
Saturday	T	–0.0376***	–0.0354**
	A	–0.0369***	–0.0377**
Sunday	T	–0.0167	–0.0307**
	A	–0.0167	–0.0250*

Note: Statistical significance at the level *** – 0.001, ** – 0.05, * – 0.10.
Source: Own study based on CEPIK data

on Saturdays was identified. In the case of the number of newly registered vehicles from production from the previous year, a statistically significant reduced sale was observed for Saturdays and Sundays. The sale of new vehicles from current production was not characterized by the occurrence of the week days effect.

CONCLUSION

The authors of the article hypothesized that analyzing the sale of new vehicles in Poland in total, without taking into account the year of production of the vehicle, was incorrect. In order to verify the hypothesis they analyzed data on the number of newly registered passenger cars in Poland in general, and by the year registered vehicle. New vehicles, new ones are vehicles that have been registered for the first time. As new, vehicles registered and manufactured in the same year as well as vehicles registered for the first time but produced in the previous year were included. The monthly data necessary to verify the research hypothesis was obtained from the Polish Central Register of Vehicles and Drivers. The data covered the years 2008–2018. To verify the hypothesis, the TRAMO-SEATS and ARIMA-X-12 methods were used, allowing the distribution of time series for trends, seasonal, irregular and random components.

Based on the results of the time series decomposition, it was found out that the total number of new vehicle sales in Poland increased. The sales of new vehicles, produced in the current year and produced in the previous year also increased. Seasonality have confirmed its occurrence. However, the nature of seasonality was different. Sales of new cars produced this year were the smallest at the beginning of the year and increased towards the end of the year. This was the result of high vehicle prices at the beginning of the year. Available reports and studies indicate that this relationship also occurred in vehicle sales in January 2019 (PZPM, 2019). The sales of vehicles manufactured in the previous year were reversed. The number of sales of these vehicles was the highest at the beginning of each year and decreased with the passage of time, making the slightest deviation from the trend towards the end of the year. During this period, the year, the vehicles are sold at a reduced price. Experts estimate that depending on the model and brand, it can range from 4% to even 25%. Seasonal fluctuations in vehicle sales by year of production influenced the shape of seasonal variations of total vehicle sales, which was characterized by positive seasonal variations in March and negative in August and September. This change was affected, among others, by changes in legislation. In addition, outliers were found, with the lowest observed for the sales of vehicles manufactured in the previous year and the highest for the sale of vehicles from the current production. Performed analyzes also allowed to identify The Weekday Effect (on Tuesdays and Fridays). Increased vehicle sales occurred on Tuesdays, and decreased on weekends. This is also confirmed by the research made by other authors (Nolan, 2010).

The differences observed in the trends in sales of new passenger cars in Poland, produced in the year of registration and in the previous year, allow for a positive verification of the hypothesis. Sales of passenger vehicles in Poland should not be presented in general, but divided into the year of vehicle production.

Summing up the presentation, considerations regarding new vehicle sales trends in Poland do not fully exhaust the essence of the issue but provide the basis for further analysis and research in order to use statistical methods in the process of identifying determinants affecting the number of new vehicle sales in Poland as well as other in Eastern European countries such as Romania, Bulgaria and the Czech Republic.

References

ATUK, O. AND URAL, B. Seasonal Adjustment in Economic Time Series. *Discussion Paper No. 1*, Central Bank of Turkey, 2002.

BARTON, E., AL-SARRAY, B., CHRETIEN, S., JAGAN, K. Decomposition of Dynamical Signals in Jumps, Oscillatory Patterns, and Possible Outliers. *Mathematics*, 2018, 6, 124, pp. 1–13.

CABRERO, A. Seasonal Adjustment in Economic Time Series: *The Experience of the Banco de Espana*, 2000.

DAGUM, E. B. AND BIANCONCINI, S. Seasonal Adjustment Based on ARIMA Model Decomposition: TRAMO-SEATS. *Seasonal Adjustment Method and Real Time Trend-Cycle Estimation*, Switzerland, Springer, 2016, pp. 115–145.
DAGUM, E. B. *The X11 ARIMA Seasonal Adjustment Method.* Ottawa: Statistics Canada Catalog, 1980, No. 12-564E.
ENDERS, W. *Applied Econometric Time Series.* 3rd Ed. New York: Wiley, 2010, pp. 97–103.
FISCHER, B. *Decomposition of time series comparing different methods in theory and practice.* Eurostat, 1995.
FRIEDMAN, M. A. comparison of alternative tests of significance for the problem of m rankings. *Annals of Mathematical Statistics* 1940, 11, pp. 86–92.
GEORGE, B. AND SARIDAKIS, CH. Brand-name effects, segment defiance, and product characteristics of an integrated model of the car market. *Journal of Product & Brand Management*, 2009, Vol. 18, Issue 2, pp. 143–151.
GHYSELS. E. AND OSBORN, D. The Econometric Analysis of Seasonal Time Series. New York: *Cambridge University Press*, 2001, pp. 93–120.
GRUDKOWSKA, S. *Demetra+ User Manual, National Bank of Poland* [online]. Warsaw, 2011. <www.cros-portal.eu/sites/default/files/Demetra+%20User%20Manual.pdf>.
GRUDKOWSKA, S. AND PAŚNICKA, E. X-12 ARIMA and TRAMO / SEATS – empirical comparison of seasonal adjustment methods in the context of sample length. *Materials and Studies*, 2007, 220, pp. 11–16.
KLIMKOWSKA, R. Client on the passenger car market in Poland. *Problems of Management, Finance and Marketing*, 2012, 25, pp. 65–76.
KRUSKAL, W. H. AND WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 1952, 47, pp. 583–621.
KUDŁAK, R., KISIAŁA, R., GADZIŃSKI, J., DYBA, W., KOŁSUT, B., STRYJAKIEWICZ. T. Socio-economic and spatial conditions of demand for new cars in Poland. *Regional and Local Studies*, 2017, 2(68), pp. 119–139.
LIAN, W., HUANG, J. Z., MCELROY, T. Time Series seasonal adjustment using regularized singular value decomposition. *Journal of Business & Economic Statistics*, 2018, 8, pp. 45–50.
MARAVALL, A., LOPEZ-PAVON, R., PEREZ-CANETE, D. Reliability of the Automatic Identification of ARIMA Model in Program TRAMO. *Empirical Economic and Financial Research*, 2014, pp. 105–122.
MICHALAK, J. AND MERŁO, P. The impact of promotional activities on the behavior of car buyers. *Internal trade*, 2015, No. 3(356), Year 61(LXI), pp. 141–151.
MUIRHEAD, C. Distinguishing Outlier Types in Time Series. *Journal of Royal Statistical Society, Series B (Methodological)*, 1986, Vol. 48, No. 1, pp. 1–2.
MUTRYNOWSKI, T. The passenger car market in Poland. *Quarterly of Business*, 2015, 1, pp. 191–201.
NOLAN, A. A dynamic analysis of household car ownership. *Transportation Research Part A: Policy and Practice*, 2010, 44(6), pp. 446–455.
POLISH ASSOCIATION OF THE AUTOMOTIVE INDUSTRY (PZPM). *Automotive industry* [online]. Warsaw, 2019. <https://www.pzpm.org.pl/en/Automotivemarket/Reports/PZPM-Automotive-Industry-Report-2019-2020>.
PRIETO, M. AND CAEMMERER, B. An exploration of factors influencing car purchasing decisions. *International Journal of Retail & Distribution Management*, 2013, Vol. 41, Issue 10, pp. 738–736.
SAX, CH. AND EDDELBUETTEL, D. Seasonal Adjustment by X-13ARIMA-SEATS in R. *Journal of Statistical Software*, December 2018, Vol. 87, Issue 11, pp. 2–8.
SOŁTYSIAK, M. Consumer purchase preferences on the car market. *Modern Management Review. Research Journal*, 2015, Vol. XX, 22, pp. 3–4.
STRYJAKIEWICZ, T., KUDŁAK, R., GADZIŃSKI, J., KOŁSUT, B., DYBA, W., KISIAŁA, W. Time-space analysis of the new car market in Poland. *Works of the Commission of Industrial Geography of the Polish Geographical Society*, 2017, 31(3), pp. 64–79.
URBAN, D. J., HOFFER, G. E., PRATT, M. D. The used-vehicle superstore: a flawed concept. *Journal of Consumer Marketing*, 2000, Vol. 17, 5, pp. 420–438.

Papers
We publish articles focused at theoretical and applied statistics, mathematical and statistical methods, conception of official (state) statistics, statistical education, applied economics and econometrics, economic, social and environmental analyses, economic indicators, social and environmental issues in terms of statistics or economics, and regional development issues.

The journal of *Statistika* has the following sections:

The *Analyses* section publishes high quality, complex, and advanced analyses based on the official statistics data focused on economic, environmental, and social spheres. Papers shall have up to 12 000 words or up to twenty (20) 1.5-spaced pages.

Discussion brings the opportunity to openly discuss the current or more general statistical or economic issues, in short what the authors would like to contribute to the scientific debate. Contribution shall have up to 6 000 words or up to 10 1.5-spaced pages.

The *Methodology* section gives space for the discussion on potential approaches to the statistical description of social, economic, and environmental phenomena, development of indicators, estimation issues, etc. Papers shall have up to 12 000 words or up to twenty (20) 1.5-spaced pages.

Consultation contains papers focused primarily on new perspectives or innovative approaches in statistics or economics about which the authors would like to inform the professional public. Consultation shall have up to 6 000 words or up to 10 1.5-spaced pages.

The *Book Review* section brings reviews of recent books in the fieled of the official statistics. Reviews shall have up to 600 words or one (1) 1.5-spaced page.

The *Information* section includes informative (descriptive) texts, information on latest publications (issued not only by the CZSO), recent and upcoming scientific conferences. Recommended range of information is 6 000 words or up to 10 1.5-spaced pages.

Language
The submission language is English only. Authors are expected to refer to a native language speaker in case they are not sure of language quality of their papers.

Recommended Paper Structure
Title (e.g. On Laconic and Informative Titles) — Authors and Contacts — Abstract (max. 160 words) — Keywords (max. 6 words / phrases) — JEL classification code — Introduction — (chapters: 1, 2, …) — Conclusion — Acknowledgments — References — Annex — Tables and Figures (for print, for review process in the text)

Authors and Contacts
Rudolf Novak*, Institution Name, Street, City, Country
Jonathan Davis, Institution Name, Street, City, Country
* Corresponding author: e-mail: rudolf.novak@domain-name.cz, phone: (+420) 111 222 333

Main Text Format
Times 12 (main text), 1.5 spacing between lines. Page numbers in the lower right-hand corner. *Italics* can be used in the text if necessary. *Do not* use **bold** or <u>underline</u> in the text. Paper parts numbering: 1, 1.1, 1.2, etc.

Headings
1 FIRST-LEVEL HEADING (Times New Roman 12, bold)
1.1 Second-level heading (Times New Roman 12, bold)
1.1.1 Third-level heading (Times New Roman 12, bold italic)

Footnotes
Footnotes should be used sparingly. Do not use endnotes. Do not use footnotes for citing references.

References in the Text
Place reference in the text enclosing authors' names and the year of the reference, e.g. "White (2009) points out that…", "… recent literature (Atkinson et Black, 2010a, 2010b, 2011; Chase et al., 2011, pp. 12–14) conclude…". Note the use of alphabetical order. Include page numbers if appropriate.

List of References
Arrange list of references alphabetically. Use the following reference styles: [for a book] HICKS, J. *Value and Capital: An inquiry into some fundamental principles of economic theory*. 1st Ed. Oxford: Clarendon Press, 1939. [for chapter in an edited book] DASGUPTA, P. et al. Intergenerational Equity, Social Discount Rates and Global Warming. In: PORTNEY, P. AND WEYANT, J., eds. *Discounting and Intergenerational Equity*. Washington, D.C.: Resources for the Future, 1999. [for a journal] HRONOVÁ, S., HINDLS, R., ČABLA, A. Conjunctural Evolution of the Czech Economy. *Statistika: Statistics and Economy Journal*, 2011, 3 (September), pp. 4–17. [for an online source] CZECH COAL. *Annual Report and Financial Statement 2007* [online]. Prague: Czech Coal, 2008. [cit. 20.9.2008]. <http://www.czechcoal.cz/cs/ur/zprava/ur2007cz.pdf>.

Tables
Provide each table on a separate page. Indicate position of the table by placing in the text "<u>insert Table 1 about here</u>". Number tables in the order of appearance Table 1, Table 2, etc. Each table should be titled (e.g. Table 1 Self-explanatory title). Refer to tables using their numbers (e.g. see Table 1, Table A1 in the Annex). Try to break one large table into several smaller tables, whenever possible. Separate thousands with a space (e.g. 1 528 000) and decimal points with a dot (e.g. 1.0). Specify the data source below the tables.

Figures
Figure is any graphical object other than table. Attach each figure as a separate file. Indicate position of the figure by placing in the text "<u>insert Figure 1 about here</u>". Number figures in the order of appearance Figure 1, Figure 2, etc. Each figure should be titled (e.g. Figure 1 Self-explanatory title). Refer to figures using their numbers (e.g. see Figure 1, Figure A1 in the Annex).

Figures should be accompanied by the *.xls, *.xlsx table with the source data. Please provide cartograms in the vector format. Other graphic objects should be provided in *.tif, *.jpg, *.eps formats. Do not supply low-resolution files optimized for the screen use. Specify the source below the figures.

Formulas
Formulas should be prepared in formula editor in the same text format (Times 12) as the main text and numbered.

Paper Submission
Please email your papers in *.doc, *.docx or *.pdf formats to statistika.journal@czso.cz. All papers are subject to double-blind peer review procedure. Articles for the review process are accepted continuously and may contain tables and figures in the text (for final graphical typesetting must be supplied separately as specified in the instructions above). Please be informed about our Publication Ethics rules (i.e. authors responsibilities) published at: http://www.czso.cz/statistika_journal.

Managing Editor: Jiří Novotný
phone: (+420) 274 054 299
fax: (+420) 274 052 133
e-mail: statistika.journal@czso.cz
web: www.czso.cz/statistika_journal
address: Czech Statistical Office | Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscription price (4 issues yearly)
CZK 372 (incl. postage) for the Czech Republic,
EUR 117 or USD 174 (incl. postage) for other countries.
Printed copies can be bought at the Publications Shop of the Czech Statistical Office (CZK 66 per copy).
address: Na padesátém 81 | 100 82 Prague 10 | Czech Republic

Subscriptions and orders
MYRIS TRADE, s. r. o.
P. O. BOX 2 | 142 01 Prague 4 | Czech Republic
phone: (+420) 234 035 200,
fax: (+420) 234 035 207
e-mail: myris@myris.cz

Design: Toman Design
Layout: Ondřej Pazdera
Typesetting: Družstvo TISKOGRAF, David Hošek
Print: Czech Statistical Office

100th year of the series of professional statistics and economy journals of the State Statistical Service in the Czech Republic: ***Statistika*** (since 1964), ***Statistika a kontrola*** (1962–1963), ***Statistický obzor*** (1931–1961) and ***Československý statistický věstník*** (1920–1930).

Published by the Czech Statistical Office
ISSN 1804-8765 (Online)
ISSN 0322-788X (Print)
Reg. MK CR E 4684