STATISTIKA

STATISTICS
AND ECONOMY
JOURNAL

VOL. **101** (1) 2021

CZECH
STATISTICAL
OFFICE

CONTENTS

ANALYSES

5 **Luboš Marek, Richard Hindls, Stanislava Hronová**
Non-Stochastic Argumentation in Predicting Economic Indices

24 **Agnieszka Sompolska-Rzechuła**
Selection of the Optimal Way of Linear Ordering of Objects: Case of Sustainable Development in EU Countries

37 **Hanen Ragoubi, Zouheir Mighri**
Environmental Kuznets Curvefor CO_2 Emissions in Middle-Income Countries: a Dynamic Spatial Panel Data Analysis

66 **Zulfiya Ibragimova, Marina Frants**
Measuring Inequality of Opportunity: Does Inequality Index Matter?

91 **Habibeh Sherafatmand, Ali Akbar Baghestany**
Dynamics Almost Ideal: Demand System Application of Kalman Filter

101 **Cristina Boboc, Magdalena Ciobanu**
Modus Operandi of Actors Involved in the Illicit Tobacco Trade in EU Countries

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

Non-Stochastic Argumentation in Predicting Economic Indices

Luboš Marek[1] | *Prague University of Economics and Business, Prague, Czech Republic*
Richard Hindls[2] | *Prague University of Economics and Business, Prague, Czech Republic*
Stanislava Hronová[3] | *Prague University of Economics and Business, Prague, Czech Republic*

Abstract

This paper studies the use of statistical prognostics in predictions of short-term year-to-year evolution of GDP and other aggregate indices of the national accounts. It shows the utilisation of a non-stochastic prediction range to be used as a prediction tool that, to a certain extent, overcomes the validity of the *ceteris paribus* principle, on which most of the currently used stochastic approaches are based. The non-stochastic range is a resultant outcome of a wide assortment of time-series models; at the same time, a point forecast for short-term evolution is derived from the said assortment of models. We illustrate our methodology on a year-to-year evolution of GDP indices in France as a time series with a sufficiently large number of observations.

Keywords	***JEL code***
Statistical prognostics, non-stochastic point forecast, non-stochastic prediction range, GDP	*C10, C53*

INTRODUCTION

The most significant economic indices that sensitively respond to the prevailing economic climate include, first and foremost, the gross domestic product (hereinafter GDP), but also other related aggregate indices of the national accounts; in particular, final consumption expenditure, gross capital formation, and exports and imports of goods and services. Monitoring these values statistically not only provides information on the current situation of the national economy, but can also be used in analysing its evolution as a basis for deriving short- and medium-term predictions.

Regarding the subsequent utilisation of such data for the purposes of estimating the performance of the economy, as well as providing a basis for creating the state budget, particular importance are predictions of short- and medium-term evolution of the said indices. The usual methods utilised in economically developed countries when estimating the GDP evolution for two to three years ahead are mainly based on predictions put forth by relevant expert groups, combined with econometric and statistical models. When creating state budgets, regression model approaches are also employed; either the concept of extrapolating the prevailing trend, or deriving the national accounts' aggregates

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

from a tightly related index that is already known, is closely monitored by the respective statistical office, and can be viewed as a suitable "prompt" to reveal the anticipated behaviour of the aggregates to be estimated.

The results achieved in this area entitle us to conjure that, under such circumstances, certain statistical methods become very important; namely those which are capable of providing the necessary information about the expected GDP evolution. The techniques for obtaining such predictions undoubtedly include methods used in time series analysis. Developments in statistical forecasting and its applications in economics extended the range of tools that can be used in macroeconomic decision-making. In the present paper, we study one of such approaches, which can be distinguished by a high success rate in short- and medium-term predictions.

1 METHODS OF TIME SERIES ANALYSIS AND PREDICTION PROCEDURES STUDIED IN THE LITERATURE

Approaches that employ time series models in predicting a given index occur quite frequently in the literature. Some of these approaches have found their ways to being used in the practice of economics, for example, in the procedures used by governments and other authorities; others have not proved their worth. To a large extent, the success rate of a prediction model is implied by that model's ability to overcome difficulties related to the *ceteris paribus* principle (saying that the future will, under unchanged circumstances, be a continuation of the past). In the reality we encounter in economics, the *ceteris paribus* principle can hardly be maintained; or rather, it is nearly completely unmaintainable in practice. As a matter of fact, economic processes are subject to many different types of interventions, whether legislative (therefore non-stochastic, such as amendments to tax and other laws) or natural (such as the stages of the economic cycle, etc.). Frequently we encounter combinations and even mutual interactions between stochastic and non-stochastic interventions. All such aspects de facto deny the *ceteris paribus* principle. That is why we often encounter in the literature routines and tools aimed at overcoming that extrapolation principle, which is too unrealistic in practice.

Of course, we can find in the literature a great number of suggestions for utilising statistical techniques in predictions. The foundations for extrapolation techniques were laid by many authors and, even though some of their concepts date back quite a few years, they have not been abandoned. Let us mention, as just a few of many possible examples, the monographs by Theil (1966), Granger and Newbold (1986), and Pankratz (1991); in the Czech Republic it was Cipra (1986). We can generally observe that the use of non-stochastic deterministic approaches is less frequent than those based on probability theory and entropy. Even if a non-stochastic concept is employed, it often represents a certain modification of brainstorming methods. This approach is not new, as a matter of fact; its first occurrence dates back to the post-World-Word-II decade; for example, Osborn (1963) described a creative group technique as a specific method for predicting and decision-making. We will go in this direction below, but we will mutually confront not personal opinions, but results of potentially acceptable models of time series.

The developments go on; newer – and, in sense, perhaps more efficient – methods are sought to mutually intertwine the deterministic and stochastic principles. The present paper is also an attempt at an original combination of deterministic and stochastic concepts. We can take lessons from the literature, in which different non-traditional approaches can be encountered. Let us recall a recent concept published in Lui (2020), which provides an in-detail study of the relationship between deterministic and stochastic-based interval predictions and attempts at bridging the gap between them with the aid of a certain hybrid approach. An Israeli geophysicist Eppelbaum (2013) derived estimates in the predictive model so that they were, predominantly, highly correlated with historic data, i.e., a reference variable with correlated evolution in time. Yachao et al. (2016) brought deterministic and probabilistic interval predictions (namely, for short-term prediction of electricity generation from wind) based on a decomposition of the variation mode.

A list of examples found in the literature indicates that the applications of the proposed techniques are, to a great extent, illustrated on data coming from the natural sciences. Economics is, however, a social science; its phenomenology is therefore completely different from that of the natural sciences. Moreover, it is marked with the existence of extensive behavioural elements. In this respect, the range of data-based experiments found in the literature is rather less abundant.

That is why we will apply our approach to an economic time series. Conceptual inspiration can be found even here: in order to gradually improve the quality of prediction, Tribbia and Baumhefner (2013) recommend that the facts of the situation be examined in general at first, and then set up the goal of the particular prediction from the phenomenological point of view. Later on, the phenomenological and non-deterministic aspects of the prediction should be intertwined at the given time horizon. We will try below to follow a similar line of thought: we will employ an assortment of models, thus introducing into our prediction uncertainty pertaining to each of those models. Subsequently we will reduce the uncertainty by deriving a compound solution based on all of the primarily used models. One of the main characteristics of such a prediction will be the fact that the uncertainty pertaining to each of the models will be reduced in the compound prediction range.

As can also be seen in the literature, the common denominator of such approaches is to derive a plausible prediction and, at the same time, to eliminate to the maximum possible extent the non-realistic *ceteris paribus* assumption. The main point here is to apply stochastic modelling in as wide a sense as possible, directed towards reflecting the effects of intervention attacks of diverse origin and nature. The weakest point of a similar approach is that the resulting prediction interval is usually too broad; consequently, its usefulness for the decision-making processes is dubious. Because of that, we will suggest a procedure below that should provide a more useful interval of the prediction.

2 TRADITIONAL REGRESSION APPROACH

Now we will derive the statistical prediction for the year-to-year evolution of any index. This technique is based on statistical tools inherent to analysis of one-dimensional time series. Let us begin with an overview of necessary basic notions.

The traditional regression-based approach to extrapolations of a time series y_t, $t = 1, 2, \ldots, n$, where n is the number of the (past) observations of the time series, is generally formulated as a requirement to begin the prediction with a suitable estimate for the future values y_{n+i}, $i = 1, 2, \ldots, N$, where N is a selected positive integer characterising the length of the prediction. Of course, we can resolve this prediction problem with the aid of a model supposedly governing the behaviour of the relevant time series, y_t. This is the so-called *point prediction*.

It is, however, a well-known fact that a point prediction is too authoritative about the future evolution of the index to be predicted. It is difficult to find a specific reason why a particular model should be singled out (incidentally, good results in interpolation – i.e., modelling the past evolution of the time series – need not be a sufficient guarantee for a good prediction). Moreover, the point prediction is utterly incapable of overcoming the *ceteris paribus* condition (of the future being a continuation of the past under circumstances that otherwise remain unchanged). This problem is viewed as very serious in economics. From the factual viewpoint, economic indices are very unstable variables; assuming that factors affecting their future evolution remain unchanged is in its substance absurd and stands in a deep contradiction with the substance of economics as a social and political science.

There are techniques which try to cope with the authoritative character of the point prediction; such techniques lead to interval predictions of future developments. In a vast majority of instances, such techniques are based on stochastically formulated and interpreted predictions with respect to a pre-set level of confidence. Another problem arises at this moment. In addition to the fact that our prediction interval stems from a point prediction, encumbered with all its weaknesses (as mentioned above), two additional requirements must be met: the confidence level of the prediction must be sufficiently high (in practice not smaller than 90%); and, at the same time, the prediction interval must be reasonably narrow. Unless both these requirements are met, the prediction interval will be more or less useless in economic practice. It is well known that requirements for a high confidence level and a narrow prediction interval are in a mutual conflict. Not even an interval prediction is able to overcome the *ceteris paribus* principle (which will also be seen from our application below). Setting up an interval, we create a funnel trough which additional possibilities are "drawn" into the prediction. The usual price to pay for this aspect is an excessive breadth of possible values, which is difficult to interpret.

If we construct an estimate for the future value y_{n+i} at time $(n + i)$ as:

$$P_{n+i} - \Delta < y_{n+i} < P_{n+i} + \Delta \,, \tag{1}$$

where P_{n+i} is the point prediction for the time period $n + i$ as estimated by any model, and Δ is the admissible error of the prediction; the latter depends on the selected confidence level of the prediction interval, as well as on the number and variability of the real observations from the past, y_t. Inequality (1) is valid for the pre-set level of confidence (that is, with a certain – sufficiently large – value of probability); that is why we call it *stochastic interval prediction*. An interval prediction defined in this way is symmetric. Since the stochastic (probabilistic) interval prediction stems from three point predictions for which inequality (1) holds, the conditions determining the success of the interval prediction is the quality of the original point prediction P_{n+i}. The actual value of the index to be predicted for time $n + i$, i.e., y_{n+i}, is "enveloped" by the stochastic prediction interval.

The particular form of the admissible error (that is, the stochastic confidence interval) of the prediction can, for example, be described by the following Formula for the linear trend model – cf. Cipra (1986):

$$\Delta = T_{1-\alpha/2}[n-2] \cdot s \cdot f_i \,, \tag{1a}$$

where T is the quantile of Student's distribution with $n - 2$ degrees of freedom, n is the number of the past observations in the time series; the length (horizon) of the prediction is $i = 1, 2, ..., N$ (we set $N = 2$ for the purposes of macroeconomic prediction in our case for practical reasons),

$$s = \sqrt{\frac{\sum_{t=1}^{n} y_t^2 - \sum_{t=1}^{n} \hat{y}_t^2}{n-2}} \,,$$

$$f_i = \sqrt{1 + \frac{1}{n} + \frac{(N - \overline{t}\,)^2}{\sum_{t=1}^{n} t^2 - n \cdot \overline{t}^2}}$$

and $\hat{y}_t$ = the model estimate for the value of the time series y_t, $t = 1, 2, \ldots, n$.

There is, however, one significant drawback encumbering the construction of a stochastically argued prediction interval: even if the number of observations is sufficient and the selected model has a good quality (from the interpolation point of view), the prediction interval on an adequate level of confidence is too broad. This is bad news regarding the practical requirements we have put on predictions.

3 CONSTRUCTION OF NON-STOCHASTICALLY ARGUED PREDICTION RANGE

Experience with predicting values of economic indices thus generally indicates that the resolution of the prediction problem needs more than a purely "regression-like" formulated short-term prediction. We have already mentioned several reasons for misgivings pertaining to a traditional regression concept to be applied in predicting macroeconomic evolution indices: a large (and, consequently, conflicting) assortment of point predictions for the same index; a practically useless breadth of the stochastically constructed interval predictions according to Formula (1) (or (1a) for the linear trend); and the traditional statistical assumptions are usually not valid for real economic data; etc.

There exists a certain generally positive outcome regarding these user-unfriendly situations; they, to a considerable extent, determine the acceptance of statistical predictions in the areas of macroeconomic studies, conjunctural surveys, etc. We must take into account a non-traditional approach to setting up our predictions, an approach within which the prediction interval (1) is argued not in a probabilistic, traditionally regression-driven way, but as follows:

$$P_{n+i} - \delta_1 < y_{n+i} < P_{n+i} + \delta_2, \tag{2}$$

where the prediction range $P_{n+i} - \delta_1$ and $P_{n+i} + \delta_2$, $i = 1, 2, ..., N$, is understood as a resultant outcome of different point predictions on the basis of a large number of factually admissible models for the past behaviour of the respective time series (see Formulae 3 and 4 for the construction of deviations δ_1 and δ_2). An interval prediction defined in this way will not be symmetric with respect to y_{n+i}.

Determination of the prediction according to general Formula (2) will be called a **non-stochastic prediction range**. Such a range is based on the idea that we can set up several (or many) models for a given time series, which may all properly describe the past behaviour of the respective series and be admissible from the factual and formal points of view. We must still keep in mind this important fact: a model providing a good-quality description of the past behaviour of a time series need not provide a good prediction of its future behaviour, due to possible changes in the conditions to which the time series is subject.

A strong point of the non-stochastic prediction range is its universal nature. In fact, in addition to stochastic models we can also utilise econometric models and combinations of both types for the construction of that range. The only condition is that, for the primary models, factual and formal admissibility should be guaranteed for the underlying problem; and there should be a realistic option to set up a higher number of such models. Extensive use of software enables us to set up many models, compute estimates in them, and compare their prediction outcomes.

Conceptually, such a construction of a non-stochastic prediction range is similar to the usual economic practice, in which different opinions concerning the future evolution are confronted with each other. Here we confront "opinions" ensuing from different models for the underlying time series. The prognosis is then a resultant outcome of all such "components". We will illustrate the construction of a non-stochastic prediction range on an example of year-to-year GDP indices time series in France.

The construction itself goes in two steps; this approach can be understood as an algorithm and programmed as such.

Step 1: *Selection of and estimates in models*

First we set up a certain high number of statistical models of adequate quality and with good factual interpretations (in our case, models for the time evolution of the year-to-year GDP indices time series in France; for the sake of clarity, we sum up these models in Table A2 in the Appendix, where the estimated model shapes, denoted by M1, M2, …, through M22, are also shown). Our basic idea here is that each model represents an opinion ("winnowing our facts") – in an analogy to the "normal human thinking under uncertainty", which may take into account several admissible variants. Our prediction task can certainly be classified as such a situation. The selection of the models is the key stage of the prediction process, because these models lay the foundations for the prediction range. Hence we must responsibly consider factual and formal statistical viewpoints during the selection.

In our experience, it is purposeful to select between 10 and 30 models from different categories, such as smooth analytical trend functions for different types of exponential fitting, from the Box-Jenkins methodology, etc.; our criterion should be based on the quality and interpretation of the models. The character of the index to be predicted should also be reflected. No other restrictions should be considered. Of course, we must always keep in mind that a primary model that is "interpolation-good" is no guarantee for a good quality of prediction due to the unrealistic assumptions hidden in the *ceteris paribus* principle.

Step 2: *Construction of prediction range and point prediction*
Let us derive three values for each year in each of the models set up and tested according to Step 1 (in our case, M1, M2, …, and M22: namely, point prediction P_{n+i} (that is, a year-to-year indices of the GDP growth for the years 2018 and 2019); the stochastic lower bound for the prediction, Lower 95.0% Limit; and the stochastic upper bound for the prediction, Upper 95.0% Limit, both bounds at the 95% confidence level. For example, the mentioned three values will look as follows for Model M1:

Table 1 Model M1. Random walk with a drift

Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
2018	1.022 060	0.984 296	1.059 830
2019	1.021 130	0.967 717	1.074 540

Note: Similarly for the remaining models, M2, M3, …, through M22, cf. Table A2 in the Appendix.
Source: Authors' own calculations

Having selected 22 models, we can see that we would get 22 point predictions P_{n+i} for each of the years to be predicted, i.e., 2018 and 2019; and 22 stochastic interval predictions at the same time. Each of these predictions can be viewed as relevant, but they are numerically different from each other. This aspect is rather indeterminate with regard to subsequent considerations. That is why we will now show a way to arrive at non-stochastic predictions, while adequately using the specific information provided by each of the models we computed.

As regards point predictions P_{n+i}, we will derive a sole aggregated value of the point prediction based on all 22 models M1, M2, …, M22; namely, we take the average of those 22 values. In a way, this approach is an analogy to colloquia, in which opinions of relevant members are comprised. Here the 22 models stand for such members, and the result is described in the part 4.2.

When setting up the interval prediction corresponding to Formula (2), we will first derive a sole aggregated lower bound δ_1 from the Lower 95.0% Limit values of all models M1, M2, …, M22; namely, we will take their maximum, that is,

$$\delta_1 = \max\{\text{Lower 95.0\% Limit of M1, M2, …, M22}\}. \tag{3}$$

In a similar way, we then derive a sole aggregated upper bound δ_2 from the Upper 95.0% Limit values as their minimum:

$$\delta_2 = \min\{\text{Upper 95.0\% Limit of M1, M2, …, M22}\}. \tag{4}$$

In other words, this particular application of Formula (2) is based on the maximum stochastic lower bound and the minimum stochastic upper bound of the traditional prediction intervals (our result is again described in the part 4.2). As already mentioned above, an interval prediction defined in this way can no longer be symmetric, in contrast with the stochastic interval prediction.

4 APPLICATION OF THE MODEL

In order to verify our model of non-stochastic prediction range, we have chosen the time series of year-to-year GDP indices in France in the period from 1950 to 2019. This is a highly aggregated index occurring on French national accounts. Our main reason for this selection is the long time series we can use to illustrate our approach.

The developments in the French national economy over the course of nearly 70 years represent a diverse mixture of various influences (post-war recovery, cold war, adoption of euro, etc.), internal political decisions (taxation, monetary/fiscal interventions, changes in the play of political forces), as well as international economic interventions (oil crisis, local military conflicts, etc.).

4.1 Economic developments in France since the end of World War II

As already explained, we will make use of the French GDP data to illustrate our method of setting up the prediction.[4] A factual description of the economic developments in France since the late 1940s until this date is important to help us provide effective interpretations and utilise the derived estimates and predictions.

The post-war period of the so-called French Fourth Republic (1945–1958) was characterised by relatively high economic growth, a high inflation rate and low unemployment rate. In the beginning of the post-war period, the Marshall Plan played the main role in the economic recovery. On the other hand, the Fourth Republic was politically rather volatile (the Prime Minister was changed 28 times in 12 years); this aspect did not contribute to the country's economic stability.

Figure 1 Year-to-year GDP indices, France

Source: <*www.insee.fr*>

[4] The input data – year-to-year French GDP indices – is depicted in Figure 1 and numerically listed in Table A1 in the Appendix.

Due to the destruction prevailing after the war, the government undertook the task to recover the French economy. Electric power plants, the coal industry, big bank and insurance companies, Renault, aerospace industry, etc., were nationalised. The nationalised sector was an important tool for implementing the government's economic policies. The first Monnet Plan was begun in 1947 (a programme for investments into the key industries and reduction of the economic dependency on abroad). The second Monnet Plan (1954–1957) was focused on public investments and ensuring higher productivity in material and human resources. A higher GDP growth rate in 1950 brought about a higher inflation rate. The subsequent slowdown in the economic performance after 1950 meant a gradual return to the normal production potential; adoption of a deflation policy also contributed to the slowdown. From 1954, the GDP regular growth by more than 5% a year prevailed. Another slowdown and an onerous economic situation in the late 1950s were mainly caused by the colonial war in Algeria, the imminent civil war, and growing expenses on nuclear armament. Under the aggravated conditions, a new Constitution was written, the Fifth Republic was born and Charles de Gaulle was elected President of France for seven years in the office. In addition to the war in Algeria, another reason for the economic slowdown was the overall decline of the conjuncture.

After the new Constitution was approved by a referendum, a new plan was developed to put the French economy back on its feet, and substantial savings reduced the budget deficit but also the expenses incurred on the social care. The economic measure brought the French economy to a recovery in the early 1960s and the economic unbalance was eliminated. A faster economic growth stated in 1960, also in consequence of franc devaluation, which favoured sales of French goods on foreign markets.

In the mid-1960s, the French GDP dynamics went down because of decreasing wages and consumption. This situation later (in 1968) caused extensive strikes because employees' economic standing was worsening. A strike lasted several weeks and the economy was paralysed; the consequent drop in the production led to the smallest GDP growth rate value in the entire 1960s (4.5% year-to-year in 1968). Charles de Gaulle resigned his presidency at noon, April 28, 1969. The new political elite, headed by President Pompidou, adopted a new plan to stabilise the economy – the primary aspects included the support to exports and restriction on imports, reduction of the state budget deficit, and a substantial increase in taxes. Nonetheless, France did not avoid the financial crisis connected with the termination of the Bretton Woods system, which caused a monetary crisis of the franc.

The GDP growth had a descending trend after 1973. The French economy was hit by a recession and went to the bottom in 1975 (a deep drop caused by the Yom Kippur War and the subsequent oil crisis was accompanied by a high level of inflation and unemployment rates). In 1976, a new plan aimed at stabilising the franc and recovering the budget balance was adopted. The plan worked as expected and the inflation was temporarily stopped. The balance was recovered, but the high unemployment rate continued to prevail. The second oil crisis and a return of the high inflation rate (nearly 14% in 1980) were negative factors. All these facts and other economic-crisis phenomena affected the presidential elections in 1981. The political establishment increased the minimum wages, the lowest pensions, and the family benefits. Nearly all banks, insurance companies and key industries were nationalised. In the late 1980s, the state-intervention policies were being abandoned, with a continuing liberalisation of prices, and decreasing taxes. The average economic growth (measured by the GDP growth rate) achieved more than 3% per year.

The international situation (the Gulf War) with the world trade cooling off, growing oil prices, as well as procrastination of the necessary reforms, were the factors that caused another slowdown in the French growth rate. The economic developments in France after 1990 are characterised by stable year-to-year GDP growth, at an average rate of about 1.5%, a low inflation rate and consistently high unemployment rate. A critical point occurred at the 1992/1993 year break (the recession began in autumn of 1992 and was relatively short).

The beginning of the new millennium was marked with attenuation of the dynamics, caused by the drop in performance of the American economy and stock markets, while the oil prices were growing. A positive turn came in 2004, but the worldwide economic crisis of the years 2008 and 2009 hit the economies of many countries heavily, including that of France. The French GDP was going down for five consecutive calendar quarters, the general government deficit was growing, and the drop in demand attenuated the price growth (it is called the 2008/2009 deflation). The consequences of the global crisis were not so bad for France as in the majority of big European economies thanks to the growing consumption by households and the fiscal stimuli for exports together with the moderate devaluation of the euro. The French economy recovered from the recession in the third quarter of 2009. The French economy again stagnated in 2012 (the GDP growth rate was 0.3% in that year), and the recovery was coming slowly. The year-to-year GDP growth got above 2% as late as in 2017. But it went back down to 1.7% in 2018 and to 1.3% in 2019.

4.2 Model of non-stochastic prediction range for estimated GDP evolution in France

The French GDP time series is long enough; hence we have been able to set up a large number of models and their variants. We have calculated our estimates of polynomial curves (including exponential ones), moving averages, stochastic models and exponential smoothing models, always verifying their statistical properties; if a model has been found to be statistically or factually unsuitable, it has immediately been discarded. In the end, 22 models have remained.

This collection of a large number of models has enabled us to gather different "statistical opinions" about the series to be predicted, including models that can "discount their memories" (in the sense of older observations having lower weights, such as the previously mentioned exponential smoothing). Table A2 in the Appendix sums up an overview of the models processed, including their parameters and statistical properties, as well as the predictions derived from them for the years 2018 and 2019 (for all models, both the point and interval predictions at the 95% level of confidence).

All these models were identified with the aid of only 68 items in the time series (from 1950 to 2017) – we "stored away" the actual values for the years 2018 and 2019. For each of the models we have, based on 68 observations, predicted the 2018 and 2019 values to compare such prediction results with the actual values (as the subsequent assessments of the predictions). In other words, we have thus "tested" each model's ability to predict.

Table A2 in the Appendix further states each model's estimated (modelled, theoretical) value for the latest actual observation, that is, 2017. The data listed in Table A1 in the Appendix says that the actual value of the year-to-year GDP index in France was $y_{68} = 1.0123$ in 2017; that is, the year-to-year growth value was 1.23%.

A cursory glance at Table A2 in the Appendix reveals a paradox occurring when we use an isolated model from our selection to set up a point prediction for the GDP index in 2018 or 2019. Individual values of point predictions listed in Table A2 show that each model naturally leads to a different prediction for the last "known" period (y_{68}, i.e., 2017; as already stated, we have "stored away" the y_{69} and y_{70} values to be checked later.) Each of the models used pertains to its own dynamics. Judging from past behaviour, it would be very difficult to assess which model is more or less "acceptable" in comparison with others; to put it bluntly: anybody could choose anything.

Let us have a look at Model 18 in Table A2 in the Appendix – ARIMA (2, 1, 1), which was, by software Statgraphics Centurion software, assessed as the best among all of our 22 models. The models' quality levels were verified with the aid of the usual statistics, whose list and more detailed descriptions are given in the Appendix prior to Table A2.

Table 2 Model M18

Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
2018	1.015 570	0.982 584	1.048 560
2019	1.012 110	0.976 179	1.048 050

Source: Authors' own calculations

For the year 2018 or 2019, the 95%-level interval prediction of the year-to-year GDP growth index is approximately between 0.982 584 and 1.048 560, or between 0.976 179 and 1.048 050, respectively; in both instances, the span between the upper and lower bounds amounts to more than 6.6 percentage points. Expressed in absolute volumes, e.g., the French GDP was 3 108.7 billion EUR in 2018, and the 6.6% corresponds to 205 billion EUR. This means nearly 40% of the French gross fixed capital formation (which was 537.9 billion EUR in 2018). It does not make much practical sense to set up a prediction interval whose "uncertainty" amounts to nearly two-fifths of French annul investment volume.

As previously mentioned, we have "stored away" the actual values of the year-to-year French GDP growth index for the years 2018 and 2019. In 2018, this actual value was 1.017, meaning an increase in the GDP of 1.7%. In 2019, the actual value of the year-to-year index was 1.013, i.e., representing an increase of 1.3%. In both instances, the 95%-level stochastic confidence interval we have created is "successful" (and similar observations can be made about other models – cf. Table A2 in the Appendix). However, this interval is too broad for subsequent decision-making.

The non-stochastic prediction range is based on the selected models and their estimated year-to-year French GDP growth indices. Namely, we have the year-to-year indices expressed by the prediction intervals of the 22 models (the Lower 95.0% Limit and Upper 95.0% Limit, values in Table A2 in the Appendix). Let us now look up the *maximum lower bound* and the *minimum upper bound* of the year-to-year index prediction intervals in Table A2 in the Appendix (separately for 2018 and 2019). These values are shown in Table 3 (as well as in Table A2 in the Appendix); they come from Model 5 (Exponential Trend: the minimum value of the Upper 95.0% Limit among all 22 models); and Model 13 (ARIMA (0, 0, 1): the maximum value of the Upper 95.0% Limit among all 22 models):

Table 3 Max lower and min upper bounds

Period	Max (Lower 95.0% Limit)	Min (Upper 95.0% Limit)
2018	0.993 472	1.037 310
2019	0.989 743	1.036 560

Source: Authors' own calculations

Comparing the lower and upper bounds for the prediction ranges in Tables 2 and 3, we can see that the span between them is smaller for the non-stochastically argued prediction range. In fact, this span is just 4.4 percentage points, as compared with 6.6 percentage points (at a 95% confidence level) valid for the best 2018 model, i.e., ARIMA (2, 1, 1). This is an improvement by one-third.

In this way, we have obtained a prediction range (as a difference between the upper and lower bounds) for the expected values of the year-to-year GDP indices in France for the years $n + 1 = 2018$ and $n + 2 = 2019$. Let us denote by P_{69} the prediction for 2018, and by P_{70} that for 2019:

prediction range for P_{69} → minimum upper bound: 1.0373
→ maximum lower bound: 0.9934,

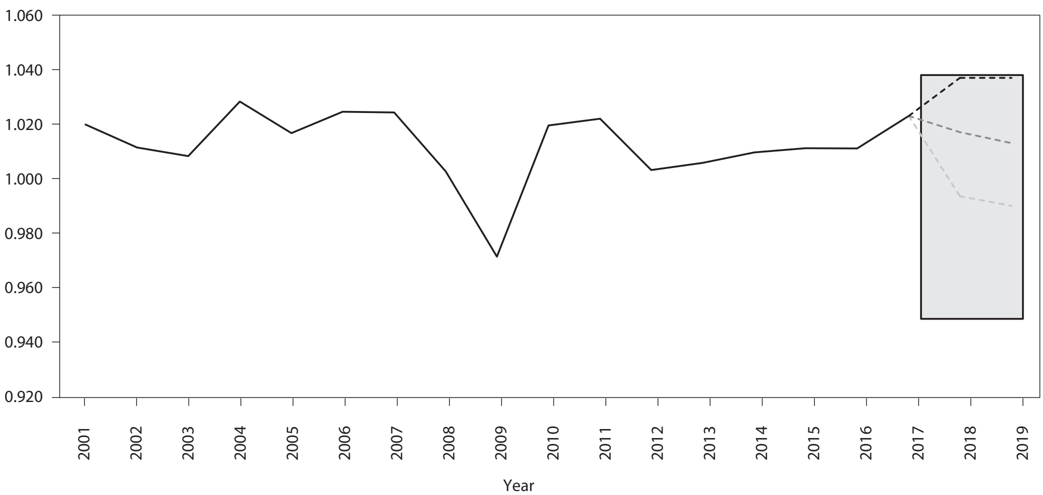
or

prediction range for P_{70} minimum upper bound: 1.0366
maximum lower bound: 0.9897.

We have, for the sake of clarity, graphically enhanced a short end section in Figure 2, which depicts the year-to-year French GDP growth indices from 2005 to 2017; this enhancement helps us see the non-stochastic prediction rage for the years 2018 and 2019 (the lower and upper bounds of the non-stochastic predictions are marked with dashed lines). At the same time, the actual year-to-year evolution of the French GDP indices is easier to see (the solid line).

The 2018 and 2019 data is represented by the actual values (the solid line again) – it is covered by the non-stochastic prediction range (the dashed lines).

Figure 2 Non-stochastic prediction range – a segment of the French GDP time series (year-to-year volume indices from 2001)



Source: Authors' own calculations, <*www.insee.fr*>

From the pragmatic point of view, it is clear that the concept of the prediction range set up and argued in a non-stochastic way is more efficient than the traditional interval predictions, based on an isolated single model, whether best or just "good" – in our case, on the ARIMA (2, 1, 1) process. The non-stochastically interpreted concept sets out the future evolution of the index to be predicted in a band that is much narrower; this reduces uncertainty in the user's decision-making.

In conclusion, let us point out one interesting phenomenon. It is known that many structural relationships are valid among different indices (such as the macroeconomic aggregates). In the case of macro-aggregates, it is of extraordinary importance to consider the relationship corresponding to the expenditure-based method for estimating the GDP:

$$GDP = FCE + GCF + E - I, \tag{5}$$

where FCE stands for the final consumption expenditure, GCF for the gross capital formation, E for exports of goods and services, and I for imports of goods and services.[5] A question arises whether the described method could also be used if we are interested not only in individual indices but also in their sum, e.g., according to Formula (5). It has turned out that the application of the prediction range is also useful for additive relationships. In other words, our approach is also consistent in structural or balance issues, in which aggregation/decomposition of individual indices plays a role.

In the end, let us address a logical question: what is the average value of the point predictions derived within all of the admissible 22 models? The data shown in Table A2 in the Appendix provide the average value for the 2018 prediction as $P_{69} = 1.015$ (an increase in the GDP by 1.5%), and for 2019 it is $P_{70} = 1.014$ (an increase in the GDP by 1.4%). Let us compare these values with the actual values "stored" for the purposes of the prediction assessment: $y_{69} = 1.017$ for 2018, and $y_{70} = 1.013$ for 2019. This result indicates a very good fit; for the sake of interest, the values of the Theil coefficient, cf. Theil (1966), for the estimates in 2018 and 2019 as compared with the actual values equal $T_H = 0.15\%$.

CONCLUSIONS

Having in mind the current empirical results, the utilisation of the prediction range in economics can be viewed as purposeful. General experience with the efficiency of the prediction range based on processing a large number of economic time series has revealed the fact that the success rate of this method is relatively high. Nearly 80% of all the ranges we have set up (mostly time series of financial and macroeconomic indices) were successful when later compared with the actual data. That is why the methodology for the prediction range can become a good corroborative technique in setting up estimates for indices of this type.

Of course, open questions to be addressed in the future remain in the presented outline of the setting up prediction range argued in a non-probabilistic way. In our example we considered a series of annual values. But series with seasonal components may be predicted as well, for example, quarterly aggregates from the national accounts, or completely different time series subject to seasonal fluctuations. Other open question is a methodology for choosing the models on which the non-stochastic prediction range is based. An ideal solution would be the creation of an algorithmic tool to automate, at least to a certain extent, the primary process of model selection and verification. Another important problem to be resolved is the question of evaluating the lower and upper bounds of the non-stochastic prediction range. For more complex traditional approaches, such as those considered by Theil (1966), an obstacle s implied by the non-stochastic character of such bounds; hence simple applications of Theil's processes may be disabled. For the moment, we have to put up with a simple way of evaluation by comparing the results with the actual values when assessing the predictions.

Another option would be to set up the non-stochastically argued prediction, whether a point or interval one, with the aid of the results of the primary models (here M1, M2, etc.) weighted, for example, with the interpolation quantity of individual models (even though we are aware that a suitable description of the past behaviour is only partly reflected in successful predictions).

It will, indisputably, be a great challenge to process the reflection of the COVID-19 pandemic in the 2020 models, as well as the applications to the future years of 2022, 2023, etc., when the economic situation will be getting back to its normal state. It is obvious that it is impossible to predict the economic evolution for 2020. The pandemic intervention in the economic relationships is so extensive that there are no known models which would be able to cope with such predictions. When we get to the economic-recovery stage, it will be interesting to observe to what extent the non-stochastically argued predictions will be capable of estimating the degree of the economic recovery. This paper has been written in the period of massive manifestation of the coronavirus crisis, which is currently the dominant intervention process of the highest intensity.

[5] Cf., for example, Hronová et al. (2019).

References

ANDĚL, J. *Statistické metody* [*Statistical methods*]. Prague: MFF UK, 1993.

BOX, G. E. P., JENKINS, G. M., REINSEL, G. C. *Time series analysis, forecasting and control.* 3rd Ed. New Jersey: Prentice-Hall, Inc. Englewood Cliffs, 1994.

BOOTS, J. C. C., FEIBES, W., LISMAN, H. H. C. Further Methods of Derivation of Quarterly Figures from Annual Data. *Applied Statistics*, 1967, 16, pp. 65–75.

BOURNAY, J. AND LAROQUE, G. Réflexions sur la méthode d´élaboration des comptes trimestriels. *Annales de l'INSEE*, 1979, 36, pp. 3–30.

CIPRA, T. *Analýza časových řad s aplikacemi v ekonomii* [*Time series analysis with economic applications*]. SNTL, Prague and ALFA, Bratislava, 1986.

DEMPSTER, A. P., LAIRD, N. M., RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 1977, 39, pp. 1–38.

DENTON, F. T. Adjustment of monthly and quarterly series to annual tools: an approach based on quadratic minimization. *Journal of American Statistical Association*, 1991, Vol. 66, No. 333.

DUREAU, G. *Les comptes nationaux trimestriels.* INSEE – Méthodes, 1991, No. 13.

EPPELBAUM, LEV V. Non-stochastic long-term prediction model for US tornado level. *Natural Hazards*, 2013, 69, pp. 2269–2278.

EUROSTAT. *European System of Accounts (ESA 1995).* Luxembourg: Eurostat, 1996.

EUROSTAT. *European System of Accounts (ESA 2010).* Luxembourg: Eurostat, 2010.

FERREIRA, L. C. E. AND CROSS, N. J. G. New Insights into Time Series Analyses. *A&A*, 2017, 604, A121.

FRIEDMAN, M. The Interpolation of Time Series by Related Series. *Journal of American Statistical Association*, 1962, 57, pp. 729–757.

GRANGER, C. W. J. AND NEWBOLD, P. *Forecasting Economic Time Series.* New York: Academic Press, 1986.

HINDLS, R. AND HRONOVA, S. Les prévisions trimestrielles du PIB en République Tchèque. *Bulletin de l´ACN*, 1996, 33, pp. 2–13.

HRONOVA, S., HINDLS, R., MAREK, L. Does the Signal Gap Indicate a Growing Risk of Economic Recession? In: *12th Annual Conference of the EuroMed Academy of Business*, Thessaloniki: EuroMed Press, 2019, pp. 388–399.

HRONOVÁ, S., FISCHER, J., HINDLS, R., SIXTA, J. *Národní účetnictví – od výroby k bohatství* [*National accounting ‑ from production to wealth*]. Prague: C. H. Beck, 2019.

CHOW, G. AND LIN, A. L. Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series. *Review of Economics and Statistics*, 1971, 43, pp. 372–375.

KOZÁK, J., HINDLS, R., HRONOVÁ, S. Some Remarks to the Methodology of the Allocation of Yearly Observations into Seasons. *Statistics in Transition*, 2000, 4, pp. 815–826.

KOZÁK, J. Možnosti využití expertních informací při konstrukci předpovědí na bázi klasického lineárního modelu časové řady [Possible utilisation of expert information when constructing predictions based on a classical linear model for a time series]. *Acta Oeconomia Pragensia*, 1996, pp. 117–142.

KUHN, M. AND JOHNSON, K. *Applied Predictive Modeling.* New York: Springer, 2013.

LIU, H. *Smart Cities: Big Data Prediction Methods and Applications.* Springer Nature Singapore Pte Ltd. and Science Press, 2020.

NASSE, PH. Le système des comptes nationaux trimestrieles. *Annales de l´INSEE*, 1973, 5, pp. 119–165.

NEUSSER, K. *Time Series Econometrics.* New York: Springer, 2016.

OCDE. *Etudes Economique de l´OCDE: France 2019.* Paris: Editions OCDE, 2019.

PANKRATZ, A. *Forecasting with dynamic regression models.* New York: John Wiley & Sons, 1991.

OSBORN, F. *Applied Imagination: Principles and Procedures of Creative Problem Solving.* 3rd Ed. New York: Charles Scribner's Sons, 1963.

SCA STATISTICAL SYSTEM. *Reference manual for general statistical analysis.* Scientific Associates Corp., Oak Brook, 1991.

THEIL, H. *Applied Economic Forecasting.* Amsterdam ‑ Chicago: North Holland Publishing & R. McNally, 1996.

TRIBBIA, J. AND BAUMHEFNER, D. On the Problem of Prediction beyond the Deterministic Range. In: SHUKLA, J. (eds.). Prediction of Interannual Climate Variations, *NATO ASI Series*, 1993, 6, pp. 251–264.

UN. *System of National Accounts 1993 (SNA 1993).* New York: United Nations, 1993.

WASSENAN, L. *All of Statistics: a Concise Course in Statistical Inference.* New York: Springer, 2004.

WEI, W. W. S. *Time series analysis – Univariate and multivariate methods.* New York: Addison Wesley Publishing, 2006.

WOODWARD, W. A., GRAY, H. L., ELLIOTT, A. C. *Applied Time Series Analysis.* CRC Press, 2012.

YACHAO, Z., KAIPEI, L., LIANG, Q., XUELI, A. Deterministic and probabilistic interval prediction for short-term wind power generation based on variational mode decomposition and machine learning methods. *Energy Conversion and Management*, 2016, 112, pp. 208–219.

APPENDIX

Table A1 GDP France (year-to-year volume indices)

Year	y/y	Year	y/y	Year	y/y	Year	y/y
1950	1.086	1970	1.061	1990	1.029	2010	1.019
1951	1.058	1971	1.053	1991	1.010	2011	1.022
1952	1.031	1972	1.045	1992	1.016	2012	1.003
1953	1.035	1973	1.063	1993	0.994	2013	1.006
1954	1.056	1974	1.043	1994	1.024	2014	1.010
1955	1.053	1975	0.990	1995	1.021	2015	1.011
1956	1.050	1976	1.044	1996	1.014	2016	1.011
1957	1.055	1977	1.035	1997	1.023	2017	1.023
1958	1.027	1978	1.040	1998	1.036	**2018**	**1.017**
1959	1.027	1979	1.036	1999	1.034	**2019**	**1.013**
1960	1.080	1980	1.016	2000	1.039		
1961	1.050	1981	1.011	2001	1.020		
1962	1.068	1982	1.025	2002	1.011		
1963	1.062	1983	1.012	2003	1.008		
1964	1.067	1984	1.015	2004	1.028		
1965	1.049	1985	1.016	2005	1.017		
1966	1.053	1986	1.023	2006	1.024		
1967	1.049	1987	1.026	2007	1.024		
1968	1.045	1988	1.047	2008	1.003		
1969	1.071	1989	1.043	2009	0.971		

Source: *<www.insee.fr>*

Table A2 Overview and comparison of models estimated for prediction purposes

Table A2 in the Appendix lists the estimates within the models we have applied to the given time series, characteristics of their "interpolation quality", and the point prediction P_{n+i} for each model, together with the 95%-level interval prediction of the year-to-year GDP indices (or rather, Lower 95.0% Limit and Upper 95.0% Limit), for the years 2018 and 2019, followed by:

- the root mean squared error (RMSE),
- the mean absolute error (MAE),
- the mean absolute percentage error (MAPE),
- the mean error (ME),
- the mean percentage error (MPE).

M1. Random Walk with Drift
Forecast model selected: Random Walk with Drift = –0.000935166

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0189166	2018	1.02206	0.984296	1.05983
MAE	0.0135242	2019	1.02113	0.967717	1.07454
MAPE	1.31152				
ME	–4.63974E-17				
MPE	–0.0146858				

M2. Constant Mean
Forecast model selected: Constant Mean = 1.03188

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0222124	2018	1.03188	0.987221	1.07654
MAE	0.0181363	2019	1.03188	0.987221	1.07654
MAPE	1.75631				
ME	–3.31434E-16				
MPE	–0.0456116				

M3. Linear Trend
Forecast model selected: Linear Trend = 1.0591 – 0.000788999 · t

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0159303	2018	1.00466	0.971913	1.03741
MAE	0.0118108	2019	1.00387	0.971083	1.03666
MAPE	1.14736				
ME	–2.36739E-16				
MPE	–0.023199				

M4. Quadratic Trend
Forecast model selected: Quadratic Trend = 1.06239 – 0.0010706 · t + 0.00000408112 · t^2

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0159879	2018	1.00795	0.973848	1.04205
MAE	0.0117821	2019	1.00744	0.973088	1.04180
MAPE	1.14392				
ME	–2.72658E-16				
MPE	–0.0230321				

M5. Exponential Trend
Forecast model selected: Exponential Trend = $e^{(0.0574661 - 0.000762605 \cdot t)}$

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0159219	2018	1.00486	0.973425	1.03731*
MAE	0.0118242	2019	1.00409	0.972644	1.03656*
MAPE	1.14847				
ME	0.000119753				
MPE	–0.011639				

* = the minimum value among Upper 95.0% Limit values of all 22 models

M6. S-Curve Trend = exp (0.0263113 + 0.0685776 / t)

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0200434	2018	1.02768	0.988234	1.06870
MAE	0.0161746	2019	1.02767	0.988219	1.06869
MAPE	1.56663				
ME	0.000188524				
MPE	–0.0183384				

M7. Simple Moving Average of three terms Forecast model selected: Simple Moving Average of three terms					
Statistic		**Period**	**Forecast P_{n+i}**	**Lower 95.0% Limit**	**Upper 95.0% Limit**
RMSE	0.0170942	2018	1.01504	0.976356	1.05373
MAE	0.0129897	2019	1.01504	0.976356	1.05373
MAPE	1.26549				
ME	–0.000988606				
MPE	–0.115514				

M8. Simple Exponential Smoothing Forecast model selected: Simple Exponential Smoothing with alpha = 0.2456					
Statistic		**Period**	**Forecast P_{n+i}**	**Lower 95.0% Limit**	**Upper 95.0% Limit**
RMSE	0.0165784	2018	1.01318	0.980931	1.04544
MAE	0.0120414	2019	1.01318	0.979972	1.04640
MAPE	1.17237				
ME	–0.00259263				
MPE	–0.273068				

M9. Brown's Linear Exp. Smoothing Forecast model selected: Brown's Linear Exp. Smoothing with alpha = 0.1095					
Statistic		**Period**	**Forecast P_{n+i}**	**Lower 95.0% Limit**	**Upper 95.0% Limit**
RMSE	0.0168608	2018	1.00960	0.976794	1.04240
MAE	0.0123323	2019	1.00911	0.975525	1.04269
MAPE	1.1999				
ME	–0.0014339				
MPE	–0.157938				

M10. Holt's Linear Exp. Smoothing Forecast model selected: Holt's Linear Exp. Smoothing with alpha = 0.1296 and beta = 0.0413					
Statistic		**Period**	**Forecast P_{n+i}**	**Lower 95.0% Limit**	**Upper 95.0% Limit**
RMSE	0.0166783	2018	1.00785	0.975643	1.04005
MAE	0.0119991	2019	1.00722	0.974722	1.03972
MAPE	1.16662				
ME	–0.000206347				
MPE	–0.0420116				

M11. Brown's Quadratic Exp. Smoothing Forecast model selected: Brown's Quadratic Exp. Smoothing with alpha = 0.0764					
Statistic		**Period**	**Forecast P_{n+i}**	**Lower 95.0% Limit**	**Upper 95.0% Limit**
RMSE	0.017078	2018	1.00915	0.975922	1.04237
MAE	0.0125613	2019	1.00865	0.974560	1.04273
MAPE	1.22117				
ME	0.000030422				
MPE	–0.0152579				

M12. ARIMA (1, 0, 0)
Forecast model selected: ARIMA (1, 0, 0) with a constant. AR(1) = 0.671835; Constant = 0.338781

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0171491	2018	1.02607	0.991490	1.06065
MAE	0.01233	2019	1.02813	0.986473	1.06979
RMSE	0.0171491				
ME	−0.00059993				
MPE	−0.0837374				

M13. ARIMA (0, 0, 1)
Forecast model selected: ARIMA (0, 0, 1) with a constant. MA(1) = −0.514135; Constant = 1.03217

RMSE	0.0187211	Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
MAE	0.0143691	2018	1.03120	0.993472**	1.06893
MAPE	1.39234	2019	1.03217	0.989743**	1.07459
ME	−0.000302896				
MPE	−0.0638256				
MPE	−0.023199				

** = the maximum value of the Lower 95.0% Limit among all 22 models

M14. ARIMA (1, 0, 1)
Forecast model selected: ARIMA (1, 0, 1) with a constant.
AR(1) = 0.9447; MA(1) = 0.605663; Constant = 0.0573209

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0168699	2018	1.01816	0.984265	1.05206
MAE	0.0120234	2019	1.01918	0.983386	1.05497
MAPE	1.16948				
ME	−0.00177662				
MPE	−0.196771				

M15. ARIMA (1, 1, 1)
Forecast model selected: ARIMA (1, 1, 1) with a constant.
AR(1) = 0.26381; MA(1) = 0.967834; Constant = −0.000579074

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0155851	2018	1.01001	0.978402	1.04162
MAE	0.0114333	2019	1.00601	0.973042	1.03897
MAPE	1.11316				
ME	−0.00149397				
MPE	−0.163421				

M16. ARIMA (1, 1, 0)
Forecast model selected: ARIMA (1, 1, 0) with a constant
AR(1) = −0.296046; Constant = −0.00117087

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0181952	2018	1.01828	0.981894	1.05466
MAE	0.0130038	2019	1.01850	0.974011	1.06300
MAPE	1.26397				
ME	−0.0000630604				
MPE	−0.0224724				

M17. ARIMA (0, 1, 1)
Forecast model selected: ARIMA (0, 1, 1) with a constant.
MA(1) = 0.975448; Constant = –0.000799191

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0160511	2018	1.00513	0.972545	1.03772
MAE	0.0116463	2019	1.00434	0.971736	1.03693
MAPE	1.13506				
ME	–0.00186452				
MPE	–0.200961				

M18. ARIMA (2, 1, 1) – the best model
Forecast model selected: ARIMA (2, 1, 1).
AR(1) = 0.217349; AR(2) = –0.153627; MA(1) = 0.785371

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0161257	2018	1.01557	0.982584	1.04856
MAE	0.0116326	2019	1.01211	0.976179	1.04805
MAPE	1.13346				
ME	–0.0032293				
MPE	–0.332023				

M19. ARIMA (1, 1, 2)
Forecast model selected: ARIMA (1, 1, 2) with a constant.
AR(1) = –0.167067; MA(1) = 0.501519; MA(2) = 0.458159; Constant = –0.000904754

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0155657	2018	1.01042	0.978757	1.04209
MAE	0.0115389	2019	1.00423	0.970864	1.03759
MAPE	1.12252				
ME	–0.00141871				
MPE	–0.156373				

M20. ARIMA (2, 1, 2)
Forecast model selected: ARIMA (2, 1, 2) with a constant.
AR(1) = –0.166424; AR(2) = 0.0439958; MA(1) = 0.495085; MA(2) = 0.462624; Constant = –0.000867036

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.015673	2018	1.01067	0.978763	1.04258
MAE	0.0114815	2019	1.00499	0.971303	1.03867
MAPE	1.11703				
ME	–0.00130787				
MPE	–0.145634				

M21. ARIMA (2, 1, 0)
Forecast model selected: ARIMA (2, 1, 0) with a constant.
AR(1) = –0.404877; AR(2) = –0.336535; Constant = –0.0013624

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0172156	2018	1.01682	0.982146	1.05150
MAE	0.01284	2019	1.01392	0.973570	1.05428
MAPE	1.24743				
ME	–0.000187674				
MPE	–0.0350347				

M22. ARIMA (0, 1, 2)
Forecast model selected: ARIMA (0, 1, 2) with a constant.
MA(1) = 0.610061; MA(2) = 0.351027; Constant = –0.000776596

Statistic		Period	Forecast P_{n+i}	Lower 95.0% Limit	Upper 95.0% Limit
RMSE	0.0154547	2018	1.01136	0.979949	1.04278
MAE	0.0115043	2019	1.00499	0.971277	1.03871
MAPE	1.11889				
ME	–0.00128165				
MPE	–0.142689				

Source: *<www.insee.fr>*

Selection of the Optimal Way of Linear Ordering of Objects: Case of Sustainable Development in EU Countries

Agnieszka Sompolska-Rzechuła[1] | *West Pomeranian University of Technology Szczecin, Szczecin, Poland*

Abstract

The aim of the article was to assess selected methods of linear ordering of objects and choosing the optimal method. The measures based on different properties of the synthetic variable were selected for evaluation. The selection of the optimal linear ordering procedure is the last step in creating a synthetic variable and is often not included in the research. The analysis was based on data from the EUROSTAT database (2017) countries. The level of socio-economic development in the context of sustainable development for 28 European Union was adopted as the ordering countries. The paper proposes a comparison of results in various methods, e.g. due to the way of normalization of diagnostic features or type of methods (based on a pattern object or a non-pattern object). Out of all the selection methods for this study, the TOPSIS methods based on zero unitarization proved to be the optimal.

Keywords

Linear ordering of objects, selection of method of linear ordering, level of socio-economic development, sustainable development

JEL code

C38, Q01, O57

INTRODUCTION

The methods of ordering objects make possible to determine the order of objects depending on the degree of intensity of specific features. Linear ordering methods are included among the ordering methods. Linear ordering is based on a feature called synthetic, but objects are multidimensional. A synthetic feature aggregates partial information contained in simple features that make up the evaluation criterion (Wysocki, 2010). The first proposal for a synthetic variable was presented by Hellwig (1968).

Among others, Hartigan (1975), Pluta (1977), Hwang and Yoon (1981), Anderson (1984), Seber (1984), Morrison (1990), Grabiński (1992), Chen (2000), Kukuła (2000) had a significant share in the development of these methods.

The idea of linear ordering of multidimensional objects is based on the concept of ordering binary relations (reflexive, non-symmetrical, transitive and coherent). The axioms of this relation show

[1] West Pomeranian University of Technology Szczecin, Department of Mathematics Applications in Economy, Janickiego 31, 71-270 Szczecin, Poland. E-mail: asompolska@zut.edu.pl.

that it is possible to state which of any two objects of the set is the first (better) and which is the second (worse), as well as whether they are identical (Bąk, 2015). The subject of linear ordering can be objects such as countries, enterprises, products or people. With many methods of linear ordering of objects available, it is not always clear which procedure to choose. The quality assessment of linear ordering procedures is the last step in creating a synthetic variable. In the literature on the subject, one can find mainly works whose final result is the ranking of objects without assessing the optimality of the results obtained.

The issue of choosing the optimal method of linear object ordering was taken up in the works Bąk (2015, 2018), and Sompolska-Rzechuła (2020). In the latter, an evaluation of selected methods of linear ordering was proposed, adopting various methods of normalizing diagnostic features and one method of aggregating variables. However, this work is a continuation and extension of research on the selection of final results obtained with the methods of linear ordering of objects, because both the methods based on a pattern object and a non-pattern object were used. Moreover, the measures of correctness of the linear ordering procedures from each group of measures presented in Table 1 were used and some modifications to the measures of correctness of the linear ordering procedures have been proposed. Both articles used data from the EUROSTAT database (2017) for 28 European Union countries. The information is pertaining to the level of socio-economic development in the context of sustainable development.

The level of socio-economic development was adopted as the criterion for ordering countries, which was presented with the use of indicators reflecting the concept of sustainable development. Sustainable development is implemented in three dimensions: economic, social and ecological and is based on the pursuit of the best economic result while respecting the natural environment and social development. It is therefore a social and economic development that ensures that the needs of modern society are met without hindering future generations from meeting their needs. The 3 × P abbreviation is often presented, *planet*, *people*, with *profit* at the very end. This order suggests an emphasis put above all on preserving Earth's resources, not threatening the environment, and the profit is seen only at the very end (Latoszek, 2016). Therefore, the essence of the concept of sustainable and permanent development is to meet the needs of present generation without reducing the chances of future generation to meet them. This definition was included in the 1987 report of the World Commission on the Environment and Development entitled: *Our Common Future* (United Nations General Assembly, 1987). Although there are many definitions, the most commonly used definition of sustainable development is the one proposed by the Brundtland Commission: "Sustainable development is development that meets the needs of the present without compromising the ability of future generation to meet their own needs" (Cerin, 2006). The concept of sustainable development was designed in the 1980s and is one of the most important contemporary concepts for economic development.

1 RESEARCH METHODOLOGY

Linear ordering methods are used to evaluate multi-feature objects (e.g. countries) allowing them to be ranked, according to a specific general criterion, from "best" to "worst". This criterion is treated as a property of the examined objects and is a complex phenomenon. Socio-economic research very often examined phenomena that are not directly measured. Sets of diagnostic features are used then, measured on various measuring scales. Linear ordering of objects is obtained on the basis of a feature called aggregate or synthetic, which is created by aggregating the initial features describing the tested objects.

The synthetic feature creation procedure is a multi-step process and includes (Wysocki, 2010; Sompolska-Rzechuła, 2020):

1) gathering preliminary information about potential diagnostic variables;
2) selecting of diagnostic variables;
3) determining the type of variables: stimulant, destimulant or nominant;
4) normalising of diagnostic variables using the selected normalizing method;

5) assigning weights to standardized features;
6) calculating values of the aggregation of features, i.e. creating a synthetic variable;
7) construction of the linear ordering of facilities due to the level of the complex phenomenon in question;
8) quality assessment of rankings using partial quality assessment criteria and aggregate measures were calculated.

The first step in creating a synthetic feature is to establish a set of diagnostic features. There are two approaches to this issue – non-statistical (theoretical) and statistical (Wysocki, 2010). The substantive approach is based on the qualitative assessment of the studied phenomenon, taking into account economic knowledge and theory. The statistical approach is designed to limit the set of diagnostic features and exclude those features that do not fully characterize the examined objects in terms of the adopted criterion. Analysis of variability and correlations between features is often used. Another condition that diagnostic features should meet is the lack of correlation between features. Therefore, from the set of potential features, features strongly correlated with others should be eliminated, because they are a carrier of similar information. In the literature on the subject, you can find many methods used in the selection of features. These include methods based on, e.g., the analysis of correlation coefficients between features. One of these methods is the procedure proposed by Hellwig and known as the parametric method of feature selection. A detailed description of this method can be found, e.g. in the works by (Wysocki, 2010, pp. 146–147), and (Sompolska-Rzechuła, 2018, pp. 74–76). The algorithm of the parametric Hellwig method is as follows:

1. Calculating the correlation matrix **R** of k variables.
2. Determining the threshold value of the correlation coefficient (r^*) e.g. based on the formula:

$$r^* = \min_i \max_j |r_{ij}| \; i, j = 1, \ldots, k, \tag{1}$$

where: r_{ij} – Pearson linear correlation coefficients between features, k – number of features. The threshold value of the coefficient can also be taken arbitrarily, often as $r^* = 0.5$.

3. Calculating the sum of the absolute values of the correlation coefficients for each column of the matrix **R**:

$$R_j = \sum_{i=1}^{k} |r_{ij}|. \tag{2}$$

4. Determining the column number (m), for which the sum R_j is the largest:

$$R_m = \max_j \{R_j\}. \tag{3}$$

5. Classification of variables: the variable with the number (k) is the central variable, and the variables for which $|r_{ik}| < r^*$ are satellite variables (they form a cluster of highly correlated variables, thus causing information redundancy).
6. Removing rows and columns from the matrix **R** corresponding to satellite variables and the column corresponding to the central variable.
7. The procedure is repeated until the set of variables is exhausted.
8. Variables that are not in any cluster are isolated variables (they form one-element clusters).
9. Central and isolated variables are included in the analysis (satellite variables are discarded).

The set of diagnostic features is the basis for further analysis, in which the nature of the features should be determined, i.e. stimulants, destimulants and nominants should be distinguished.

After recognizing the nature of the features, they must be transformed; most often destimulants are converted into stimulants by means of difference or quotient transformations.

The concept of stimulants and destimulants was introduced by Hellwig (1968).

The stimulant means a feature which higher value indicates a better condition of the object in a given context. Thus, the maximum value of the stimulant is considered the most favourable, and the minimum – the least favourable for the examined objects. While the destimulant is a feature which lower values mean a better situation of the object in a given respect. Therefore, the maximum value of destimulants is considered the least favourable, and the lowest – the most favourable for the examined objects. While the neutral variable is characterized by the existence of an optimal value (for this reason sometimes called a nominal value), below which such a variable has the character of a stimulant (and therefore larger values are more favourable), and above the destimulant (which means that after exceeding the optimal value, a further increase in the value of the feature becomes unfavourable) or the other way round. Neutral variables are often overlooked in empirical studies due to difficulties in establishing nominal values. If it is difficult to determine the nature of the characteristics, specific substantive criteria or correlation analysis should be used. It is also possible to evaluate the nature of the features after determining the value of the synthetic feature, then stimulants should be positively correlated with the synthetic feature and destimulants – negatively.

The next stage of building the synthetic feature is the normalization of features. It leads to deprivation of physical units of measurement results and unification of orders of magnitude. The literature contains many proposals for these methods and discussions on the criteria for their selection. The rest of the work will present those normalizing formulas that relate to stimulus traits.

The following standardizing formulas have been used in this work (Kukuła, 2000):

- zero unitarization:

$$z_{ij} = \frac{x_{ij} - \min\limits_{l} x_{lj}}{\max\limits_{l} x_{lj} - \min\limits_{l} x_{lj}} \quad (\max_{l} x_{lj} \neq \min_{l} x_{lj}), \tag{4}$$

- quotient transformation:

$$z_{ij} = \frac{x_{ij}}{\bar{x}_j} \quad (\bar{x}_j \neq 0), \tag{5}$$

where: z_{ij} – standardized value of the j-th feature ($j = 1, 2, \ldots, k$) for the i-th object ($i = 1, 2, \ldots, n$), n – number of object.

In the zero unitarization method, a constant pattern point is assumed – the range of the normalized variable. The use of this method makes the range of the normalized feature constant and amounts to one. The normalized feature assumes values in the range [0,1]. Moreover, this method makes possible to normalize the features taking positive, negative and zero values.

In the next step of creating a linear ordering of objects, the values of the synthetic feature are determined. There are many methods of constructing a synthetic development measure that can be divided into non-pattern and pattern (Grabiński, 1992). The main difference between the pattern and non-pattern methods lies in the fact that in pattern methods the basis of analyses takes on the form of a concept of a development pattern, which is understood as a certain artificially constructed object, characterized by some optimal properties expressed in properly defined values of diagnostic features.

A non-pattern object methods rely on the operation of averaging the values of standard features (Wysocki, 2010):

$$\mu_i = \frac{1}{k} \sum\nolimits_{j=1}^{k} z_{ij}, \tag{6}$$

where: μ_i – the value of the synthetic feature for the i-th object.

The idea of pattern methods of aggregation of features is to determine the distance of individual objects from a certain pattern object. Among the pattern methods, the method proposed by Hellwig (1968) deserves attention. It is based on standardized values of diagnostic features $X_1, X_2, \ldots, X_k$ which are treated as equally important. The Euclidean distances of each object are calculated from the pattern according to the formula:

$$d_i = \sqrt{\sum_{i=1}^{k} (z_{ij} - z_{0k})^2} \; (i = 1, 2, ..., n), \tag{7}$$

where: $z_{0k} = \max_i \{z_{ik}\}$ – standardized value of the k-th feature for the pattern object, n – number of objects. In this paper zero unitarization method was used to standardize variables.

Based on the d_i value, the relative taxonomic measure of development is constructed, defined as (Nowak, 1990):

$$\mu_i = 1 - \frac{d_i}{d_0} \; (i = 1, 2, ..., n), \tag{8}$$

where: $d_0 = \bar{d} + 2 \cdot s_d$,

wherein: $\bar{d} = \frac{1}{n} \sum_{i=1}^{n} d_i$ and $s_d = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_i - d)^2}$.

The synthetic Hellwig development measure almost always takes values from [0,l]. The smaller the difference of the μ_i measure from one, the less the level of object development differs from the level of object development recognized as the standard. The synthetic measure of development is a resultant of all the features characterizing the examined objects, it allows to determine the "average" level of the value of the features achieved at some time.

The linear ordering method using the pattern and non-pattern is the TOPSIS method (*Technique for Order Preference by Similarity to an Ideal Solution*; Hwang et Yoon, 1981). It consists in calculating the Euclidean distances of each assessed object from both the pattern and non-pattern of the development, which distinguishes it from the Hellwig method, which only takes into account the distance from the development pattern (Wysocki, 2010). The coordinates of the model units are set – development pattern and non-pattern. The values of the pattern (A^+) and non-pattern of development (A^-) are defined as (Wysocki, 2010):

$$A^+ = \left(\max_i(z_{i1}), \max_i(z_{i2}), \ldots, \max_i(z_{ik})\right) = (z_1^+, z_2^+, \ldots, z_k^+), \tag{9}$$

$$A^- = \left(\min_i(z_{i1}), \min_i(z_{i2}), \ldots, \min_i(z_{ik}),\right) = (z_1^-, z_2^-, \ldots, z_k^-). \tag{10}$$

If zero unitarization is used as the normative formula, it is:

$$z^+ = \underbrace{(1, 1, ..., 1)}_{k} \quad z^- = \underbrace{(0, 0, ..., 0)}_{k}. \tag{11}$$

Calculating the Euclidean distances of each object from the pattern and non-pattern is made according to the formulas:

$$d_i^+ = \sqrt{\sum_{j=1}^{k} (z_{ij} - z_j^+)^2}, \; d_i^- = \sqrt{\sum_{j=1}^{k} (z_{ij} - z_j^-)^2}, \; i = 1, 2, ..., n. \tag{12}$$

While the value of the synthetic feature is determined as follows (Hwang and Yoon, 1981):

$$\mu_i = \frac{d_i^-}{d_i^+ + d_i^-}, \tag{13}$$

wherein: $0 \leq \mu_i \leq 1$, $i = 1, 2, \ldots, n$.

The smaller the distance of a given object from the development pattern, and thus greater than the development non-pattern, the closer the value of the synthetic feature.

The final stage in the construction of the synthetic variable is the assessment of the correctness of the procedure of linear ordering of objects. This stage is often overlooked in the analysis of the linear ordering of objects. Perhaps this is due to the lack of publicly available computer software capable of performing this type of analysis. Assessing the optimality of the procedure of linear ordering of objects is a time consuming and quite complicated issue.

In the final stage, measures are used to characterize the effectiveness of individual methods for determining synthetic variables. These measures can be divided into five groups, each of which includes measures related to different properties of synthetic variables (Grabiński et al., 1989; Bąk, 2018; Sompolska-Rzechuła, 2020):

1) mapping compatibility ($m_1 - m_3$),
2) linear correlation of the synthetic variable with diagnostic variables ($m_4 - m_5$),
3) rank correlation of the synthetic variable with diagnostic variables ($m_6 - m_8$),
4) variability and concentration of the synthetic variable ($m_9 - m_{10}$),
5) taxonomic distance of the synthetic variable from the original variable ($m_{11} - m_{12}$).

The optimality measures for linear ordering procedures are shown in Table 1.

Table 1 The optimality measures for linear ordering procedures

<table>
<tr><th>Group of measures</th><th>Measure</th><th>Comments</th></tr>
<tr><td>(1)</td><td>$m_1 = \frac{\sum_{i=1}^{n-1}\sum_{j>i}^{n}(d_{ij} - \tilde{d}_{ij})^2}{\sum_{i=1}^{n-1}\sum_{j>i}^{n} d_{ij}^2}$
$m_2 = \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j>i}^{n}\left(\frac{d_{ij} - \tilde{d}_{ij}}{d_{ij}}\right)^2$
$m_3 = \frac{\sum_{i=1}^{n-1}\sum_{j>i}^{n}(d_{ij} - \tilde{d}_{ij})^2 / \tilde{d}_{ij}}{\sum_{i=1}^{n-1}\sum_{j>i}^{n} d_{ij}}$</td><td>$\tilde{d}_{ij}$ – average distance between the i-th and j-th object in the k-dimensional space of diagnostic variables,
d_{ij} – distance between the i-th and j-th object in the one-dimensional space of the synthetic variable,
n – number of objects</td></tr>
<tr><td>(2)</td><td>$m_4 = 1 - \frac{1}{k}\sum_{j=1}^{k} r_{qx_j}$
$m_5 = \frac{1}{k}\sum_{j=1}^{k} l_j$</td><td>$r_{qx_j}$ – correlation coefficient between the j-th diagnostic variable and synthetic variable,
k – number of features and
<table><tr><td></td><td>l_j</td><td>0</td><td>1</td><td>2</td><td>3</td></tr><tr><td rowspan="2">r_{qx_j}</td><td>max</td><td>1.0</td><td>0.5</td><td>0.0</td><td>–0.5</td></tr><tr><td>min</td><td>0.5</td><td>0.0</td><td>–0.5</td><td>–1.0</td></tr></table></td></tr>
<tr><td>(3)</td><td>$m_6 = 1 - \frac{1}{k}\sum_{j=1}^{k} \rho_{qx_j}$
$m_7 = \frac{1}{k}\sum_{j=1}^{k} l_j$
$m_8 = \frac{2}{ml}\sum_{i=1}^{n}\sum_{j=1}^{k} \lvert x_{ij}^{\cdot} - q_i^{\cdot} \rvert$</td><td>$\rho_{qx_j}$ – Spearman's rank correlation coefficient between the j-th diagnostic variable and the synthetic variable,
$x_{ij}^{\cdot}$ – rank of the i-th object due to the j-th primary variable,
$q_i^{\cdot}$ – rank of the i-th object due to synthetic variable,
$l = n^2$ for n even and
$l = n^2 - 1$ for odd n</td></tr>
</table>

Table 1 (continuation)

Group of measures	Measure	Comments
(4)	$m_9 = \frac{s_\mu}{\bar{\mu}}$ $m_{10} = \frac{s_\Delta}{\bar{\Delta}}$	$\bar{\mu}$ and s_μ – mean and standard deviation of the synthetic variable, $\bar{\Delta}$, s_Δ – mean and standard deviation for: $\Delta_i = \tilde{\mu}_i - \tilde{\mu}_{i-1}$ $(i = 1, \ldots, n-1)$ $\tilde{\mu}_i$ – ordered non-descending values of the synthetic variable
(5)	$m_{11} = \frac{1}{nk} \sum_{i=1}^{n} \sum_{j=1}^{k} \lvert x_{ij}' - \mu_i' \rvert$ $m_{11} = \left[\frac{1}{nk} \sum_{i=1}^{n} \sum_{j=1}^{k} \lvert x_{ij}' - \mu_i' \rvert^2 \right]^{1/2}$	x_{ij}' – standardized value of j-th primary variable for i-th object, μ_i' – standardized value of the synthetic variable for i-th object

Source: Own elaboration based on Grabiński et al. (1989)

In this paper, some modifications were made that relate to the determination of the value $\tilde{d}_{ij}$ as distance between the i-th and j-th object in the k-dimensional space of diagnostic variables, compared to the information contained in (Grabiński et al., 1989, pp. 122–123). The calculations include the values of diagnostic variables after normalization (zero unitarization). The distance $\tilde{d}_{ij}$ was determined as the average distance in the k-dimensional space. The introduced modifications made it possible to obtain comparable values of $\tilde{d}_{ij}$ and d_{ij} and partial measures m_l. In this paper was made some modification of measure m_9 too. Grabiński et al. (1989) introduced a "minus" sign for the measure m_9. In this paper, the value of the measure m_9 was calculated without the "minus" sign. It was assumed that the lower value of the measure m_9 indicates a lower diversity of objects in terms of the phenomenon under consideration.

The following measures from individual groups were used in this study: m_1, m_4, m_6, m_9 and m_{11}.

The aggregation of partial measures was performed according to the following formula (Bąk, 2015):

$$M_q = \sqrt{\sum_{l=1}^{g} m_l^2}\,, \tag{14}$$

where: M_q – aggregate measure, m_l – partial measure ($l = 1, \ldots, g$), g – number of partial measures.

2 RESULTS AND DISCUSSION

To achieve the goal, data from the EUROSTAT database (2017) for 28 European Union countries was used (Sompolska-Rzechuła, 2020):

X_1 – live births per 1 000 population,
X_2 – deaths per 1 000 population,
X_3 – infant deaths rate per 1 000 population,
X_4 – natural increase per 1 000 population,
X_5 – age dependency (population aged 0–14 and 65 and more per 100 persons aged 15–64),
X_6 – activity rate in %,
X_7 – employment rate in %,
X_8 – unemployment rate in %,
X_9 – at-risk-of poverty rate in %,
X_{10} – severely materially deprived people in %,
X_{11} – GDP per capita in thous euro,
X_{12} – investment rate in %,
X_{13} – industrial production (2015 = 100),
X_{14} – obtaining primary energy per 1 000 inhabitants from renewable energy sources (in tone),
X_{15} – final energy consumption per capita (in thous. kgoe),

X_{16} – share of high-tech exports in total exports in %,

X_{17} – net current account balance in % of GDP.

The parametric Hellwig method was used to eliminate strongly correlated features and the final set of diagnostic features was obtained, taking into account: X_8, X_{10}, X_{12}, X_{14}, X_{16}, X_{17}.

Table 2 presents the values of the basic descriptive parameters of the features finally adopted for the study.

Table 2 Summary statistics

Statistics	Variables					
	X_8	X_{10}	X_{12}	X_{14}	X_{16}	X_{17}
Mean	8.65	8.93	20.01	0.53	12.26	2.16
Median	7.70	5.65	19.85	0.37	10.50	1.85
Minimum	4.00	0.70	11.40	0.04	3.80	–5.30
Maximum	23.60	31.90	29.30	1.92	24.20	8.40
Standard deviation	4.48	7.44	3.45	0.47	6.18	3.74
Variation coefficient (%)	51.82	83.35	17.22	87.91	50.43	172.76
Skewness coefficient	1.94	1.56	0.19	1.85	0.46	0.07

Source: Own elaboration based on Eurostat (2017)

All features are characterized by strong or very strong volatility, in addition, X_8, X_{10} and X_{14} are characterized by strong right-sided asymmetry.

In 2016, the lowest unemployment rate was recorded in the Czech Republic and the highest – in Greece. In many countries (Belgium, Cyprus, Finland, Spain, France, Croatia, Ireland, Italy, Lithuania, Latvia, Portugal, Slovenia and Slovakia), the unemployment rate was higher than 7.70%, i.e. the median. However, in countries such as: Austria, Bulgaria, Germany, Denmark, Estonia, Hungary, Luxembourg, Malta, Netherlands, Poland, Romania, Sweden and United Kingdom the unemployment rate in 2016 did not exceed the median value.

The highest level of the deeper material deprivation rate was recorded in Bulgaria (31.9%), and the lowest – in Sweden (0.7%). In addition, the value of this indicator higher than the median was observed in Cyprus, Spain, Ireland, Hungary, Greece, Croatia, Italy, Lithuania, Latvia, Poland, Portugal, Romania and Slovakia. In countries such as: the Czech Republic, Germany, Denmark, Estonia, Finland, France, Luxembourg, Malta, Netherlands, Sweden, Slovenia and United Kingdom level of the deeper material deprivation rate was lower as median (5.7%).

The greatest diversity of countries is due to the current account balance as a percentage of GDP. Some countries, such as Belgium Cyprus, Greece, Finland, France, Lithuania, Poland, Romania, Slovakia and the United Kingdom recorded a negative balance in 2016. In Cyprus, the balance was the lowest and amounted to –5.3%. The highest positive balance of 8.4% was recorded in the Netherlands.

The average investment rate for the 28 EU countries amounted to 20.01% in 2016. In eleven countries, the investment rate above the average value was observed, and in Ireland its level was the highest and amounted to 29.30%. However, the lowest value (11.40%) occurred in Greece.

Malta (24.2%) has the highest percentage of high technology exports in total exports, while Portugal has the lowest – 3.8%. In addition, countries such as Austria, Cyprus, the Czech Republic, Germany, Estonia, France, Hungary, Luxembourg, the Netherlands, Sweden and the United Kingdom recorded a percentage of high technology exports in total exports above the average.

The non-pattern and pattern methods were used for comparison of the results of the linear ordering of European Union countries by socio-economic situation in 2016. In case of non-pattern methods, they were based on the zero unitarization to standardize variables and quotient transformation with an arithmetic mean. While in the analysis using standard methods at the stage of feature standardization, for the Hellwig and TOPSIS methods – the zero unitarization was used. The features: X_8 and X_{10} were recognized as destimulants and they were transformed into stimulants by means of quotient transformation as the inverse of the feature's value.

Table 3 presents the results of the linear ordering of EU countries by socio-economic situation in 2016.

Table 3 Results of the linear ordering of EU countries by socio-economic situation in 2016

Country	Method			
	a non-pattern object based on		a pattern object	
	zero unitarization (1)	quotient transformation (2)	Hellwig (3)	TOPSIS (4)
Austria (AT)	7	7	2	7
Belgium (BE)	17	20	16	17
Bulgaria (BG)	20	15	23	19
Croatia (HR)	21	17	22	21
Cyprus (CY)	26	28	27	22
Czech Republic (CZ)	6	13	7	6
Denmark (DK)	10	5	9	10
Estonia (EE)	8	8	3	8
Finland (FI)	11	9	11	11
France (FR)	13	16	13	13
Germany (DE)	3	3	4	3
Greece (EL)	28	27	28	28
Hungary (HU)	12	12	12	12
Ireland (IE)	4	10	10	4
Italy (IT)	24	18	24	26
Latvia (LV)	15	14	14	16
Lithuania (LT)	23	23	21	24
Luxembourg (LU)	9	6	8	9
Malta (MT)	2	4	6	2
Netherlands (NL)	5	2	5	5
Poland (PL)	19	21	18	20
Portugal (PT)	27	22	26	27
Romania (RO)	18	26	19	18
Slovakia (SK)	22	24	20	23
Slovenia (SI)	14	11	15	15
Spain (ES)	25	19	25	25
Sweden (SE)	1	1	1	1
United Kingdom (UK)	16	25	17	14

Source: Own elaboration based on Eurostat (2017)

Table 4 presents the evaluation of order compliance with selected methods measured by the Kendall rank correlation coefficient (1948, p. 82).

Table 4 Values of the Kendall rank correlation coefficients according to individual methods

Method	(1)	(2)	(3)	(4)
(1)	1.000	0.688	0.868	0.958
(2)	0.688	1.000	0.683	0.656
(3)	0.868	0.683	1.000	0.825
(4)	0.958	0.656	0.825	1.000

Source: Own elaboration based on Table 3

Assessment of order compliance with selected methods, measured by Kendall rank correlation coefficient, indicates the existence of significant links between country positions. The strongest correlation was observed between orders made using the non-pattern with zero unitarization and TOPSIS methods, which were obtained on the basis of zero unitarization and between the non-pattern method with Hellwig method. While the weakest relationship occurs between the results according to the following methods: non-pattern using the quotient transformation with the arithmetic mean and the pattern Hellwig method and TOPSIS based on zero unitarization.

When analysing the position occupied by individual countries, it can be seen that some countries took the same or similar position in individual orders, e.g. Austria, Estonia, France, Germany or Hungary. While in case of Bulgaria, Denmark or Ireland one can notice differences in the positions occupied in the obtained orders. The question arises, the results of which ordering should be adopted as optimal? In response to this question, help is provided by partial measures of the optimality of linear ordering procedures and the aggregate measure determined on their basis, the values of which for individual procedures are presented in Table 5.

Table 5 Values of the aggregate measure of optimality of procedures for linear ordering of EU countries in terms of socio-economic situation

Measure	Method			
	a non-pattern object based on zero unitarization (1)	a non-pattern object based on quotient transformation (2)	Hellwig based on zero unitarization (3)	TOPSIS method with zero unitarization (4)
M_q	0.860	1.101	0.925	0.729
m_1	0.331	0.157	0.238	0.204
m_4	0.463	0.641	0.411	0.405
m_6	0.435	0.467	0.578	0.431
m_9	0.379	0.514	0.500	0.304
m_{11}	0.288	0.542	0.212	0.220

Source: Own elaboration based on Eurostat (2017)

The optimality assessment of the linear ordering procedures can be performed using the following methods by comparing the results obtained with:

- all methods,
- a pattern object or a non-pattern object,

- methods based on the same way of standardizing features.

In assessing the optimality of linear ordering procedures, the criteria characterized in the chapter devoted to the research method were taken into account, i.e. mapping compatibility, linear and rank correlation of the synthetic variable with diagnostic variables, and variability of the synthetic variable and taxonomic distance of the synthetic variable from the original variable.

Taking into account the results of all methods, the most correct way of linear ordering of objects is to order EU countries obtained using the TOPSIS method based on zero unitarization (4). Also the results of this method are "better" compared to the results obtained according to Hellwig pattern method based on zero unitarization (3). The ordering obtained according to non-pattern method based on zero unitarization (1) gave more correct results compared to the ordering using non-pattern method based on quotient transformation (2).

Of the methods based on zero unitarization, the results obtained using the TOPSIS method (4) and then method non-pattern (1) were the most correct.

Taking into account the information obtained on the basis of the results included in Table 5, including the above-mentioned criteria, the results obtained using the TOPSIS method based on zero unitarization as the normative formula were considered the most correct synthetic variable (4).

In this order, Sweden came first, followed by Malta and Germany. The last place was occupied by Greece, Italy and Portugal occupied only slightly better places.

Sweden obtained its first place due to the favourable values of many features adopted in the study. The unemployment rate was 6.6% (only in the Czech Republic a lower value was observed – 4%). The in-depth deprivation rate was lowermost – 0.7%. In addition, the value referring to obtaining primary energy from renewable energy sources (thousand tonnes) per 1 000 inhabitants was at a high level. In 2016, Germany recorded a positive balance of the current balance of payments account in % GDP.

Greece occupied the last place in the linear ordering of countries, with the highest unemployment rate (23.6%) and the lowest investment rate (11.4%), a high deep deprivation rate of 22.2%, which is almost two and a half times higher than the average for all countries. Primary energy extraction from renewable energy sources (thousand tonnes) per 1 000 inhabitants was in Greece at one of the lower levels (18.95) and constituted only 16% of the highest value of this indicator, which concerned Austria. The share of exports of high technology products in total exports was also very low, at 4.6% with a maximum value of 24.2% for Malta. A negative value was recorded for the current account balance of payments in % GDP of – 0.6%.

CONCLUSION

The paper presents the results of a comparative analysis of four methods of linear ordering and two methods of variable standardization (quotient transformation and zero unitarization). The study was conducted on the basis of data on 28 European Union countries due to the level of socio-economic development in the context of sustainable development. The paper proposes a comparison of the results obtained in various methods concerning, e.g. the method of normalizing diagnostic features, the type of methods (based on a pattern object or a non-pattern object) and including all procedures. The assess of selected methods of linear ordering and selection of the optimal method was carried out on the basis of measures of correctness of the procedures of linear ordering of objects. Measures based on different properties of the synthetic variable were used in this paper. Some modifications of the measures were proposed regarding the measures of the mapping compliance and variability of the synthetic variable. The least correct procedure of linear ordering of objects among the methods selected for the study was non-pattern method based on quotient transformation. Comparing the results of two non-pattern methods based on different methods of standardization, the results obtained with the use of zero unitarization proved to be better. Taking into account all methods, the TOPSIS method with zero unitarization proved to be the most correct.

References

ANDERSON, T. W. *An Introduction to Multivariate Statistical Analysis.* 2nd Ed. New York: Wiley and Sons, 1984.

BĄK, A. Application of multivariate analysis methods comparative assessment of the state of the environment in the Dolnośląskie Voivodeship. *The Polish Statistician*, 2018, 1(680), pp. 7–20 (in Polish: Zastosowanie metod wielowymiarowej analizy porównawczej do oceny stanu środowiska w województwie dolnośląskim, *Wiadomości Statystyczne*).

BĄK, A. Problem of Choosing the Optimal Linear Ordering Procedure in the Pllord Package. *Research Papers of Wrocław University of Economics*, 2015, 384, pp. 33–41 (in Polish: Zagadnienie wyboru optymalnej procedury porządkowania liniowego w pakiecie pllord, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*). DOI: 10.15611/pn.2015.384.03

CERIN, P. Bringing Economic Opportunity into Line With Environmental Influence: a Discussion on the Coase theorem and the Porter and van der Linde hypothesis. *Ecological Economics*, 2006, pp. 209–225.

CHEN, C.-T. Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment. *Fuzzy Sets and Systems*, 2000, 114, pp. 1–9.

EUROSTAT. *Database – Eurostat* [online]. Eurostat, 2017. [cit. 10.9.2019]. <https://ec.europa.eu/eurostat/data/database>.

GRABIŃSKI, T. *Taxonometry Methods.* Cracow: Cracow University of Economics, 1992 (in Polish: *Metody taksonometrii*, Wydawnictwo Akademii Ekonomicznej).

GRABIŃSKI, T., WYDYMUS, S., ZELIAŚ, A. *Numerical Taxonomy Methods in Modeling Socio-Economic Phenomena.* Warsaw: PWN, 1989 (in Polish: *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*).

HARTIGAN, J. A. *Clustering Algorithms.* New York: Wiley and Sons, 1975.

HELLWIG, Z. Application of the Taxonomic Method to the Typological Division of Countries by Level of Their Development as well as Resources and Structure of Qualified Staff. *Statistical Review*, 1968, 4, pp. 307–327 (in Polish: Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr, *Przegląd Statystyczny*).

HWANG, C. L. AND YOON, K. *Muliple Attribute Decision Making. Methods and Applications.* Berlin: Springer, 1981.

KENDALL, M. G. *Rank Correlation Methods.* Londyn: Charles Griffin & Company Limited, 1948.

KUKUŁA, K. *Zero Unitarization Method.* Warsaw: Scientific Publisher PWN, 2000 (in Polish: *Metoda unitaryzacji zerowanej*, Wydawnictwo Naukowe PWN).

LATOSZEK, E. The Concept of Sustainable Development in UN Theory and Practice. In: LATOSZEK, E., PROCZEK, M., KRUKOWSKA, M. (eds.) *Sustainable Development and Global Public Goods in the Theory and Practice of International Organizations*, Warsaw: Warsaw School of Economics Press, 2016. (in Polish: Koncepcja zrównoważonego rozwoju w teorii Ii praktyce ONZ, *Zrównoważony rozwój a globalne dobra publiczne w teorii i praktyce organizacji międzynarodowych*).

MORRISON, D. F. *Multidimensional Statistical Analysis.* Warsaw: PWN, 1990 (in Polish: *Wielowymiarowa analiza statystyczna*).

NOWAK, E. *Taxonomic methods in the classification of socio-economic objects.* Warsaw: PWE, 1990 (in Polish: *Metody Taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*).

PLUTA, W. *Multidimensional Comparative Analysis in Economic Research.* Warsaw: PWE, 1977 (in Polish: *Wielowymiarowa analiza porównawcza w badaniach ekonomicznych*).

SEBER, G. A. F. *Multivariate Observations.* New York: Wiley and Sons, 1984.

SOMPOLSKA-RZECHUŁA, A. Selection of the method of linear ordering using the example of assessing the level of socio-economic development of European Union countries. *Research Papers of Wrocław University of Economics*, 2020, 64(7), pp. 118–129 (in Polish: *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*).

SOMPOLSKA–RZECHUŁA, A. *The Measurement and Assessment of the Quality of Life in the Region.* Szczecin: University Publishing House of the West Pomeranian University of Technology in Szczecin, 2018 (in Polish: *Pomiar i ocena jakości życia w ujęciu regionalnym*, Wydawnictwo Uczelniane Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie).

UNITED NATIONS GENERAL ASSEMBLY. *Report of the World Commission on Environment and Development: Our Common Future* [online]. Oslo, Norway: United Nations General Assembly, Development and International Cooperation Environment, 1987. [cit. 26.10.2019]. <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>.

WYSOCKI, F. *The Methods of Taxonomy for Recognition of Economic Types in Agriculture and Rural Areas*. Poznan: University Publisher Poznan University of Life Sciences, 2010 (in Polish: *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*, Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu).

Environmental Kuznets Curve for CO_2 Emissions in Middle-Income Countries: a Dynamic Spatial Panel Data Analysis

Hanen Ragoubi[1] | *University of Sousse, Sousse, Tunisia*
Zouheir Mighri[2] | *University of Jeddah, Jeddah, Saudi Arabia*

Abstract

This paper examines the carbon dioxide (CO_2) Environmental Kuznets curve (EKC) hypothesis of a balanced panel of 50 middle-income countries over the period 1996–2013 using a dynamic spatial panel data model with country and time-period fixed effects. Using a Bayesian comparison approach, we systematically searched for the most suitable spatial weights matrix describing the spatial arrangement of the countries in the sample. We found substantial spatial dependence effect in CO_2 emissions across the sample of middle-income countries, highlighting the influence these countries exert on their neighbors. Besides, the empirical results showed that the relationship between economic growth and CO_2 emissions shaped as an inverted-U trajectory. Furthermore, it has been found that trade openness and energy intensity are the main factors on slightly increasing CO_2 emissions, while the urbanization contributes to relative decrease in CO_2 emissions.

Keywords

CO_2 emission, EKC hypothesis, dynamic spatial panel, Bayesian comparison, spillover effects

JEL code

C21, C23, P25, Q53, Q56

INTRODUCTION

Over the three last decades, global warming, and particularly increasing temperatures, have a significant deep impact on economic productivity (Burke et al., 2015). Indeed, economic production has warmed the earth by releasing mass emissions of greenhouse gas in the atmosphere. In particular, the ever-increasing global emissions of CO_2 appear to be aggravating this issue. Accordingly, both the global environmental change and sustainable development become the critical challenge for human beings today (Roy Chowdhury and Moran, 2012). Exploring the potential relationship between economic growth

[1] Research Laboratory for Economy, Management and Quantitative Finance (LaREMFiQ), University of Sousse, Rue Khalifa Karoui Sahloul 4, Sousse, Tunisia. E-mail: hragoubi2205@gmail.com.

[2] College of Business, Department of Finance and Economics, University of Jeddah, Hamzah Ibn Al Qasim Street, Al Sharafeyah, Jeddah 23218, Kingdom of Saudi Arabia. Also Higher Institute of Finance and Taxation of Sousse, University of Sousse, Rue Khalifa Karoui Sahloul 4, Sousse, Tunisia. E-mail: zmighri@gmail.com.

and environmental degradation is also becoming a necessity in order to provide policy recommendations for taking a sustainable development trend in countries.

To explore the way of sustainable development, Grossman and Krueger (1991 and 1995) put forward the EKC theory to depict the relationship between economic growth and environmental degradation. Different econometric methodologies[3] have been used to investigate the CO_2 EKC hypothesis in different countries and regions. However, mixed empirical results are reported (Richmond and Kaufmann, 2006; Aldy, 2006; Galeotti et al., 2006; Kaika and Zervas, 2013a, 2013b, among others). Scholars have shown that the formulation of the EKC hypothesized multiple shaped EKC such as U, inverted-U, N, etc. For instance, Grossman and Krueger (1991) pointed out that economic growth can improve environmental quality after an economy has reached an adequate level of development. Furthermore, there were pieces of evidence that the testing results depended on the specific econometric models (Roy Chowdhury and Moran, 2012).

The mixed results further confirm that studies based on traditional cross-sectional panel data or time series techniques would provide incorrect inferences because of ignoring the spatial correlations dimension. Compared with traditional econometric methods, the spatial econometric techniques can be used to explore whether the local regional economic performances depend on the neighbors or not. While conventional econometric approaches have been used in most EKC studies, there is little evidence in the context of the nexus between economic growth and CO_2 emissions using spatial econometric techniques (Zhao et al., 2014; Kang et al., 2016; Meng et al., 2017; Meng and Huang, 2018; You and Lv, 2018).

As acknowledged by LeSage and Pace (2009), ignoring spatial dependence would lead to biased estimated parameters. Besides, Roy Chowdhury and Moran (2012) argued that spatial effects represent an important factor influencing the impact of economic growth on CO_2 emissions since several environmental problems, including CO_2 emissions, are inherently spatial. Furthermore, Anselin (2001) argues that spatial units (countries, states, counties, provinces, cities, etc.) can interact strongly with one another via channels such as trade, technological spillover, capital inflow, and common political, economic, and environmental policies. Recent research suggests that the closer the two countries are in terms of geographic distance, the more likely the economic activities and environmental degradation within each country will affect one another (You and Lv, 2018). In other words, economic growth and CO_2 emissions across countries are not independent. If such dependencies are not considered, some bias will be produced when estimating the EKC. As argued by Elhorst (2010a, 2010b), spatial econometric techniques provide ways to test and accommodate many forms of dependence among observations.

This study contributes to the empirical literature in several ways. First, it offers a more rigorous examination of the relationship between CO_2 emissions and economic growth for middle-income countries. The influence factors of CO_2 emissions are not only per capita real income but also other social, economic and industrial variables such as trade openness, urbanization, energy intensity and population which will be incorporated in the economic model to improve the accuracy of EKC fitting. Second, this paper uses the recently developed dynamic spatial panel models with controls for spatial and time-specific effects in order to capture the spatial interactions between explanatory variables and CO_2 emissions focusing on the middle-income countries. Specifically, this study seeks to explore the CO_2 emissions Kuznets curve in middle-income countries, and a comparative analysis between the non-spatial panel data model and the dynamic spatial panel data model is conducted to validate the spatial spillovers effects of variables in order to provide more rigorous references for policymakers. Third, using a Bayesian comparison approach developed by LeSage (2014, 2015), this study tests and compares simultaneously four frequently used dynamic spatial panel data models and twelve spatial weight matrices describing the mutual relationships among the middle-income countries, all within a common framework, which helps clarify the impact of neighboring countries on CO_2 emissions.

[3] A large strand of empirical literature is summarized in Table 1.

The remainder of this paper is organized as follows. Section 1 outlines the theoretical framework of the empirical model specification, the conventional spatial autocorrelation measures and the methodology of dynamic spatial panel data models. Section 2 provides a description of the data. Section 3 is devoted to the empirical estimation results and discussions. Final section concludes this paper and provides some policy suggestions.

Table 1 Summary of previous EKC studies on CO_2 emissions

Authors	Time period	Regions	Econometric methodology	Shaped EKC
Holtz-Eakin and Selden (1995)	1951–1986	130 countries	panel data	no EKC relationship
Carson et al. (1997)	1990	US states	cross-sectional data	inverted-U-shaped relationship
Roberts and Grimes (1997)	1962–1991	low-medium-high income countries	time series	inverted-U-shaped relationship for rich countries no EKC relationship for low/medium income countries
Lim (1997)	1980s onwards	South Korea	time series	no EKC relationship
Moomaw and Unruh (1997)	1950–1992	16 industrial OECD countries	panel data	N-shaped relationship
Schmalensee et al. (1998)	1950-1990	141 countries	panel data	inverted-U-shaped relationship
De Bruyn et al. (1998)	1960–1993 intervals	Netherlands, W. Germany, UK, USA	time series	no EKC relationship
Galeotti and Lanza (1999)	1970–1996	110 countries	panel data	inverted-U-shaped relationship
Agras and Chapman (1999)	various years	34 countries	panel data	no EKC relationship
Perrings and Ansuategi (2000)	1990	114 countries	panel data	no EKC relationship
Lindmark (2002)	1870–1997	Sweden	time series	inverted-U-shaped relationship
Friedl and Getzner (2003)	1960–1999	Austria	time series	N-shaped relationship
Cole (2004)	1980–1997	21 countries	panel data	inverted-U-shaped relationship
Dijkgraaf and Vollebergh (2005)	1960–1997	OECD countries	panel data	inverted-U-shaped relationship
Aldy (2005)	1960–1999	US states	panel data	inverted-U-shaped relationship in few states no EKC relationship (consumption model)
Azomahou et al. (2006)	1960–1996	100 countries	panel data	no EKC relationship
Richmond and Kaufmann (2006)	1973–1997	36 countries	panel data	no EKC relationship
Lantz and Feng (2006)	1970–2000	5 Canadian regions	panel data	no EKC relationship
Kunnas and Myllyntaous (2007)	1800–2003	Finland	time series	no EKC relationship
Coondoo and Dinda (2008)	1960–1990	88 countries	panel data	inverted-U-shaped relationship for Europe no EKC relationship for whole
Lee et al. (2009)	1960–2000	89 countries	panel data	N-shaped relationship for the whole panel inverted-U-shaped relationship in middle-income, American and European countries
Aslanidis and Iranzo (2009)	1971–1997	77 Non-OECD countries	panel data	no EKC relationship

Table 1 (continuation)

Authors	Time period	Regions	Econometric methodology	Shaped EKC
Dutt (2009)	1960–2002	124 countries	panel data	no EKC relationship (1960-1980) Inverted-U-shaped relationship (1984-2002)
Halicioglu (2009)	1960–2005	Turkey	time series	no EKC relationship
Jalil and Mahmud (2009)	1971–2005	China	time series	inverted-U-shaped relationship
Aslanidis and Iranzo (2009)	1971–1997	non-OECD countries	panel data	no EKC relationship
Narayan and Narayan (2010)	1980–2004	43 developing countries	panel data and time series	inverted-U-shaped relationship in 15 countries (time series) inverted-U-shaped relationship in Middle Eastern and South Asian countries (panel data)
Acaravci and Ozturk (2010)	1960–2005	19 European countries	time series	inverted-U-shaped relationship in 2 countries
Iwata et al. (2011)	1960–2003	28 countries (17 OECD, 11 non-OECD countries)	panel data	no EKC relationship
Wang et al. (2011)	1995–2007	28 China's provinces	panel data	U-shaped relationship
Jaunky (2011)	1980–2005	36 high-income countries	panel data	inverted-U-shaped relationship in 5 countries no EKC relationship for whole panel
Fosten et al. (2012)	1830–2003	United Kingdom	time series	inverted-U-shaped relationship
Esteve and Tamarit (2012)	1857–2007	Spain	time series	inverted-U-shaped relationship
Du et al. (2012)	1995–2009	29 China's provinces	panel data	no EKC relationship
Ahmed and Long (2012)	1971–2008	Pakistan	time series	inverted-U-shaped relationship
Saboori et al. (2012)	1980–2009	Malaysia	time series	inverted-U-shaped relationship
Saboori and Sulaiman (2013)	1980–2009	Malaysia	time series	no EKC relationship
Ozturk and Acaravci (2013)	1960–2007	Turkey	time series	inverted-U-shaped relationship
Burnett et al. (2013)	1970–2009	48 US states	spatial panel data	inverted-U-shaped relationship
Onafowora and Owoye (2014)	1970–2010	8 countries	time series	inverted-U-shaped relationship in two of the eight countries N-shaped relationship in six of the eight countries
Shahbaz et al. (2014a)	1971–2010	Tunisia	time series	inverted-U-shaped relationship
Farhani and Ozturk (2015)	1971–2012	Tunisia	time series	no EKC relationship
Apergis and Ozturk (2015)	1990–2011	14 Asian countries	panel data	inverted-U-shaped relationship
Yin et al. (2015)	1999–2011	China (29 provinces)	panel data	inverted-U-shaped relationship
Wang et al. (2016b)	1995–2011	30 China's provinces	spatial panel data	N-shaped relationship
Kang et al. (2016)	1997–2012	30 China's provinces	spatial panel data	inverted-N-shaped relationship
Li et al. (2016)	1996–2012	28 China's provinces	spatial panel data	inverted-U-shaped relationship
Wang and Liu (2017a)	1992–2013	341 China's cities	panel data and dynamic panel data	inverted-U-shaped relationship
Meng and Huang (2018)	1995–2012	331 China's cities	spatial panel data	no EKC relationship
You and Lv (2018)	1985–2013	83 developed and developing countries	spatial panel data	inverted-U-shaped relationship

Source: Created by the authors

1 THEORETICAL FRAMEWORK AND METHODOLOGY

1.1 EKC Hypothesis

Originally, EKC is an empirical hypothesis that characterizes an inversely U-shaped curve for the relationship between economic growth and environmental quality. Several indices of environmental quality degenerate with economic growth. As suggested by Grossman and Krueger (1995), the environment deterioration starts to decrease after reaching a threshold. Furthermore, Maddison (2006) pointed out that development may promote environmental quality as a result of economies of scale from pollution reduction, technological upgrade, industrial structure escalation, and public's demand for a clean environment. In this paper, the considered model for the EKC is a polynomial function type that is expressed as follows:

$$Y_{it} = \alpha_i + \beta_1 X_{it} + \beta_2 X_{it}^2 + \beta_3 Z_{it} + \varepsilon_{it} , \tag{1}$$

where Y stands for the indices of environmental degradation, while X refers to the economic growth level, usually measured by per capita Gross Domestic Product (GDP), and Z includes other influential factors for the environment. The polynomial function form of EKC offers to us an adequate tool to estimate the nonlinear relationship (if it exists) between economic growth and CO_2 emission.

1.2 STIRPAT Model

In this paper, we use the STIRPAT model (Dietz and Rosa, 1997; York et al., 2003) as our theoretical foundation to test the existence of an EKC for CO_2 emissions related to affluence. Ehrlich and Holdren (1971) first proposed the concept of IPAT (Influence, Population, Affluence, and Technology). The IPAT model relates environmental impact to population, affluence and technology. Nevertheless, this model is only an overly simplified function form and just indicates that the impact of human activities on the environment can fully be differentiated into population, affluence, and technology effects. Therefore, the IPAT model cannot estimate to what extent a specific factor affects the environment in such a framework, not to mention test any hypothesis. An additional limitation is that the IPAT model has been criticized as being primarily a mathematical equation which is not suitable for hypothesis testing, and also assuming a rigid proportionality between effects and factors.

To overcome these limitations, Dietz and Rosa (1997) proposed a stochastic version of IPAT, known as STIRPAT and later refined by York et al. (2003), expressed by the following equation:

$$I_{it} = \alpha_0 P_{it}^{\alpha_1} A_{it}^{\alpha_2} T_{it}^{\alpha_3} e_{it} , \tag{2}$$

where I denotes the environmental impact, P, A and T indicate human activities, i.e., respectively, population, affluence (per capita), and technological influences (per unit of economic activity). α_0, α_1, α_2 and α_3 are coefficients to be estimated and e denotes the random disturbance (the proportionality of IPAT model pre-assume $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 1$). The subscript i refers to the ith country and vary across observations.

The regression form of the STIRPAT model for estimation and hypothesis testing is obtained by logarithmic transformation of the variables in Formula (2). In this case, the coefficients α_1, α_2, and α_3 stand for the Ecological Elasticity (EE) which measures the sensitivity of environmental impacts to a change occurring in a driving force. It is defined as the proportion of change in environmental impacts due to its significant determinants. Using natural logarithms, the STRIPAT model can be converted to a convenient linear specification for panel estimation:

$$lnI_{it} = a_0 + \alpha_1 lnP_{it} + \alpha_2 lnA_{it} + \alpha_3 lnT_{it} + lne_{it} . \tag{3}$$

The above basic model analyses the impacts of population (P), economic development (A) and industrial structure (T) on the environmental impacts, but ignores other important factors influencing CO_2 emissions. According to the EKC hypothesis, CO_2 emissions is a function of par capita GDP and square of per capita GDP (Kasman and Duman, 2015; Kang et al., 2016; Meng and Huang, 2018; You and Lv, 2018, among others). Therefore, a quadratic or higher term of affluence can enter the STIRPAT specification. Besides, we further investigate the effects of additional factors on CO_2 emissions such as urbanization, energy intensity and trade openness (Martínez-Zarzoso et al., 2007; Pao and Tsai, 2011; Madlener and Sunak, 2011; Zhang et al.,2014; Al-Mulali et al., 2015; Kang et al., 2016; You and Lv, 2018; Lv and Xu, 2019, among others). Accordingly, we applied an augmented STIRPAT for our study purpose:

$$lnCO_{2it} = a_0 + \alpha_1 ln(POP_{it}) + \alpha_2 ln(RGDP_{it}) + \alpha_3 ln(RGDP_{it})^2 + \alpha_3 ln(TECH_{it}) + \alpha_4 ln(TRO_{it}) + \alpha_5 ln(URBA_{it}) + \alpha_6 ln(EI_{it}) + \alpha_7 CV_{it} + \mu_i + \eta_t + \varepsilon_{it}\,, \tag{4}$$

where CO_2 denotes per capita carbon dioxide emissions; *TRO* represents the trade openness; *POP* is the total population and measures the impact of demographic factors on CO_2 emissions; *RGDP* stands for per capita real GDP, which is seen as a proxy for economic factors; *URBA* denotes the urbanization level, which is typically associated with increased economic activity resulting in high energy consumption, and thus accelerating the emission of CO_2 (Martínez-Zarzoso and Maruotti, 2011; Adams and Klobodu, 2017); *TECH* is the technological improvement, measured by percentage of industrial activity with respect to total production, and represents a proxy for the level of environmentally damaging technology (Martínez-Zarzoso et al., 2007); *EI* refers to the energy intensity[4] per unit of GDP and can be considered as a proxy for energy consumption (Martínez-Zarzoso et al., 2007); μ_i is the individual fixed effect, which controls for all space-specific time-invariant variables that if omitted could potentially bias the coefficient estimates; η_t denotes the time period effects; ε is the standard error term; and CV_{it} stands for the potential control variables that could influence the CO_2 emissions.

In general, the estimation of the empirical model, i.e., Formula (4), tests the statistical significance of the coefficients α_2 and α_3. The following cases may occur (Dinda, 2004; Kaika and Zervas, 2013a):

i. If $\alpha_2 = \alpha_3 = 0$, then there is no relationship between economic growth and CO_2 emissions.
ii. If $\alpha_2 > 0$ and $\alpha_3 = 0$, then a monotonic increasing or linear relationship exists between economic growth and CO_2 emissions.
iii. If $\alpha_2 < 0$ and $\alpha_3 = 0$, then a monotonic decreasing or linear relationship exists between economic growth and CO_2 emissions.
iv. If $\alpha_2 > 0$ and $\alpha_3 < 0$, then an inverted-U-shaped relationship (EKC) exists between economic growth and CO_2 emissions.
v. If $\alpha_2 < 0$ and $\alpha_3 > 0$, then a U-shaped relationship exists between economic growth and CO_2 emissions.

Note that only the (iv) case indicates an EKC-relationship. Accordingly, the EKC is a specific form of the CO_2-income relationship. If the (iv) case holds, then the turning point is calculated as follows:

$$RGDP^{*} = exp(-(\alpha_2/2\alpha_3))\,. \tag{5}$$

[4] Energy intensity was measured as energy use divided by GDP at purchasing power parity (PPP) prices, where energy use refers to apparent consumption (production + imports – exports).

1.3 Spatial autocorrelationl

Spatial autocorrelation is a spatial data analysis method which is used to examine the degree of spatial dependence or autocorrelation in spatial data. It includes i) the global spatial autocorrelation which is used to estimate the overall degree of spatial dependence, and ii) the local indicators of spatial association (LISA) which is used to assess the impact of individual locations on the magnitude of the global statistic and to identify the locations and types of clusters. The spatial weights were created by rook contiguity rule and applied to describe the spatial relationships among countries. We explored the spatial distribution of per capita CO_2 emissions from 50 middle-income countries by calculating the Global Moran's I (Moran, 1950) and LISA (Anselin, 1995) using GeoDa software. The Global Moran's I statistic can be specified as follows:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \cdot \frac{\sum_{j=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})}\,; i \neq j\,, \tag{6}$$

where $-1 \leq I \leq 1$; y_i and y_j are the values of the per capita CO_2 emissions of countries i and j, respectively; $\bar{y}$ is equal to the average of the per capita CO_2 emissions of all countries; w_{ij} is the element in row i column j of a spatial weights matrix and denotes the spatial weight between country i and country j; and n is the number of countries.

At a given level of statistical significance, $I > 0$ points to positive spatial autocorrelation, and the greater the value of I, the more obvious the spatial correlation. $I < 0$ refers to negative spatial autocorrelation, and the smaller the value of I, the greater the spatial difference. Otherwise, $I = 0$ points to a random spatial distribution. As argued by Anselin and Florax (1995), a significant positive Moran's I value indicates spatial clustering, while a significative negative Moran's I value indicates spatial dispersion across the sample of geographical units.

To evaluate the statistical significance of the Global Moran's I, both a z-score and p-value can be calculated. The z_I-score for the statistic I is computed as follows:

$$z_I = \frac{I - E(I)}{\sqrt{V(I)}} \longrightarrow N(0,1)\,, \tag{7}$$

where $E(I) = -1/(n-1)$; $V(I) = E(I^2) - E^2(I)$.

Alternatively, LISA is calculated as follows:

$$I_i = \frac{z_i}{\sum_{i=1}^{n} z_i^2} \times z_i^{\circ}\,, \tag{8}$$

where z_i denotes the observation for country i on per capita CO_2 emissions as a deviation from the mean, and z_i° is the spatial lag for location i, obtained as follows:

$$z_i^{\circ} = \sum_{j=1}^{n} w_{ij} z_j\,. \tag{9}$$

1.4 Dynamic spatial panel data models

A spatial econometric model is a linear regression model extended to include spatial interaction effects among the dependent variable, the explanatory variables, the error terms, or some combination thereof. Including all spatial lags yields a so-called general nesting spatial (GNS) model (Elhorst, 2014a, 2014b). When accounting for the dependent variable lagged one period, such a specification is known as a dynamic GNS model. The econometric counterpart of the dynamic GNS model reads, in vector form, as:

$$\boldsymbol{Y}_t = \tau\boldsymbol{Y}_{(t-1)} + \delta\boldsymbol{W}\boldsymbol{Y}_t + \eta\boldsymbol{W}\boldsymbol{Y}_{(t-1)} + X_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\,\boldsymbol{\theta} + \boldsymbol{\mu} + \lambda_t\,\boldsymbol{\iota}_N + \boldsymbol{v}_t\,, \tag{10}$$

$$\boldsymbol{v}_t = \lambda\boldsymbol{W}\boldsymbol{v}_t + \boldsymbol{\varepsilon}_t\,, \tag{11}$$

where $\boldsymbol{Y}_t$ is an $N \times 1$ vector consisting of one observation of the dependent variable for every spatial unit ($i = 1, \ldots, N$) in the sample at a particular point in time $t(t = 1, \ldots, T)$, which for this study is the CO_2 emissions ; $\boldsymbol{X}_t$ denotes an $N \times K$ matrix of exogenous or predetermined explanatory variables. Note that a vector or a matrix with subscript $t - 1$ stands for its serially lagged value, while a vector or a matrix premultiplied by $\boldsymbol{W}$ denotes its spatially lagged value. Moreover, the $N \times N$ matrix $\boldsymbol{W}$ denotes a non-negative matrix of known constants describing the spatial arrangement of the spatial units in the sample. It should be stressed that the diagonal elements of the matrix $\boldsymbol{W}$ are set to zero by assumption, since no spatial unit can be viewed as its own neighbor. Furthermore, the parameters τ, δ and η denote the response parameters of successively the dependent variable lagged in time, $\boldsymbol{Y}_{t-1}$, the dependent variable lagged in space, $\boldsymbol{W}\boldsymbol{Y}_t$, and the dependent variable lagged in both space and time, $\boldsymbol{W}\boldsymbol{Y}_{t-1}$. The variables $\boldsymbol{W}\boldsymbol{Y}_t$ and $\boldsymbol{W}\boldsymbol{Y}_{t-1}$ stand for contemporaneous and lagged endogenous interaction effects among the dependent variables. The symbols $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ stand for $K \times 1$ vectors of the response parameters of the exogenous explanatory variables. Furthermore, the error term specification consists of different components: the vector $\boldsymbol{v}_t$ that is assumed to be spatially correlated with autocorrelation coefficient λ; the $N \times 1$ vector $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ that consists of i.i.d. disturbance terms, which have zero mean and finite variance σ^2; the $N \times 1$ vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)'$ that contains spatial specific effects μ_i and is meant to control for all spatial-specific, time-invariant variables whose omission could bias the estimates in a typical cross-sectional study; and the time specific effects $\lambda_t(t = 1, \ldots, T)$, where $\boldsymbol{\iota}_N$ is a $N \times 1$ vector of ones meant to control for all time-specific, unit-invariant variables whose omission could bias the estimates in a typical time-series study.

It should be mentioned that spatial- and time period-specific effects can be treated as fixed or random effects. Otherwise, direct interpretation of the coefficients in the dynamic GNS model is not straightforward since they do not represent true partial derivatives (LeSage and Pace, 2009). Elhorst (2012, 2014a, 2014b) show that the matrix of (true) partial derivatives of the expected value of the dependent variable with respect to the kth independent variable for $i = 1, \ldots, N$ in year t for the long-term is given by the $N \times N$ matrix:

$$\left[\frac{\partial E(\boldsymbol{Y})}{\partial x_{1k}} \cdots \frac{\partial E(\boldsymbol{Y})}{\partial x_{Nk}}\right] = [(1-\tau)\boldsymbol{I} - (\delta+\eta)\boldsymbol{W}]^{-1}[\beta_k\boldsymbol{I}_N + \theta_k\boldsymbol{W}]\,, \tag{12}$$

whose diagonal elements represent long-term impacts on the dependent variable of unit 1 up to N if the kth explanatory variable in the own country changes, while its off-diagonal elements represent the long-term impacts on the dependent variable if the kth explanatory variable x_k in other countries changes. The average diagonal element of this matrix can be used as a summary indicator for the direct effect, whereas the average row sum of its off-diagonal elements represents a summary indicator of the spillover effect. Furthermore, these impacts are independent of t since the spatial weight matrix W is not time-varying, and error terms drop out due to the use of expectations.

As acknowledged by LeSage and Page (2009), the direct effect is defined as the average diagonal element of the full $N \times N$ matrix expression on the right-hand side of Formula (12); the indirect effect (i.e. country spillover effects) is the average row or column sum of the off-diagonal elements. Moreover, short-term direct and country spillover effects can be obtained by setting $\tau = \eta = 0$.

It should be stressed that the dynamic GNS model is problematic since is its parameters are not identified (Anselin et al., 2008; Elhorst, 2014a, 2014b). Indeed, the interaction effects among the dependent variable and the error terms cannot be distinguished formally, if the interaction effects among the explanatory variables are also included. Therefore, one of the two spatial interaction effects should be excluded. If the spatial interaction effects for the dependent variable are excluded ($\delta = \eta = 0$), the dynamic SDEM specification results, while the spatial multiplier matrix $[(1 - \tau)\boldsymbol{I} - (\delta + \eta)\boldsymbol{W}]^{-1}$ reduces to $1/(1 - \tau)\boldsymbol{I}$.

If the spatial interaction effects among the error terms is left aside ($\lambda = 0$), a dynamic spatial Durbin model (SDM) results. Although the SDM specification does not account for interaction effects among the error terms, which reduces the efficiency of the parameter estimates, it does not affect the consistency of the parameter estimates. Besides, it does not influence the direct or spillover effects derived from Formula (12).

As pointed out by Anselin et al. (2008), LeSage and Pace (2009), and Elhorst (2014a, 2014b), among others, an important difference between the SDEM and SDM specifications is that the country spillover effects in the first model are local, whereas in the second model they are global in nature. Local spillovers occur at other countries only if they are connected to each other. In other words, local spillovers occur when $\delta = 0$ and $\theta \neq 0$, and countries are connected. If two countries i and j are unconnected, such that $w_{ij} = 0$, a change in x_{ik} of country i cannot affect the dependent variable of country j, and vice versa. Global spillovers instead occur when $\delta \neq 0$ and $\theta = 0$, regardless of whether countries are connected, so a change to x_{ik} of country i due to the spatial multiplier matrix $(\boldsymbol{I} - \delta\boldsymbol{W})^{-1}$ gets transmitted to all other countries, even if the two countries are unconnected, i.e., $w_{ij} = 0$.

If CO_2 emissions at a local level can spread to other countries across the continent or around the world, even if they are not directly connected, then the SDM or SAR specifications make more sense, due to their ability to capture such global spillovers. If other countries are connected to each other, the SDEM specification may be more appropriate since it captures only local country spillovers. Otherwise, the choice between local and global spillovers depends on the specification of the spatial weight matrix W. It should be stressed that a sparse spatial weight matrix with only a limited number of non-zero elements, such as a binary contiguity matrix, is more likely to occur in combination with a global spillover model ($\delta \neq 0$, $\theta = 0$), while a dense spatial weight matrix in which many off-diagonal elements are non-zero (e.g. inverse distance matrix) is more likely in combination with a local spillover model ($\delta = 0$, $\theta \neq 0$). Therefore, the choice of spatial model and spatial weight matrix might be improved if they take place within a common framework.

In this paper, we employ a Bayesian comparison approach (LeSage, 2014; LeSage, 2015) in order to choose between a global spillover model, i.e., SDM, and a local spillover model, i.e., SDEM, as well as to choose between different potential specifications of the spatial weight matrix W. It should be noted that this approach allows determining the Bayesian posterior model probabilities of the SDM and SDEM specifications given a particular spatial weight matrix, as well as the Bayesian posterior model probabilities of different spatial weight matrices given a particular spatial panel model specification. These probabilities are based on the log marginal likelihood of a spatial panel model obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. If the log marginal likelihood value of one spatial panel model or of one spatial weight matrix W is higher than that of another model or another W, the Bayesian posterior model probability is also higher. It should be stressed that the classical LR, Wald and/or LM statistics compare the performance of one spatial model against another spatial model based on specific parameter estimates within the parameter space. However, the main strength of the Bayesian comparison approach is that it compares the performance of one spatial model against another spatial model on their whole parameter space (LeSage, 2014; LeSage, 2015). Furthermore, statistical inferences drawn on the log marginal likelihood function values for the SDM and SDEM models are further justified since they have the same set of explanatory variables, i.e., X_t and WX_t, and are based on the same uniform prior for δ and λ. This prior takes the following form:

$$p(\delta) = p(\lambda) = 1/D\,, \tag{13}$$

where:

$$D = 1/\omega_{max} - 1/\omega_{min}\,, \tag{14}$$

and ω_{max} and ω_{min} denote respectively the largest and the smallest (negative) eigenvalue of the spatial weight matrix W. Note that this prior requires no subjective information on the part of the practitioner since it relies on the parameter space ($1/\omega_{min}$, $1/\omega_{max}$) on which δ and λ are defined, where $\omega_{max} = 1$ if W is row normalized. Finally, and depending on the outcomes of the Bayesian comparison approach, either the SDM or the SDEM model is estimated using maximum likelihood estimation (MLE). Then, the estimation results could serve to test the following null hypotheses:

$$H_0 : \theta = 0 \text{ and } = 0\,, \tag{15}$$

$$H_0 : \theta + \delta\beta = 0 \text{ and } + \delta\tau = 0\,. \tag{16}$$

That is, it is possible to test whether the dynamic SDM might be reduced to a dynamic SAR model or dynamic SEM. Both tests follow a chi-squared distribution with $K + 1$ degrees of freedom (i.e., the number of spatially lagged explanatory variables and the spatially lagged dependent variable) and take the form of a Wald test, since the simplified models have not been estimated.

2 DATA AND VARIABLES

In this paper, we use a balanced panel sample of 50 countries[5] over the period 1996–2013. In contrast to high income countries, time series data on energy use in many middle-income countries are very limited. Therefore, we limited our sample to 50 middle-income countries due to the availability of reliable data. Furthermore, the beginning of the sample period is motivated by the fact that the transition of several middle-income countries from socialism to capitalism has likely led to a structural break in environmental policy in general. The dependent variable is CO_2 emissions (metric tons of per capita carbon dioxide emissions), which are considered as the primary greenhouse gas responsible for global warming and proxies for overall environmental pollution in a country.

In our empirical analysis, affluence is the natural log of per capita real GDP (real GDP divided by population at the end of the year), population is the natural log of total population in a country, technology is the natural log of the weight of the industry in economic activity (the proportion of the added value of industry to GDP), energy intensity is the natural log of total energy use per dollar of GDP (kg of oil equivalent per capita), trade openness is the natural log of trade openness (exports plus imports as percent of GDP) and urbanization is the natural log of urbanization (% urban population in the total population).

All data except per capita real GDP are obtained from World Development Indicators (WDI) online database. The series of real GDP (at constant 2011 national prices in millions 2011 US$) is obtained

[5] Table A1 in the Appendix provides the list of sample countries.

Table 2 Summary statistics

	$\ln CO_2$	*ln POP*	*ln RGDP*	*ln EI*	*ln TECH*	*ln TRO*	*ln URBA*
Mean	0.5759	16.8724	11.9313	6.7971	3.4205	–0.4238	3.9542
Median	0.5483	16.8309	11.7698	6.6574	3.3771	–0.3901	4.0414
Maximum	2.7502	20.9690	15.7035	8.5501	4.3492	0.6021	4.4900
Minimum	–1.9926	13.9235	9.1564	4.8820	0.9909	–1.9393	2.8726
Std. Dev.	1.0152	1.4326	1.5232	0.7266	0.3087	0.5127	0.3680
Skewness	–0.2151	0.2950	0.2372	0.2305	–0.4442	–0.4432	–1.0374
Kurtosis	2.5662	2.6978	2.1477	2.5733	9.8599	2.9691	3.6196
Observations	900	900	900	900	900	900	900

Source: Own estimates

Table 3 Correlation coefficient matrix and VIF test

	VIF	$\ln(CO_2)$	*ln*(*RGDP*)	*ln*(*TRO*)	*ln*(*URBA*)	*ln*(*POP*)	*ln*(*TECH*)	*ln*(*EI*)
$\ln(CO_2)$		1.0000						
ln(*RGDP*)	1.65	0.2650***	1.0000					
		(0.0000)						
ln(*TRO*)	1.91	0.1807***	–0.4337***	1.0000				
		(0.0000)	(0.0000)					
ln(*URBA*)	1.92	0.6076***	–0.0255	0.1150***	1.0000			
		(0.0000)	(0.4441)	(0.0005)				
ln(*POP*)	1.48	0.0678**	0.7499***	–0.5907***	–0.2078***	1.0000		
		(0.0420)	(0.0000)	(0.0000)	(0.0000)			
ln(*TECH*)	2.15	0.3000***	–0.0110	0.2331***	0.2682***	–0.0040	1.0000	
		(0.0000)	(0.7411)	(0.0000)	(0.0000)	(0.9053)		
ln(*EI*)	1.94	0.9263***	0.2049***	0.1745***	0.5841***	0.0100	0.2767***	1.0000
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.7656)	(0.0000)	

Notes: * denotes p<0.1. ** denotes p<0.05. *** denotes p<0.01.
Source: Own estimates

from the Penn World Table version 9.1.[6] Table 2 summarizes the descriptive statistics of the above-mentioned variables.

The correlation coefficients of the variables are displayed in Table 3. CO_2 emissions have a relatively low and significant correlation with per capita real GDP and trade openness. While the correlation between CO_2 emissions and urbanization is moderate, it is rather strong and significant between CO_2 emissions and energy intensity. However, the correlation between CO_2 emissions and population is weak and statistically significant. To test for multi-collinearity issue, a variance inflation factor (VIF) test is used over a data range of 1.48–2.15, with a mean value of 1.842. As shown in Table 3, the VIF values are all less than the cut-off value of 10, indicating that there is no multi-collinearity.

6 *<https://www.rug.nl/ggdc/productivity/pwt>.*

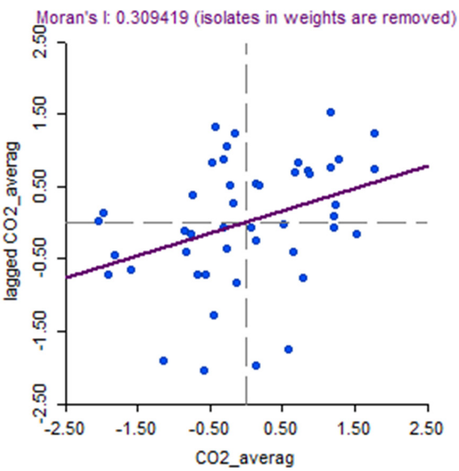
3 EMPIRICAL RESULTS AND DISCUSSIONS

3.1 Exploratory spatial data analysis

Following Abreu et al. (2005), among others, we examined the spatial dependence and spatial heterogeneity in our dataset using the exploratory spatial data analysis (ESDA) approach. To further test whether spatial dependence exists or not, we computed the global Moran's I to identify spatial dependence among the observations, where a significant positive Moran's I value indicates spatial clustering and a negative Moran's I value with statistical significance indicates spatial dispersion across the sample countries (Anselin and Florax, 1995; Anselin, 2006). Furthermore, global Moran's I is a measure of the geographical concentration of a distribution. Generally, the larger the global Moran's I index, the more significant the spatial dependence among countries. A trend of rapid spatial autocorrelation can be clearly seen in Figure 1.

The results of the global spatial autocorrelation for the CO_2 variable by using global Moran's I statistic are summarized in Table 4. Using both the z test and its corresponding p value, we test the statistical significance of the Moran's I values. As shown in Table 4, the Moran's I index values are positive and statistically significant at the 5% level or better. This means that air pollution in middle-income countries exhibits significant positive spatial autocorrelation, which ranges from 0.2627 to 0.3875. Note that the high positive values signal the occurrence of similar attribute values over space, and hence spatial clustering. This means that CO_2 emissions in middle income countries are spatially autocorrelated between 1996 and 2013. They also appear to be less spatially clustered in 2013 than in 1996.

Figure 1 Moran's I Scatter Plot for country-level CO_2 emissions in middle income countries, 1996–2013



Source: Own construction

Table 4 Statistical tests of global Moran's *I* of CO_2 emissions in middle-income countries

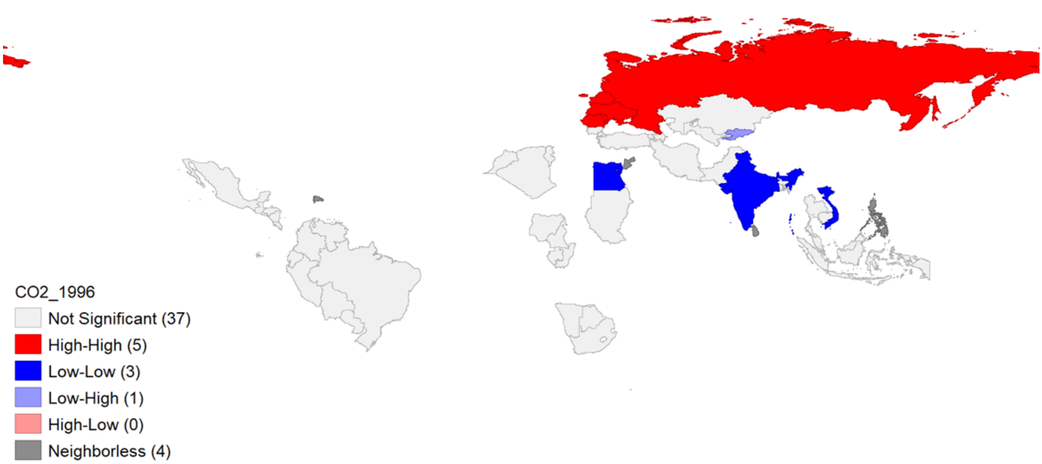
Year	Moran's *I*			Year	Moran's *I*		
	Statistic	*Z* score	*p*-value		Statistic	*Z* score	*p*-value
1996	0.3875***	2.7119	0.0090	2005	0.2871**	2.1404	0.0220
1997	0.3594**	2.5717	0.0120	2006	0.3087**	2.2915	0.0190
1998	0.3299**	2.3633	0.0160	2007	0.3106**	2.2977	0.0190
1999	0.2933**	2.1327	0.0220	2008	0.3375**	2.4963	0.0160
2000	0.2627**	1.9544	0.0280	2009	0.2929**	2.1792	0.0200
2001	0.2715**	2.0171	0.0230	2010	0.2956**	2.2065	0.0220
2002	0.2725**	2.0214	0.0280	2011	0.3166**	2.3489	0.0160
2003	0.2670**	1.9838	0.0260	2012	0.3396**	2.5056	0.0140
2004	0.2825**	2.0987	0.0250	2013	0.3159**	2.3438	0.0160
Average	0.3094**	2.2721	0.0180				

Notes: * denotes p<0.1. ** denotes p<0.05. *** denotes p<0.01. The null hypothesis is no global spatial autocorrelation.
Source: Own estimates

In a second step, we turn to spatiotemporal patterns of country level CO_2 emissions. To visually explore the spatial dependence of the middle-income countries' CO_2 emissions, we undertook a local LISA analysis with the aim of identifying local spatial autocorrelations. The results of the LISA allowed us to identify a detailed local pattern of spatial clustering in relation to changes in per capita CO_2 emission levels. The resulting LISA cluster maps of the countries for which the local Moran's I statistics are statistically significant at the 5% level are displayed in Figures 2, 3 and 4. These figures reveal characteristics of significant local spatial autocorrelation in the distribution of initial CO_2 level in 1996, CO_2 level in 2013 and the average annual CO_2 level over the study period. Spatially, countries with high levels of per capita CO_2 emissions are clustered with neighboring countries that have similar values. Besides, countries with low values of per capita CO_2 emissions clustered with neighboring countries with similar values. The red color denotes the High-High (H-H) clusters (i.e., high values surrounded by high values), while the blue represents Low-Low (L-L) clusters (i.e., low values surrounded by low values). Note that H-H and L-L clusters are the main types of spatial distribution. Furthermore, the pink areas indicate H-L associations and the blue-gray areas denote Low-High (L-H) correlations (i.e., low values surrounded by high values). The gray clusters represent countries that are not associated in a spatially significant manner.

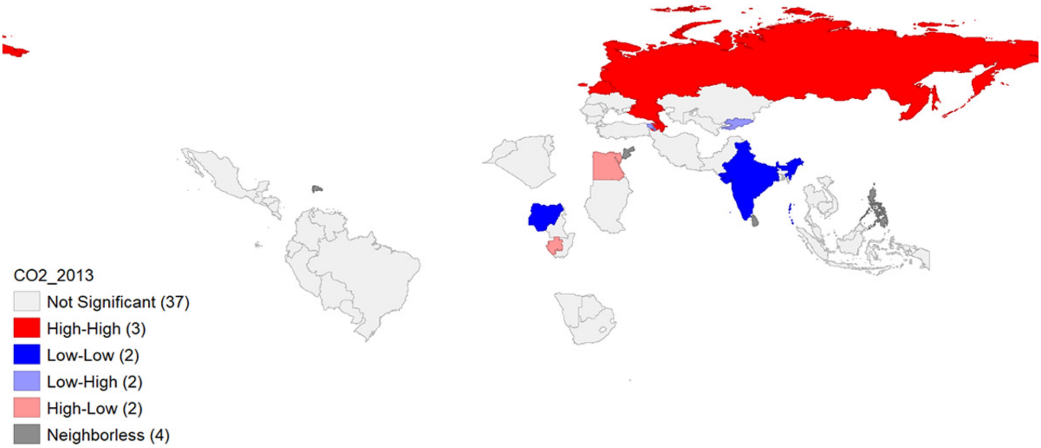
The number and the distribution of each cluster of countries also display regional dynamic characteristics. For instance, in 1996, the numbers of countries belonging to H-H and L-L cluster were 5 and 3 respectively, accounting for 16% of the sample of middle-income countries. This phenomenon is consistent with the situation revealed by a relatively large global Moran's I (0.3875). Correspondingly, only 2% of all countries conformed to the remaining High-Low (H-L) and L-H classifications. These results indicate the existence of a significant dual structure in the spatial distribution of country's per capita CO_2 emissions in 1996. However, by 2013, the number of H-H and L-L countries had decreased by 3 and 2, respectively, indicating that the spatial extent of dependence of per capita CO_2 emissions had weakened markedly between 1996 and 2013. The corresponding global Moran's I index also decreased (0.3159). These results imply that, for geographic data, it is almost inevitable that "close things are more related than distant things," a phenomenon that can be described in terms of "spatial dependence." In addition, the computed findings confirm our previous analysis of spatial dependence in per capita CO_2

Figure 2 Local Moran Scatter Plot map for *ln*(CO_2) in 1996



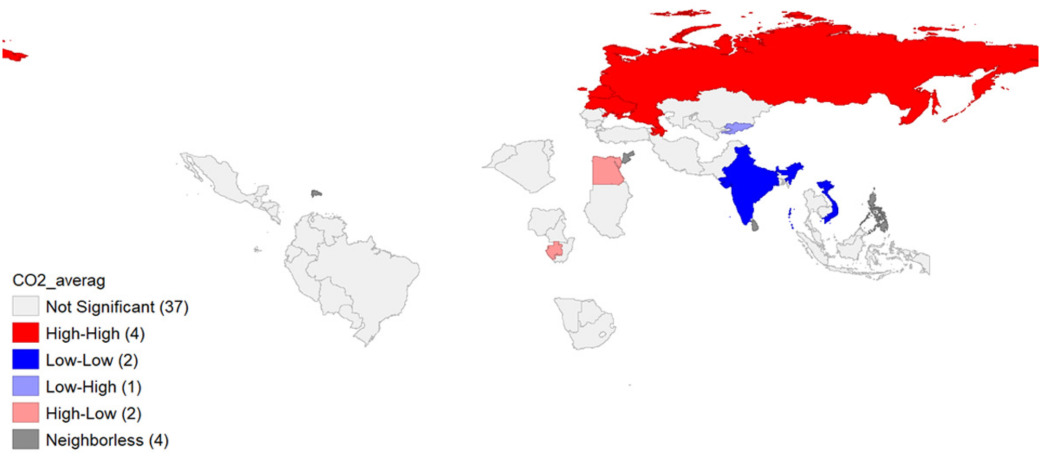
Source: Own construction

Figure 3 Local Moran Scatter Plot map for *ln*(CO_2) in 2013



Source: Own construction

Figure 4 Local Moran Scatter Plot map for *ln*(CO_2) over 1996 to 2013



Source: Own construction

emissions at the country level. Note that if such dependence is ignored, standard econometric models risk being biased in ways that conceal the impact of the determinants they purport to study – in our case, changes in per capita CO_2 emissions in middle income countries. Therefore, we empirically test whether the spatial panel econometrics models are better than conventional econometrics and chose the appropriate model to analyze the impact factors of per capita CO_2 emissions in middle income countries.

3.2 Spatial econometric regression results

To decide which type of model (spatial vs. non-spatial) best fits the data, we begin our investigation by testing several different model specifications. This testing procedure is a mixture of a specific-to-general approach and general-to-specific approach (Elhorst, 2012). Note that the procedure begins by testing the non-spatial panel model against the spatial lag and spatial error models. If the non-spatial panel models are rejected, then the spatial Durbin model (SDM) is tested to determine if it can be simplified to either the spatial lag or spatial error model. It should be stressed that this step seeks corroborating evidence from the first step.

Table 5 reports the estimation results for the non-spatial panel data models: pooled OLS only (no fixed or time-period effects), spatial fixed effects only (no time-period effects), time-period fixed effects only (no fixed effects) and both spatial fixed effects and time-period fixed effects, respectively.

To investigate the null hypothesis that the spatial fixed effects and time-period effects are jointly insignificant, we performed a likelihood ratio (LR) test. The null hypothesis that the spatial fixed effects are jointly insignificant is rejected at the 1% significance level (1 806.2527; 50 degrees of freedom; $P = 0.0000<0.01$). Likewise, the null hypothesis that the time-period fixed effects are jointly insignificant is rejected at the 1% significance level (41.7798; 18 degrees of freedom; $P = 0.0012<0.01$). These findings justify the extension of the model with fixed effects and time-period effects.

It should be stressed that if the country-level fixed effects term is correlated with the explanatory variables, but it is not controlled for within the model, then ordinary least squares (OLS) estimates will result in omitted variable bias (OVB). The pooled OLS estimates (column 2 in Table 5) for all the coefficients in the model are all highly statistically significant ($p < 0.01$), except for TECH variable, which arguably results from the OVB. Given the joint significance of the fixed and time-period effects from the LR test, we focus on the estimation results in column 5 in Table 5.

Table 5 Estimation results without spatial interaction effects

	Pooled OLS	Spatial fixed effects	Time-period fixed effects	Spatial and time-period fixed effects
lnRGDP	0.7598***	0.4199***	0.7857***	0.5241***
	(0.0000)	(0.0023)	(0.0000)	(0.0001)
lnRGDP^2	–0.0293***	–0.0095	–0.0300***	–0.0060
	(0.0000)	(0.1049)	(0.0000)	(0.2981)
lnTRO	0.2147***	0.1406***	0.2515***	0.1486***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
lnURBAN	0.4014***	0.1585	0.4256***	0.2485**
	(0.0000)	(0.1811)	(0.0000)	(0.0339)
lnPOP	0.0718***	–0.1250	0.0752***	0.0848
	(0.0000)	(0.1351)	(0.0000)	(0.3394)
lnTECH	0.0577	–0.0215	0.0530	–0.0417
	(0.1559)	(0.5124)	(0.1894)	(0.2093)
lnEI	1.1189***	0.6014***	1.1120***	0.6357***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Intercept	–14.7637***	-	-	-
	(0.0000)	-	-	-

Table 5				(continuation)
	Pooled OLS	**Spatial fixed effects**	**Time-period fixed effects**	**Spatial and time-period fixed effects**
R^2	0.8871	0.5120	0.8888	0.3327
$\bar{R}^2$	0.8862	0.5087	0.8881	0.3283
σ^2	0.1173	0.0161	0.1142	0.0154
FE R^2		0.9845	0.8899	0.9852
Log Likelihood	–308.6256	584.9733	–297.2632	605.8631
LM spatial lag	64.6114***	5.6159**	70.5684***	3.3388*
	(0.0000)	(0.0180)	(0.0000)	(0.0680)
Robust LM spatial lag	6.1237**	14.5566***	10.7820***	60.3123***
	(0.0130)	(0.0000)	(0.0010)	(0.0000)
LM spatial error	153.7638***	0.6179	139.7624***	4.2464**
	(0.0000)	(0.4320)	(0.0000)	(0.0390)
Robust LM spatial error	95.2760***	9.5585***	79.9760***	61.2198***
	(0.0000)	(0.0020)	(0.0000)	(0.0000)

Notes: All variables are in natural logarithms. Numbers in the parentheses represent P values. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

It should be mentioned that all the non-spatial panel data models may suffer from misspecification if spatial dependence exists within the data. To test for the presence of spatial dependence, we begin by conducting the classical Lagrange Multiplier (LM) tests and their robustness to examine whether non-spatial panel data models ignore the spatial interaction effects of data or not (Anselin et al., 2008; Burridge, 1980). These tests' results are presented in the bottom part in Table 5. For the classical LM test (labeled "LM spatial lag"), the hypothesis of no spatially lagged dependent variable is strongly rejected at the 5% significance level or better for each of the specifications. In addition, and for the classical LM test (labeled "LM spatial error"), the hypothesis of no spatially autocorrelated error term is rejected for each of the specifications except for spatial fixed effects model (although the hypothesis of no spatially lagged dependent variable is rejected at the 1% significance level with this specification). Regarding the results of their robustness tests (Debarsy and Ertur, 2010), both hypotheses are rejected at the 5% significance level or better for each of the specifications. These findings imply the existence of spatial dependence among the panel data, which is consistent with the results of Moran's I index (see Table 4). Besides, they imply that a model specification with a spatially lagged dependent variable may be favored over a non-spatial panel model since we find consistent rejection of the hypothesis of no spatially lagged dependence. However, if the robust LM tests reject a non-spatial panel data model in favour of the SAR model or SEM model, one of these models must be carefully endorsed.

To further test which spatial panel data model specification is more appropriate, LeSage and Pace (2009), and Elhorst (2014b) recommend estimating the SDM, and then conducting both LR and Wald tests to verify whether it can be simplified to the SAR model or to the SEM (see also Burridge, 1981).

In this paper, we take a broader view and apply a Bayesian comparison approach. First, the Bayesian posterior model probabilities of the SDM and SDEM specifications are calculated, as well as the simpler SAR and SEM specifications, to identify which model specification best describes the data. Second, this analysis is repeated for several specifications of the neighbourhood matrices, to find the specification of W that best describes the data.

For this empirical study, we use the following principles to construct twelve spatial weight matrices:

i. Sharing a common land or maritime border implies the first-order binary contiguity matrix, W_1. Maritime borders are based on the United Nations Convention on the Law of the Sea and additional sources further explaining this convention.
ii. The influence of a country might go beyond its immediate neighbors, as implied by the inverse distance matrix and the different cut-off points. Hence, we also consider a second order binary contiguity matrix, $W_2 = W_1 \times W_1$.
iii. A country may respond to the threat of even more distant countries, which is also the main reason that elements of the weight matrix within a certain radius of a country are not always set to 0. Therefore, we include a third-order binary contiguity matrix, $W_3 = W_2 \times W_1$.
iv. Except for the matrix based on the common border countries, the spatial weight matrix could be based on the calculation of distances using the spherical distance between geographic centroids of the countries. Therefore, we create a distance based spatial weight matrix, labeled as W_4, using latitude and longitude coordinates and the Great Circle distance formula.[7]
v. Inverse distance matrix based on the geographical distance between the centroids of every pair of countries. This matrix is labeled as W_5.
vi. k-nearest neighbours matrix for k = 5, 6, 7, 8, 9, 10 and k = 20: it is a binary matrix of the k-nearest neighbour, where the weight $w_{ij} = 1$ if the country j is within the k-nearest neighbour of the country i and $w_{ij} = 1$ if otherwise. Therefore, we create seven additional spatial weight matrices, which are labeled as W_6 for $k = 5$, W_7 for $k = 6$, W_8 for $k = 7$, W_9 for $k = 8$, W_{10} for $k = 9$, W_{11} for $k = 10$, and W_{12} for $k = 20$.

Finally, all the matrices are row normalized, which is standard in spatial econometrics literature when the elements of W have a binary (0/1) character.

Table 6 Simultaneous Bayesian comparison of dynamic spatial panel data model specifications and spatial weight matrices

W matrix	Statistics	SAR	SDM	SEM	SDEM
	Log marginal	538.7932	544.9566	541.0575	545.3341
W_1	Model probabilities	0.0008	0.4031	0.0082	0.5879
	Posterior model probabilities	0.0000	0.0001	0.0000	0.0002
	Log marginal	539.0006	549.7389	538.3043	549.4675
W_2	Model probabilities	0.0000	0.5674	0.0000	0.4326
	Posterior model probabilities	0.0000	0.1690	0.0000	0.0129
	Log marginal	538.8849	551.5624	539.2391	550.9163
W_3	Model probabilities	0.0000	0.6561	0.0000	0.3438
	Posterior model probabilities	0.0000	0.1695	0.0000	0.0550
	Log marginal	538.6144	539.9027	539.2829	539.8787
W_4	Model probabilities	0.0988	0.3584	0.1929	0.3499
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000

[7] Formally, the spherical distance (in kilometers) between the centroids of two countries is defined as follows: $d_{ij} = 6\,366.2 \times Arccos\{\{\cos|Y_i - Y_j| \times \cos X_i \times \cos X_j\}+\{\sin X_i \times \sin X_j\}\}$. X_i denotes the latitude of the centroid of country i, while Y_i is the longitude of the centroid of country i.

Table 6 (continuation)

W matrix	Statistics	SAR	SDM	SEM	SDEM
W_5	Log marginal	−861.2964	−2 079.6651	−1 280.7416	−2 277.1929
	Model probabilities	1.0000	0.0000	0.0000	0.0000
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000
W_6	Log marginal	538.7246	549.0644	538.7472	548.8082
	Model probabilities	0.0000	0.5637	0.0000	0.4363
	Posterior model probabilities	0.0086	0.0086	0.0000	0.0067
W_7	Log marginal	539.3243	549.1309	538.6578	548.9059
	Model probabilities	0.0000	0.5560	0.0000	0.4440
	Posterior model probabilities	0.0000	0.0092	0.0000	0.0074
W_8	Log marginal	540.2980	548.3807	538.6815	547.7051
	Model probabilities	0.0002	0.6626	0.0000	0.3372
	Posterior model probabilities	0.0000	0.0044	0.0000	0.0022
W_9	Log marginal	540.7850	542.3514	538.6779	542.1279
	Model probabilities	0.1027	0.4917	0.0125	0.3932
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000
W_{10}	Log marginal	539.9414	548.1252	538.6742	547.2536
	Model probabilities	0.0002	0.7049	0.0001	0.2949
	Posterior model probabilities	0.0000	0.0034	0.0000	0.0014
W_{11}	Log marginal	539.6121	543.9585	538.6546	543.6084
	Model probabilities	0.0075	0.5806	0.0029	0.4090
	Posterior model probabilities	0.0000	0.0001	0.0000	0.0000
W_{12}	Log marginal	538.8998	553.1569	540.5124	552.4275
	Model probabilities	0.0000	0.6747	0.0000	0.3253
	Posterior model probabilities	0.0000	0.1571	0.0000	0.2493

Notes: The highest posterior model probability in each row is highlighted in italics and the probabilities in each block sum to 1.
Source: Own estimates, based on LeSage (2014, 2015)

The results displayed in Table 6 show that both the dynamic SAR and SEM models are generally outperformed by either the dynamic SDM or dynamic SDEM specifications. In terms of the log marginal likelihood value, the worst-performing spatial neighbourhood matrix is the inverse distance matrix (W_5). This matrix corroborates the point that decomposing market potential variables into their underlying components and considering the spatially lagged values of these components creates a much greater degree of empirical flexibility. If the neighbourhood matrix is specified as a p-order binary contiguity matrices for p = 2, 3, as either a distance neighbourhood matrix, or as k-nearest neighbours matrices for k = 5, 6, 7, 8, 9, 10, 20, then the Bayesian posterior model probabilities point to the dynamic SDM specification. Conversely, if the neighbourhood matrix is specified as a first-order binary contiguity matrix, the Bayesian posterior model probabilities point to the dynamic SDEM specification. Alternatively, if neighbourhood matrix based on the inverse distance is adopted, the Bayesian posterior model probabilities provide further evidence in favour of the dynamic SAR specification.

Table 6 also contains the Bayesian posterior model probabilities of the different spatial models (SAR, SDM, SEM, SDEM), in combination with the twelve proposed spatial weight matrices. These probabilities are calculated for dynamic versions of the spatial panel data model specifications. With these probabilities, we can simultaneously identify the most likely spatial econometric model and the most likely spatial weight matrix. Note that the probabilities are based on the log-marginal likelihood obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. Furthermore, they are normalized such that the probabilities of all 48 combinations sum to 1. Following LeSage (2014, 2015), this normalization is based on the (non-linear) property that the Bayesian posterior model probability increases if the log-marginal likelihood value of one model or one W exceeds that of another model or W.

The results in Table 6 show that by considering the log-marginal values and Bayesian posterior model probabilities of the different specifications of the neighbourhood matrix, it is to be noted that the third-order binary contiguity matrix, i.e., W_3, and the SDM specification achieve the best performance of all 48 combinations, in line with the initial robust LM test statistics for the nonspatial panel data model, which pointed to a SAR rather than a SEM. Accordingly, spatially lagged explanatory variables (WX) are necessary and should be included in the empirical model.

Furthermore, we decided to estimate the dynamic SDM specification using the bias-corrected maximum likelihood (ML) estimator developed by Elhorst (2010a, 2010b), and Lee and Yu (2010a).[8] Note that the results without the bias correction are almost identical.[9] Nevertheless, since the dynamic SDM model produces global country spillover effects, it is more likely to occur in combination with a sparse spatial weight matrix. Therefore, we determined the average number of neighbors of each country in the sample based on these two spatial weight matrices. It equals 6.16 for the W_2 matrix, 9.46 for the W_3 matrix, and 6.00 for the W_4 matrix. Alternatively, the average number of adjacent neighbors based solely on land or maritime borders, i.e., W_1, is 3.080. Based on the principle of sparsity, the W_3 matrix thus seems to offer a better choice than W_1 , W_2 and W_4 matrices.

The estimation results of the dynamic SDM with fixed and time-period effects specification, based on the W_3 matrix, are reported in Table 7. Then, the results could serve to test whether the dynamic SDM might be simplified to a dynamic SAR model or to a dynamic SEM. The empirical findings reject both hypotheses and show that the dynamic SDM is preferred over the dynamic SAR model or the dynamic SEM. Otherwise, a necessary and sufficient condition for stationarity (stability), i.e., $\tau + \delta + \eta = 0.7923 < 1$, is also satisfied. This result is confirmed by the Wald test, across which the null hypothesis $\tau + \delta + \eta = 1$ is strongly rejected at the 1% level of significance.

3.3 Analysis of estimation results

Since the diagnostic results suggest that the dynamic SDM with spatial and time-period fixed effects in Table 6 is the best fitting, we will limit the interpretation of coefficient estimates on it. It should be mentioned that our results are in line with some of the results of previous empirical studies. As shown in Table 7, the CO_2 emissions strongly depend on their value in the previous year, or internal habit persistence (Korniotis, 2010); its coefficient amounts to 0.1595 and is highly significant at the 1% level.

[8] This bias correction is necessary since the dependent variables lagged in time and in both space and time on the right-hand side of Formula (10) are correlated with the spatial fixed effects , which is the spatial counterpart of the Nickell bias, as shown by Yu et al. (2008), and Lee and Yu (2010a) for a dynamic spatial panel data model with and without time-period fixed effects, respectively. In addition, the bias correction is needed because the demeaning procedure to wipe out the country and time-period fixed effects in a standard panel data model (Baltagi, 2005) produces a singularity among the transformed error terms if the model is augmented with a spatial lag in the dependent variable, causing the asymptotic distributions of the parameters not to be properly centered.

[9] To save space, the estimation results of the dynamic SAR model without the bias correction are not reported in this paper, but they are available upon request.

The significant positive estimated coefficient δ indicates that CO_2 emissions in neighboring countries have a positive effect on local CO_2 emissions. Besides, we find evidence of what Korniotis (2010) labels external habit persistence; the coefficient of the CO_2 emissions observed in neighboring countries in the previous period is negative and statistically insignificant (–0.1060, p-value <0.01). Otherwise, countries respond to the CO_2 emissions set in neighboring countries in the same year, such that the coefficient τ takes a positive value of 0.7388 and is highly significant (p-value <0.01), in line with the common feature of horizontal interaction among countries (Brueckner, 2003).

Focusing on the estimated coefficient of per capita real GDP, the elastic coefficient is 0.3319 and statistically significant at the 1% level, which indicates that per capita real income has a negative effect on CO_2 reduction. In addition, the estimated coefficients of the quadratic polynomial of real per capita GDP are highly significant indicating that the relationship between CO_2 emissions and economic growth validate the traditional EKC hypothesis. Our results corroborate the view of other authors (e.g., You and Lv, 2018). The turning point of EKC for CO_2 emissions in the dynamic SDM model is approximately \$ 1 849 516.4465. While it is difficult to estimate the specific year when the turning point has been occurred, governments should abandon the pattern of treatment after pollution, develop the economy and cure the environmental issues at the same time.

Concentrating on the estimated coefficient of trade openness, the elastic coefficient is 0.0509 and significant at the 5% level. All else being equal, higher trade openness increases CO_2 emissions. This result indicates that import and export trade have a negative effect on CO_2 reduction. This result accepts the pollution haven hypothesis (PHH), or pollution haven effect, that polluting countries will relocate to jurisdictions with less stringent environmental regulations. Our results are not consistent with the views of Kearsley and Riddel (2010), Dong et al. (2010), and Kang et al. (2016).

The estimated coefficient on both population and technology are respectively positive and negative but statistically insignificant. Accordingly, we can ignore their impact on per capita CO_2 emissions. Otherwise, the estimated coefficient on energy intensity is positive and highly significant. It indicates that a 1% increase on the total energy use per dollar of GDP will lead to a 0.2247 % increase in CO_2 emissions. This result implies that, all else equal, higher energy intensity, increased CO_2 emissions in a given country. This result is consistent with the view of Shahbaz et al. (2015).

Finally, urbanization has a negative and significant effect on CO_2 emissions in middle-income countries. This finding is not consistent with the views of You and Lv (2018). The elastic coefficient of urbanization is –0.2509, which means a 1% increase in urban population will result in a 0.2509% decrease in CO_2 emissions. In other words, urbanization has a positive impact on CO_2 reduction in middle-income countries. This result indicates that middle-income countries considered in this paper promote low-carbon urbanization progress and spread the application of green architecture technology with the topic of energy-saving and environmental protection to develop green city. Overall, the results of this study show that per capita real income, trade openness and energy intensity have significant positive effects on CO_2 emissions, while urbanization has a significant negative effect on CO_2 emissions. Considering these results, policymakers should realize an integrated policy with the aim at reducing CO_2 emissions based on the determinants.

3.4 Direct and spillover effects

It is noteworthy that the coefficients of the dynamic SDM do not directly reflect the marginal effects of the corresponding explanatory variables on the dependent variable. Therefore, we report both the short-term and long-term impacts of the direct and spillover effects of the explanatory variables. Table 7 displays both short and long-term estimates of the direct and spillover effects, derived from the parameter estimates using Formula (12). To draw inferences regarding the statistical significance of these effects, the variation of 1 000 simulated parameter combinations is used, drawn from the variance-covariance matrix implied by the ML estimates.

To get better fitting effects, we conduct a comparative analysis between the dynamic SDM with spatial and time-period fixed effects in Table 7 and non-spatial panel data model with two-way fixed effects in Table 5. The results indicate that most coefficients in non-spatial panel data model are larger than those in dynamic spatial panel data model. Two main reasons could explain this difference. The first one is mainly attributed to ignoring the spatial spillover effect of data. The second reason is due to the feedback

Table 7 Results of the dynamic SDM with $W = W_3$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t$: δ	0.1595***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}$: τ	0.7388***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}$: η	–0.1060***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.3319***	0.0017	0.1510	1.5542	–1.1431	–1.5933	0.8889	0.6858	–2.9471	–0.2954
$ln(RGDP)_t^2$	–0.0115**	0.0133	–0.0055	–1.2788	0.0384	1.3563	–0.0315	–0.7068	0.1000	0.2700
$ln(TRO)_t$	0.0509**	0.0165	0.0578***	2.7957	0.0405	0.3912	0.1976	0.9516	0.0292	0.0442
$ln(URBA)_t$	–0.2509**	0.0222	–0.2259**	–2.2668	0.1795	0.3599	–0.9323	–0.7773	0.2357	0.0390
$ln(POP)_t$	0.0495	0.5680	0.0336	0.4904	–0.1329	–0.4346	0.1741	0.3284	–0.2097	–0.1015
$ln(TECH)_t$	–0.0068	0.6391	–0.0102	–0.4476	–0.0218	–0.1849	–0.0290	–0.1567	–0.0506	–0.0810
$ln(EI)_t$	0.2247***	0.0000	0.2499***	8.1734	0.1548	0.9899	0.8927***	2.9320	0.1922	0.1439
$W * ln(RGDP)_t$	0.0719	0.2074	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	–0.0023	0.3038	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	–0.0129	0.1559	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0246	0.4998	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	0.0043	0.9443	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	0.0040	0.7317	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	–0.0530***	0.0001								
Observations	850	-	-	-	-	-	-	-	-	-
R^2	0.9940	-	-	-	-	-	-	-	-	-
σ^2	0.0081	-	-	-	-	-	-	-	-	-
Log-likelihood	943.4051	-	-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	0.7923	-	-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	59.2319***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	59.4501***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	233.9187***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.

Source: Own estimates

effects that arise CO_2 emissions of local country as a result of influencing the CO_2 emissions of adjacent countries. In addition, one part of the feedback effects is from spatially lagged dependent variable, while the other part comes from the spatially lagged independent variables.

The coefficient estimates and short-term direct effects estimates derived from the parameter estimates using Formula (12) exhibit a plausible model structure. The direct effect of trade openness on CO_2 emissions is positive and highly significant but lesser than 1. A one percentage point increase of the trade openness has an adverse effect on CO_2 emissions, equal to 0.0578 percentage points. The impact of urbanization on a country's CO_2 emissions is negative and statistically significant at the 5% level. The direct effect of the energy intensity variable is positive and highly significant. This finding indicates that the CO_2 emissions increases with a higher level of energy intensity. However, only the energy intensity variable exhibits significant long-term direct effects, but its magnitude almost a fourth.

Spatial spillover effects are local in nature and cannot be observed directly from the estimated coefficients reported in Table 5. Alternatively, we report the average values of the short and long-term spillover effects of Formula (12) in Table 7. The observed spillover effects in the short term or in the long term are not statistically significant. Therefore, the considered explanatory variables observed in neighboring countries do not have impacts on CO_2 emissions.

3.5 Robustness checks

We report and discuss the results of two robustness checks, thereby focusing on short-term direct and country spillover effects. First, we re-estimate the dynamic SDM specification by replacing the spatial weight matrix by the second-order binary contiguity matrix W_2, in line with the results in Table 6. The results reported in Table 8 show that changes are somewhat tiny for almost the independent variables whether in terms of statistical significance or magnitude, which further confirms the robustness of our main findings with model specification.

With our second robustness check, we follow You and Lv (2018) by exploring whether the results are changed when ruling out population explanatory variable and expressing the main variables as population weighted values. As acknowledged by You and Lv (2018), the rationale behind of this model is that it factors out the impacts of population on each of these variables. To do so, we also repeated the Bayesian comparison approach and selected simultaneously the best spatial econometric model as well as the best spatial weight matrix. The Bayesian comparison approach allows selecting simultaneously both the dynamic SDM model and W_2 as the most likely spatial panel model and the most likely spatial weight matrix, respectively.[10] The results from Table 9 further support the robustness of the previous findings.

CONCLUSIONS AND POLICY IMPLICATIONS

In this paper, we contributed to the existing literature by performing a more rigorous analysis of the relationship between economic growth and CO_2 emissions in middle income countries. We firstly examined the EKC hypothesis for CO_2 emissions at the country level using a dynamic SDM model with country and time period fixed effects. We also computed the short- and long-term spillover effects of explanatory variables for CO_2 emissions in neighboring countries. Our results imply a positive, nonlinear relationship between economic growth and CO_2 emissions. In other words, we found evidence for the EKC hypothesized, inverted U-shaped relationship between CO_2 emissions and economic growth in middle-income countries. Moreover, trade openness and energy intensity were the major drivers of increasing CO_2 emissions, while urbanization effect plays a crucial role in carbon reduction. The results were generally hold when robustness checks were performed.

[10] To save space, the results of the Bayesian comparison approach are not reported but are available upon request.

Based on the empirical findings of this study, the following policy recommendations are put forward to further mitigate CO_2 emissions in middle-income countries. First, the results of this paper showed evidence of an inverted U-shaped relationship between CO_2 emissions and economic growth, suggesting that CO_2 emissions increases at the early stages of development, but goes down at later stage of development. In this vein, the policies should be device in a way to reduce CO_2 emissions at the later stages of economic development. The PHH stipulates that, when big industrialized countries seek to set up factories abroad, they will often search for the cheapest option in terms of resources and labor that offers the land

Table 8 First robustness check: results of the dynamic SDM with $W = W_2$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t$: δ	0.1753***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}$: τ	0.7317***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}$: η	–0.1504***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.2185**	0.0170	0.2551**	2.2990	0.3468	0.3230	0.7127	0.1053	–1.9641	–0.0257
$ln(RGDP)_t^2$	–0.0059	0.1788	–0.0086*	–1.7435	–0.0234	–0.5390	–0.0142	–0.0571	0.1294	0.0463
$ln(TRO)_t$	0.0451**	0.0253	0.0374	1.5106	–0.0712	–0.3850	0.1734	0.1384	0.0822	0.0059
$ln(URBA)_t$	–0.2229**	0.0142	–0.2325**	–2.3990	–0.0597	–0.0779	–0.5854	–0.1701	3.2875	0.0861
$ln(POP)_t$	0.0039	0.6216	–0.0232***	–0.2692	–0.2673	–0.6168	0.1450	0.0460	1.2978	0.0368
$ln(TECH)_t$	0.0073	0.9561	–0.0023	–0.0831	–0.0829	–0.3538	0.0870	0.0561	0.8105	0.0458
$ln(EI)_t$	0.2368***	0.0000	0.2375***	6.8674	–0.0041	–0.0145	0.6395	0.1838	–3.0480	–0.0764
$W * ln(RGDP)_t$	–0.0625	0.5022	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	0.0027	0.4646	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	–0.0033	0.5496	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0466	0.5393	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	0.0162	0.2813	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	0.0050	0.6336	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	–0.0417***	0.0022	-	-	-	-	-	-	-	-
Observations	850		-	-	-	-	-	-	-	-
R^2	0.9940		-	-	-	-	-	-	-	-
σ^2	0.0079		-	-	-	-	-	-	-	-
Log-likelihood	953.2061		-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	0.7566		-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	88.0738***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	66.4035***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	280.4242***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.

Source: Own estimates

and material access they require. This hypothesis surpasses the income per capita that noticeably increases CO_2 emissions in the middle-income countries. The existence of EKC in the middle-income countries gives food-for-thought for the environmentalist to establish environmentally friendly and sustainable policies. Besides, the world is in fierce competition which can damage the natural flora of the world's resources that is considered the brazen growth for the economies. Thus, there is a strong need to set an optimistic target for economic growth that would easily be achieved without the cost of environmental degradation. Second, middle-income countries should decrease the amount of trade for lower pollution. However, this decision may deteriorate the economic situation of these countries. Although trade openness in conjunction with economic growth may cause environmental worsening, it is an important contributor to economic growth of several middle-income countries. Accordingly, policymakers should use trade openness to stimulate non-polluted industries by imposing taxes on polluted industries and creating incentives on non-polluted industries in order to encourage producers to shift toward cleaner and more environmentally friendly industries. Third, the positive impact of energy intensity on CO_2 emissions emphasizes the importance of re-structuring the energy use in middle-income countries such that increase in energy intensity does not necessarily translate into higher CO_2 emissions. As an adequate solution for these countries, governments should promote renewable energy technologies. Finally, urban planners should use efficient urbanization to curb the CO_2 emissions, especially for the countries with high density of population. Particularly, they should take thoughtful action on climate change by improving the public transportation systems and the energy efficiency of buildings and increasing the share of renewable energy sources in energy supplies.

Table 9 Second robustness check: results of the dynamic SDM with $W = W_2$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t$: δ	0.1799***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}$: τ	0.7311***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}$: η	−0.1579***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.2268***	0.0082	0.2467**	2.4668	0.1996	0.3230	0.6230	0.1343	−3.5816	−0.0663
$ln(RGDP)_t^2$	−0.0063*	0.0965	−0.0080*	−1.8229	−0.0166	−0.5390	−0.0121	−0.0497	0.1664	0.0592
$ln(TRO)_t$	0.0429**	0.0216	0.0400*	1.8849	−0.0304	−0.3850	0.2199	0.1146	0.9094	0.0328
$ln(URBA)_t$	−0.2023**	0.0218	−0.2488***	−2.9100	−0.4525	−0.0779	−0.2300	−0.0279	6.3815	0.0699
$ln(POP)_t$	0.0089	0.9485	−0.0025	−0.1027	−0.0990	−0.6168	0.1867	0.0528	2.0634	0.0418
$ln(TECH)_t$	0.2356***	0.0000	0.2354***	7.6816	−0.0037	−0.3538	0.6551	0.1722	−2.6891	−0.0610
$ln(EI)_t$	−0.0542	0.7221	-	-	-	−0.0145	-	-	-	-
$W * ln(RGDP)_t$	0.0024	0.6364	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	−0.0055	0.3288	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	0.0717**	0.0135	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0068	0.5174	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	−0.0421***	0.0017	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	850	-	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	0.9940	-	-	-	-	-	-	-	-	-

Table 9 (continuation)

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Observations	0.0080	-	-	-	-	-	-	-	-	-
R^2	952.56153	-	-	-	-	-	-	-	-	-
σ^2	0.7531	-	-	-	-	-	-	-	-	-
Log-likelihood	90.9290***	0.0000	-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	13.3232**	0.0382	-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	14.9861**	0.0101	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	66.4035***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	280.4242***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

References

ABREU, M., DE GROOT, H. L. F., FLORAX, R. J. G. M. Space and growth: a survey of empirical evidence and methods. *Region et Development*, 2005, 21, pp. 13–44.

ACARAVCI, A. AND OZTURK, I. On the relationship between energy consumption, CO_2 emissions and economic growth in Europe. *Energy*, 2010, 35, pp. 5412–5420.

ADAMS, S. AND KLOBODU, E. K. M. Urbanization, Democracy, Bureaucratic Quality, and Environmental Degradation. *Journal of Policy Modeling*, 2017, 39(6), pp. 1035–1051.

AGRAS, J. AND CHAPMAN, D. A dynamic approach to the Environmental Kuznets Curve hypothesis. *Ecological Economics*, 1999, 28, pp. 267–277.

AHMED, K. AND LONG, W. Environmental Kuznets curve and Pakistan: an empirical analysis. *Proc. Econ. Financ.*, 2012, 1, pp. 4–13.

ALDY, J. E. An Environmental Kuznets Curve analysis of U.S. state-level carbon dioxide emissions. *The Journal of Environment Development*, 2005, 14, pp. 48–72.

ALDY, J. E. Per capita carbon dioxide emissions: convergence or divergence? *Environ. Resour. Econ.*, 2006, 33(4), pp. 533–555.

AL-MULALI, U., OZTURK, I., LEAN, H. H. The Influence of Economic Growth, Urbanization, Trade Openness, Financial Development, and Renewable Energy on Pollution in Europe. *Natural Hazards*, 2015, 79(1), pp. 621–644.

ANSELIN, L. Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 1995, 27(2), pp. 93–115.

ANSELIN, L. Spatial econometrics. In: BALTAGI, BADI H. (eds.) *A companion to theoretical econometrics*, Massachusetts: Blackwell Publishers Lte., 2001.

ANSELIN, L. Spatial econometric. In: MILLS, T. C. AND PATTERSON, K. (eds.) *Palgrave Handbook of Econometrics Econometric Theory*, Vol. 1, pp. 901–969, Basingstoke: Palgrave Macmillan, 2006.

ANSELIN, L. AND FLORAX, R. J. *Small sample properties of tests for spatial dependence in regression models: some further results. New Directions in Spatial Econometrics.* Berlin, Heidelberg: Springer, 1995, pp. 21–74.

ANSELIN, L., LE GALLO, J., JAYET, H. Spatial Panel Econometrics. In: MATYAS, L. AND SEVESTRE, P. *The Econometrics of Panel Data. Fundamentals and Recent Developments in Theory and Practice.* 3rd Ed. Springer-Verlag, 2008, pp. 627–662.

APERGIS, N. AND OZTURK, I. Testing environmental Kuznets curve hypothesis in Asian countries. *Ecological Indicors*, 2015, 52, pp. 16–22.

ASLANIDIS, N. AND IRANZO, S. Environment and development: is there a Kuznets curve for CO_2 emissions? *Applied Economics*, 2009, 41, pp. 803–810.

AZOMAHOU, T., LAISNEY, F., VAN PHU, N. Economic development and CO_2 emissions: a non-parametric panel approach. *Journal of Public Economics*, 2006, 90, pp. 1347–1363.
BALTAGI, B. H. *Econometric analysis of panel data*. 3rd Ed. Wiley, Chichester, 2005.
BRUECKNER, J. K. Strategic Interaction Among Governments: an Overview of Empirical Studies. *International Regional Science Review*, 2003, 26(2), pp. 175–188.
BURKE, M., HSIANG, S. M., MIGUEL, E. Global non-linear effect of temperature on economic production. *Nature*, 2015, 527, pp. 235–239.
BURNETT, J. W., BERGSTROM, J. C., DORFMAN, J. H. A spatial panel data approach to estimating US state-level energy emissions. *Energy Economics*, 2013, 40, pp. 396–404.
BURRIDGE, P. On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society. Series B-Statistical Methodology*, 1980, 42(1), pp. 107–108.
BURRIDGE, P. Testing for a common factor in a spatial autoregression model. *Environment and Planning A*, 1981, 13, pp. 795–800.
CARSON, R. T., JEON, Y., MCCUBBIN, D. R. The relationship between air pollution emissions and income: US Data. *Environment and Development Economics*, 1997, 2, pp. 433–450.
COLE, M. A. Trade, the pollution haven hypothesis and the Environmental Kuznets Curve: examining the linkages. *Ecological Economics*, 2004, 48, pp. 71–81.
COONDOO, D. AND DINDA, S. Carbon dioxide emissions and income: a temporal analysis of cross-country distributional patterns. *Ecological Economics*, 2008, 65, pp. 375–385.
DE BRUYN, S. M., VANDENBERGH, J. C. J. M., OPSCHOOR, J. B. Economic growth and emissions: reconsidering the empirical basis of Environmental Kuznets Curves. *Ecological Economics*, 1998, 25, pp. 161–175.
DEBARSY, N. AND ERTUR, C. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 2010, 40, pp. 453–470.
DIETZ, T. AND ROSA, E. A. Effects of Population and Affluence on CO_2 Emissions. *Proceedings of the National Academy of Sciences*, 1997, 94(1), pp. 175–179.
DIJKGRAAF, E. AND VOLLEBERGH, H. R. J. A test for parameter homogeneity in CO_2 panel EKC estimations. *Environ. Resour. Econ.*, 2005, 32, pp. 229–239.
DINDA, S. Environmental Kuznets Curve hypothesis: a survey. *Ecological Economics*, 2004, 49, pp. 431–455.
DONG, Y., ISHIKAWA, M., LIU, X., WANG, C. An analysis of the driving forces of CO_2 emissions embodied in Japan-China trade. *Energy Policy*, 2010, 38, pp. 6784–6792.
DU, L., WEI, C., CAI, S. Economic development and carbon dioxide emissions in China: Provincial panel data analysis. *China Economic Review*, 2012, 23, pp. 371–384.
DUTT, K. Governance, institutions and the environment-income relationship: a cross-country study. *Environment, Development and Sustainability*, 2009, 11, pp. 705–723.
EHRLICH, P. R. AND HOLDREN, J. P. Impact of Population Growth. *Science*, 1971, 171(3977), pp. 1212–1217.
ELHORST, J. P. Spatial panel data models. In: FISCHER, M. AND GETIS A. (eds.) *Handbook of Applied Spatial Analysis*, Berlin: Springer, 2010a, pp. 377–407.
ELHORST, J. P. Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 2010b, 5(1), pp. 9–28.
ELHORST, J. P. Dynamic spatial panels: models, methods, and inferences. *Journal of Geographical Systems*, 2012, 14(1), pp. 5–28.
ELHORST, J. P. Matlab Software for Spatial Panels. *International Regional Science Review*, 2014a, 37(3), pp. 389–405.
ELHORST, J. P. *Spatial Econometrics: from Cross-Sectional Data to Spatial Panels.* Heidelberg, New York, Dordrecht, London: Springer, 2014b.
ESTEVE, V. AND TAMARIT, C. Threshold cointegration and nonlinear adjustment between CO_2 and income: the Environmental Kuznets Curve in Spain, 1857–2007. *Energy Economics*, 2012, 34, pp. 2148–2156.
FARHANI, S. AND OZTURK, I. Causal relationship between CO_2 emissions, real GDP, energy consumption, financial development, trade openness, and urbanization in Tunisia. *Environmental Science and Pollution Research*, 2015, 22(20), pp. 15663–15676.
FOSTEN, J., MORLEY, B., TAYLOR, T. Dynamic misspecification in the environmental Kuznets curve: Evidence from CO_2 and SO_2 emissions in the United Kingdom. *Ecological Economics*, 2012, 76, pp. 25–33.
FRIEDL, B. AND GETZNER, M. Determinants of CO_2 emissions in a small open economy. *Ecological Economics*, 2003, 45, pp. 133–148.
GALEOTTI, M. AND LANZA, A. Richer and cleaner? A study on carbon dioxide emissions in developing countries. *Energy Policy*, 1999, 27, pp. 565–573.
GALEOTTI, M., LANZA, A., PAULI, F. Reassessing the environmental Kuznets curve for CO_2 emissions: a robustness exercise. *Ecol. Econ.*, 2006, 57(1), pp. 152–163.
GROSSMAN, G. M. AND KRUEGER, A. B. *Environmental impacts of a North American free trade agreement.* Cambridge, Mass., USA: National Bureau of Economic Research, 1991, pp. 1–39.

GROSSMAN, G. M. AND KRUEGER, A. B. Economic growth and the environment. *The Quarterly Journal of Economics*, 1995, 110(2), pp. 353–377.
HALICIOGLU, F. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy*, 2009, 37, pp. 1156–1164.
HOLTZ-EAKIN, D. AND SELDEN, T. M. Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics*, 1995, 57, pp. 85–101.
IWATA, H., OKADA, K., SAMRETH, S. A note on the Environmental Kuznets Curve for CO_2: a pooled mean group approach. *Applied Energy*, 2011, 88, pp. 1986–1996.
JALIL, A. AND MAHMUD, S. F. Environment Kuznets curve for CO_2 emissions: a cointegration analysis for China. *Energy Policy*, 2009, 37, pp. 5167–5172.
JAUNKY, V. C. The CO_2 emissions-income nexus: evidence from rich countries. *Energy Policy*, 2011, 39, pp. 1228–1240.
KAIKA, D. AND ZERVAS, E. The environmental Kuznets curve (EKC) theory. Part A: Concept, causes and the CO_2 emissions case. *Energy Policy*, 2013a, 62, pp. 1392–1402.
KAIKA, D. AND ZERVAS, E. The environmental Kuznets curve (EKC) theory. Part B: Critical issues. *Energy Policy*, 2013b, 62, pp. 1403–1411.
KANG, Y.Q., ZHAO, T., YANG, Y. Y. Environmental Kuznets curve for CO_2 emissions in China: a spatial panel data approach. *Ecological Indicators*, 2016, 63, pp. 231–239.
KASMAN, A. AND DUMAN, Y. S. CO_2 emissions, economic growth, energy consumption, trade and urbanization in new EU member and candidate countries: a panel data analysis. *Econ. Model.*, 2015, 44, pp. 97–103.
KEARSLEY, A. AND RIDDEL, M. A further inquiry into the pollution haven hypothesis and the environmental Kuznets curve. *Ecological Economics*, 2010, 69(4), pp. 905–919.
KORNIOTIS, G. M. Estimating Panel Models with Internal and External Habit Formation. *Journal of Business & Economic Statistics*, 2010, 28(1), pp. 145–158.
KUNNAS, J. AND MYLLYNTAUS, T. The Environmental Kuznets Curve hypothesis and air pollution in Finland. *Scandinavian Economic History Review*, 2007, 55(2), pp. 101–127.
LANTZ, V. AND FENG, Q. Assessing income, population, and technology impacts on CO_2 emissions in Canada: Where's the EKC? *Ecological Economics*, 2006, 57, pp. 229–238.
LEE, C-C., CHIU, Y-B., SUN, C-H. Does one size fit all? Are examination of the Environmental Kuznets Curve using the dynamic panel data approach. *Review of Agricultural Economics*, 2009, 31(4), pp. 751–778.
LEE, L. F. AND YU, J. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 2010, 154(2), pp. 165–185.
LESAGE, J. P. Spatial econometric panel data model specification: a Bayesian approach. *Spatial Statistics*, 2014, 9, pp. 122–145.
LESAGE, J. P. Software for Bayesian cross section and panel spatial model comparison. *Journal of Geographical Systems*, 2015, 17(4), pp. 297–310.
LESAGE, J. P. AND PACE, R. K. *Introduction to Spatial Econometrics.* New York: CRC Press, 2009.
LI, T., WANG, Y., ZHAO, D. Environmental Kuznets Curve in China: New evidence from dynamic panel analysis. *Energy Policy*, 2016, 91, pp. 138–147.
LIM, J. Economic Growth and Environment: Some Empirical Evidences from South Korea. *Seoul Journal of Economics*, 1997, 10(3), pp. 273–292.
LINDMARK, M. An EKC-pattern in historical perspective: carbon dioxide emissions, technology, fuel prices and growth in Sweden 1870–1997. *Ecological Economics*, 2002, 42, pp. 333–347.
LV, Z. AND XU, T. Trade openness, urbanization and CO_2 emissions: Dynamic panel data analysis of middle-income countries. *The Journal of International Trade & Economic Development*, 2019, 28(3), pp. 317–330.
MADDISON, D. Environmental Kuznets curves: a spatial econometric approach. *J. Environ. Econ. Manag.*, 2006, 51(2), pp. 218–230.
MADLENER, R. AND SUNAK, Y. Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management? *Sustainable Cities and Society*, 2011, 1, pp. 45–53.
MARTÍNEZ-ZARZOSO, I. AND MARUOTTI, A. The impact of urbanization on CO_2 emissions: evidence from developing countries. *Ecol. Econ.*, 2011, 70(7), pp. 1344–1353.
MARTÍNEZ-ZARZOSO, I., BENGOCHEA-MORANCHO, A., MORALES-LAGE, R. The impact of population on CO_2 emissions: evidence from European countries. *Environmental and Resource Economics*, 2007, 38, pp. 497–512.
MENG, B., WANG, J., ANDREW, R., XIAO, H., XUE, J., PETERS, G. P. Spatial spillover effects in determining China's regional CO_2 emissions growth: 2007–2010. *Energy Economics*, 2017, 63, pp. 161–173.
MENG, L. AND HUANG, B. Shaping the relationship between economic development and carbon dioxide emissions at the local level: Evidence from spatial econometric models. *Environmental and Resource Economics*, 2018, 71(1), pp. 127–156.
MOOMAW, W. R. AND UNRUH, G. C. Are Environmental Kuznets Curves misleading us? The case of CO_2 emissions. *Environment and Development Economics*, 1997, 2, pp. 451–463.

MORAN, P. A. P. Notes on Continuous Stochastic Phenomena. *Biometrika*, 1950, 37(1/2), pp. 17–23.

NARAYAN, P-K. AND NARAYAN, S. Carbon dioxide emissions and economic growth: panel data evidence from developing countries. *Energy Policy*, 2010, 38, pp. 661–666.

ONAFOWORA, O. A. AND OWOYE, O. Bounds testing approach to analysis of the environment Kuznets curve hypothesis. *Energy Economics*, 2014, 44, pp. 47–62.

OZTURK, I. AND ACARAVCI, A. The long-run and causal analysis of energy growth, openness and financial development on carbon emissions in Turkey. *Energy Economics*, 2013, 36, pp. 262–267.

PAO, H. T. AND TSAI, C. M. Multivariate Granger causality between CO_2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. *Energy*, 2011, 36(1), pp. 685–693.

PERRINGS, C. AND ANSUATEGI, A. Sustainability, growth and development. *Journal of Economic Studies*, 2000, 27(1/2), pp. 19–54.

RICHMOND, A. K. AND KAUFMANN, R. K. Is there a turning point in the relationship between income and energy use and/or carbon emissions? *Ecological Economics*, 2006, 56(2), pp. 176–189.

ROBERTS, J. T. AND GRIMES, R. Carbon intensity and economic development 1962–1991: a brief exploration of the Environmental Kuznets Curve. *World Development*, 1997, 25(2), pp. 191–198.

ROY CHOWDHURY, R. AND MORAN, E. F. Turning the curve: a critical review of Kuznets approaches. *Appl. Geogr.*, 2012, 32(1), pp. 3–11.

SABOORI, B. AND SULAIMAN, J. Environmental degradation, economic growth and energy consumption: Evidence of the environmental Kuznets curve in Malaysia. *Energy Policy*, 2013, 60, pp. 892–905.

SABOORI, B., SULAIMAN, J., MOHD, S. Economic growth and CO_2 emissions in Malaysia: a cointegration analysis of the environmental Kuznets curve. *Energy Policy*, 2012, 51, pp. 184–191.

SCHMALENSEE, R., STOKER, T. M., JUDSON, R. A. World carbon dioxide emissions: 1950–2050. *The Review of Economics and Statistics*, 1998, 80, pp. 15–27.

SHAHBAZ, M., KHRAIEF, N., UDDIN, G. S., OZTURK, I. Environmental Kuznets curve in an open economy: a bounds testing and causality analysis for Tunisia. *Renewable and Sustainable Energy Reviews*, 2014, 34, pp. 325–336.

SHAHBAZ, M., SOLARIN, S. A., SBIA, R., BIBI, S. Does energy intensity contribute to CO_2 emissions? A trivariate analysis in selected African countries. *Ecological Indicators*, 2015, 50, pp. 215–224.

WANG, S. AND LIU, X. China's city-level energy-related CO_2 emissions: Spatiotemporal patterns and driving forces. *Applied Energy*, 2017, 200, pp. 204–214.

WANG, S., FANG, C., WANG, Y. Spatiotemporal variations of energy-related CO_2 emissions in China and its influencing factors: an empirical analysis based on provincial panel data. *Renewable and Sustainable Energy Reviews*, 2016, 55, pp. 505–515.

WANG, S. S., ZHOU, D. Q., ZHOU, P., WANG, Q. W. CO_2 emissions, energy consumption and economic growth in China: a panel data analysis. *Energy Policy*, 2011, 39, pp. 4870–4875.

YIN, J., ZHENG, M., CHEN, J. The effects of environmental regulation and technical progress on CO_2 Kuznets curve: an evidence from China. *Energy Policy*, 2015, 77, pp. 97–108.

YORK, R., ROSA, E. A., DIETZ, T. STIRPAT, IPAT and ImPACT: Analytic Tools for Unpacking the Driving Forces of Environmental Impacts. *Ecological Economic*s, 2003, 46(3), pp. 351–365.

YOU, W. AND LV, Z. Spillover effects of economic globalization on CO_2 emissions: a spatial panel approach. *Energy Economics*, 2018, 73, pp. 248–257.

YU, J., DE JONG, R., LEE, L. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 2008, 146(1), pp. 118–134.

ZHANG, Y. J., LIU, Z., ZHANG, H., TAN, T.-D. The impact of economic growth, industrial structure and urbanization on carbon emission intensity in China. *Nat Hazards*, 2014, 73, pp. 579–595.

ZHAO, X., BURNETT, J. W., FLETCHER, J. J. Spatial analysis of China province-level CO_2 emission intensity. *Renewable and Sustainable Energy Reviews*, 2014, 33, pp. 1–10.

APPENDIX

Table A1 Country list

Lower middle-income countries ($ 996 to $ 3 895)		Upper middle-income countries ($ 3 896 to $ 12 055)	
Country Name	**Country Code**	**Country Name**	**Country Code**
Bangladesh	BGD	Algeria	DZA
Bolivia	BOL	Armenia	ARM
Cambodia	KHM	Azerbaijan	AZE
Cameroon	CMR	Belarus	BLR
Congo, Rep.	COG	Botswana	BWA
Egypt, Arab Rep.	EGY	Brazil	BRA
El Salvador	SLV	Bulgaria	BGR
Honduras	HND	Colombia	COL
India	IND	Costa Rica	CRI
Indonesia	IDN	Dominican Republic	DOM
Kyrgyz Republic	KGZ	Ecuador	ECU
Moldova	MDA	Gabon	GAB
Morocco	MAR	Guatemala	GTM
Nicaragua	NIC	Iran, Islamic Rep.	IRN
Nigeria	NGA	Jordan	JOR
Pakistan	PAK	Kazakhstan	KAZ
Philippines	PHL	Malaysia	MYS
Sri Lanka	LKA	Mexico	MEX
Sudan	SDN	Namibia	NAM
Tunisia	TUN	Paraguay	PRY
Ukraine	UKR	Peru	PER
Uzbekistan	UZB	Romania	ROU
Vietnam	VNM	Russian Federation	RUS
		South Africa	ZAF
		Thailand	THA
		Turkey	TUR
		Venezuela, RB	VEN

Source: World Bank Country Classifications by income level (2018–2019)

Measuring Inequality of Opportunity: Does Inequality Index Matter?

Zulfiya Ibragimova[1] | *Bashkir State University, Ufa, Russia*
Marina Frants[2] | *Ufa State Aviation Technical University, Ufa, Russia*

Abstract

The equal opportunity theory is based on the idea that it is important to distinguish between two sources of inequality: the inequality caused by factors outside an individual's control (inequality of opportunity) and the inequality generated by factors within an individual's control (inequality of effort). The aim of this study is to assess the impact of choosing an inequality index on the results of measuring the inequality of opportunity. The empirical analysis was carried out based on the data from Life in Transition III sociological survey. Important findings suggest that: 1) the choice of inequality measure has a significant impact on the outcome of measuring the inequality of opportunity; 2) within the methodology under consideration, when using the Gini index, the contribution of inequality of opportunity to the inequality in labor income turns out to be much greater than when using other measures of inequality with the direct method of assessment, and vice versa, noticeably less with the indirect method of assessment; 3) the L-Theil and T-Theil indices look more preferable to use; 4) a country's ranking in terms of absolute and relative inequality of opportunities changes depending on the choice of the measure of inequality and on the choice of the assessment method, sometimes quite significantly; 5) the ranking position for absolute inequality of opportunity may differ significantly from the ranking position for relative inequality.[3]

Keywords

Inequality, inequality of opportunity, inequality indices

JEL code

D31, D63

INTRODUCTION

The subject of socio-economic inequality is currently a very popular line of research. This is due to the fact that the level of inequality, which was gradually declining in the developed capitalist countries after the Second World War until the 80s of the 20th century, started to grow steadily again (Atkinson, 2018), creating conditions for the growth of social tension.

---

1 Bashkir State University, Institute of Economics, Finance and Business, 32 Zaki Validi, 450076 Ufa, Russia. Corresponding author: e-mail: badertdinova@mail.ru, phone: (+7)3472299608.

2 Institute of Economics and Management, Ufa State Aviation Technical University, 12 Karl Marx, 450000 Ufa, Russia. E-mail: tan-marina@mail.ru.

3 This manuscript follows the prior research "Transition report 2016–17. Inequalty of opportunity" conducted and published by the European Bank for Reconstruction and Development – online at: <*https://2016.tr-ebrd.com/inequality-of-opportunity*>.

The term «inequality» is rather ambiguous. One can understand it in a mathematical, emotionless sense simply as a statement of the fact that the amount of benefit an individual has (e.g. income) differs from subject to subject. In this case, the terms «inequality» and «variation» can be regarded as synonyms. But the concept of inequality is far more likely to hold a negative narrative, reflecting inconsistencies between the actual distribution of some benefit in society and the distribution of the benefit according to a certain kind of «ideal of justice». This in mind, the «ideal of justice» does not necessarily imply that all individuals should have equal amount of benefit; on the whole, there exists a public consensus that some people deserve a bigger share of the public pie. For example, with respect to income, there exist at least two arguments: the needs and the merit. It is commonly agreed, for example, that in order to achieve the same quality of life, families with a large number of children or people with disabilities are in need of higher income than single or healthy people; therefore, as a matter of common justice, a higher income should be made available to the former than to the latter. Besides, a public consensus also exists in respect of prominent figures being eligible to more benefit than others – because they «deserve it». Said otherwise, in the context of variation, all differences are equal to each other, while in terms of inequality some differences look natural and even fair, whereas others are unfair or even shameful.

An innovative vision of the «ideal of justice» as regards the distribution of benefit was put forward by the theory of equal opportunity originated in the Western social philosophy at the end of the 20th century. This theory stems from the idea that an individual should be responsible for what he is in full control of. Therefore, the differences in the amount of benefit arising from the factors dependent on different individuals (called effort-factors) are reasonable. On the contrary, the differences in the amount of benefit arising from the factors beyond an individual's control (called circumstance-factors) are unjust and subject to compensation. Thus, according to the theory of equal opportunities, the «ideal of justice» is to ensure that the inequality caused by circumstances beyond an individual's control (called inequality of opportunity) should be fully compensated, and the efforts, on the contrary, should be adequately rewarded.

Currently, a wide range of methods for measuring inequality of opportunity has been developed and tested on the microdata from various countries. In terms of describing the relationship between circumstances, efforts and achievements, the distinguished parametric and non-parametric approaches are applied. A specific type of the relationship function is selected in the parametric approach, of which the parameters are evaluated by means of regression analysis, whereas the function in the nonparametric approach is considered unknown. As to the way the equality of opportunity is understood, there exist two approaches: ex-ante and ex-post. The ex-ante approach builds upon the idea that equality of opportunity is considered achieved if the average achievement is the same for individuals in all groups that are homogeneous by circumstance-factors. The ex-post approach implies that the equality of opportunity is considered achieved when the achievements of individuals making equal efforts are the same.

Finally, methods for assessing inequality of opportunity vary depending on the measures of inequality they use. There is a wide choice of inequality indices in case of a continuous variable of achievement, given a plethora of measures developed and applied for measuring inequality, the best known of which being the Gini index, the Atkinson and Dalton indices families, and generalized entropy measures.

The aim of this work is to fill a gap in the empirical research on the assessment of inequality of opportunity related to the study of the impact of the choice of an inequality measure on the measurement result. The reason to do this work was necessitated by the results of the assessment of opportunity inequality in the Russian Federation presented in the EBRD's work Transition report (2016–17), performed by the European Bank for Reconstruction and Development (EBRD) based on the microdata of the sociological survey LiTS III (Life in Transition). This study measures inequality of opportunity in 33 countries (mainly economies in transition countries). Parents' educational background, their membership in the Communist Party, gender, nationality and place of birth of the respondent are taken as circumstance-factors. As reflected in the results obtained the contribution of inequality of opportunity to income inequality in the Russian Federation accounts for 34.5%. Our estimates of inequality of opportunity in the Russian Federation (Ibragimova and Frants, 2019; Ibragimova and Frants, 2020; Pauhofová et al., 2020), received on the basis of the RLMS data (wave 2011), are almost twice as small (19.2%) despite the fact that a set of circumstance-factors taken into consideration in our work is broader than in the EBRD study. Such a wide gap encouraged us to compare the EBRD results with other researchers' studies on the same countries, carried out using similar methods. The results of the comparison are shown in Table 1.

Table 1 Comparative analysis of the EBRD results with other researchers' studies on the same countries

Comparison	Transition report 2016–17	Marrero and Rodriguez (2012)	Brzeziński (2015)	Checchi et al. (2010)
Assessment methodology	Parametric, based on ex-ante approach	Parametric, based on ex-ante approach	Parametric, based on ex-ante approach	Parametric, based on ex-ante and ex post approaches
Applied measures of ineaquality	Gini index	L-Theil index	L-Theil index	L-Theil index
Countries	33 post-socialist states	23 European countries	23 European countries; the list of the countries is identical (Marrero and Rodriguez, 2012)	25 European countries
Assessment results	From 19.5% (Montenegro) up to 47.1 (Estonia), as well as Germany – 23.0%, Slovakia – 40.4%, Czech Republic – 41.9%	From 1.89% (Denmark) up to 22.22% (Portugal), as well as Germany – 2.07%, Slovakia – 3.60%, Czech Republic – 5.85%	2004: from 2.00% (Denmark) up to 18.00% (Portugal), as well as Germany – 2.5%, Slovakia – 2.5%, Czech Republic – 6.0% 2010: from 1.88% (Sweden) up to 18.00% (Greece), including Germany –3.1%, Slovakia – 7.3%, Czech Republic – 8.0%	Ex-ante: from 3.0% (Norway) up to 35.5% (Cyprus), ex post: from 15.8% (Slovenia) up to 48.6% (UK), including ex-ante: Germany – 21%, Slovakia – 14%, Czech Republic – 13%, ex-post: Germany – 33.3%, Slovakia – 31.4%, Czech Republic – 39.20%
Circumstance- factors	Parents' educational background and membership in Communist Party, birth place, gender, nationality	Parents' educational background, father's occupational status, the country of an individual's birth, welfare of the family where the individual was raised	Parents' educational background, father's occupational status, the country of an individual's birth	Gender, parents' educational background, father's occupational status, the country of an individual's birth, type of locality where the individual lives
Income measure	Self-reported incomes over the last 12 months, which come from a variety of different types of employment	Disposable equivalent income	Disposable equivalent income	Post-tax individual earnings
Informational basis	LITS III (end of 2015–beginning of 2016)	EU-SILC (2005)	EU-SILC (2005, 2011)	EU-SILC (2005)
Size of the sample serving as the basis for making calculations	≈ 500	From 1 159 up to 8 638	From 1 133 up to 8 338	From 2 104 up to 15 562

Source: Own construction

As can be seen from Table 1, a similar wide gap both between our estimates and those of the EBRD can often be found when comparing the estimates of the EBRD for a number of countries with the estimates made in other researchers' studies on the same countries. A comparative analysis indicates that the gap may arise due to a number of causes. Firstly, the EBRD study uses, unlike other works, the Gini index rather than the L-Theil index as a measure of inequality. Secondly, various measures of income are applied. Thirdly, the gap in the assessments may be related to the amount and time of data collection. Fourthly, the set of the circumstance-factors involved differs.

The inequality of opportunity assessment problems related to imperfect data are well known – empirical studies on assessing the inequality of opportunity use the data from ready-made sociological surveys, hence, the choice of circumstance-factors, effort-factors, as well as individual achievement measures is limited by the availability of data. We could not find any works which would carry out a data collection tailored for assessing the inequality of opportunity, or at least, would theoretically design a sociological survey focused on this task.

Neither have we seen any work dedicated to a detailed study of the extent to which the choice of a measure of inequality affects the assessment result. Most works on inequality of opportunity apply one measure of inequality, without strong reasoning of advantages of the index used. Those studies which apply a number of indices (they are discussed in detail below in Section 1.2), do it as a robustness check of the main result and do not set it as the main task of the study. In this paper, we aim at in-depth studying of this very aspect – to what extent the choice of inequality index affects the outcome of assessing the inequality of opportunity.

The work is structured as follows. The initial overview section contains, first, a brief description of concepts and approaches to assessing the inequality of opportunity, which enables to clarify the role of inequality indices in those calculations. Second, it reviews the works raising the issue of sensitivity of estimation results to the choice of an inequality measure. Third, it gives an overview of inequality indices, including their origins and properties. The second section describes the methods used to measure inequality of opportunity and the research information base. The third section presents the results of assessing the inequality of opportunity in terms of income for 16 countries based on the LiTS III survey, and discusses the effect of inequality index on the assessment result and a country's ranking by inequality of opportunity.

1 THEORETICAL OVERVIEW

1.1 Measuring the inequality of opportunity: conceptual approaches and the role of inequality indices

The research on inequality of opportunity is based on the 'equal opportunities for all members of society' concept, which at the end of the 20th century resulted from development of egalitarian theories of social justice. According to this concept, the achievements that are significant for everyone or the majority, such as income, material well-being, and the state of health, depend on two groups of factors – the circumstances for which an individual should not be responsible, and efforts, which, on the contrary, are in the area of personal responsibility. The inequality of achievement arising from the inequality of effort is treated as ethically acceptable, while the inequality arising from circumstances is unfair and therefore must be eliminated.

As scholars began attempting a mathematical formalization of this concept of equal opportunity, it quickly became clear that many related problems appear. An excellent overview of the issues of incompatibility between the principles of compensation and natural reward, as well as the incompatibility between ex-ante and ex-post approaches to determining the equality of opportunity based on the principle of compensation, is given in Ramos (2016).

The compensation principle implies that the inequality caused by inequality of circumstances must be eliminated. To date, two criteria have been proposed to assess whether the inequality of opportunity has been compensated – the ex-ante and the ex-post. The ex-ante approach, proposed by van der Gaer, provides that the equality of opportunity is achieved if the average achievement is the same for individuals in all groups that are homogeneous by circumstance-factors. The ex-post approach proposed by Roemer is based on the idea that the equality of opportunity is achieved when the achievement of individuals with the same effort is the same. As shown in Fleurbaey and Peragine (2013), the ex-ante and ex-post approaches are incompatible.

The natural reward principle implies that inequality of achievement caused by the inequality of effort must retain. The literature discusses two approaches to the implementation of this principle – the liberal and the utilitarian ones. The liberal approach is based on the idea that the achievements of individuals with the same circumstances should not be redistributed, because they are solely due to differences in effort. The utilitarian approach says «inequalities due to unequal effort do not matter», advocating a sum-maximizing policy among subgroups with identical circumstances. The same work (Fleurbaey and Peragine, 2013) proved that the liberal and the utilitarian approaches are incompatible with each other and with the ex-post approach to compensation.

Methods of assessing the inequality of opportunity are based on the principle of compensation and can be based on both the ex-ante and the ex-post approaches.

From the ex-ante point of view, the equality of opportunity is achieved if the average achievement of individuals in all the circumstance-factor homogeneous groups is the same. Therefore, the ex-ante-assessed inequality of opportunities is based on calculating v_i scalar measure for each individual which estimates the individual's particular set of circumstances. Understandably, the v_i measures will be identical for all individuals with the same circumstances. In the case of equality of opportunities, the vi values should be the same for all individuals in general. If this is not the case, then the inequality of opportunity can be estimated through calculating some inequality index *I* by distribution $\{v_i\}$ (hereinafter, curly braces will be used to denote the distribution). Thus, when assessing inequality of opportunity based on the ex-ante-approach, it is necessary to answer two questions: what is to be used as the v_i measure, and which inequality index *I* is to be chosen for assessing the inequality of distribution $\{v_i\}$. In practice, the average value of achievement is almost always used as v_i for all individuals, for which the set of circumstance-factor values coincides with the set of circumstance-factor values for the *i*-th individual. The L-Theil index is most often used as a measure of inequality, but there are many examples of using other measures of inequality.

The ex-post approach to determining the equality of opportunity is based on the idea that the achievements of individuals with the same efforts should be the same. Within this approach, it is necessary to assess an individual's efforts, but not the circumstances. That does complicate things, since the efforts are much less observable than circumstances. However, Roemer has proposed the way to bypass the problem known as the Roemer's identification assumption. According to his idea, the efforts of individuals from the same percentile of intragroup distribution (a group is defined as a set of individuals with the same circumstances) are the same. Therefore, it is possible to form effort-homogeneous groups from groups that are homogeneous in circumstances. In the literature on inequality of opportunity, these groups are commonly referred to as tranches. According to the ex-post approach, the equality of opportunity is achieved if the intra-tranche variation of achievement is equal to 0. If this is not the case, then to assess the inequality of opportunity we must, first, calculate the inequality indices *I* for each tranche and, second, aggregate them to obtain a generalized estimator of the intra-tranche component of variation, which is a measure of inequality of opportunity in a population. Accordingly, in the case of assessing the inequality of opportunity based on the ex-post approach, it is again necessary to decide on which particular inequality index to use. As in the case of the ex-ante approach, the L-Theil index is most often used, but other measures of inequality are also used quite frequently.

Thus, the above review shows that the existing approaches to measuring the inequality of opportunity are based on the principle of compensation and can rely on either the ex-ante or the ex-post interpretation of this principle. In both the ex-ante and ex-post assessments, there exists the problem of choosing an inequality index.

1.2 A review of works addressing the sensitivity of opportunity inequality assessments to the choice of inequality measure

Empirical studies on the contribution of inequality of opportunity to economic inequality are of a wide geography, and the assessment methods applied are constantly being improved and supplemented. However, they rarely consider the effect of choosing a measure of inequality on the outcome. Most works concentrated on the assessing of inequality of opportunity, make use only of one measure of inequality, and in most cases this is the L-Theil index. In this section, we will consider several works that made use of different measures of inequality and discussed their impact on the outcome of the assessment and countries' ranking.

The paper Bourguignon et al. (2003) is concentrated on assessing the inequality of opportunity in Brazil supported by the PNAD data, National Household Survey, 1996. The evaluation technique can be described as parametric basing on the ex-post approach. The study considers the following set of circumstance-factors: race, parents' educational background, occupational status of the father. Two indicators were used as income measures: the actual hourly rate of pay and equivalent household income. The calculation was performed for seven age cohorts. The results for the measure of income, namely equivalent household income, are listed in Table 2.

Table 2 Inequality of opportunity in Brazil (based on the results from Bourguignon et al., 2003)

Years of birth		1936–1940	1941–1945	1946–1950	1951–1955	1956–1960	1961–1965	1966–1970
General inequality	Gini index	0.605	0.602	0.588	0.591	0.597	0.594	0.573
	T-Theil index	0.750	0.736	0.682	0.720	0.709	0.691	0.635
Residual inequality	Gini index	0.474	0.492	0.481	0.489	0.490	0.478	0.482
	T-Theil index	0.434	0.475	0.455	0.468	0.465	0.429	0.437
Absolute inequality of opportunity	Gini index	0.131	0.110	0.107	0.102	0.107	0.116	0.091
	T-Theil index	0.316	0.261	0.227	0.252	0.244	0.262	0.198
Relative inequality of opportunity	Gini index	21.7%	18.3%	18.2%	17.3%	17.9%	19.5%	15.9%
	T-Theil index	42.1%	35.5%	33.3%	35.0%	34.4%	37.9%	31.2%

Source: General and residual inequality (Bourguignon et al., 2003), we added the calculation of the absolute and relative inequality of opportunity in Table 3

As we can see from the Table 2, the use of the Theil index results in a higher contribution of inequality of opportunity to household income inequality than the one obtained through the Gini index, the differences being very significant.

Pistolesi (2009) explores the inequality of opportunity in the USA during the period from 1968 to 2001 based on the Michigan Panel Study of Income Dynamics. The evaluation technique is parametric and based on the ex-post approach in two modifications – with direct and indirect estimation methods. The work takes into account the following set of circumstance-factors: age and educational background of the parents, occupational status of the father, region of birth, and individual's affiliation with a black minority group. As a measure of income, the annual labor income was used: the average annual labor income over 3 years, the average annual labor income over 5 years. A whole range of inequality measures used by the author include: T-Theil index, L-Theil index, GE (2), the standard logarithmic deviation, the Gini index. The results are presented in Table 3.

Table 3 Relative inequality of opportunity (based on the results Pistolesi, 2009)

	Direct method					Indirect method				
	T-Theil	L-Theil	GE(2)	SDI	Gini	T-Theil	L-Theil	GE(2)	SDI	Gini
Annual labour income										
Average value	23.6%	23.3%	17.2%	35.3%	32.2%	27.6%	29.4%	34.3%	27.9%	24.7%
Minimum	15.0%	16.4%	7.5%	23.9%	24.4%	7.3%	10.0%	16.9%	18.3%	14.6%
Maximum	33.7%	34.2%	25.6%	47.2%	41.9%	41.8%	48.2%	44.5%	35.6%	33.6%
Average annual labour income for 3 years										
Average value	26.0%	27.1%	19.2%	41.3%	34.6%	30.5%	32.8%	40.0%	31.7%	25.8%
Minimum	17.2%	19.0%	9.9%	31.9%	26.7%	7.9%	11.3%	20.6%	20.5%	15.0%
Maximum	37.1%	38.8%	28.7%	52.4%	44.4%	43.8%	56.5%	53.0%	40.8%	35.0%
Average annual labour income for 5 years										
Average value	27.0%	28.4%	20.1%	43.0%	35.5%	31.6%	34.3%	41.8%	32.7%	26.3%
Minimum	18.6%	20.8%	11.5%	34.9%	28.1%	8.1%	12.8%	21.9%	22.0%	15.3%
Maximum	37.5%	39.4%	29.2%	53.0%	44.7%	44.6%	55.4%	53.8%	41.1%	35.9%

Note: Average, minimum and maximum values for the years of 1968–2001.
Source: Pistolesi (2009)

As indicated in Table 3, the contribution of inequality of opportunity to income inequality significantly varies depending on the choice of the inequality measure. Comparing the results obtained with the help of the L-Theil and Gini indices, we can see that when applying the direct method, the relative inequality of opportunity obtained with the Gini index is substantially greater than with the L-Theil index. On the contrary, if the indirect method is used, the relative inequality of opportunity turns out to be considerably larger when obtained with the L-Theil index rather than with the Gini index.

Björklund et al. (2011) is another work that addresses the impact of inequality measures on the outcome of measuring inequality of opportunity. This study explores the inequality of opportunity in Sweden. The evaluation technique used can be characterized as parametric based on the ex-ante approach. The total market income, averaged over 7 years when the person was aged 37–43, is used as a measure of income. The analysis included only men born in the period of 1955–1967. The authors used a comprehensive set of circumstance-factors: income and parents' educational background, type of the family in which the individual grew up, the number of siblings, IQ and the body mass index of an individual at the age of 18 years. Four measures of inequality were used for the calculation: the Gini index, the L-Theil index, the T-Theil index, and GE (2). The results are presented in Table 4.

As follows from Table 4, the contribution of inequality of opportunity depends heavily on the choice of the measure of inequality. In this work, the result obtained with the help of the Gini index is considerably higher than the result obtained through the L-Theil index.

Hederos et al. (2017) that also deals with the assessment of inequality in Sweden, was carried out using the same evaluation technique as Björklund et al. (2011). The difference lies in the fact that the analysis included both men and women, and a somewhat different set of circumstance-factors: income and educational background of parents, type of the family, the number of brothers and sisters, IQ of an individual and his non-cognitive skills. The results are shown in Table 5.

Table 4 Inequality of opportunity (based on the results Björklund et al., 2011)

Measure of inequality	Gini index	L-Theil index	T-Theil index	GE(2)
All individuals (born in the period of 1955–1967)				
General inequality	0.263	0.158	0.183	3.946
Contribution of residual inequality	71.8%	86.9%	79.3%	58.9%
Contribution of inequality of opportunities	28.2%	13.1%	20.7%	41.1%
Individuals born in the period of 1955–1959				
General inequality	0.231	0.122	0.111	0.379
Contribution of residual inequality	73.4%	90%	85.3%	81.3%
Contribution of inequality of opportunities	26.6%	10%	14.7%	18.7%
Individuals born in the period of 1960–1967				
General inequality	0.279	0.181	0.236	6.976
Contribution of residual inequality	68.7%	82.6%	62.5%	5.7%
Contribution of inequality of opportunities	31.3%	17.4%	37.5%	94.3%

Source: Björklund et al. (2011)

Table 5 Inequality of opportunity (based on the results Hederos et al., 2017)

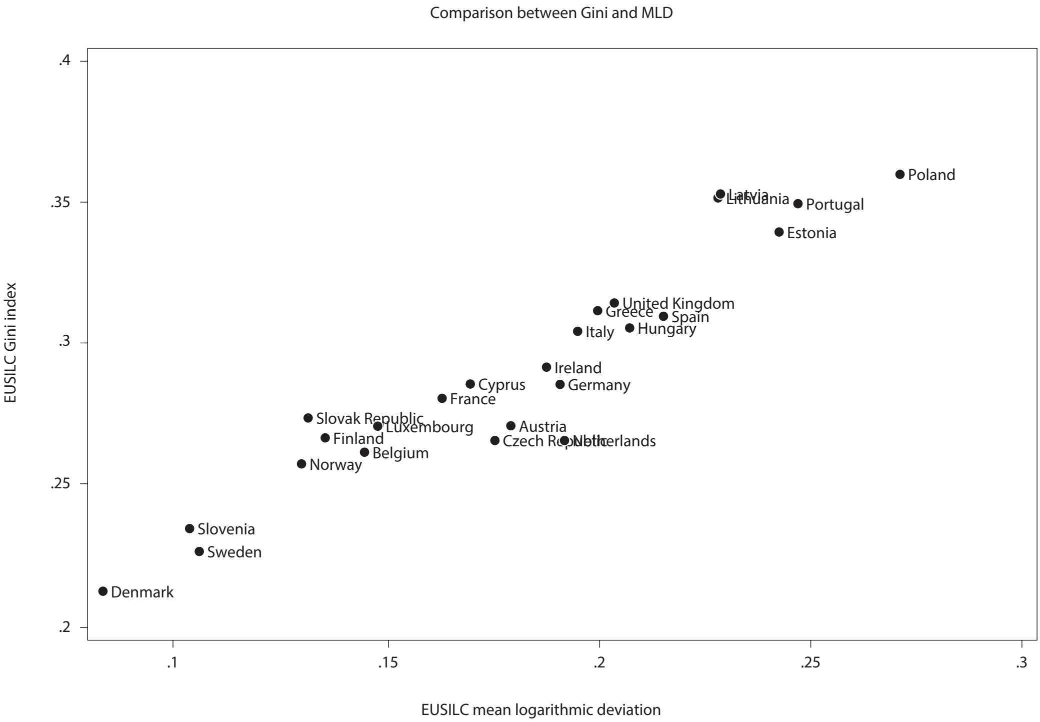
Measure of inequality	Gini index	L-Theil index	T-Theil index	GE(2)
Men				
General inequality	0.303	0.197	0.226	1.754
Contribution of residual inequality	69.0%	84.1%	78.9%	70.6%
Contribution of inequality of opportunity	31.0%	25.9%	21.2%	29.4%
Women				
General inequality	0.240	0.136	0.122	0.476
Contribution of residual inequality	75.0%	90.7%	85.5%	77.2%
Contribution of inequality of opportunity	25.0%	9.3%	14.5%	22.8%
All				
General Inequality	0.296	0.186	0.204	1.450
Contribution of Residual Inequality	62.9%	80.1%	73.5%	62.1%
Contribution of Inequality of Opportunity	37.1%	19.9%	26.5%	37.9%

Source: Hederos et al. (2017)

As shown in Table 5, the conclusions that can be made about the variation of the estimates depending on the measure of inequality are similar to the conclusions made in the previous work.

The issue of the extent to which the choice of inequality index affects a country's ranking is poorly studied. We are aware of only one work on inequality of opportunity in which some attention is paid to this aspect – Checchi (2010) compares inequality assessments using the Gini and L-Theil indices (Figure 1) in European countries. As the figure shows, the correlation between these indices is high, but the ranking of countries is different. We do not know of any studies in which inequality of opportunity would be assessed for a number of countries using several inequality indices, so we believe that this work is the first study of this kind.

Figure 1 Correlation of the Gini and L-Theil indexes



Source: Checchi (2010)

In general, the analysis of works in which the impact of a measure of inequality on the result of assessing the inequality of opportunity was addressed shows that the choice of an inequality measure affects the result of the assessment. Comparison of the results shows that sometimes the contribution of inequality of opportunity to income inequality when using the L-Theil index turns out to be much larger, and sometimes much less than when using the Gini index. What this is due to – peculiarities of assessment methodology, or information base, or income distribution in the country – remains unclear.

1.3 Inequality measures, their origin and properties

Due to our study being focused on the dependence of the inequality of opportunity assessment result on the choice of an inequality index, in this section we give a short but multi-aspect overview of inequality indices, reflecting the theoretical prerequisites for their occurrence, as well as the properties and features that are important for practical application. The below overview of measures of inequality does not claim to be complete – in writing this section we focused on the measures of inequality used to assess income inequality, since the purpose of our study is to measure inequality of opportunity in terms of earned income.

The measures of inequality being used in practice have three «sources of origin» (Cowell, 2009):

1. The inequality measures «borrowed» from the list of measures of variation used in statistics

Variation in statistics means the changeability of a parameter, its ability to take different values; accordingly, the variation measures are the indicators of variability. As mentioned previously, inequality is a concept that is similar to variation, albeit not identical; therefore, the idea of adoptig statistical measures of variation to measure inequality is plain to see. Measures of inequality borrowed from statistics comprise:

- Measures based on the percentiles of the distribution of the continuous variable, such as decile and quintile coefficients of funds, the Palma ratio, and the percentiles. The deciles (quintile) coefficients of funds is calculated as the ratio of the average income of the 10 (20%) of the richest individuals to the average income of 10 (20%) of the poorest ones. The Palm coefficient is defined as the ratio of the total income of 10% of the richest individuals to the total income of 40% of the poorest ones. The percentiles ratio, as the name implies, is calculated by dividing one percentile of distribution by another. In practice, the ratio of 90% percentile to 10% is commonly used; however, there other options are also possible.
- Measures based on the formula of variance, including the variance itself (V), coefficient of variation (c), logarithmic variance (v), variance of logarithms (v1). Formulas (1)–(4) demonstrate how the listed indicators are calculated:

$$V = \frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2 , \tag{1}$$

$$c = \sqrt{V / \bar{y}} , \tag{2}$$

$$v = \frac{1}{n} \cdot \sum_{i=1}^{n} (\ln (y_i) - \ln (\bar{y}))^2 , \tag{3}$$

$$v1 = \frac{1}{n} \cdot \sum_{i=1}^{n} (\ln (y_i) - \ln (\tilde{y}))^2, \tilde{y} = \sqrt[n]{y_1 \cdot y_2 \cdot \ldots \cdot y_n} . \tag{4}$$

- Relative mean deviation (M), calculated using Formula (5):

$$M = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_i}{\bar{y}} - 1 \right| . \tag{5}$$

- The Gini index, calculated using Formula (6):

$$G = \frac{1}{2 \cdot n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \frac{y_i - y_j}{\bar{y}} \right| . \tag{6}$$

The notation in the formulas stated above and below is as follows: y_k is the income of the k-th individual, $\bar{y}$ is the average income, n is the population size.

The Gini Index, popularized by Italian statistician Corrado Gini, is undoubtedly the most well-known measure of inequality used in practice. In spite of the fact that the index is named Gini, Gini himself recognized that it was German scientists Karl Christopher von Andre and Friedrich Robert Helmert (Atkinson, 2018) who developed the basic statistics (mean difference).

2. Measures of inequality based on the theory of social welfare functions (SWF)

A SWF is a function that links the public's satisfaction with its social well-being to the characteristics of that state. Social well-being is simply the vector of values of individual characteristics of the members of society, reflecting their economic sustainability. The idea of measuring inequality on the basis of social welfare functions emerged as an attempt to measure inequality as the degree of satisfaction (or dissatisfaction) of society with the distribution of any individual benefit, in particular, income.

The following two families of measures of inequality stem from the SWF-functions: the Dalton and Atkinson indices. The Dalton inequality indices family is determined through Formula (7). As Formula (7) states, the Dalton index is equal to 0, if the income of all members of the society are the same; the Dalton index will increase with the growth of income inequality:

$$D(\beta) = 1 - \frac{SWF(y_1, y_2, \dots, y_n)}{SWF(\bar{y}, \bar{y}, \dots, \bar{y})} = 1 - \frac{\sum_{i=1}^{n} y_i^{1-\beta} - 1}{n \cdot (\bar{y}^{1-\beta} - 1)}. \tag{7}$$

The shortcoming of the Dalton indices lies in the fact that they are sensitive to multiplying by a number. Therefore, Atkinson suggested the following modification of the inequality index (Formula 8). Atkinson's idea was to calculate the average utility of an individual's income, then to obtain an income corresponding to that average utility and then to compare it with the actual average income. As it follows from Formula (8), the Atkinson indices, unlike the Dalton indices, are insensitive to the multiplication by a number:

$$A(\beta) = 1 - \frac{U^{-1}(\bar{U})}{\bar{y}} = 1 - \left[\frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{y_i}{\bar{y}}\right)^{1-\beta}\right]^{\frac{1}{1-\beta}}. \tag{8}$$

Thus, the social welfare functions theory is the source of origin for two families of inequality indices – the Dalton and the Atkinson ones.

It is interesting to note that the Gini index was interpreted in terms of the SWF theory in Sen (1973). He has shown that the Gini index can be calculated by Formula (9) which is equivalent to Formula (6):

$$G = 1 - \frac{1}{n^2 \cdot \bar{y}} \sum_{i=1}^{n} \sum_{j=1}^{n} Min(y_i, y_j). \tag{9}$$

This form of mathematical notation leads to the following definition: if the income level of any pair of individuals is equaled to the income level of the poorer, and if the public welfare is the total of the welfare of all possible pairs, then the inequality index for such public welfare function will be determined according to the Gini index formula.

3. Measures of inequality based on the information theory

One of the key tasks of the information theory is to measure the value of information on the likeliness of a certain event among many possible ones to occur. If the probability of the event is close to 1, then the value of such information is obviously small, but if, on the contrary, the event is unlikely to happen, then the value of information should increase. Thus, the value of information h to the effect that an event will take place should decrease monotonically as its probability increases. The information theory proposes a measure of disorder of a system, which is a weighted sum of the values of the information of the events. The weights in this construction are the probabilities of the events under consideration (Formula 10):

$$\text{Entrophy} = \sum_{i=1}^{h} h(p_i) \cdot p_i. \tag{10}$$

In 1967, Theil made the following proposal: if we reformulate the problem then the concept of entropy will provide a powerful tool for measuring inequality. Consider n individuals instead of n events, use the individual's share in the «overall pie» instead of the probability of an event, interpret the value of the information about the event as the «coordinate» of the individual that allows you to «measure the distance» between two individuals depending on their shares in the total income. Inequality measures based on the theory of information are constructed so that the change in the measure of inequality in the transfer of a part of the income from a richer individual to a poorer one should depend, firstly, on the size of the transfer, and secondly, on the «distance» between the individuals participating in a small transfer (Formula 11):

$$\Delta Inf = \Delta s \cdot (h(s_R) - h(s_P)) . \tag{11}$$

In Formula (11), *Inf* – measure of inequality generated by the theory of information, ΔInf – change in the measure of inequality due to the transfer, s_j – share of the j-th individual in the «overall pie», s_R, s_P are respectively shares of the «rich» and «poor» individuals that participate in the transfer, $\Delta s = s_R - s_P$, $\Delta s < (s_R - s_P) / 2$, $h(s_R) - h(s_P)$ – «distance» between the individuals under consideration.

The choice of function $h(s)$ reflects different ways of determining the distance between two individuals. Formula (12) below determines a class of the functions appropriate for use as $h(s)$:

$$h(s, \beta) = \begin{cases} \frac{1-s^\beta}{\beta}, \beta <> 0 \\ -\ln(s), \beta = 0 \end{cases} . \tag{12}$$

Index of inequality *Inf*(β) is formed as the difference between the entropy that would occur if the shares of income of all individuals were equal and actually takes place (Formula 13):

$$Inf(\beta) = \frac{1}{1+\beta}\left[\sum_{i=1}^{n} \frac{1}{n} h\left(\frac{1}{n}\right) - \sum_{i=1}^{n} s_i \cdot h(s_i)\right]. \tag{13}$$

As regards the practice of measuring inequality, a number of properties are distinguished which are important for measuring inequality, whereas the first four properties are considered basic and supported by almost all researchers, the fifth one is more likely to be that of a recommendation.

Independence from multiplication by a number: if all individuals' income is multiplied by the same number, then the measure of inequality should not change.

The principle of population: if a new population is obtained by combining with the second similar one, then the measure of inequality should remain unchanged.

Decomposability: if we divide the population into several constituent subgroups, then the inequality in the entire population can be represented as a function of intra-group inequalities, intra-group averages and group sizes.

Additive decomposability is singled out separately: when the inequality index for the entire population can be represented as the algebraic sum of the intra-group and inter-group components.

Weak transfer principle: if there is a transfer of the income from the richer to the poorer individual, so that the rich individual is still richer after the transfer than the poor one, then the inequality index should decrease.

In the case of 1–4 properties, a very important theorem has been proved: for an inequality measure to exhibit properties 1–4, it can be represented in the form (14) or take the form of an ordinally equivalent transformation (14). Simple transformations result in Formula (15) in which $\beta = \alpha - 1$.

$$GE(\alpha) = \frac{1}{\alpha^2 + \alpha}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{y_i}{\bar{y}}\right)^\alpha - 1\right], \tag{14}$$

$$Inf(\beta) \cdot n^{\beta} = GE(\alpha). \tag{15}$$

A family of measures $GE(\alpha)$ is called generalized entropy measures. Two measures are most popular in this family, especially in the works on inequality of opportunity: $GE(0)$, or the Mean Logarithmic Deviation, or the L-Theil index, and $GE(1)$, or the T-Theil index, which are «special cases» that are not calculated according to Formula (14), but according to Formulas (16) and (17), respectively.

$$GE(0) = \frac{1}{n} \cdot \sum_{i=1}^{n} ln\left(\frac{\bar{y}}{y_i}\right), \tag{16}$$

$$GE(1) = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{y_i}{\bar{y}} ln\left(\frac{y_i}{\bar{y}}\right). \tag{17}$$

Finally, a strong principle of transfers: a decrease of inequality index due to transfers should depend only on the «distance» between the rich and poor individuals. This principle is desirable because it allows us to formalize the intuitive idea: the greater the difference between the two individuals in terms of income, the greater should be the leveling effect provided by a transfer between them. If the measure of inequality satisfies the strong principle of transfers, then it can be interpreted as the average distance between the individual's actual share in the «overall pie» and his share in a society of 100% income equality. It is clear from the foregoing, that the idea of distance is definitely laid down in the measures resulting from the theory of information, and therefore they satisfy the strong principle of transfer.

As to the Gini index, it is definitely the «atypical reaction» of it to the transfer that gives rise to the criticism of this measure of inequality. The matter is that the change in the Gini index due to a transfer between two individuals depends on the ranks of these individuals in an ascending order of incomes: the more significant difference between the ranks results in a greater reduction of the Gini index. Therefore, a transfer between two individuals in the middle results in a more significant decrease in the Gini index than a transfer between the individuals at its beginning or the end.

The properties of different measures of inequality are given in Table 6.

Thus, there is a broad range of inequality measures of different origin, set of properties, and degree of popularity among the researchers. Choosing a specific measure of inequality from the list of possible

Table 6 Properties of inequality measures

Measure of inequality	Independence from multiplying by a number	Principle of population	Decomposability	Weak transfers principle	Strong transfers principle
V	–	+	+	+	+
C	+	+	+	–	+
M	+	+	–	±	–
V	+	+	–	–	–
v1	+	+	–	–	–
G	+	+	–	+	–
D(β)	–	+	+	+	–
A(β)	+	+	+	+	–
GE(α)	+	+	+	+	+

Source: Cowell (2019)

ones is a challenging task. In this article, we will try to find out how strongly the choice of the measure of inequality affects the results of measuring inequality of opportunity and which measures of inequality are more preferable for the given task. For a more detailed study of the issue, in our calculations we will use a wide palette of inequality indices from all the sources of origin listed above: the Gini index (GINI), the Atkinson indices A(2), A(1), A(0.5), the indices from the generalized entropy measures family GE(−1), GE(0), GE(1), GE(2).

Thus, with our research we contribute to the existing publications on the assessment of inequality of opportunity by way of: 1) applying a large number of inequality indices and exploring their influence on the outcomes of the assessment; 2) performing calculations not for one country, but for a number of countries using the same assessment methods and one set of circumstance-factors; and 3) exploring the impact of selecting a measure of inequality on a country's ranking by the absolute and relative inequality of opportunity.

2 METHODOLOGY AND INFORMATION BASE OF RESEARCH

The evaluation technique we have used in this study can be characterized as parametric and based on an ex-ante approach to the interpretation of equal opportunities. The choice of evaluation technique was determined by the fact that we wanted to proceed from the results obtained in Transition report (2016–17), which encouraged us to carry out this work. Unfortunately, in Transition report (2016–17) there is no detailed description of the calculation procedure, therefore, we cannot argue that our methodology is totally identical, but in general, the brief description, that is presented there, is sufficient to understand how the calculation was done.

The method under consideration was first proposed Ferreira and Gignoux (2011), and, in our opinion, is currently the most popular tool used in measuring the inequality of opportunity. This methodology uses predicted achievement values $\hat{y}_i$ calculated on the basis of regression of individual achievements on circumstance-factors.

The use of regression analysis requires selecting a functional form of relationship between the factors and the achievement. In Transition report (2016–17) there is no information on the specification used, but, actually, all studies on inequality of opportunity that use personal income or an individual's earnings as the indicator of achievement apply the semilogarithmic form of relationship (18).

$$ln(y_i) = C_i \cdot \alpha + u_i \,. \tag{18}$$

In Formula (18) y_i is the achievement of the i-th individual, α is the vector of regression coefficients, C_i is the vector of circumstance-factor values, u_i is a random error encapsulating the influence of unobservable factors, including the efforts for individual achievement.

Variation of $\hat{y}_i$ is determined only by variation of the circumstance-factors included in the model, while variation of $\hat{u}_i$ is the variation determined by other factors affecting the individual achievement, including the effort-factors and random factors. In this regard, there exist two options for assessing the inequality of opportunity – the direct and the indirect ones. In case of the direct method using a measure of inequality I, the inequality in distribution of $\hat{y}_i$ is measured and the $I(\{\hat{y}_i\})$ value is used as an absolute measure of inequality of opportunity. To assess the contribution of inequality of opportunity to the inequality of achievement, a relative measure of inequality of opportunity is calculated by formula $\theta_d = I(\{(\hat{y}_i\}) / I(\{y_i\})$. In case of the indirect method using a measure of inequality I, the inequality in $\{\hat{u}_i\}$ is measured and the $I(\{y_i\}) - I(\{\hat{u}_i\})$ value is used as an absolute measure of inequality of opportunity. We note that Ferreira and Gignoux (2011) describe the direct method only. The possibility of using the indirect method of assessment is discussed in Ramos (2016).

So, compared to Transition report (2016–17), we have expanded the methodological aspect by, firstly, having carried out a calculation using a range of inequality indices (A(2), A(1), A(0.5), GE(−1), GE(0),

Table 7 Descriptive statistics

Indicator	BLR	BUL	CRO	CPR	CZK	EST	DEN	HAN	ITA	KAZ	LAT	LIT	MON	MNG	RUS	SVK
Father's educational background																
Lower secondary education or less	8.94	35.36	49.78	60.74	61.55	33.88	49.79	57.31	55.10	21.36	25.34	38.72	40.88	37.00	15.42	64.00
(Upper) secondary education	32.78	49.46	34.51	26.67	30.32	20.45	19.36	32.95	35.35	33.15	18.95	18.61	26.37	40.36	26.39	29.65
Post-secondary non-tertiary education	38.91	6.79	1.11	1.98	0	28.31	15.11	4.18	2.07	25.97	33.11	16.35	10.99	8.97	33.23	0
Tertiary education and more advanced	19.37	8.39	14.60	10.62	8.12	17.36	15.74	5.57	7.48	19.52	22.60	26.32	21.76	13.68	24.96	0.35
Mother's educational background																
Lower secondary education or less	10.26	35.89	55.75	63.46	66.97	31.40	54.26	61.25	58.12	22.65	25.11	37.59	42.64	45.07	14.79	65.88
(Upper) secondary education	32.62	49.46	30.31	25.43	27.80	20.66	18.94	28.31	35.67	31.68	19.86	17.67	27.69	38.79	26.87	28.71
Post-secondary non-tertiary education	36.59	4.64	1.33	2.47	0.18	26.86	8.94	6.03	2.39	27.81	29.91	13.35	10.33	8.07	30.68	0.94
Tertiary education and more advanced	20.53	10.00	12.61	8.64	5.05	21.07	17.87	4.41	3.82	17.86	25.11	31.39	19.34	8.07	27.66	4.47
Gender																
Male	44.70	51.25	47.35	48.15	45.49	43.18	58.94	52.44	54.94	41.25	42.47	44.17	50.77	51.57	38.31	43.76
Female	55.30	48.75	52.65	51.85	54.51	56.82	41.06	47.56	45.06	58.75	57.53	55.83	49.23	48.43	61.69	56.24
Birthplace																
Urban	69.04	79.64	86.50	70.62	86.64	70.04	56.81	81.21	84.87	41.99	78.31	59.77	26.15	77.58	71.70	81.88
Rural	30.96	20.36	13.50	29.38	13.36	29.96	43.19	18.79	15.13	58.01	21.69	40.23	73.85	22.42	28.30	18.12
Ethnic																
Minority	9.11	13.57	2.65	9.38	31.95	24.59	5.96	2.78	2.07	38.31	25.57	12.59	16.04	45.29	8.90	7.06
Majority	90.89	86.43	97.35	90.62	68.05	75.41	94.04	97.22	97.93	61.69	74.43	87.41	83.96	54.71	91.10	92.94
Age																
Average	40.13	43.88	41.49	41.21	42.28	44.77	39.18	43.00	42.17	40.38	42.37	43.43	38.75	39.28	38.97	43.60
Number of observations	604	560	452	405	554	484	470	431	628	543	438	532	455	446	629	425

Source: Author's calculations

GE(1), GE(2), GINI), and secondly, making calculations by both direct (similar to Transition report (2016–17), and indirect methods.

The study is based on the data of «Life in Transition III» (LITS III) sociological survey (wave 2016), conducted by the European Bank for Reconstruction and Development. The LITS III survey is conducted in 34 countries (Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, the Czech Republic, Estonia, Macedonia, Georgia, Germany, Greece, Hungary, Italy, Kazakhstan, Kosovo, Kyrgyzstan, Lithuania, Latvia, Moldova, Mongolia, Montenegro, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Tajikistan, Turkey, Ukraine, and Uzbekistan). The following factors were used in our work as circumstance-factors: parents' educational background, place of birth, gender, individual's nationality. The set of circumstance-factors used by us differs from the one used in Transition report (2016–17), our work not including the membership of respondents' parents in the Communist Party to the analysis, because firstly, this is not relevant for all the countries under consideration, and secondly, this factor turned out to be insignificant in Transition report (2016–17).

The total number of the respondents in the LITS III survey for each country amounts approximately to 1 500 people. However, there are a lot of omissions in the data. After the removal of the respondents with gaps in the data and limiting the sample to the respondents aged 18–65, the sample size decreased threefold, in some countries the sample size decreased very significantly. In this regard, we limited ourselves to the countries the number of observations in which comprised minimum 400 respondents. Descriptive statistics for these 16 countries are shown in Table 7.

As we can see in Table 7, there are noticeable differences in the distribution of variables. Firstly, it concerns the level of the parents' education. Perhaps these differences are due to the fact that the LITS survey makes an attempt to create some universal categories of the levels of education for a whole range of countries, the educational systems of which differ significantly. Quite big differences across the countries are also noted regarding the place of respondents' birth. In most countries, there are more people born in urban areas than those born in rural areas, but there are also exceptions (Kazakhstan, Montenegro).

3 RESULTS AND DISCUSSION

Table 8 presents the results of the evaluation of Formula (18) for all the countries. As follows from Table 8, the gender factor is highly significant in relation to labor income: women earn less than men in all the countries. A higher educational level of parents makes a positive impact on the individual's income; in many countries, the coefficients of the parents' educational level often tend to be positive and significant. In this case, mother's educational status is obviously more important than the educational status of the father (since mother's educational level results in more significant coefficients than those of the father). In most countries birth in rural areas does not appear to make a significant effect on the level of labor income, but there are exceptions: birth in rural areas has a significant negative effect on the level of labor income in the Czech Republic, Estonia, Kazakhstan, Mongolia, and Russia.

Figure 2 presents a comparative analysis of the contribution of opportunity inequality to labor income inequality across the countries when direct and indirect methods as well as various measures of inequality are used.

As we can see in Figure 2, the choice of the inequality measure has a marked impact on the measurement results and consequently on the interpretation of the results. When the Gini index is used, the contribution of opportunity inequality to labor income inequality is much larger than when using other measures of inequality with the direct method of assessment, and vice versa, it is markedly smaller when the indirect method of assessment is applied. The estimates obtained through other considered measures of inequality are more uniform.

Table 8 Results of the OLS-regression of the labor income logarithm on circumstance-factors

Indicator	BLR	BUL	CRO	CPR	CHE	EST	DEN	HAN	ITA	KAZ	LAT	LIT	MON	MNG	RUS	SVK
Father's educational background																
Lower secondary education or less																
(Upper) secondary education	.046	.009	.099**	.058	.111***	.083	.074	.075	.063	.054	.108	–.015	.113	.088	.142	.189***
Post-secondary non-tertiary education	–. 033	.149	–.189	.268		–.063	.163**	.246**	.103	.033	–.018	.028	–.155	.045	.163	
Tertiary education and more advanced	.182	.133	.224***	.526***	.033	.115	.107	.032	.092	.313**	.090	.126	.068	.080	.187	.304**
Mother's educational background																
Lower secondary education or less																
(Upper) secondary education	.217**	.265**	.121***	–.131	.051	.202***	.023	.113**	.040	.158	.192**	.088	.081	.053	.043	–.021
Post-secondary non-tertiary education	.305***	.273*	.338**	–.542***	.790***	.187***	.073	.285***	.036	.037	.142	.099	.354***	.153*	.143	.140*
Tertiary education and more advanced	.362***	.307**	.141**	–.285**	.226***	.340***	.024	.300**	.047	.127	.421***	.197**	.354***	.143	.193	–.013
Gender																
Female	–.124**	–.232***	–.092***	–.252***	–.217***	–.372***	–.171***	–.217***	–.212***	–.123**	–.309***	–.303***	–.211***	–.202***	–.249***	–.211***
Birthplace																
Rural	–.167	–.095	–.063	.018	–.102**	–.076*	.005	–.055	–.006	–.163**	.044	–.032	–.189***	.051	–.197***	–.081
Ethnicity																
Minority	–.000	.263**	–.117	–.266***	–.031	–.221***	–.244**	–.225	.083	–.002	–.077	–.049	–.073	.013	.149*	–.075
Constant	15.59***	6.62***	8.41***	7.20***	9.77***	6.72***	7.68***	11.71***	7.27***	11.27***	6.35***	6.42***	13.52***	6.02***	10.44***	6.46***
N	**604**	**560**	**452**	**405**	**554**	**484**	**470**	**431**	**628**	**543**	**438**	**532**	**455**	**446**	**629**	**425**
R^2	0.089	0.087	0.150	0.153	0.135	0.262	0.071	0.137	0.110	0.061	0.162	0.107	0.133	0.086	0.094	0.139
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: * p<0.1, ** p<0.05, *** p<0.01.
Source: Author's calculations

Figure 2 Contribution of inequality of opportunity to income inequality across the countries (in %)

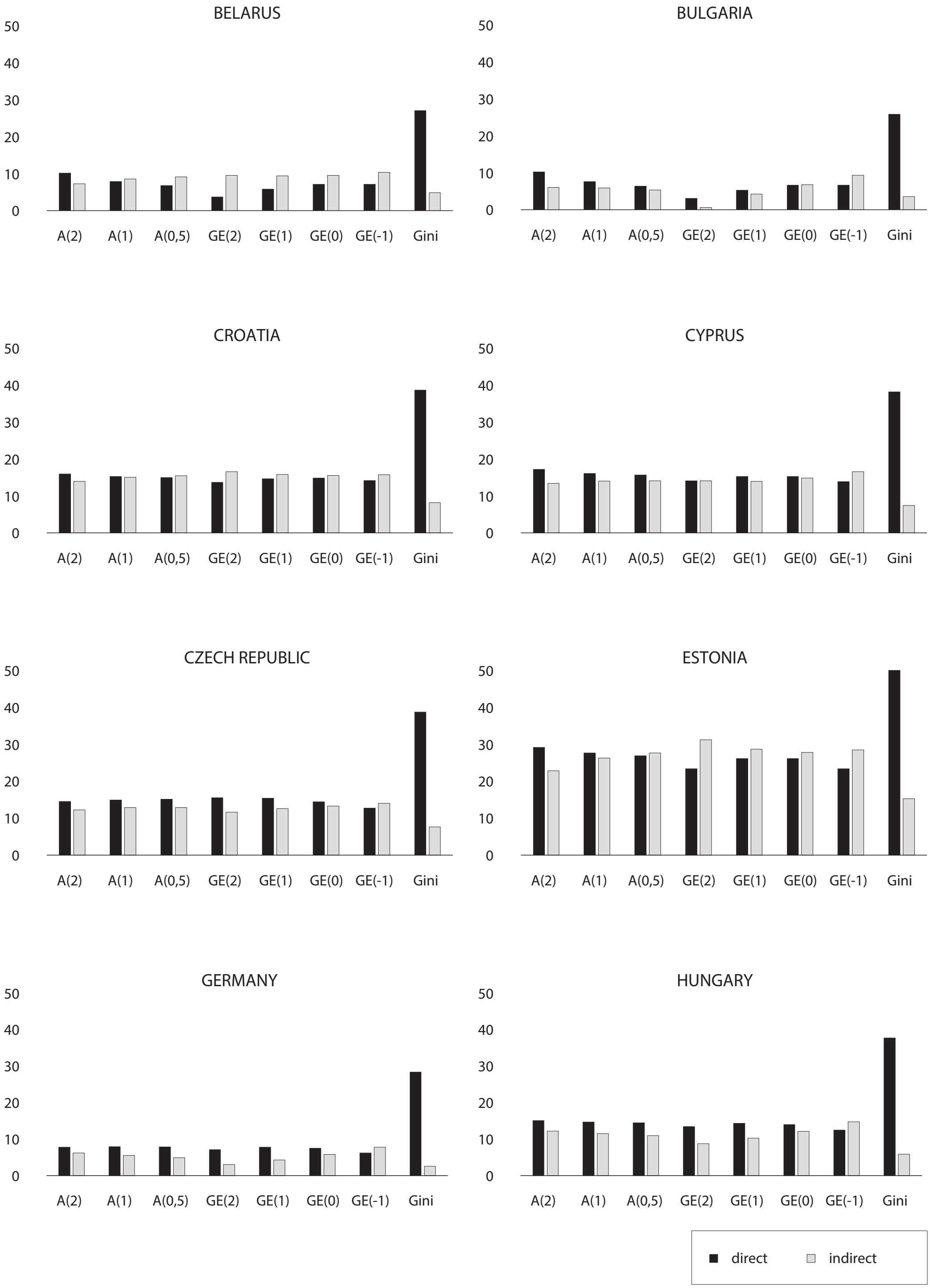
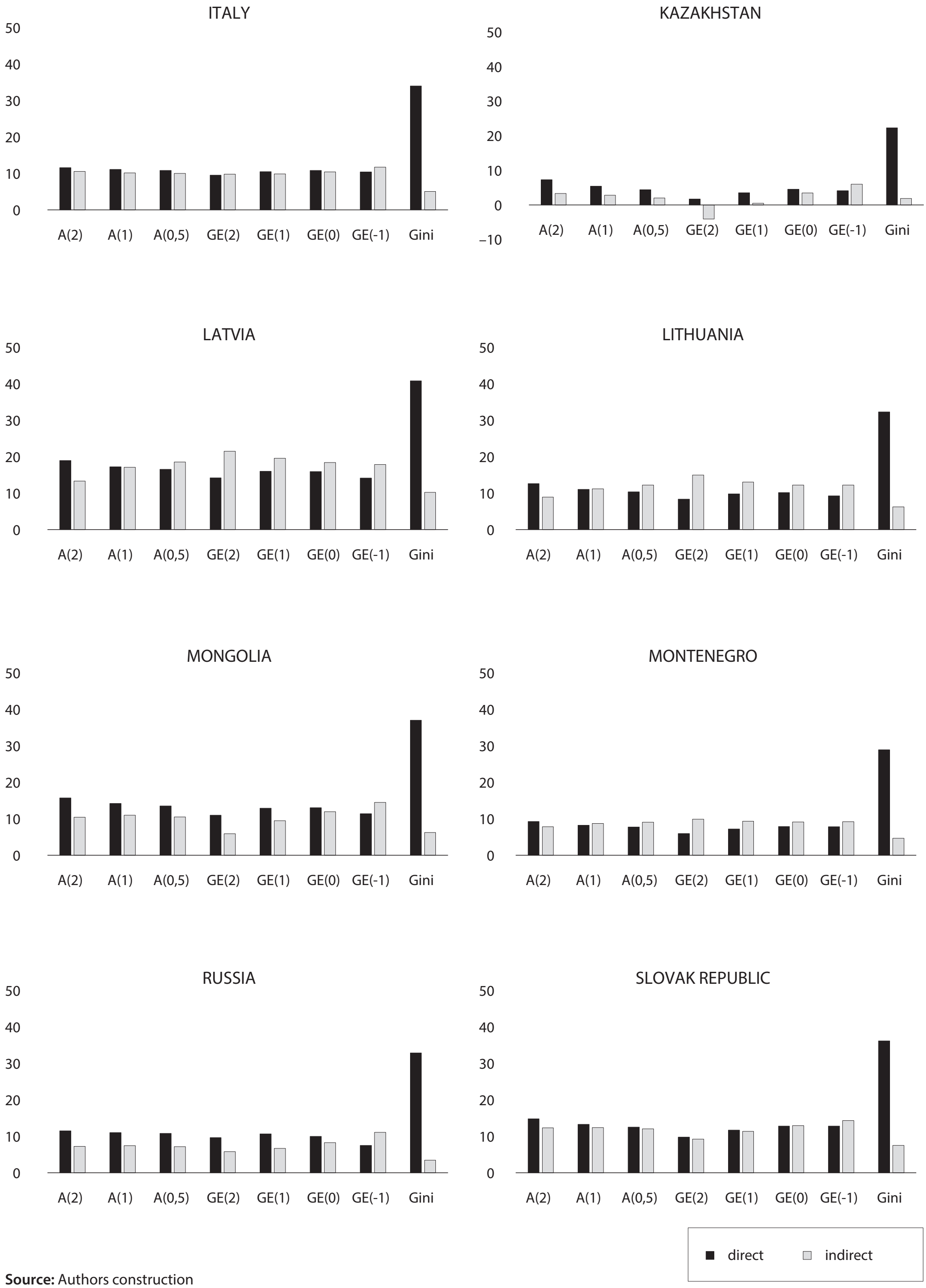


Figure 2 (continuation)



Source: Authors construction

In our view, the Gini index is a poor choice for the following reasons: firstly, the estimates based on this index obtained using the direct method are much higher than those obtained with other measures of inequality; secondly, when using the Gini index, the discrepancy between the estimates obtained with the help of the direct and indirect methods is enormous. The use of the L- and T-Theil indices is more preferable: the estimates obtained through the direct and indirect methods using these indices are close to each other; besides, these estimates are comparable to those obtained through other measures of inequality.

The correlation of inequality indices is shown in Table 9; the ranking of the countries according to the level of inequality obtained through different inequality indices is given in Table 10.

Table 9 Correlation of inequality indices

	A(2)	A(1)	A(0,5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini
A(2)	1	0.975157	0.944105	0.817974	0.905825	0.966176	0.989257	0.975224
A(1)		1	0.993271	0.916432	0.975362	0.998701	0.984187	0.993019
A(0,5)			1	0.955099	0.994259	0.99695	0.965879	0.979455
GE(2)				1	0.980792	0.93435	0.876497	0.878689
GE(1)					1	0.983948	0.94018	0.953935
GE(0)						1	0.982415	0.986783
GE(−1)							1	0.969477
Gini								1

Source: Author's calculations

Table 10 Ratings of the countries according to the level of inequality

Country	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini	min	max
BLR	4	3	3	3	3	3	4	3	3	4
BUL	2	2	2	2	2	2	2	2	2	2
CRO	15	15	14	14	14	15	15	14	14	15
CPR	9	9	9	9	9	9	9	9	9	9
CZE	14	14	15	15	15	14	14	15	14	15
EST	8	8	8	8	8	8	8	8	8	8
DEU	10	11	11	13	12	11	10	12	10	13
HAN	11	10	10	12	10	10	11	10	10	12
ITA	16	16	16	16	16	16	16	16	16	16
KAZ	1	1	1	1	1	1	1	1	1	1
LAT	6	7	7	6	7	7	6	7	6	7
LIT	7	6	6	5	5	6	7	6	5	7
MON	5	5	5	4	4	5	5	4	4	5
MNG	12	12	12	10	11	12	12	11	10	12
RUS	3	4	4	7	6	4	3	5	3	7
SVK	13	13	13	11	13	13	13	13	11	13

Source: Author's calculations

Table 11 Ratings of the countries by absolute inequality of opportunity

Country	Direct method										Indirect method									
	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini	min	max	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini	min	max
BLR	9	9	9	10	9	9	9	9	9	10	8	5	4	3	3	5	8	6	3	8
BUL	5	5	6	7	7	5	5	6	5	7	9	7	7	14	7	6	5	11	5	14
CRO	13	13	13	13	13	13	13	12	12	13	11	11	11	7	9	12	12	8	7	12
CPR	6	6	7	6	6	6	6	7	6	7	4	6	6	5	6	7	9	5	4	9
CZE	12	12	12	12	12	12	12	13	12	13	12	12	12	12	13	13	13	9	9	13
EST	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DEU	14	14	14	14	14	14	14	15	14	15	14	16	16	15	15	16	14	16	14	16
HAN	10	10	10	9	10	10	10	10	9	10	7	9	9	11	11	9	10	10	7	11
ITA	16	16	16	16	16	16	16	16	16	16	16	15	15	13	14	15	16	14	13	16
KAZ	8	8	8	8	8	8	8	8	8	8	13	13	14	16	16	10	7	15	7	16
LAT	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
LIT	7	7	5	5	5	7	7	5	5	7	6	4	3	4	4	4	6	4	3	6
MON	3	3	3	3	3	3	3	3	3	3	3	3	5	6	5	3	4	3	3	6
MNG	15	15	15	15	15	15	15	14	14	15	15	14	13	8	12	14	15	13	8	15
RUS	4	4	4	4	4	4	4	4	4	4	5	8	8	9	8	8	3	12	3	12
SVK	11	11	11	11	11	11	11	11	11	11	10	10	10	10	10	11	11	7	7	11

Source: Author's calculations

Table 12 Ratings of the countries by relative inequality of opportunity

Country	Direct method										Indirect method									
	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini	min	max	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini	min	max
BLR	13	13	14	14	14	14	13	14	13	14	12	12	11	9	11	11	12	11	9	12
BUL	12	15	15	15	15	15	14	15	12	15	15	14	14	15	15	14	13	13	13	15
CRO	4	4	5	5	5	4	2	4	2	5	2	3	3	3	3	3	4	3	2	4
CPR	3	3	3	4	4	3	4	5	3	5	3	4	4	5	4	4	3	6	3	6
CZE	8	5	4	2	3	5	6	3	2	8	6	5	5	6	6	5	8	4	4	8
EST	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
DEU	15	14	12	12	12	13	15	13	12	15	14	15	15	14	14	15	15	15	14	15
HAN	6	6	6	6	6	6	7	6	6	7	7	7	8	11	8	8	5	9	5	11
ITA	10	9	9	10	10	9	9	9	9	10	8	10	10	8	9	10	10	10	8	10
KAZ	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
LAT	2	2	2	3	2	2	3	2	2	3	4	2	2	2	2	2	2	2	2	4
LIT	9	10	11	11	11	10	10	11	9	11	10	8	6	4	5	7	9	7	4	10
MON	5	7	7	7	7	7	8	7	5	8	9	9	9	12	10	9	6	8	6	12
MNG	14	12	13	13	13	12	11	12	11	14	11	11	12	7	12	12	14	12	7	14
RUS	11	11	10	9	9	11	12	10	9	12	13	13	13	13	13	13	11	14	11	14
SVK	7	8	8	8	8	8	5	8	5	8	5	6	7	10	7	6	7	5	5	10

Source: Author's calculations

As it follows from Tables 9–10, the country ranking depends on the choice of the measure of inequality despite the coefficients of the inequality indices correlation being high and positive. In some cases, the difference can be significant: 4 positions (Russia), 3 positions (Germany). This outcome seems to be related to the fact that inequality measures are not ordinary equivalent.

Table 11 shows the ranking of countries by the absolute inequality of opportunity depending on the measure of inequality and the method of assessment.

As can be seen from Table 11, a country's ranking by the absolute inequality of opportunity varies depending on the choice of both the measure of inequality and the assessment method. Given that, when using the direct method the rating position looks more resilient to the choice of the measure of inequality than when the indirect method is used.

The ranking of the countries according to the relative inequality of opportunity depending on the measure of inequality and the method of assessment is presented in Table 12.

As can be seen from Table 12, the country's ranking by relative inequality of opportunity also changes depending on the choice of the measure of inequality, and on the choice of the assessment method. The ranking position by the absolute inequality of opportunity can significantly differ from the ranking position by the relative inequality, which raises the question of what should be focused on when comparing the countries in terms of inequality of opportunity.

Table 13 shows the correlation between the indicators of inequality and the absolute inequality of opportunity when different measures of inequality are used.

Table 13 Correlation of inequality indices

				Direct method of assessment				
	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini
R	0.519	0.420	0.353	0.164	0.286	0.390	0.431	0.546
				Indirect method of assessment				
	A(2)	A(1)	A(0.5)	GE(2)	GE(1)	GE(0)	GE(−1)	Gini
R	0.251	0.235	0.166	−0.509	0.018	0.316	0.565	0.086

Source: Author's calculations

As follows from Table 13, in most cases the relationship between the inequality and the inequality of opportunity is direct (except for the case when the GE(2) index and the indirect assessment method were used), but the strength of relationship varies sensibly. The strength of relationship is usually higher with the direct assessment method (except for the case when the GE(−1) index was used).

The conducted analysis of the relationship between inequality and inequality of opportunity allows the important conclusion that the general level of inequality does not predetermine the level of inequality of opportunity. This is important with regard to the ongoing discussion in economics about the relationship between inequality and economic growth. No consensus on the issue has been reached yet; a lot of arguments have been proposed by the scientists in support of both the positive and negative relationship between these indicators (Bradbury and Triest, 2016). Numerous empirical studies also produce inconsistent results, fueling a long-standing debate (Henderson et al., 2015; Babu et al., 2016; Fang et al., 2015; Rubin and Segal, 2015). The equal opportunity theory gave rise to the hypothesis that the inconsistency of results from assessing the effect of inequality on economic growth is due to the «multi-component origin» of inequality being disregarded. The inequality attributable to inequality of opportunity adversely affects the economic growth, while the inequality resulting from inequality of effort, on the contrary, shows a positive influence. The negative impact of inequality of opportunity on the economic growth as a whole is explained by the fact that the barriers it forms result in incomplete realization of the human potential, which reduces the resources and lowers the aggregated economic performance. The inequality in income due to inequality of effort, on the contrary, stimulates individuals to self-realization, thereby contributing to the growth of aggregated economic achievements. The absence of a strong relationship between inequality of opportunity and general inequality allows the conclusion that the level of inequality of opportunity is an independent socio-economic indicator and a potentially significant factor affecting the aggregated economic and social results.

On the whole, among various lines of economic research in the area of inequality, the analysis from the equal opportunity standpoint is interesting because it takes into consideration the ethical categories of justice and responsibility, unusual and alien to economists. Traditionally, economic models are built on the assumption that the behavior of economic agents is determined exclusively by their own selfish preferences. At the same time, real people are not alien to justice, responsibility, sympathy, compassion, and pity. These components affect the behavior of real people along with their preferences, and therefore the real economic behavior can differ greatly from that predicted by the theoretical model. Hence, the inclusion of such categories as justice and responsibility in the analysis may allow economists to move towards building models that are more adequate to the real world and therefore better explain the processes taking place in it.

CONCLUSION

The calculations made have confirmed the proposed hypothesis that the choice of the measure of inequality makes a significant contribution to the variance in assessments of inequality of opportunity. The resulting differences are so significant that they lead to different meaningful interpretation of the outcome. For example, the 33% share of the contribution made by inequality of opportunity to the inequality of labor income in the Russian Federation, as obtained using the Gini index, looks very impressive, given that far from all of the circumstance-factors are taken into account, and pushes to the conclusion that there is little that can be achieved through personal efforts in Russia. The 10% contribution of inequality of opportunity to the inequality of earned income, as obtained using the L-Theil index, on the contrary, inspires some optimism - after all, it turns out that 90% of everything depends on one's own efforts.

If the described calculation method is applied, the use of the Gini index in case of direct method of measuring the inequality of opportunity results in the assessments being much higher than when other measures are used. Besides, it has been established that when the Gini index is used, the differences in assessments obtained through the direct and indirect methods are much greater than those obtained using other measures of inequality. The estimates obtained through direct and indirect methods are closest to each other when the L- and T-Theil indices are used.

Also, our calculations show that the selection of inequality index may significantly affect a country's ranking by the inequality of opportunity. Given the popularity of all kinds of ratings in the modern world, this conclusion further emphasizes the relevance of the issue under consideration and proves the necessity of further efforts by the scientific community to solve it.

ACKNOWLEDGMENTS

The reported study was funded by RFBR according to the research project No. 19–010–00453.

References

ATKINSON, E. *Inequality: What about It?* Moscow: Delo, RANKhiGS, 2018.

BABU, M., BHASKARAN, V., VENKATESH, M. Does inequality hamper long run growth? Evidence from emerging economies. *Economic Analysis and Policy*, 2016, 52(4), pp. 99–113.

BJÖRKLUND, A., JÄNTTI, M., ROEMER, J. E. Equality of opportunity and the distribution of long-run income in Sweden. *Social Choice and Welfare*, 2011, 39(2–3), pp. 675–696.

BOURGUIGNON, F., FERREIRA, F. H. G., MENENDEZ, M. *Inequality of Outcomes and Inequality of Opportunities in Brazil.* William Davidson Institute Working Paper No. 630, World Bank Working Paper No. 3174, 2003.

BRADBURY, K. AND TRIEST, R. Inequality of Opportunity and Aggregate Economic Performance. *The Russell Sage Foundation Journal of the Social Sciences*, 2016, 2(2), pp. 178–201.

BRZENZIŃSKI, M. *Inequality of opportunity in Europe before and after the Great Recession.* Working Paper No. 2/2015 (150), University of Warsaw, 2015.

CHECCHI, D., PERAGINE, V., SERLENGA, L. *Fair and Unfair Income Inequalities in Europe.* IZA Discussion Papers 5025, Institute of Labor Economics (IZA), 2010.

COWELL, F. *Measuring Inequality. Handbook of Economic Inequality.* Oxford: Oxford University Press, 2009.

FANG, W., MILLER, S., YEH, C. The effect of growth volatility on income inequality. *Economic Modelling*, 2015, 45(1), pp. 212–222.

FERREIRA, F. AND GIGNOUX, J. The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. *Review of Income and Wealth*, 2011, 57(4), pp. 622–657.

FLEURBAEY, M. AND PERAGINE, V. Ex Ante Versus Ex Post Equality of Opportunity. *Economica*, 2013, 80(317), pp. 118–130.

HEDEROS, K., JÄNTTI, M., LINDAHL, L. Gender and inequality of opportunity in Sweden. *Social Choice and Welfare*, 2017, 49(3–4), pp. 605–635.

HENDERSON, D., QIAN, J., WANG, L. The inequality–growth plateau. *Economics Letters*, 2015, 28(1), pp. 17–20.

IBRAGIMOVA, Z. F. AND FRANTS, M. V. Inequality of opportunities in the Russian Federation: factor analysis (in Russian). *Economics and management*, 2019, 149(5), pp. 10–15.

IBRAGIMOVA, Z. F. AND FRANTS, M. Inequality of opportunity: theory and practice of measurement using micro-data of RLMS-HSE. *Economic Policy*, 2020, 15(1), pp. 64–89.

MARRERO, G. AND RODRÍGUEZ, J. *Macroeconomic determinants of inequality of opportunity and effort in the US: 1970–2009.* ECINEQ Working Paper 249, Society for the Study of Economic Inequality, 2012.

PAUHOFOVÁ, I., BUKHARBAEVA L., IBRAGIMOVA, Z., FRANTS, M. Measuring of Inequality of Opportunity: Parametric Approach. *Ekonomický časopis/Journal of Economics*, 2020, 68(5), pp. 455–476.

PISTOLESI, N. Inequality of opportunity in the land of opportunities, 1968–2001. *The Journal of Economic Inequality*, 2009, 7(4), pp.411–433.

RAMOS, X. AND VAN DE GAER, D. Approaches to Inequality of Opportunity: Principles, Measures and Evidence. *Journal of Economic Surveys*, 2016, 30(5), pp. 855–883.

RUBIN, A. AND SEGAL, D. The effects of economic growth on income inequality in the United States. *Journal of Macroeconomics*, 2015, 45, pp. 258–273.

SEN, A. *On Economic Inequality.* Oxford: Clarendon Press, 1973.

Transition report 2016–17 [online]. London: European Bank for Reconstruction and Development, 2016. [cit. 7.5.2020]. <http://2016.tr-ebrd.com/inequality-of-opportunity>.

Dynamics Almost Ideal: Demand System Application of Kalman Filter

Habibeh Sherafatmand[1] | *Islamic Azad University, Tehran, Iran*
Ali Akbar Baghestany[2] | *Economic and Rural Development Research Institute, Tehran, Iran*

Abstract

The demand elasticity for a product is the basis of its price determination. The ratio in which a product demand will fall with the rise in its price and, vice versa, can be known as demand elasticity. With increasing population and increasing demand for meat, it is important to accurately estimate price and income demand elasticities. This paper used almost ideal demand systems (AIDS) with log linear analogue of the Paasche price index, referred to as the corrected Stone index to model consumer demand system. The study employs the Kalman filter estimation strategy, which is based on state-space models that are applied to linear regressions with stochastically time-varying parameters, to determine the evolution of price and income elasticities of red meat and fish demand for monthly data 1997–2017. Variables stationary is tested with Hegy test. Results show that Price elasticity for fish is elastic. Elasticity results indicate that the two products are strong substitutes. Income elasticity indicated that fish considered to be luxury good.

Keywords	***JEL code***
AIDS model, fish, red meat, Kalman Filter	*C22, Q11*

INTRODUCTION

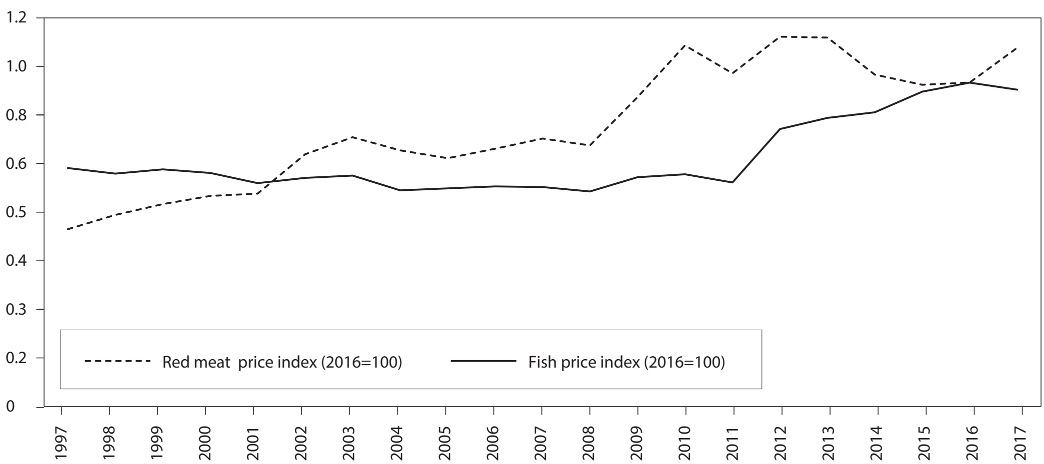
Total of 68 300 tons of red meat were produced in Iran during the first quarter of the current fiscal year 2019 (March 21–June 21) to register a 29% decline compared with the similar period of the year before. The Statistical Center of Iran's latest report shows beef accounted for 37 000 tons or 54.2% of the overall production, indicating a decrease of 31% year-on-year. The production of lamb reached 25 100 tons (down 27% YOY), goat meat 4 700 tons (down 28% YOY) and meat of other types of livestock amounted to 1 500 tons during the three-month period, accounting for 36.7%, 6.8% and 2.3% of the total output, respectively, SCI reported on its website. The Iranians consume around 920 000 tons of red meat per year, 90% of which are supplied from local sources. Imports are made from CIS countries as well as from Brazil and Australia. With rapid population growth and improved per capita income and lifestyle changes resulting from urbanization, it is predicted that there will be further increases in demand

[1] Ph.D. researcher in Agricultural economics, Islamic Azad University, Science and research branch, Hesarak blvd., Daneshgah Square, 1477893855, Tehran, Iran. Corresponding author: e-mail: sherafatmandm@gmail.com.
[2] Assistant Professor of Agricultural Planning, Economic and Rural Development Research Institute (APERDRI), No. 5 Roudsar St., South Shahid Azodi (Aban) Ave., Karimkhan-e-Zand Blvd., Tehran, Iran. E-mail: a.baghestany@agri-peri.ac.ir.

for meat products in the country. Average capital consumption of red meat in developed countries is 27 kg per year while it is about 11 kg in Iran. Average capital consumption of fish in developed countries is 21.6 kg per year while it is about 10 kg in Iran in 2016. According to FAO estimates in 2015, the average per capita consumption of meat in the world was 41.3 kg which is estimated to reach 45.3 kg by 2030 as the world's population grows. Currently in Iran the average per capita consumption of meat is 35.5 kg comprising of 12.5 kg of red meat and 23 kg of poultry meat (faostat.org). High level retail price of these two kinds of products with also low purchase power, are most important causes for low capital consumption and demand. The self-sufficiency level for meat has been declining over the years. Efforts have been made to increase meat production; nevertheless the progress has been slow. Currently, Iran imports about 10–15% of its red meat requirement from various countries to meet local beef demand. Hence, knowing the demand for meat products is thus an important concern for policymakers due to its impact on self-sufficiency, changing food prices, and the nation's trade balance. Therefore, understanding meat demand and its characteristics is important in order to give a more accurate evaluation of the factors that govern consumers' behavior for meat products. Meats are important component of Iranian diet.

Figure 1 shows the increasing but fluctuation trend in the price index of red meat and fish at the retail level in Iran from 1997 to 2017. Red meat price index starts from about 0.6 in 1997 and during the time passes it increased smoothly till 2012. In 2012 red meat price index reached its maximum amount 1.2. In 2015 both red meat and fish price indexes are equal and 0.97 and after that red meat price index again increased but fish price decrease. It is concluded that during the whole period fish price index is lower than red meat price index. Main reasons for high red meat price index are: high cost of animal food and rising cost of inputs, importing live sheets to neighbor's country, some companies which are allocated dollars to import meat, put these dollars in their pockets instead of importing red meats.

Figure 1 Annual price index of red meat and fish in Iran during 1997–2017



Source: Central Bank of Iran (CBI)

This study determined consumer demand for red meat and fish. The growing population and consequent rise in demand for meat has increased the importance of estimating the demand function and significant factors affecting demand. It is crucial to estimate the demand function to identify consumer preferences, develop coherent policies for consumption, and to forecast and plan for future consumer needs.

The results may help policymakers to predict demand and control the prices of these two important products.

However a number of studies have used the dynamic AIDS specification, including Thamae et al. (2015); Nzuma and Sarker (2010); Iootty et al. (2009); Basmann et al. (2009); Barnett and Serlertis (2008); Barnett and Seck (2008); Li et al. (2006); Taljaard et al. (2006); Eakins and Gallagher (2003); Katranidis and Velentzas (2000); Poray et al. (2000); Kremers, Ericsson, Dolado (1992); Goddard and Akiyama (1989); Banerjee et al. (1986); Anderson and Blundell (1983, 1984); and Blancifiorti and Green (1983). Barnett and Kanyama (2013), assesses the ability of the Rotterdam model and of three versions of the almost ideal demand system (AIDS) to recover the time-varying elasticities of a true demand system and to satisfy theoretical regularity. They find that the Rotterdam model performs better than the linear-approximate AIDs at recovering the signs of all the time-varying elasticities. Motallebi and Pendell (2013) present a dynamic form of the almost ideal demand system (AIDS). The static AIDS model was employed to determine the long-run equilibrium model and represents the short-run dynamics by an error correction mechanism. This estimation procedure is applied to estimate three kinds of popular meats (red meat, chicken and fish) demand function in Iran. The estimated elasticities of red meat and chicken are found to be price elastic in the long run. While fish is price inelastic in the long run. Iranian government will remove all indirect and direct goods subsidies. It is suggested that government should be careful about chicken and red meat pricing policy to decreasing malnutrition after subsidy removal. Kilungu et al. (2012), estimate a dynamic version of an almost ideal demand system (AIDS) model for U.S.A. imports of fresh tropical fruits: bananas, pineapples, avocadoes, papayas, mangoes/guavas, grapes and other fresh fruit imports. Estimated income elasticities show that fresh grapes and other fresh fruit imports appear to be considered luxury commodities. All own-price elasticities were negative and significant. While imported bananas, pineapples, U.S.A. grapes and other fresh fruit were quite inelastic, demand for papayas and mangoes/guavas were elastic. Carew et al. (2005) employ a source differentiated almost ideal demand system (AIDS) model with time-varying parameters to estimate the demand for premium quality wines using scanner sales data from the British Columbia wine market. The empirical findings reveal that consumers' response to foreign-produced wines differs from that of wine produced locally. It is evident that the expenditure elasticities for British Columbia, European and Rest-of-the-World white wines are bigger than those for red wines. The high expenditure elasticities associated with British Columbia white wines may suggest that these wines are associated with higher quality. We reject the hypotheses of block separability and product aggregation. There is no evidence of structural change from the tests employed in this paper. Mario Mazzocchi (2003) provides a generalisation of the structural time series version of the Almost Ideal Demand System (AIDS) that allows for time-varying coefficients (TVC/AIDS) in the presence of cross-equation constraints. An empirical appraisal of the TVC/AIDS is made using a dynamic AIDS with trending intercept as the baseline model with a data set from the Italian Household Budget Survey (1986–2001). The assessment is based on four criteria: adherence to theoretical constraints, statistical diagnostics on residuals, forecasting performance and economic meaningfulness. No clear evidence is found for superior performance of the TVC/AIDS, apart from improved short-term forecasts.

1 OBJECTIVE OF THE STUDY

Due to changes in prices of agricultural products, the consumers behave differently against price changes over time so this study focuses on:

- estimation time varying price elasticities of red meat and fish,
- estimation time varying income elasticities of red meat and fish,
- comparison between averages of time varying elasticities with their fixed ones.

1.1 Data

Data on red meat and fish expenditure in Iran were provided by monthly Data from 1997–2017. All of the following data is from the statistical office of the Central Bank of Iran:

- red meat expenditure in Rials using a constant price of 2016 = 100,
- fish expenditure in Rials using a constant price of 2016 = 100,
- red meat and fish price indexes using a constant price of 2016 = 100.

1.2 The full AIDS model

The AIDS model in budget shares is:

$$W_i = \alpha_i + \sum y_{ij} \log P_j + \beta_i \log \left(\frac{X}{P^*}\right), \tag{1}$$

in which, W_i is share of budget that allocate to commodity i from total budget, P_j is the price of commodity j, X is total expenditure and *P** is the price index or price deflator. The loglinear analogue of the Paasche price index, referred to as the corrected stone index, is written as:

$$\log P_t^* = \sum_{i=1}^{n} W_{it} \log \frac{P_{it}}{P_i^0}. \tag{2}$$

In applications, the nonlinearity of the AIDS model is usually viewed as a technical problem to be circumvented by a linearizing approximation to income's price deflator. Deaton and Muellbauer (1980a, 1980b) suggest Stone's price index. The restrictions on the demand functions are deduced from the cost function, using Shephard's duality lemma. The following are the resulting conditions imposed during estimation of the constrained model:

$\sum_{i=1}^{n} \alpha_i = 1$ for adding up, $\sum_{i=1}^{n} y_{ij} = 0$ and $\sum_{i=1}^{n} \beta_i = 0$ for linear Homogeneity, $y_{ij} = y_{ji}$ for symmetry.

1.3 Kalman Filter estimation strategy

The study employs the Kalman (1960, 1963) filter estimation strategy, which is based on state-space models that are applied to linear regressions with stochastically time-varying parameters, to determine the evolution of price and income elasticities of red meat and fish demand. The use of this technique compared to other conventional econometric methods is based on the following advantages. First, this approach is considered an ideal model for estimating regressions with variables whose impact changes over time (Slade, 1989). Second, the Kalman filter is believed to be superior to the least squares models, especially in the presence of parameter instability (Morisson and Pike, 1977). Third, this procedure can be used with non-stationarity data and it is predictive and adaptive (Inglesi-Lotz, 2011). The formal representation of this dynamic model (assuming its parameters are functions of time) is given by the following observation and state equations, respectively:

$$\begin{aligned} x_t &= \alpha(z_t) + [\beta(z_t)]' \varepsilon_t + w_t, \\ \varepsilon_{t+1} &= H(z_t)\varepsilon_t + v_{t+1}, \end{aligned} \tag{3}$$

where α is a constant parameter, β and H are matrices of parameters, x_t is a vector of observations and z_t is a vector of exogenous variables. Furthermore, $\alpha(z_t)$ and $\beta(z_t)$ are vector and matrix valued functions, respectively, and $H(z_t)$ is a matrix with elements that are functions of x_t. A vector of unobserved variables is then given by ε_t while w_t and v_t are the disturbance vectors that are assumed to be independent and white noise. The estimating equations which allow for stochastically time varying parameters:

$$W_{it} = \alpha_i + \sum y_{ijt} \log P_{jt} + \beta_{it} \log \left(\frac{X}{P^*}\right). \tag{4}$$

Above equation is then specified below as a state-space model following the Eviews software notation in order allow for time-varying coefficients:

$$
\begin{aligned}
&@\ signal \quad w_{it} = sv_1 \log P_{it} + sv_2 \log P_{jt} + sv_3 \log(x/p) + [\text{var} = \exp c(1)] \\
&@\ state \quad sv_1 = sv_1(-1) \\
&@\ state \quad sv_2 = sv_2(-1) \\
&@\ state \quad sv_3 = sv_3(-1), \\
&@\ signal \quad w_{jt} = sv_4 \log P_{it} + sv_5 \log P_{jt} + sv_6 \log(x/p) + [\text{var} = \exp c\ (3)] \\
&@\ state \quad sv_4 = sv_4\ (-1) \\
&@\ state \quad sv_5 = sv_5\ (-1) \\
&@\ state \quad sv_6 = sv_6\ (-1),
\end{aligned}
\tag{5}
$$

where sv_1 $i = 1, ..., 6$ are, the final estimates for price and income elasticity. c_i are the constant parameters of estimation. The evolution of price and income elasticities over time is therefore shown to follow a random walk process. After finding time varying parameters of AIDS system with Kalman state space model, time varying elasticities are calculated using below formulas:

$$\text{Marshallian price elasticity } \varepsilon_i^M = \frac{y_{ij}}{w_i} - \beta_i\left(\frac{w_j}{w_i}\right) - \delta_{ij},$$

where δ_{ij} is the Kronecker delta, defined as: $\delta_{ij} = 1$ if i=j, $\delta_{ij} = 0$ otherwise (Buse, 1996; Barnett and Ousmane, 2007).

$$\text{Hicks price elasticity } E_{ij} = -\delta_{ij} + \frac{y_{ij}}{w_i} + w_j,$$

$$\text{Income elasticity } \eta_i = 1 + \frac{\beta_i}{w_i}.$$

2 ESTIMATION RESULTS

All variables used in this study were time series; so the augmented Dickey-Fuller (ADF) test was used to test the stationarity of variables (Table 1). The ADF unit root test results allowed for acceptance of the null hypothesis of non-stationarity of the red meat price index (p red meat) and the fish price index

Table 1 Result of unit root test (ADF test) for variables

Variable		ADF at level	ADF with one difference
(p red meat)		0.7	–3.1*
(P fish)		–0.6	–3.7*
(x red meat)		–5.4	
(x fish)		–1.6	–3.7*
X/P		–1.3	–10
	1%	–3.6	–4.3
Critical value	5%	–2.9	–3.7
	10%	–2.6	–3.2

Note: * indicate significance at the 5% level.
Source: Research finding

Table 2 Result of Hegy unit root test for variables

Variable	Null hyphotesis	Test statistic	Simulated P-value
LOG(PR)	Non seasonal unit root (zero frequency)	–0.659314	0.882998
	Seasonal unit root (2 months per cycle)	–2.773849	0.111750
	Seasonal unit root (4 months per cycle)	15.25249	0.000000
	Seasonal unit root (2.4 months per cycle)	12.16319	0.000000
	Seasonal unit root (12 months per cycle)	11.52706	0.000000
	Seasonal unit root (3 months per cycle)	8.850001	0.000000
	Seasonal unit root (6 months per cycle)	12.07574	0.000000
LOG(PF)	Non seasonal unit root (zero frequency)	–2.035364	0.541567
	Seasonal unit root (2 months per cycle)	–3.415713	0.054486
	Seasonal unit root (4 months per cycle)	12.14653	0.000000
	Seasonal unit root (2.4 months per cycle)	16.78231	0.000000
	Seasonal unit root (12 months per cycle)	10.78571	0.000000
	Seasonal unit root (3 months per cycle)	17.17169	0.000000
	Seasonal unit root (6 months per cycle)	13.90648	0.000000
LOG(WR)	Non seasonal unit root (zero frequency)	–3.311438	0.162262
	Seasonal unit root (2 months per cycle)	–4.668780	0.054486
	Seasonal unit root (4 months per cycle)	30.03822	0.000000
	Seasonal unit root (2.4 months per cycle)	24.70168	0.000000
	Seasonal unit root (12 months per cycle)	26.59229	0.000000
	Seasonal unit root (3 months per cycle)	32.25468	0.000000
	Seasonal unit root (6 months per cycle)	26.77380	0.000000
LOG(WF)	Non seasonal unit root (zero frequency)	–3.039070	0.209804
	Seasonal unit root (2 months per cycle)	–6.821745	0.054486
	Seasonal unit root (4 months per cycle)	46.13639	0.000000
	Seasonal unit root (2.4 months per cycle)	46.00186	0.000000
	Seasonal unit root (12 months per cycle)	46.82536	0.000000
	Seasonal unit root (3 months per cycle)	46.04062	0.000000
	Seasonal unit root (6 months per cycle)	46.36882	0.000000
LOG(WP)	Non seasonal unit root (zero frequency)	1.437834	1.000000
	Seasonal unit root (2 months per cycle)	–0.971411	0.789852
	Seasonal unit root (4 months per cycle)	4.457963	0.077572
	Seasonal unit root (2.4 months per cycle)	30.70673	0.000000
	Seasonal unit root (12 months per cycle)	6.421286	0.021544
	Seasonal unit root (3 months per cycle)	12.15545	0.000000
	Seasonal unit root (6 months per cycle)	4.341545	0.069994

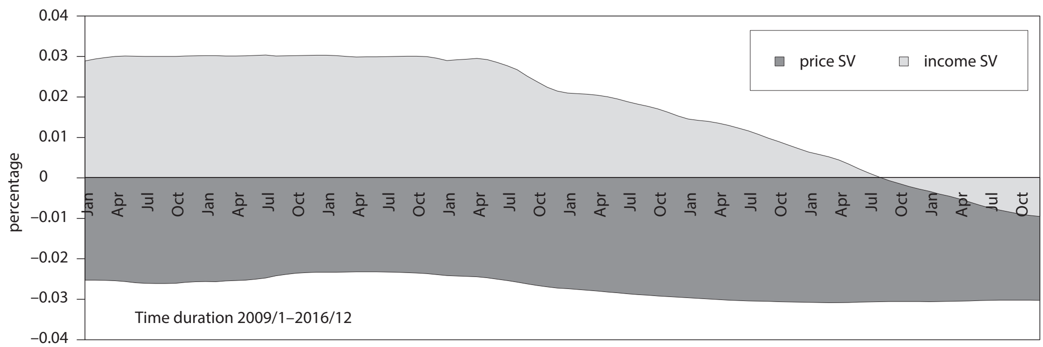
Source: Research finding

Table 3 Kalman state space model

Intercept	Estimated coefficient	Prob statistic
C(1)	0.06	00
C(2)	–147	00
Red meat share expenditure equation	Final state	P value
Sv1(red meat price coefficient)	–0.03*	00
Sv2(fish price coefficient)	–0.001	00
Sv3(income coefficient)	0.001*	00
Fish share expenditure equation	Final state	P value
Sv4(fish price coefficient)	0.00006	00
Sv5(red meat price coefficient)	–0.00002	00
Sv6(income coefficient)	0.000007*	00

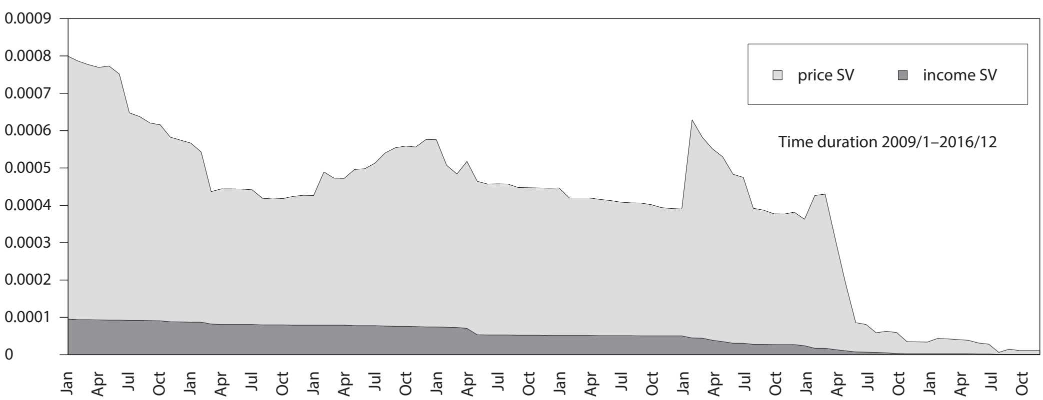
Note: * indicate significance at the 5% level.
Source: Research finding

Figure 2 Trends of red meat price and income states



Source: Research findings

Figure 3 Trends of fish price and income states



Source: Research findings

Table 4 Marshalian price elasticity and income elasticity

Income elasticity	Fish	Red meat	Product
Red meat	–0.03* (–0.9*)	0.01 (0.13*)	–0.017*
Fish	1.12* (1.15*)	–2.51* (-0.3*)	1.34*

Notes: Compensated elasticities in parenthesis (). * indicate significance at the 1% level.
Source: Research finding

(p fish). The two indexes were non-stationary at level, but they were stationary after the first difference. Red meat expenditure (x red meat) was stationary at level. Results of Hegy test also show that all variables in logarithm form are stationary.

For linearized AIDS model, Paasche, price index was used. Table 3 displays the results from the Kalman filter estimation technique. Table 4 shows the final (average) estimates for price and income elasticities being significant and having the values of –0.113 (–0.155) and 1.009 (1.797), respectively. These show that, on average, increases (decreases) in real red meat prices have resulted in less than proportionate fall (rise) in red meat consumption, implying that the demand is price inelastic (in agreement with Zeranezhad and Saadatmehr, 2007 results). However, if the real red meat prices become too high over time, consumers might change their behavior and sensitivity to price and hence, policymakers will need to reconsider their impact in the long-run. The increase in the coefficient of disposable income in the red meat equation (budget coefficient or X/P) increased red meat share expenditure. In both budget share equations, red meat and fish were regarded as substitutes, as indicated by the positive cross-price coefficients. The cross-price coefficient of red meat equation was positive and significant; a 1% increase in the fish price index increased the red meat budget share 0.02%. A negative and significant own-price coefficient was found for red meat and fish in AIDS, which satisfies the law of demand. Since the income elasticity of fish was more than unity, fish is considered a luxury and elastic good.

DISCUSSION AND CONCLUSION

The aim of this paper was to evaluate the ability of the AIDS to recover time varying elasticities. A structural time series model was specified for red meat and fish demand specification and the time varying elasticities estimated using Kalman filter. Results of Hegy unit root test shows that, logarithm of all variables include price indexes and expenditures for both red meat and fish are stationary. Results also show that state space coefficient in both red meat and fish equations are significant and in accordance with theory. Red meat price elasticity is –0.03 and significant. Negative price elasticities of these two products indicate, as their prices increased, expenditure on them decreased. Since fish is elastic, an increase in price led to a larger-than-proportional decrease in value demanded and a decrease in sales revenue. A decrease in income does not result in uniform changes in expenditure for all goods. Expenditure on fish decreased in a proportionately larger amount than any other demand; this may have been the result of fish being considered a luxury good. An increase in the price of fish, increased the consumption of red meat because they were substitutes for one another.

Successive drought, conflicting trade policies, the absence of subsidies for inputs, and lack of government support for producers has raised the price of inputs. This has led to increases in the price of products. Red meat and fish prices have grown sharply, which has decreased per capita consumption of these products. This has led to insufficient consumption of protein to maintain nutritional health. Also, the increase in population and subsequent rise in demand for meat has emphasized the importance of estimating the demand function and significant factors affecting demand. Demand function identifies consumer preferences and helps to develop coherent policies on consumption, forecast future consumer needs, and plan for the future.

Since red meat and fish are elastic goods, to modify consumption patterns, the price tool can be recommended as effective. In addition, government policies such as decreasing subsidies should be carefully considered because it has resulted in unacceptably high prices for red meat and fish. The results of this study may help policymakers to predict demand and to control the prices of these two important products.

References

ANDERSON, G. J. AND BLUNDELL, R. Consumer Non-Durables in the U.K. A Dynamic Demand System. *Economic Journal*, 1984, 94, pp. 35–44.

ANDERSON, G. J. AND BLUNDELL, R. Testing Restrictions in a Flexible Dynamic Demand System: an Application to Consumers' Expenditure in Canada. *Review of Economic Studies*, 1983, 50, pp. 397–410.

BANERJEE, A., DOLADO, J. J., HENDRY, D. F., SMITH, G. W. Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence. *Oxford Bulletin of Economics and Statistics*, 1986, 48, pp. 253–277.

BARNETT, W. A. AND KANYAMA, I. K. Time-varying parameters in the almost ideal demand system and the Rotterdam model: will the best specification please stand up? *ERSA working paper 335*, 2013.

BARNETT, W. A. AND SECK, O. Rotterdam vs Almost Ideal Models: Will the Best Demand Specification Please Stand Up? *MPRA Paper No. 417*, 2007.

BARNETT, W. A. AND SECK, O. Rotterdam model versus almost ideal demand system: will the best specification please stand up? *Journal of Applied Econometrics*, 2008, 23, pp. 795–824.

BARNETT, W. A. AND SERLERTIS, A. Consumer preferences and demand systems. *Journal of Econometrics*, 2008, 147, pp. 210–24.

BASMANN, R. L., HAYES, K., MCALEER, M., MCCARTHY, I., SOLTTJE, D. J. The GFT utility function. In: SLOTTJE, D. J. (eds.) *Quantifying Consumer Preferences*, Cambride, MA: Contributions to Economic Analysis, Emerald Group Publishing, 2009, pp. 119–48.

BLANCIFIORTI, L. AND GREEN, R. An Almost Ideal Demand System Incorporating Habits: an Analysis of Expenditures on Food and Aggregate Community Groups. *Review of Economics and Statistics*, 1983, 65, pp. 511–15.

CAREW, R., FLORKOWSKI, W. J., HE, S. Demand for Domestic and Imported Table Wine in British Columbia: a Source-differentiated Almost Ideal Demand System Approach. *Canadian journal of agricultural economics*, 2005.

DEATON A. S. AND MUELLBAUER, J. An Almost Ideal Demand System. *American Economic Review*, 1980a, No. 70, pp. 312–26.

DEATON A. S. AND MUELLBAUER, J. *Economics and consumer behavior.* Cambridge, UK: Cambridge University Press, 1980b.

EAKINS, J. M. AND GALLAGHER, L. A. Dynamic almost ideal demand systems: an empirical analysis of alcohol expenditure in Ireland. *Applied economics*, 2003, Vol. 35(9), pp. 1025–1036.

GODDARD, E. W. AND AKIYAMA, T. United States Demand for Coffee Imports. *Agricultural Economics*, 1989, 3, pp. 147–159.

INGLESI-LOTZ, R. The Evolution of Price Elasticity of Electricity Demand in South Africa: a Kalman Filter Application. *Energy Policy*, 2011, Vol. 39, No. 6, pp. 3690–3696.

KALMAN, R. E. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 1960, Vol. 82, No. 1, pp. 35–45.

KALMAN, R. E. New Methods in the Wiener Filtering Theory. In: *Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability*, 1963, p. 270.

KATRANIDIS, S. D. AND VELENTZAS, K. The Markets of Cotton Seed and Maize in Greece: Welfare Implications of the Cotton Agricultural Policy. *Agricultural Economics Review*, 2000, 1, pp. 82–95.

KILUNGU, N., HOUSTON, J. E., ESENDUGUE G. F. A Dynamic Application of the AIDS Model to Import Demand for Tropical Fresh Fruits in the USA. *International association of agricultural economists triennial conference*, Foz do Brazil, 2012.

KREMERS, J., ERICSSON, N., DOLADO, J. The power of cointegration tests. *Bull. Econ. Stat.*, Oxford, 1992, 48, p. 325.

LI, G., SONG, H., WITT, S. F. Time varying parameter and fixed parameter linear AIDS: an application to tourism demand forecasting. *International Journal of Forecasting*, 2006, 22, pp. 57–71.

LOOTTY, M., PINTO, JR., H., EBELING, F. Automotive fuel consumption in Brazil: applying static and dynamic systems of demand equations. *Energy Policy*, 2009, 37, pp. 5326–33.

MAZZOCCHI, M. Time-varying coefficients in the Almost Ideal Demand System: an empirical appraisal [online]. *European Review of Agricultural Economics*, 2003, Vol. 30, Iss. 2, pp. 241–270. <*https://doi.org/10.1093/erae/30.2.241*>.

MORISSON, G. W. AND PIKE, D. H. Kalman Filter Applied to Statistical Forecasting. *Management Sciences*, 1977, Vol. 23, No. 7, pp. 768–774.

MOTALLEBI, M. AND PENDELL, D. L. Estimating an Almost Ideal Demand System Model for Meats in Iran. *Agricultural & Applied Economics Association's 2013 AAEA&CAES Joint Annual Meeting*, Washington, D.C., 2013.

NZUMA, J. M. AND SARKER, R. An error corrected almost ideal demand system for major cereals in Kenya. *Agricultural Economics*, 2010, 41, pp. 43–50.

PORAY, M. C., FOSTER, K. A., DORFMAN, J. H. Measuring an Almost Ideal Demand System with Generalized Flexible Demand System. *American Agricultural Economists Association (AAEA) Annual Meeting*, Tampa, Florida, 2000.

SLADE, M. E. Modelling Stochastic and Cyclical Components of Technical Change: an Application of the Kalman Filter. *Journal of Econometrics*, 1989, Vol. 41, No. 3, pp. 363–383.

TALJAARD, P., VAN SCHALKWYK, H., ALEMU, Z. Choosing between the AIDS and Rotterdam models: a meat demand analysis case study. *Agrekon*, 2006, 42, pp. 158–72.

THAMAE, R. I., THAMAE, L. Z., THAMAE, T. M. Dynamics of electricity demand in lesotho: a kalman filter approach. *Studies in Business and Economics*, 2015, No. 10(1), pp. 130–139.

ZERANEZHAD, M. AND SAADATMER, M. Estimation of demand for red meat in Iran. *Journal of the faculty of humanities*, 2007, Vol. 7, No. 6, pp. 63–82.

Modus Operandi of Actors Involved in the Illicit Tobacco Trade in EU Countries

Cristina Boboc[1] | *Bucharest Academy of Economic Studies, Bucharest, Romania*
Magdalena Ciobanu[2] | *National Institute of Pneumology, Bucharest, Romania*

Abstract

The goal of the present study is to contribute to the improvement of cooperation between countries in global efforts to eliminate illicit trade in tobacco products (ITTP), by identifying common gaps and potential solutions using modern statistical instruments. For each of the 30 European reference countries, the first objective of this paper is to identify models of ITTP *modus operandi*. Empirical and individual observations suggest that such models exist, but no rigorous statistical evidence is available. The second objective of this paper is to assess the similarities and differences between various components of governance in countries for each ITTP model identified. The paper demonstrates that countries sharing common patterns of *modus operandi* in ITTP, also share common strengths and weaknesses in their governance status. Reinforcing governance with shared instruments and common goals across countries sharing common ITTP *modus operandi*, can potentially improve the control of illicit trade in these products. The current study presents evidence for the need to tailor cooperation between countries and the significant role of non-fiscal measures in fighting ITTP.

Keywords	***JEL code***
Illicit tobacco trade, European Union countries, cluster analysis	*C38, I18*

INTRODUCTION

The European Commission's 2nd Action Plan to Fight the Illicit Tobacco Trade 2018–2022, based on the recommendations of the WHO Framework Convention on Tobacco Control (FCTC) Protocol to Eliminate Illicit Trade in Tobacco Products, as well as reports of other international organizations (the World Bank, Centre for Disease Prevention and Control, etc.) emphasizes the essential role of bilateral and multilateral cooperation between states for an effective and efficient fight against illicit tobacco trade (ITT) (WHO-FCTC, 2013; EC, 2017: *Com (2013) 324*).

It is a major concern in Europe, as stated by the European Commission in its progress report to the Council and European Parliament regarding the implementation of the European Union (EU)

---

[1] Professor, Department of Statistics and Econometrics, Bucharest Academy of Economic Studies, 15–17 Calea Dorobantilor, 010552 Bucharest, Romania. Also Senior Researcher, Institute of National Economy, Romanian Academy, 13 Calea 13 Septembrie, 050711, Bucharest, Romania. E-mail: cristina.boboc@csie.ase.ro.
[2] Coordinator of the Smoking Cessation Centre, National Institute of Pneumology "Marius Nasta", 90 Sos. Viilor, 050159 Bucharest, Romania. E-mail: magda_ciobanu@yahoo.com.

strategy against cigarette smuggling and other forms of illicit trade in tobacco products (*Com (2013) 324*). By making cigarettes more affordable and accessible to people from low-income groups, as well as to children, through lower prices than those set to discourage smoking, and by avoiding product regulation (e.g. such as labelling and control of ingredients), the illicit trade in tobacco products poses a serious threat to public health because it facilitates the uptake of tobacco use by youth and undermines tobacco control policies. According to the European Commission, there are substantial losses in government revenues: It is estimated that, if all cigarettes sold on the black market were sold legally, the budget of the EU and its Member States would receive above € 10 billion annually (*Com (2013) 324*). ITTP is also a source of revenue for organized crime groups from Europe and beyond, as well as for terrorist organisations (*UN Security Council Resolution No. 2199*). Therefore, fighting the global illicit tobacco trade is essential to protect EU public health, public revenues and public security.

The actors contributing to development and persistence of illicit tobacco trade are numerous and diverse, from individuals to transnational criminal networks. Illicit trade can be undertaken both by illicit players, not legally registered, as well as by legitimate entities with (some) business operations that do not comply with applicable laws and regulations (e.g. some duty-free zones, tobacco product manufacturers). These illicit tobacco trade activities are carried out by three main types of actors, each one adopting different *modus operandi* practices: large-scale actors, medium-scale actors and small-scale actors (Savona and Riccardi, 2015). The potential profits associated with large-scale ITTP and low levels of risks in terms of detection, seizures, penalties and criminal procedure, create incentives for participation by organized crime networks. Even if the number of large-scale actors is less numerous, they are believed to be responsible for more than 90% of illicit tobacco trade (Savona and Riccardi, 2015).

The features of ITTP vary from one region or country to another, although the main characteristics are common, largely falling under the three following categories: contraband, counterfeit or illicit whites.[3]

Three main established routes are used to bring cigarettes into Europe: the North-Eastern route, the extended Balkan route and the Maghreb route (Savona and Riccardi, 2015). The North-Eastern route is the main way by which illicit flows of cigarettes from extra-EU eastern European countries enter EU Member States. About half are illicit whites and the remainder are contraband. The actors are organized crime groups involved in large-scale cigarette trafficking.

Price and tax differences between countries create financial incentives to avoid or evade taxes. The impact of tax and price disparities on type and level of illicit trade activities has been examined extensively by economists. For example, price differences between adjacent geographical areas motivate bootlegging and legal cross-border shopping, according to studies conducted in the United States (Baltagi and Levin, 1992; DeCicca et al., 1997; Licari and Meier, 1997), multiple European countries (Joossens and Raw, 2008; Merriman, 2000), Estonia (Taal et al., 2004), the United Kingdom (Buck et al., 1994), France (Lakhdar, 2008), and in many other countries.

Despite studies and campaigns conducted by the tobacco industry promoting the message that taxes and prices have the most important impact on ITTP at a country level, independent evidence indicates that the illicit cigarette market is relatively larger in countries with low taxes and prices while being relatively smaller in countries with higher cigarette taxes and prices (National Research Council, 2015). Illicit trade in tobacco is not only inconsistent with the rule of law, but often depends on and can contribute to weakened governance (e.g. through corruption and the presence of organized criminal networks) (World Bank Group, 2019). Thus, non-price factors such as governance status, weak regulatory frameworks, social acceptance of illicit trade, and the availability of informal distribution networks appear to be far more important determinants of the size of the illicit tobacco market (Chaloupka et al., 2019).

3 'Illicit whites' (also known as 'cheap whites') refers to cigarettes produced lawfully in one jurisdiction for the sole purpose of being exported and illegally sold in a jurisdiction where they have no legitimate market. Illicit whites have emerged in ITTP channels in the EU over the past decade and several sources indicate their growing importance.

The laws, regulations, systems and effectiveness of governance that contribute to the political and regulatory environment influencing the illicit trade, were analysed by The Economist Intelligence Unit in 2018 using relevant literature and consultations with independent and tobacco industry-related advisers. The result is the Global Illicit Trade Environment Index (ITEI) Report (The Economist Intelligence Unit, 2018), which evaluates 84 countries, including EU Member States, on their structural capability to protect against illicit trade, either through action or inaction. The index is built around four main categories, each with four to seven indicators: government policy; supply and demand; transparency and trade; and customs environment.[4]

The goal of this study is to contribute to the improvement of cooperation between countries in global efforts to eliminate ITT, by identifying common gaps and common possible solutions using modern statistical instruments.

The analysis was carried out for 30 European countries (28 EU Member States, Norway and Switzerland). It comprises two phases: the identification of patterns of modus operandi for ITTP (or 'typologies') and the identification of associations between specific patterns and specific governmental policy.

For each of the 30 European reference countries, the first objective of this paper is to identify models of *modus operandi* in ITT. Empirical and individual observations suggest that such models exist, but no rigorous statistical evidence is available.

With an increasing body of evidence suggesting the substantial role of non-price factors as determinants of the size of the illicit tobacco market, the second objective of this paper is to assess the similarities and differences between different components of governance in countries for each identified ITT model.

The main sources of information for the statistical analysis (SUN Report, the N-EXUS Report and the ITEI Report) were funded by three multinational cigarette manufacturers and use data from independent but industry-related sources. Thus, the most significant limitation of the current study is the use of data belonging to the tobacco industry in the statistical analysis. Taking into account the tobacco industry's long history in manipulating research, and suggestions from different studies about the use of similar strategies in relation to ITTP (Gallagher, 2019), including the recommendations contained in the World Bank review (World Bank Group, 2019), the authors were conscious through the whole study process of the need to be very cautious in using these data. This limitation was overcome by cross-verification of data, where available, including the use of discussions developed in the framework of the World Health Organization (WHO)-EU project. Stakeholders and customs and other governmental experts from countries studied were contacted, and the data were verified for accuracy. A careful and comprehensive analysis of methodologies used in the reports was also undertaken, in order to identify potential bias and distrusted information. Efforts were made to introduce primary data instead of secondary data into the analysis, where available, in order to diminish potential subjective interpretations.

1 TYPOLOGIES OF COUNTRIES IN TERMS OF ITTP

In the first phase, defining typologies of countries in terms of ITTP, five categories of variables were used in describing ITTP (Table 1): (i) category(s) of illicit tobacco trade products; (ii) main brand(s) of illicit cigarettes; (iii) illicit tobacco trade flows; (iv) illicit tobacco trade routes; (v) main country(s) of origin

[4] Indicators included in **Government policy**: 1. Commitment to illicit trade-related treaties, 2. Compliance to Financial Action Task Force (FATF) money laundering provisions and standards, 3. Intellectual property protection, 4. Corruption, 5. Law enforcement techniques, 6. Interagency collaboration, 7. Cybersecurity preparedness; in **Transparency and Trade**: 1. Track and trace services, 2. Adoption of Annex D of Revised Kyoto Convention, 3. Free trade zones governance, 4. International reporting; in **Supply and Demand**: 1. Tax and social security burdens, 2. Quality of state institutions, 3. Labour market regulations, 4. Perception of organized crime; and in **Customs environment**: 1. Percentage of shipments physically inspected, 2. Customs clearance and inspection, 3. Automation, 4. Authorized Economic Operator programme, 5. Customs recording system.

of illicit cigarettes. One of the most common hierarchical clustering techniques, the Ward method, was used to create homogenous groups of countries. As the database contains both quantitative and binary variables, we have chosen to use the Gower and Legendre measure of dissimilarity (Gower and Legendre,1986).

The dendrogram shown in Figure 1 was derived using the Ward method. By analysing the latest ten steps of clustering history – by applying the pseudo T-square index[5] and pseudo

Table 1 Variables describing illicit tobacco trade

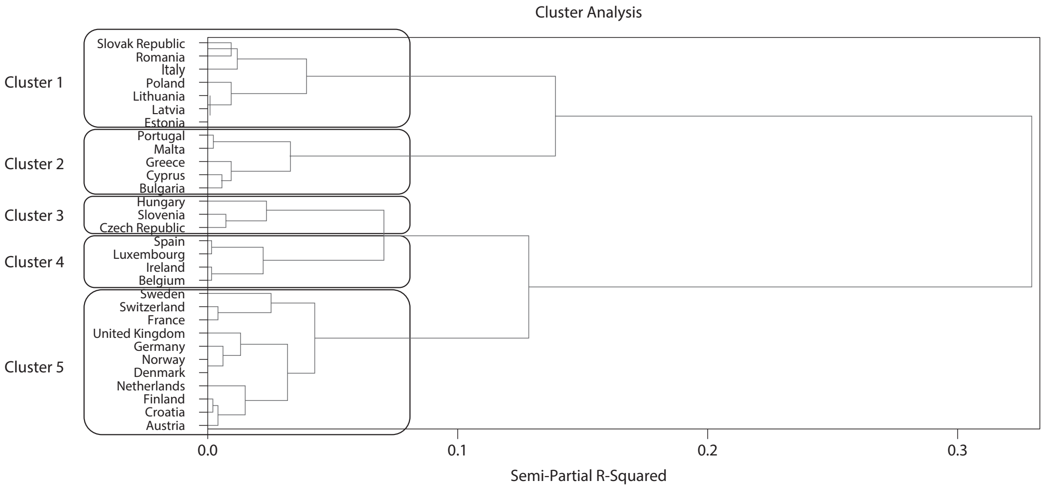
	Description	Indicators	Data source
1	2	3	4
Category(s) of illicit tobacco trade products	Percentages of cigarettes in each category in the total number of illicit cigarettes in the reference country (in 2017).	1. Illicit whites (IW, %) 2. Counterfeits (%) 3. Contraband or loose tobacco (%)	RUSI data (KPMG, 2017)
The main brand of illicit cigarettes	Indicator that specifies which are the manufacturers of the first two most frequent brands of illicit cigarettes in the reference country.	1. British American Tobacco 2. Japan Tobacco International 3. Philip Morris International 4. Other (not known or manufacturer of illicit whites)	RUSI data (KPMG, 2017)
Illicit tobacco trade flows	Indicators that specify which are the main illicit tobacco trade flows in the reference country. They are constructed as follows: - if the reference country is destination for illicit tobacco products the flows are Inwards; - if the reference country is origin for illicit tobacco products the flows are Outwards; - if the reference country is on a route between an origin and a destination country for illicit tobacco products the flows are Transit. A reference country could have Inwards, Outwards and Transit flows or any other combination of them.	1. Inwards 2. Transit 3. Outwards	RUSI data (KPMG, 2017) and NEXUS data (Aziani and Dugato, 2019) on traffic routes. To have more accurate data, they were compared and completed with information from qualitative interviews with in-country experts conducted under the WHO–EU project: Illicit Tobacco Trade in the European Union 2017–2019 – raising awareness and enhancing understanding of illicit tobacco trade among academic researchers in the European Union.
Illicit tobacco trade routes	Indicators that specify the routes used by traffickers with tobacco products in reference country.	1. North eastern route 2. Balkan route 3. Maghreb route	NEXUS data (Aziani and Dugato, 2019)
Main country(s) of origin for illicit cigarettes	Indicators that specify whether the main origin country for illegal cigarettes found in reference country is one of the non-EU neighbouring countries: Belarus, Russian Federation or Ukraine, or other known origin country (from EU or not), or unknown origin country.	1. Ukraine 2. Belarus 3. Russian Federation 4. Illicit whites with unknown country of origin 5. Other country (e.g. Algeria, Bosnia-Herzegovina, Bulgaria, Czech Republic, Estonia, Gibraltar, Poland, Romania)	RUSI data (KPMG, 2017)

Source: Authors own synthesis and computations

[5] Pseudo T-square Index quantifies the difference between two clusters that are merged at a given step. If the pseudo T-square statistic has a distinct jump at step k of the hierarchical clustering, then the clustering in step k + 1 is selected as the optimal cluster (Milligan and Cooper, 1987).

F statistic[6]– the most appropriate number of clusters to group the thirty European countries included was identified as five. The homogeneity inside the clusters is high (with semi-partial R-squared of 0.03) meaning that the countries inside a cluster are very similar from the point of view of the features of ITTP. The variation between clusters is large (with a pseudo F statistic of 13.10), meaning that they are indeed different, enabling the five different typologies of ITTP to be distinguished (Annex Table A1).

Figure 1 Dendrogram obtained by using Ward Method and Gower and Legendre measure



Source: Authors computations using SAS Studio software on RUSI data (KPMG, 2017)

The characteristics of the five clusters identified are summarised in Table A2 in the Annex.

The first typology (Cluster 1), specific to Estonia, Italy, Latvia, Lithuania, Poland, Romania and Slovakia, can be defined as follows: The main category of illicit trade products is illicit whites, representing a mean of 63.4% of total illicit trade. The contraband and loose illicit tobacco products comprise 28.9% and counterfeit illicit tobacco products 7.8%. Smuggled brands are not known brands produced by the top three global manufacturers. The country of origin is either Belarus or unknown. In all countries with this typology, illegal cigarettes enter the country through the north eastern route. Cigarettes also arrive in Italy via the Balkan and Maghreb routes. In general, the countries in this cluster are transit countries, except for Romania, which is also an origin country.

The second typology (Cluster 2), specific to Bulgaria, Cyprus, Greece, Malta and Portugal, can be defined as follows: The illicit tobacco market is divided equally between illicit whites and contraband or loose tobacco, but the main country of origin is unknown. The main brand of illicit tobacco products for all countries is not produced by one of the top three global cigarettes manufacturers. However, in the case of Malta and Portugal, the second brand of illicit tobacco is produced by PMI. All the countries from this cluster use the Balkan route in illegal tobacco trade and in the case of Malta and Portugal illegal trade also uses the Maghreb route (by passing through Spain and Italy). The flows of ITTPs is inward, outward and/or transit.

[6] The pseudo F statistic describes the ratio of between-cluster variance to within-cluster variance, meaning that if there are no significant changes in pseudo F-statistic at step k of the hierarchical clustering, then the clustering in step k + 1 is selected as the optimal cluster (Milligan and Cooper, 1987).

The third typology (Cluster 3), specific to Czech Republic, Hungary and Slovenia, can be defined as follows: The main category of illicit trade products is contraband or loose tobacco, with a mean of 68.3% of total illicit trade. The main two brands traded on the illicit tobacco market are produced by PMI and BAT. The countries are mainly transit countries on the north eastern and Balkan routes, with the main country of origin being Ukraine or one of the countries from the 'Other country' category (e.g. Bosnia-Herzegovina).

The fourth typology (Cluster 4), specific to Belgium, Ireland, Luxembourg and Spain, can be defined as follows: The main category of illicit trade products is contraband or loose tobacco, with a mean of 73.9% of total illicit trade. The two main brands traded on the illicit tobacco market are produced by PMI and BAT. These countries are mainly transit countries. Traffickers mainly use the Maghreb route and the origin country of the products is typically one of the three main source countries for illicit tobacco products in Europe: Belarus, Russian Federation or Ukraine.

The fifth typology (Cluster 5) includes the most affluent EU countries, namely: Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and United Kingdom, and can be defined as follows: The main category of illicit trade products is contraband or loose tobacco, with a mean of 80.8% of total illicit trade. The two main brands traded on the illicit tobacco market are produced by any of the top three global producers (i.e. PMI, BAT or JTI). The countries are mainly destination countries and the illegal cigarettes largely arrive through the north eastern route, mainly from Ukraine.

2 THE RELATIONSHIP BETWEEN THE IDENTIFIED TYPOLOGIES AND GOVERNANCE ENVIRONMENT

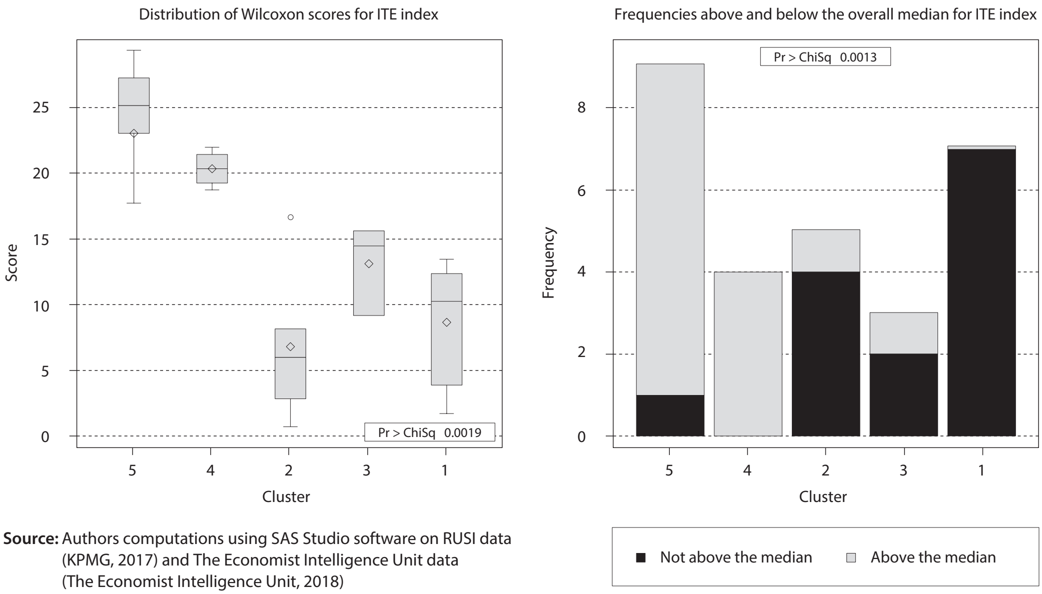
In the second phase, we used the Illicit Trade Environment Index (ITEI) developed by the Economist Intelligence Unit to identify the relationship between the identified typologies and respective governance environments.

Higher values assigned by the ITEI indicate a less favourable political environment for illegal traffic. Conversely, lower values assigned for each component denote an environment that is more favourable for ITTP. Given the small number of countries in designated clusters, a nonparametric analysis of variance was used using two nonparametric tests (the median test and the Kruskal Wallis test) to determine if there were significant differences between groups of countries.

An overview of the results derived using the Kruskal Wallis and median tests demonstrates that there are significant differences between the five clusters regarding the ITEI, government policy, as well as transparency and trade components. In relation to the customs environment component, there are significant differences between clusters in terms of distribution (meaning that the central tendency and the variability are different), with a level of significance of 0.05, but there are no significant differences between medians. Regarding the supply and demand component, there are significant differences between clusters in terms of distribution with a level of significance of 0.1, but there are no significant differences between medians (Figures 2 and 3).

In the case of *countries with the first typology* of illicit tobacco trade (i.e. Cluster 1), all the coefficients of variation are less than 15%, except of supply and demand, which has a coefficient of variation close to 30% (Annex Table A3). This means that this cluster is homogenous with respect to all variables. Moreover, all indicators have low mean levels, meaning that the general governance environment facilitates illicit tobacco trade. The countries from this cluster have made significant improvements in their customs environments, reaching almost the level of clusters with high ITEI (i.e. with an environment only slightly favourable for illegal trafficking). The number of countries above the overall median is zero for government policies and one for transparency and trade, and also the median values for these variables are very low (Figure 3) meaning that, in the countries from this cluster, improvements should be made in government policies and transparency and trade. In Italy and Romania, improvements in supply and demand policies could also be beneficial.

Figure 2 Distribution of Wilcoxon scores and the number of countries above and below overall median for each cluster for each ITEI category



Source: Authors computations using SAS Studio software on RUSI data (KPMG, 2017) and The Economist Intelligence Unit data (The Economist Intelligence Unit, 2018)

In case of *countries with the second typology* of illicit tobacco trade (i.e. Cluster 2), all variables have coefficients of variance less than 30%, but higher than in first cluster (Annex Table A3), meaning that the cluster is homogenous but less so than the previous cluster. All indicators have the lowest mean levels compared to the other clusters (Annex Table A3), meaning that the entire governance environment facilitates illicit tobacco trade. Because the frequency of countries above the overall median is zero for transparency and trade, and the lowest median value is registered for the customs environment (Figure 3), improvements mainly in transparency and trade and in the customs environment would lead to an environment less favourable for illegal trade in general, and in tobacco products in particular. Moreover, in Bulgaria, Greece and Portugal, improvements in government policies would also consistently improve the fight against ITTP.

The *countries with the third typology* of illicit tobacco trade (i.e. Cluster 3), all have coefficients of variation less than 9% (Annex Table A3), meaning that this cluster is the most homogenous with respect to all variables related to the illicit trade environment. The countries from this cluster have a medium ITEI meaning that the governance policies for limitation of ITTP are better than those from second cluster of countries but, comparing with those from the fifth cluster, improvements could be made. The lowest median value is registered for the supply and demand component (Figure 3), meaning that reducing the supply and demand for illicit cigarettes would lead to substantial reductions in ITTP.

The fourth and the fifth clusters both have coefficients of variation less than 25% (Annex Table A3), meaning that they are also very homogenous. These countries have the least favourable environment for illicit trade.

In the case of *countries with the fourth typology*, the median value for government policies is lower than that of the fifth cluster (Figure 3) meaning that improvements in government policies could be made in order to improve the illegal trade environment and therefore to reduce ITTP.

In the case of the *countries with the fifth typology*, the ITEI mean is the highest overall (Annex Table A3), meaning that the entire governance is the strongest and the most efficient in combating ITTP among the countries studied. However, the fifth cluster demonstrate lower values for the customs environment compared to the fourth cluster (Figure 3), meaning that improvements in the customs environment could lead to increased efficiency in reducing illicit tobacco trade, particularly in Croatia. Even if Croatia does not register very high values for transparency and trade or supply and demand, it has good government policies, which compensate and are making the entire environment less favourable for illicit trade.

Figure 3 Distribution of Wilcoxon scores and the number of countries above and below overall median for each cluster for Government Policy, Transparency, Supply and demand, and Customs

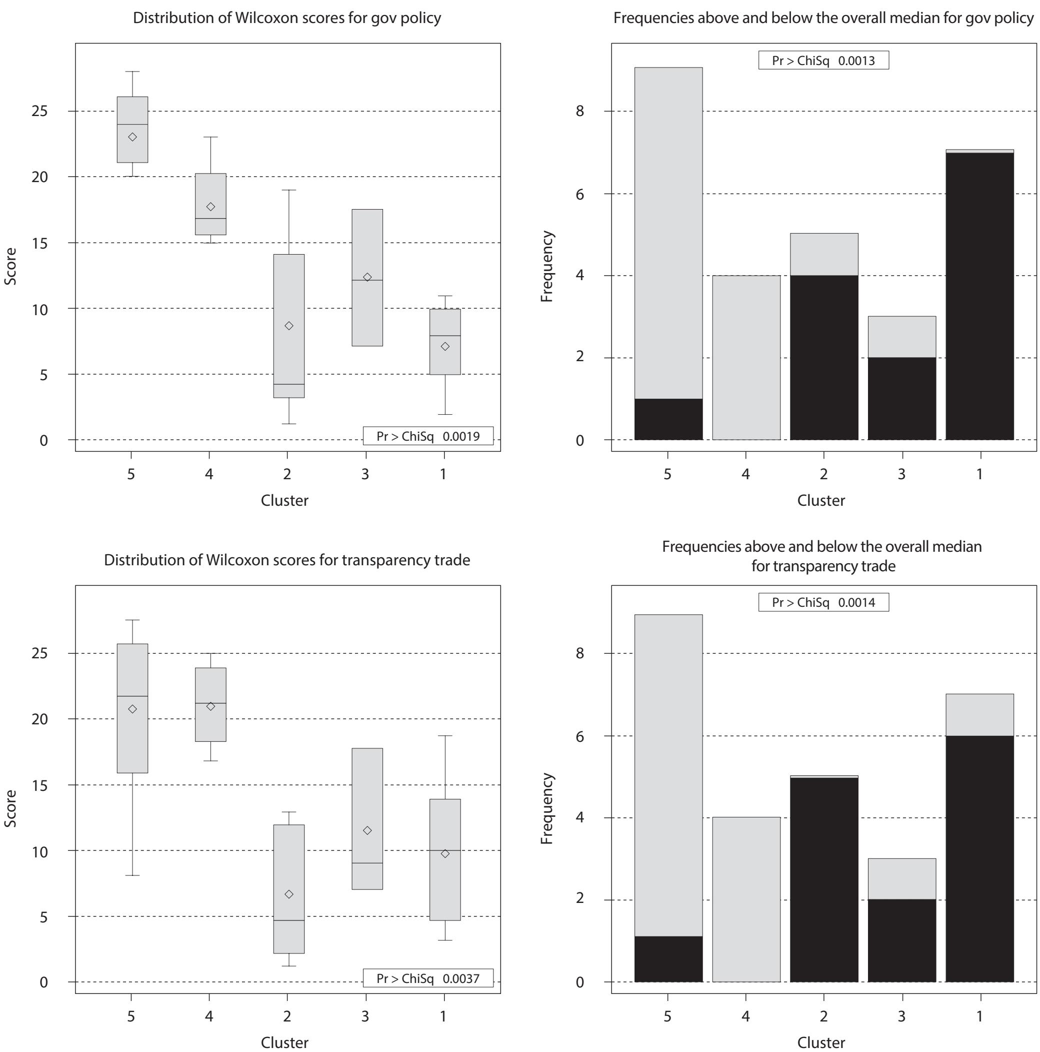
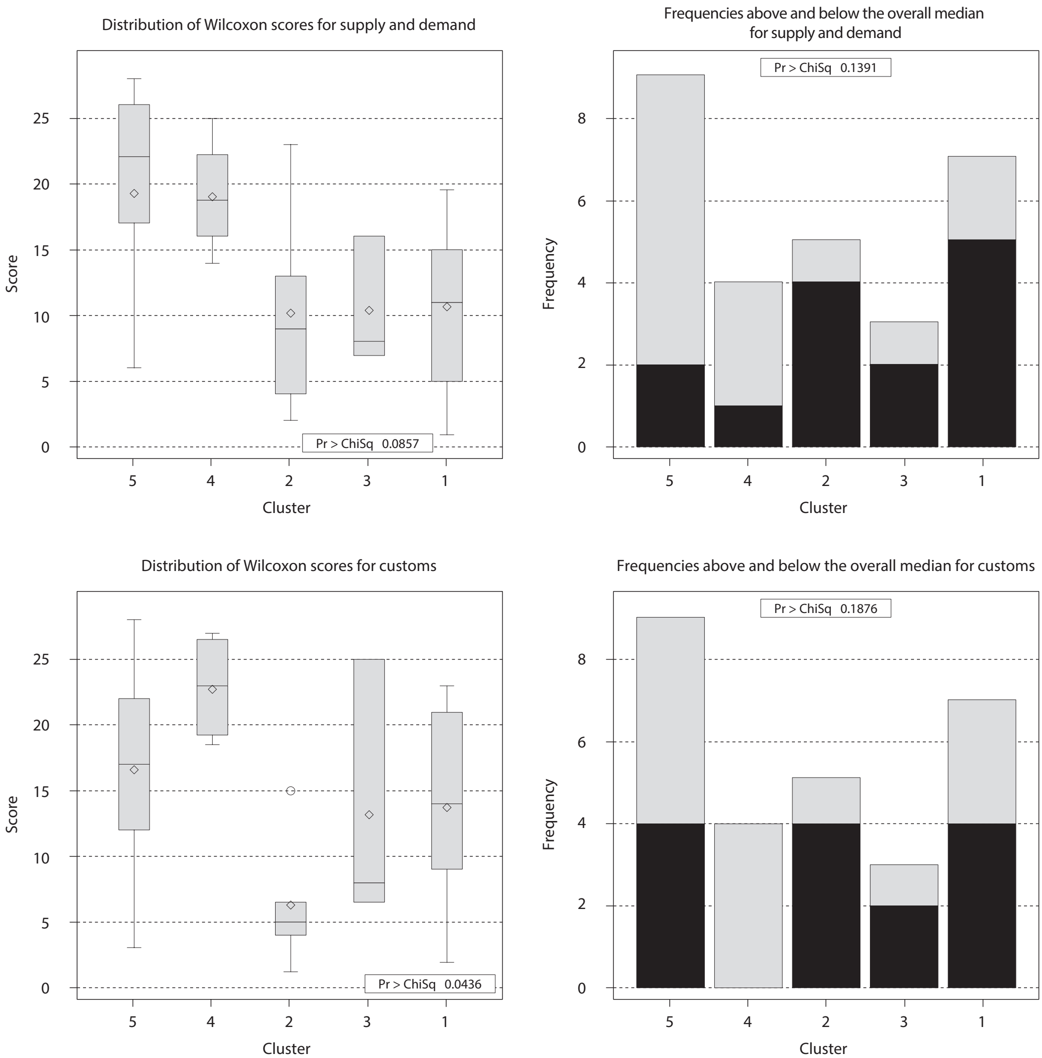
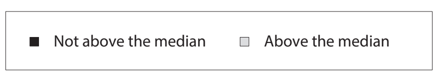


Figure 3 (continuation)





Source: Authors computations using SAS Studio software on RUSI data (KPMG, 2017) and The Economist Intelligence Unit data (The Economist Intelligence Unit, 2018)

3 DISCUSSION

The study identified five models of ITT modus operandi in EU countries plus Norway and Switzerland (Table 2).

Of the seven Model 1 countries, six (Estonia, Latvia, Lithuania, Poland, Romania and Slovakia) share a land border with former Soviet countries (Belarus, Republic of Moldova, Russian Federation and Ukraine) and a geographical position in north eastern Europe. This could explain the similarities in illicit trade of cigarettes. The results are consistent with the opinions and observational remarks of stakeholders interviewed in the framework of the EU project.

Table 2 Models of *modus operandi* for ITTP

Model of ITTP	Countries	The main features of ITTP	Governance policy to improve
Model 1	Estonia, Italy, Latvia, Lithuania, Poland, Romania and Slovakia	Illicit whites Main brand: JTI Main country of origin: IW Second main country of origin: Belarus Transit country North eastern route	Government policies Transparency and trade
Model 2	Bulgaria, Cyprus, Greece, Malta and Portugal	Illicit whites Contraband or loose tobacco Main brand: PMI Main country of origin: IW Origin country Destination country Transit country Balkan route	Transparency and trade Customs environment
Model 3	Czech Republic, Hungary and Slovenia	Contraband or loose tobacco Main brand: PMI Main country of origin: Ukraine Transit country North eastern route Route Balkan	Supply and demand
Model 4	Belgium, Ireland, Luxembourg and Spain	Contraband or loose tobacco Main brands: PMI and BAT Main country of origin: Other Transit country Maghreb route	Government policies
Model 5	Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and United Kingdom	Contraband or loose tobacco Main brand: PMI Main country of origin: Other Destination country North eastern route	Customs environment

Source: Authors own synthesis and computations

The intriguing aspect is the presence of Italy in this cluster of seven countries, despite a totally different geographical position and political background. The high statistical power of association in this cluster assures us that this grouping is not random, and that there have to be some common aspects. Analysing the variables, the common features of the *modus operandi* are that the most commonly smuggled cigarettes are illicit whites, and that all seven countries are mainly transit countries for ITTP. Italy has an accessible source of illicit whites due to its geographical position in the Mediterranean (by sea, through the Maghreb route) and also has a non-EU land border favouring the transit of illicit products.

Our analysis demonstrates the importance of geographical position to the existence of opportunities for trafficking illicit white cigarettes. If a country has borders and the geographical position favours communication with illicit white source countries, it is more prone to share Model 1 traits with other countries. These countries could cooperate to improve governmental policies in areas such as cybersecurity preparedness, money laundering provisions, developing common standards, inter-agency collaboration and international reporting – as statistics show that these policies are the weakest in the fight against ITTP in these countries.

Model 2 countries (Bulgaria, Cyprus, Greece, Malta and Portugal) are the least homogenous group of the five models. These countries share a common route for illicitly traded cigarettes: the Balkan route. This route is proximal for four out of the five countries in the cluster. In case of Portugal and Malta, the Maghreb route is also used. Moreover, the proximity of the sea (an accessible source for illicit whites) and of a non-EU land border favours the transit of illicit products. From the perspective of improvements in governance policies, possibly achieved through extensive cooperation between Model 2 countries, the most interesting potential areas are improvement of customs recording systems, the governance of free trade zones and international reporting. These improvements are achievable, as indicated by the discussions in the framework of the EU-funded project.

The countries of Model 3 (Czech Republic, Hungary and Slovenia) are land neighbours with Austria, from which cigarettes are illicitly traded to Germany, an EU state among the highest consumers of illicit tobacco products (cigarettes and loose tobacco). This geographical feature and the lack of formal borders (as all these countries are part of the Schengen area) ease the illicit trade. The weakest component in these three countries is supply and demand for illicit cigarettes. According to the latest Eurobarometer on public perception of the illicit tobacco trade, only 18% of Hungarian citizens believe that black market cigarettes provide one of the most important sources of revenue for organized crime, in Czechia this is 12% and in Slovenia 11%. These perceptions could motivate the authorities from the three countries to collaborate in enhancing perception of organized crime among their citizens in relation to ITTP, in an attempt to improve the weakest area of governance policy, namely supply and demand.

The Model 4 countries (Belgium, Ireland, Luxembourg and Spain) are all very high developed countries that in general prefer original brands of international manufacturers rather than illicit whites. However, the proximity of the sea (three of them have sea borders), and the lack of a formal borders with other EU countries as members of the Schengen Area, makes these countries accessible for illicit products, favouring the Maghreb route to France or the United Kingdom and Ireland (using the western sea borders of the EU). Even though they are very developed countries, they could cooperate more to improve governmental policies in areas such as cybersecurity preparedness and money laundering provisions.

The Model 5 countries (Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and the United Kingdom) strongly prefer original brands to illicit whites. The study finds that the most smuggled brand in these destination countries is owned by Philip Morris International. This situation could be explained by the high quality of life and revenue indicators of the population living in these countries. The products are not manufactured in the reference country but are transported via the north eastern route. As most of these countries are members of the Schengen Area, it is obvious that the weakest link in the chain of governance policies is the customs environment; however, this is difficult to improve within the EU borderless framework. Considering that the most used illicit cigarettes in these countries are the brands owned by the three big manufacturers, and the factories are in the countries along the north eastern route (the main route for transport), the collaboration between these states should be focused more on implementing the EU 'track and trace' system.

CONCLUSIONS

The paper demonstrates that countries sharing a common pattern of *modus operandi* in ITTP also share common strengths and weaknesses in their governance status. Reinforcing governance with common instruments and common goals in countries sharing a common ITTP *modus operandi*, could improve the control of illicit trade in these products. Thus, the study presents evidence for the need to tailor cooperation between countries in order to maximize the result.

The study also presents evidence for the significant role of non-fiscal measures in fighting ITTP. While the recommended fiscal measures are the same for all countries (i.e. increased taxation using comparable instruments), the non-fiscal measures must be adapted to the internal needs and particularities of each country, in order to be effective and efficient. The study supports the empirical observations and assumptions that good implementation of the EU track and trace system, part of trade and transparency policy of good governance, can diminish the illicit outflow of cigarettes from Model 2 countries and the illicit inflow to Model 5 countries. If Model 1 and Model 4 countries collaborate in improving governmental policies targeting cybersecurity preparedness, corruption and money laundering, we could expect a decrease in the illicit trade of branded and non-branded cigarettes transported through routes with both EU and non-EU origins.

This paper is not intended to support the use of industry-related data, or to encourage the use of this information. In the absence of any independent and publicly available assessments of the 3 reports, due to their recent publishing (in 2018 and 2019), and in the absence of other sources of detailed quantitative information regarding the magnitude and the modus operandi of ITTP, the authors consider that they used these sensitive data with the greatest possible precaution. Although, the statistical methods used in current paper are reliable and can be used in attaining the objectives related to a better understanding of ITTP, in future studies it is recommended the use of data from total industry-independent sources.

ACKNOWLEDGEMENTS

This paper was supported by the WHO project *Illicit Tobacco Trade in the European Union 2017–2019 – Raising awareness and enhancing understanding of illicit tobacco trade among academic researchers in the European Union*, financed by EU under the contract with WHO Reference 2018/866436-0.

References

AZIANI, A. AND DUGATO, M. (eds.) *ITTP Nexus in Europe and Beyond.* Milano: Transcrime – Università Cattolica del Sacro Cuore, 2019.

BALTAGI, B. H. AND LEVIN, D. Cigarette taxation: raising revenues and reducing consumption. *Struct. Change Econ. Dynam.*, 1992, 3(2), pp. 321–35.

BUCK, D., GODFREY, C., RICHARDSON, G. *Should cross-border shopping affect tax policy?* York: Centre for Health Economics, Yorkshire Addictions Research, Training and Information Consortium (YARTIC), 1994.

DECICCA, P., KENKEL, D. S., LIU, F. *Excise tax avoidance: the case of state cigarette taxes.* NBER working paper No. 15941, Cambridge, MA: National Bureau of Economic Research, 2010.

CHALOUPKA, F. J. ET AL. *Preventing and Reducing Illicit Tobacco Trade in the United States* [online]. Atlanta, GA: Centers for Disease Control and Prevention, 2015. [cit. 30. 12. 2019]. <https://www.cdc.gov/tobacco/stateandcommunity/pdfs/illicit-trade-report-121815-508tagged.pdf>.

EUROPEAN COMMISSION. *Communication 2nd Action Plan to fight the illicit tobacco trade 2018–2022 (COM (2018) 846 of 7. 12. 2018)* [online]. Brussels: EC, 2018. <https://www.eumonitor.eu/9353000/1/j4nvke1fm2yd1u0_j9vvik7m1c3gyxp/vku3ls1e68z9/v=s7z/f=/com(2018)846_en.pdf>.

EUROPEAN COMMISSION. *Progress report on the implementation of the Commission Communication: "Stepping up the fight against cigarette smuggling and other forms of illicit trade in tobacco products – a comprehensive EU strategy (Com (2013) 324 final of 6. 6. 2013)"* [online]. Brussels: EC, 2017. [cit. 30. 12. 2019]. <https://ec.europa.eu/anti-fraud/sites/antifraud/files/tobacco_implementation_report_12052017_en.pdf>.

GALLAGHER, A. W. A., EVANS-REEVES, K. A., HATCHARD, J. L., GILMORE, A. B. Tobacco industry data on illicit tobacco trade: a systematic review of existing assessments. *Tobacco Control*, 2019, 28, pp. 334–345.

GOWER, J. C. AND LEGENDRE, P. Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, 1986, 3, pp. 5–48.

JOOSSENS, L. AND RAW, M. Progress in combating cigarette smuggling: controlling the supply chain. *Tobacco Control*, 2008, 17, pp. 399–404.

KPMG. *Project SUN: a Study of the Illicit Cigarette Market in the European Union, Norway and Switzerland* [online]. London: KPMG, 2017. [cit. 30. 12. 2019]. <https://assets.kpmg/content/dam/kpmg/uk/pdf/2018/08/project_sun_methodology.PDF>.

LAKHDAR, C. B. Quantitative and qualitative estimates of cross-border tobacco shopping and tobacco smuggling in France. *Tob. Control*, 2008, 17, pp. 12–26.

LICARI, M. J. AND MEIER, K. J. Regulatory policy when behavior is addictive: smoking, cigarette taxes and bootlegging. *Polit. Res. Q.*, 1997, 50, pp. 5–24.

MERRIMAN, D., YÜREKLI, A., CHALOUPKA, F. How big is the worldwide cigarette smuggling problem? In: JHA, P. AND CHALOUPKA, F. J. (eds.) *Tobacco control in developing countries*, London: Oxford University Press, 2000.

MILLIGAN, G. W. AND COOPER, M. C. Methodology Review: Clustering Methods [online]. *Applied Psychological Measurement*, 1987, 11(4), pp. 329–354. <https://doi.org/10.1177/014662168701100401>.

NATIONAL RESEARCH COUNCIL. *Understanding the U.S. Illicit Tobacco Market: Characteristics, Policy Context, and Lessons from International Experiences.* Washington, DC: The National Academies Press, 2015.

SAVONA, E. U. AND RICCARDI, M. (eds.) *From illegal markets to legitimate businesses: the portfolio of organised crime in Europe.* Trento, Italy: Transcrime – Università degli Studi di Trento, 2015.

TAAL, A., KIIVET, R., HU, T. W. *The economics of tobacco in Estonia* [online]. HNP discussion papers: Economics of tobacco control paper No. 19, Washington, DC: World Bank, 2004. [cit. 30. 12. 2019]. <http://documents.worldbank.org/curated/en/940641468038054384/pdf/297730PAPER0Es1h0cover0smaller0size.pdf>.
The Global Illicit Trade Environment Index [online]. London: The Economist Intelligence Unit, 2018. [cit. 30. 12. 2019]. <http://illicittradeindex.eiu.com/documents/EIU%20Global%20Illicit%20Trade%20Environment%20Index%202018%20-%20Overall%20Results%20White%20Paper%20June%206%20FINAL.pdf>.
UN Security Council Resolution No. 2199 Adopted by the Security Council at its 7379th meeting on 12 February 2015 [online]. 2015. [cit. 30. 12. 2019]. <https://www.undocs.org/S/RES/2199%20>.
WHO-FCTC. *Protocol to eliminate illicit trade in tobacco products* [online]. Geneva: World Health Organization Framework Convention on Tobacco Control, 2013. [cit. 30. 12. 2019]. <http://www.who.int/fctc/protocol/en>.
WORLD BANK GROUP. *Confronting Illicit Tobacco Trade: a Global Review of Country Experiences* [online]. WBG Global Tobacco Control Program, Washington, DC: World Bank Group, 2019. [cit. 30. 12. 2019]. <http://documents.worldbank.org/curated/en/677451548260528135/Confronting-Illicit-Tobacco-Trade-a-Global-Review-of-Country-Experiences>.

ANNEX

Table A1 Clustering history – Ward method

Cluster history

Number of clusters	Clusters joined		Freq.	New cluster RMS (Std)	Semi-partial R-squared	R-squared	Pseudo F statistic	Pseudo T-square
10	CL18	Hungary	3	0.18	0.02	0.85	12.70	3.30
9	CL21	Sweden	3	0.17	0.03	0.83	12.50	6.10
8	CL12	CL13	8	0.14	0.03	0.79	12.10	4.70
7	CL17	CL23	5	0.16	0.03	0.76	12.20	6.10
6	CL15	CL14	7	0.15	0.04	0.72	12.40	6.40
5	CL8	CL9	11	0.17	0.04	0.68	13.10	3.80
4	CL11	CL10	7	0.21	0.07	0.61	13.30	6.30
3	CL5	CL4	18	0.22	0.13	0.48	12.20	7.60
2	CL7	CL6	12	0.22	0.14	0.34	14.10	11.70
1	CL3	CL2	30	0.26	0.33	0.00	.	14.10

Notations: RMS: Root Mean Square; Std: standard deviation.
Source: Computations using SAS Studio software on RUSI data (KPMG, 2017)

Table A2 Typologies of illicit trade of cigarettes

Cluster	No obs.	Variable	Mean	Std.	Minimum	Maximum
Cluster 1 Estonia, Italy, Latvia, Lithuania, Poland, Romania and Slovakia	7	Illicit whites (%)	63.4%	15.5%	38.9%	87.7%
		Counterfeit (%)	7.8%	5.5%	1.4%	16.7%
		Contraband or loose tobacco (%)	28.8%	10.7%	10.5%	44.4%
		Main brand of illicit cigarettes: JTI	14.3%			
		Main country of origin: IW	42.9%			
		Main country of origin: Belarus	57.1%			
		Origin country	14.3%			
		Transit country	71.4%			

Table A2 (continuation)

Cluster	No obs.	Variable	Mean	Std.	Minimum	Maximum
Cluster 1 Estonia, Italy, Latvia, Lithuania, Poland, Romania and Slovakia	**7**	North eastern route	100.0%			
		Balkan route	14.3%			
		Maghreb route	14.3%			
Cluster 2 Bulgaria, Cyprus, Greece, Malta and Portugal	**5**	Illicit whites (%)	49.7%	11.9%	38.1%	62.5%
		Counterfeit (%)	5.8%	9.4%	0.0%	22.3%
		Contraband or loose tobacco (%)	44.5%	12.3%	35.0%	61.9%
		Main brand of illicit cigarettes: PMI	40.0%			
		Main country of origin: IW	100.0%			
		Origin country	40.0%			
		Destination country	60.0%			
		Transit country	60.0%			
		Balkan route	100.0%			
		Maghreb route	40.0%			
Cluster 3 Czech Republic, Hungary and Slovenia	3	Illicit whites (%)	26.4%	13.3%	11.4%	36.6%
		Counterfeit (%)	5.3%	4.4%	0.3%	8.3%
		Contraband or loose tobacco (%)	68.3%	17.5%	56.1%	88.3%
		Main brand of illicit cigarettes: PMI	66.7%			
		Main brand of illicit cigarettes: BAT	33.3%			
		Main country of origin: Ukraine	66.7%			
		Main country of origin: other	33.3%			
		Origin country	33.3%			
		Transit country	100.0%			
		North eastern route	100.0%			
		Balkan route	100.0%			
Cluster 4 Belgium, Ireland, Luxembourg and Spain	4	Illicit whites (%)	20.8%	14.5%	3.0%	33.3%
		Counterfeit (%)	5.3%	3.8%	0.0%	9.1%
		Contraband or loose tobacco (%)	73.9%	11.6%	62.3%	87.9%
		Main brand of illicit cigarettes: PMI	50.0%			
		Main brand of illicit cigarettes: BAT	50.0%			
		Main country of origin: other	100.0%			
		Transit country	75.0%			
		Maghreb route	75.0%			
		North eastern route	25.0%			
Cluster 5 Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and United Kingdom	11	Illicit whites (%)	9.1%	6.2%	1.1%	23.1%
		Counterfeit (%)	10.0%	6.1%	0.0%	18.8%
		Contraband or loose tobacco (%)	80.9%	8.9%	68.5%	93.3%
		Main brand of illicit cigarettes: PMI	90.9%			

Table A2 (continuation)

Cluster	No obs.	Variable	Mean	Std.	Minimum	Maximum
Cluster 5 Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and United Kingdom	11	Main brand of illicit cigarettes: BAT	36.4%			
		Main brand of illicit cigarettes: JTI	27.3%			
		Main country of origin: Russian Federation	9.1%			
		Main country of origin: Ukraine	18.2%			
		Main country of origin: other	72.7%			
		Destination country	100.0%			
		Transit country	27.3%			
		North eastern route	90.9%			
		Balkan route	27.3%			
		Maghreb route	9.1%			

Note: The dummy variables with mean equal to zero are not included among the characteristics of the cluster as long as mean zero means the absence of that attribute.
Notations: No. obs.: number of observations; Std.: standard deviation; BAT: British American Tobacco; JTI: Japan Tobacco International; PMI: Philip Morris International.
Source: Computations using SAS Studio software on RUSI data (KPMG, 2017)

Table A3 Means and standard deviations for all Illicit trade environment indexes

Cluster	No obs.	Variable	Mean	Std.	Coefficient of variation	Minimum	Maximum
Cluster 1 Estonia, Italy, Latvia, Lithuania, Poland, Romania and Slovakia	7	ITEI	67.8	3.8	5.7%	60.8	71.1
		Government policy	70.1	3.5	5.0%	62.6	72.5
		Transparency trade	58.9	6.4	10.8%	50.8	68.0
		Supply and demand	51.5	13.4	26.0%	23.8	64.4
		Customs	84.8	3.2	3.8%	78.0	87.5
Cluster 2 Bulgaria, Cyprus, Greece, Malta and Portugal	5	ITEI	65.6	6.0	9.1%	57.7	73.1
		Government policy	68.5	7.8	11.4%	62.5	79.4
		Transparency trade	53.7	11.6	21.5%	37.8	65.2
		Supply and demand	53.0	13.6	25.6%	36.0	71.8
		Customs	80.9	3.3	4.1%	77.2	85.8
Cluster 3 Czech Republic, Hungary and Slovenia	3	ITEI	70.5	1.7	2.5%	68.5	71.6
		Government policy	74.3	4.0	5.4%	71.1	78.8
		Transparency trade	61.6	5.5	8.9%	57.5	67.8
		Supply and demand	55.0	4.2	7.6%	52.3	59.8
		Customs	84.5	3.2	3.8%	81.6	87.9
Cluster 4 Belgium, Ireland, Luxembourg and Spain	4	ITEI	76.9	2.1	2.7%	74.1	78.6
		Government policy	79.8	3.7	4.6%	76.9	85.2
		Transparency trade	70.5	2.3	3.3%	67.6	72.9
		Supply and demand	65.1	7.0	10.7%	58.2	74.8
		Customs	87.4	1.0	1.2%	86.5	88.5

Table A3 (continuation)

Cluster	No obs.	Variable	Mean	Std.	Coefficient of variation	Minimum	Maximum
Cluster 5 Austria, Croatia, Denmark, Finland, France, Germany, Netherlands, Norway, Sweden, Switzerland and United Kingdom	11	ITEI	78.9	6.0	7.6%	65.8	85.6
		Government policy	85.9	7.2	8.4%	73.4	96.3
		Transparency trade	69.8	5.1	7.3%	58.8	75.0
		Supply and demand	67.0	13.7	20.4%	43.9	81.3
		Customs	85.8	3.2	3.8%	78.3	90.2

Note: If coefficient of variation is less than 30% it means that the cluster is homogenous.
Notations: Obs.: observations; Std: standard deviation; ITEI: Illicit Trade Environment Index.
Source: Computations using SPSS software on RUSI data (KPMG, 2017) and The Economist Intelligence Unit data (The Economist Intelligence Unit, 2018)

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 underline 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 and 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 "insert Table 1 about here". 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 "insert Figure 1 about here". 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 66 per copy + postage.
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
Czech Statistical Office | Information Services
Na padesátém 81 | 100 82 Prague 10 | Czech Republic
Phone: (+420) 274 052 733, (+420) 274 052 783
E-mail: objednavky@czso.cz

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

101st 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