

Dębski, Wiesław; Świdorski, Bartosz; Kurek, Jarosław

Article

Scientific research activity and GDP: An analysis of causality based on 144 countries from around the world

Contemporary Economics

Provided in Cooperation with:

University of Finance and Management, Warsaw

Suggested Citation: Dębski, Wiesław; Świdorski, Bartosz; Kurek, Jarosław (2018) : Scientific research activity and GDP: An analysis of causality based on 144 countries from around the world, Contemporary Economics, ISSN 2300-8814, University of Finance and Management in Warsaw, Faculty of Management and Finance, Warsaw, Vol. 12, Iss. 3, pp. 315-336, <https://doi.org/10.5709/ce.1897-9254.280>

This Version is available at:

<https://hdl.handle.net/10419/297504>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Primary submission: 14.06.2017 | Final acceptance: 19.12.2017

Scientific research activity and GDP. An analysis of causality based on 144 countries from around the world

Wiesław Dębski¹, Bartosz Świdorski², Jarosław Kurek²

ABSTRACT

Using the Granger methodology, this paper presents the causal relationship between scientific research activity, expressed as the number of significant publications, and gross domestic product (GDP). With causality tests, this relationship is investigated from two points of view: for each individual country (144 were selected) and for each specific academic field (28 were selected). Considering annual data from 1996 to 2012, two hypotheses are tested. The first suggests that scientific research activity in a given country has a significant effect on GDP; the second verifies how much each specific field of scientific research activity affects this growth. Our research confirmed the existence of this relationship for a relatively large number of countries, especially highly developed countries and those with a high potential both in the fields of scientific research activity and in GDP. Moreover, this study identifies the most significant fields of this activity that affect GDP. Additionally, the article includes an empirical study regarding how information related to the number of significant scientific publications influenced the quality of Polish GDP forecasts for 2011-2012.

KEY WORDS:

Granger causality, weighted least squares method, number of significant publications, GDP, forecast quality

JEL Classification: O10, O40

¹ University of Finance and Management in Warsaw - Management and Finance, Poland; ² Warsaw University of Life Sciences - Faculty of Applied Informatics and Mathematics, Department of Applied Informatics, Poland

Introduction

GDP (gross domestic product) has been used in economic research for centuries and for various types of analyses, e.g., at the GDP level in economic development analysis in order to analyze the quality of life or wealth of society, at the GDP per capita level, and at the GDP growth rate level in economic cycle analysis.

In macroeconomic analyses, especially in economic

growth models or econometric models of national economies, the GDP level is used most frequently. Researchers have always been interested in the cause-effect relationship between economic growth and the factors from which it results. The factors most commonly used in models describing how a country's economic growth is determined include 1) capital, which is usually measured as the value of fixed assets or the value of productive fixed assets, 2) employment, which is measured as the number of employed people or the amount of time worked, and 3) technological—or more generally, technological and organizational—progress. The problems of specifying and measuring

Correspondence concerning this article should be addressed to: **Wiesław Dębski**, University of Finance and Management in Warsaw - Management and Finance, 55 Pawia Str., Warsaw 01-030, Poland. E-mail: wdebski@vizja.pl

these relationships have been considered in detail when developing and estimating econometric models of national economies (econometric macro-models). An extensive overview of these issues can be found in (Ashley, 1988; Bodkin, Klein, & Marwah, 1991; Welfe, 2013). The last factor, i.e., taking into account technological developments in research showing the development of GDP or production, was first considered by Solow (1957; 1962) in detail. In econometric models of national economies, the most difficult issue regarding this factor was always choosing the appropriate measure of technological development. In the most basic method, this development was expressed by the symptomatic variable t (e.g., the simple Cobb-Douglas function in which the set of exogenous variables was enlarged by the time variable t), the value of the newest machines or devices, or the expenditure incurred to acquire them. In a somewhat later period (the 1980s and 1990s) this factor was also expressed as expenditure on research and development (R&D). Examples of such work include (Coe & Helpman, 1995; Eckstein, 1984). After the year 2000, this topic generated widespread discussion; participants included (Engelbrecht, 2002; Keller, 2004; Saggi, 2002; Welfe, 2004). There were also attempts to decompose this factor into tangible development (resulting from using the newest machines and devices) and intangible development (resulting from raising the qualifications of the employed, measured as the number of graduates from technical universities, the number of patents and licenses granted, expenditure on research and higher education, for example). Works by (Lee, 2005; Welfe, 2006; Xu & Wang, 1999; 2000) can be cited as examples of such attempts. Subsequently, intellectual—or even more widely, human—capital was taken into account when determining GDP. Nonetheless, the greatest problem encountered was selecting the appropriate measure of this intellectual capital. Real or cumulated expenditure on academic research activity was often used in empirical studies (Borensztein, De Gregorio, & Lee, 1998; Cincera & Van Pottelsberghe de la Potterie, 2001; Van Pottelsberghe de la Potterie & Lichtenberg, 2001; Welfe, 2006; 2009). Thus, as time passed, explaining the cause-effect relationship between the economic growth and scientific research activity of a given country, which illustrates human capital, became increasingly important in economic models, especially

econometric models of national economies describing the development of economic growth.

This fact is connected with the primary aim of the following article, which is focused on researching the degree of causality of academic research activity in a country to determine its economic growth, a phenomenon of interest to many economists. As mentioned above, the issue of defining the influence of scientific research activity on economic growth has been the subject of much investigation; however, no study has included as wide a range of countries as this paper. The main issue that arises here is to find a “good” measure of scientific R&D activity. Because our research includes cross-sectional time data and covers 144 countries, an important issue to be addressed was the wide availability of data (a statistical database for all the countries covered) and its comparability. Taking into account these two criteria, we decided to use the number of documents or the number of citable documents referring to a given country as a measure, which can be considered a symptomatic variable of human capital or technical-organizational development in a narrower sense, or possibly non-material technological development, which influences economic growth. The main aim of the article is not to elaborate a specific econometric model describing a given country’s GDP but to verify if scientific research activity significantly, in the Granger sense (Engle & Granger, 1987), results in a country’s economic growth. Another aim of this paper is to perform a comparative study, which results from the cross-sectional and temporal nature of the research. This study compares the significance of causality of individual research fields and for individual countries. Therefore, this paper verifies the main empirical hypothesis, which states that the scientific research activity of a given country, expressed by the number of published or cited academic papers, contributes significantly to determining the GDP of this country. In the second hypothesis, the degree to which chosen academic fields contribute to GDP growth in the given countries is tested.

Due to the availability of statistical data during our investigation, the following two measures have been adopted as the indicator of a given country’s scientific research activity (“IoSA”):

- 1) Documents – the number of documents published during the selected year, usually called the country’s scientific output.

Table 1. The result of correlation coefficients (Pearson's, Spearman's, Kendall's)

	rho (lag(PKB,-1), lag(IoSA,1))		
	Pearson	Spearman	Kendall
Documents	0.9345	0.8759	0.7096
Citable documents	0.9348	0.8766	0.7105

2) Citable documents – the number of citable documents for the selected year. Only articles, reviews and conference papers are considered.

Because two alternative indicators of academic research activity were adopted, their cause-effect relationship with GDP in the Granger sense was investigated. The tests were carried out for various orders of lag between these indicators and GDP. Statistical data (in an annual framework) related to the indicators for academic research activity were drawn from www.scimagojr.com. The data concern various fields of publication and cover the 1996-2012 period. The information obtained includes the number of documents and citable documents by country and by academic field (28 different fields). Additionally, the data provide information on publications for each given country spanning all academic fields in the form of a summary “all” field.

Information regarding GDP in current prices originates from the World Bank service (data.worldbank.org). These data were recalculated to 2011 prices in order to ensure comparability (the 2011 was taken as the base year). Because such data were incomplete, the research takes into consideration only those countries that had complete indicator numbers for scientific research activity, as well as GDP spanning the years 1996-2012. This sample was the longest possible sample that could be achieved, and 144 countries fulfilled the above conditions.

To begin, we decided to check the degree of general correlation between scientific research activity and GDP in the countries being investigated by combining two vectors [lag(PKB,-1),lag(IoSA)] in one matrix with two columns, where lag(PKB,-1) indicates the GDP observation vector for a given country from the second to the last observation, and lag (IoSA,1) indicates the IoSA observation vector from the first to the penultimate observation. These vectors were then

placed into consecutive two-column matrices—one for each country—resulting in a matrix with the dimensions of 144 x (2012-1996) = 2304 x 2). The correlation coefficients (Pearson's, Spearman's, Kendall's) between the first and second columns were then calculated. These coefficients are presented in table 1.

The values presented in the above table clearly show a high degree of correlation between the adopted scientific research activity indicators and GDP in the countries covered by the investigation. An analysis of the signs of the obtained correlation coefficients between lag(PKB,-1) and lag(IoSA,1) for each country individually shows that the correlations according to Pearson, Spearman and Kendall are positive in over 80% of these cases.

1. Granger causality test

As shown in the article, the study analyzes causality in the Granger sense, where the cause is defined as scientific research activity (expressed as two indicators that signify the number of publications) in a given country, and the effect is defined as that country's GDP. The usefulness of this activity is also investigated in terms of its influence on the quality of Polish GDP forecasts complied using a linear model for the years 2011-2012.

In the primary view of examining causality, the validity of following equation is considered

$$x_t = c_1 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + \mu_t \tag{1}$$

against its alternative formulation

$$x_t = c_0 + \gamma_1 x_{t-1} + \gamma_2 x_{t-2} + \dots + \gamma_p x_{t-p} + e_t \tag{2}$$

estimated using the ordinary least squares (OLS) method.

Table 2. The rejection frequency for tests (1) and (2) in relation to the lag order p

p	RF			
	Test 1		Test 2	
	Documents	Citable documents	Documents	Citable documents
1	29.86%	29.86%	34.03%	34.72%
2	34.72%	34.03%	51.39%	52.08%
3	42.36%	43.75%	71.53%	69.44%
4	46.53%	47.22%	84.72%	84.72%
5	81.94%	83.33%	97.92%	97.22%

Examining causality is based on the general statement that variable y is the cause in relation to variable x if past values of x help explain future values y more accurately (Granger, 1969). Therefore, the following hypothesis is tested:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad (3)$$

There are many varieties of Granger causality tests (Ashley, Granger, & Schmalensee, 1980; Amblard, 2012; Berndt, 1991; Can-Zhong & Qing-Wen, 2017; Ding, Chen, & Bressler, 2006; Geweke, 1982; Geweke, Meese, & Dent, 1983; Sims, 1972; Zhou & Sornette, 2006). In this paper, two are applied, and both are considered to be among the most popular varieties (Hamilton, 1994; Sims 1980). The statistics S_1 for the first test and S_2 for the second test are as follows (the results from these tests are denoted as *test 1* and *test 2*, respectively):

$$S_1 \equiv \frac{\frac{RSS_0 - RSS_1}{p}}{\frac{RSS_1}{T - 2p - 1}} \quad (4)$$

$$S_2 \equiv \frac{T(RSS_0 - RSS_1)}{RSS_1} \quad (5)$$

where

$$RSS_1 = \sum_{t=1}^T \hat{\mu}_t^2 \quad (6)$$

$$RSS_0 = \sum_{t=1}^T e_t^2 \quad (7)$$

It is assumed that statistics S_1 and S_2 have the corresponding distributions: $F(p, T-2p-1)$ and $\chi^2(p)$, respectively, where T is the number of observations. In this study, a 5% significance level is applied as standard.

The following part of the paper analyzes the hypothesis rejection frequency (RF) H_0 at a significance level of 0.05:

$$RF = \frac{1}{K} \sum_{i=1}^K h_i \quad (8)$$

where:

$$h_i = \begin{cases} 1 & \text{if } p\text{-value}_i < 0.05 \\ 0 & \text{if } p\text{-value}_i \geq 0.05 \end{cases}$$

$p\text{-value}_i$ – denotes the significance level of the Granger test for a given country,

k – number of countries (for all research areas = 144).

The table 2 shows the rejection frequency for tests (1) and (2) in relation to the lag order p .

The value of 42.36% (item [3,1]) indicates that for the lag order of 3 (years), out of 144 countries researched, the hypothesis was rejected in 42.36% of cases (at a significance level of 5%). It should be noted that assuming full (not asymptotic) congruity of tests, a rejection of hypothesis H_0 should be observed at

Table 3. The differences between RF and *RF rand*

Lag	RF-RF _{random}			
	Test 1		Test 2	
	Documents	Citable documents	Documents	Citable documents
1	23.92%	23.93%	22.99%	23.72%
2	26.03%	25.37%	28.86%	29.57%
3	27.89%	29.25%	28.76%	26.65%
4	19.85%	20.62%	14.37%	14.47%
5	13.30%	14.68%	2.23%	1.52%

level of approximately 5% of cases for random indicators of scientific research activity.

2. Comparing causality tests with the random permutation version

Keeping in mind the asymptotic properties of the test adopted for this study and the fact that both of the above statistics have a tendency for an easier rejection of hypothesis H_0 as the number of observations decreases or by increasing the order p , we decided to compare the test results with their alternative versions. Specifically, the results obtained from the statistics were compared with their versions for a random arrangement of the IoSA vector, i.e., the Granger test was repeated for every country analogically to the tests described above, with the one difference that instead of the original citation vector (IoSA), its random permutation was applied. The idea for this approach arises from the observation that if past scientific research activity remains unrelated to future values of GDP, one may expect a similar result for the Granger causality test for randomly selected values of the IoSA vector in the statistics of the applied test. However, to maintain the same IoSA distribution in the simulation, a permuted IoSA vector (in relation to the original) was applied.

This procedure was repeated 10,000 times. As a result, 10,000 different indicators of RF_rand_i rejection were obtained. The results were averaged to obtain one value of RF_rand for a given country. The differences

between RF and RF_rand are presented in the table below. A positive difference between RF and RF_rand indicates a more frequent rejection of hypothesis H_0 for real IoSA values than for their random permutations (carried out independently for each of the countries studied).

For both of the above tests, higher values of $RF-RF_{random}$ in the table 3 are an indication in favor of hypothesis H_1 (the hypothesis of a lack of causality in comparison to randomly permuted IoSA values is more frequently rejected).

Interpreting the obtained results, one can state that the value of 27.89% (item [3.1]) suggests that if randomly permuted Documents values are applied, then the indicator for the rejection of hypothesis H_0 would nominally be 27.89 percentage points lower than the same indicator obtained from non-random (i.e., original) Documents values. The indicators shown in the above table imply that scientific research activity, measured as the amount of published articles, results in an economic effect after approximately 2-3 years (i.e., considering lag orders from 1 to 4, the relatively highest values are lag orders 2 and 3).

3. Causality in individual scientific fields

The research presented so far tested the causality of scientific research activity, measured as the general number of publications in all fields, in determining the GDP for all countries. Let us study this causality

Table 4. The results of rejection frequency of the zero hypothesis

test	RF								RF - RFrandom								NoC
	Documents				Citable documents				Documents				Citable documents				
	Lag 2		Lag 3		Lag 2		Lag 3		Lag 2		Lag 3		Lag 2		Lag 3		
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	
all	0.35	0.51	0.42	0.72	0.34	0.52	0.44	0.69	0.26	0.29	0.28	0.29	0.25	0.30	0.29	0.27	144
Agricultural and Biological Sciences	0.28	0.43	0.31	0.60	0.26	0.44	0.32	0.59	0.19	0.20	0.16	0.17	0.17	0.22	0.18	0.16	124
Arts and Humanities	0.10	0.29	0.15	0.46	0.13	0.33	0.21	0.42	0.01	0.06	-0.01	0.03	0.03	0.10	0.06	-0.02	48
Biochemistry, Genetics and Molecular Biology	0.30	0.46	0.37	0.69	0.30	0.45	0.38	0.68	0.21	0.23	0.23	0.26	0.21	0.22	0.24	0.25	105
Business, Management and Accounting	0.32	0.47	0.43	0.70	0.30	0.45	0.42	0.66	0.23	0.24	0.29	0.27	0.21	0.22	0.27	0.23	53
Chemical Engineering	0.19	0.44	0.30	0.59	0.22	0.44	0.29	0.56	0.11	0.21	0.16	0.16	0.13	0.21	0.14	0.13	73
Chemistry	0.24	0.48	0.32	0.65	0.22	0.49	0.32	0.66	0.15	0.25	0.17	0.23	0.13	0.26	0.17	0.24	92
Computer Science	0.27	0.52	0.32	0.62	0.27	0.52	0.32	0.62	0.18	0.29	0.18	0.19	0.18	0.29	0.18	0.19	71
Decision Sciences	0.16	0.41	0.22	0.55	0.14	0.41	0.24	0.51	0.07	0.19	0.07	0.12	0.05	0.19	0.09	0.08	51
Dentistry	0.15	0.37	0.34	0.63	0.15	0.34	0.34	0.59	0.06	0.14	0.20	0.21	0.06	0.12	0.20	0.16	41
Earth and Planetary Sciences	0.29	0.44	0.30	0.58	0.26	0.45	0.30	0.57	0.20	0.21	0.16	0.15	0.18	0.22	0.16	0.15	103
Economics, Econometrics and Finance	0.35	0.41	0.29	0.63	0.35	0.41	0.31	0.59	0.26	0.18	0.15	0.20	0.26	0.18	0.17	0.16	51
Energy	0.16	0.38	0.28	0.55	0.17	0.38	0.28	0.55	0.07	0.15	0.13	0.12	0.09	0.15	0.13	0.12	69
Engineering	0.18	0.42	0.25	0.53	0.18	0.40	0.26	0.52	0.09	0.19	0.10	0.10	0.09	0.18	0.11	0.09	89
Environmental Science	0.25	0.51	0.35	0.63	0.25	0.48	0.34	0.63	0.16	0.28	0.20	0.21	0.16	0.26	0.19	0.20	104
Health Professions	0.19	0.38	0.21	0.57	0.14	0.38	0.19	0.60	0.10	0.15	0.07	0.14	0.06	0.15	0.05	0.17	42
Immunology and Microbiology	0.17	0.37	0.28	0.57	0.17	0.36	0.26	0.56	0.08	0.14	0.13	0.14	0.09	0.13	0.11	0.13	109
Materials Science	0.21	0.38	0.22	0.54	0.21	0.38	0.23	0.56	0.12	0.16	0.08	0.12	0.12	0.16	0.09	0.13	81
Mathematics	0.29	0.49	0.33	0.67	0.27	0.48	0.34	0.67	0.20	0.27	0.18	0.24	0.18	0.26	0.20	0.24	79
Medicine	0.23	0.38	0.39	0.62	0.23	0.38	0.37	0.67	0.15	0.15	0.25	0.19	0.14	0.15	0.22	0.24	128
Multidisciplinary	0.18	0.33	0.20	0.63	0.20	0.31	0.22	0.53	0.09	0.10	0.06	0.20	0.11	0.08	0.08	0.10	49
Neuroscience	0.21	0.43	0.27	0.46	0.23	0.48	0.25	0.52	0.13	0.20	0.12	0.04	0.15	0.26	0.10	0.09	56
Nursing	0.39	0.47	0.33	0.64	0.33	0.44	0.36	0.58	0.30	0.25	0.19	0.21	0.25	0.22	0.21	0.15	36
Pharmacology, Toxicology and Pharmaceutics	0.20	0.33	0.24	0.53	0.19	0.28	0.25	0.49	0.12	0.10	0.10	0.10	0.10	0.05	0.11	0.07	79
Physics and Astronomy	0.20	0.45	0.30	0.61	0.20	0.46	0.31	0.60	0.11	0.23	0.15	0.18	0.11	0.24	0.16	0.17	84
Psychology	0.26	0.47	0.25	0.45	0.28	0.47	0.23	0.47	0.17	0.24	0.09	0.03	0.19	0.25	0.08	0.05	53
Social Sciences	0.30	0.54	0.32	0.65	0.29	0.47	0.27	0.62	0.21	0.31	0.18	0.22	0.20	0.25	0.12	0.20	93
Veterinary	0.22	0.54	0.25	0.58	0.24	0.51	0.24	0.61	0.13	0.32	0.10	0.16	0.15	0.29	0.09	0.18	72

Table 5. Sum of ranks (in descending order)

1	All
2	Biochemistry, Genetics and Molecular Biology
3	<i>Business, Management and Accounting</i>
4	Mathematics

with respect to particular scientific fields. The table 4 shows the results (rejection frequency of the zero hypothesis) for these fields based on the previously mentioned tests, with the lag orders 2 and 3 for the original base data (RF). Differences compared to the results obtained for randomly permuted IoSA vectors (RF – RFRandom) are also listed.

Additionally, the table includes the indicator NoC (number of countries), defined as the number of countries for which combined data were available, i.e., both regarding GDP and the number of publications for each country for the 1996-2012 period. Thus, in further calculations, only those countries that had complete data were taken into account. Because some observations were missing from the database, different scientific fields of study had data from different numbers of countries. For example, data from the field of chemistry was available for only 92 countries (which could also be the result of natural inactivity of some countries in this field). A smaller number of countries results in a smaller number of observations used in the calculations. It also often leads to a change in the structure among the researched countries. Ultimately, it has a negative impact on the representativeness of the results. In the table below, where the number of countries used for calculations was less than half, i.e., $144 / 2 = 77$ (144 is the number of countries with full data used to study the relationship between GDP and the number of publications in any and all fields – all), the results are shown in italics.

The interpretation of the results shown in the table above is the same as for the previous study, e.g., for item [5, 3] the value 0.43 indicates that for the lag order of 3 (years) in the field of business, management and accounting, for 53 countries studied, hypothesis H_0 was rejected in 43% of cases (at the significance level

of 5% for test 1, testing Documents). In turn, item [5, 11] with a value of 0.29 means that if, in the same test, randomly permuted IoSA values are used, the result would be 29 percentage points lower; therefore, the rejection frequency for hypothesis H_0 would equal $43\% - 29\% = 14\%$. It should be emphasized that the above results are aggregate estimates, and the real effect of scientific research activity on economic growth may vary depending on the specifics of a given country.

If the sum of ranks (in descending order, meaning that the largest value would be given the rank of 1) is calculated for each column (except the last, which lists the indicators of the number of countries) and subsequently sorted in descending order, the following fields appear in positions 1 to 4:

The above scientific fields (shown in the table 5) can be seen (from the point of view of the approach applied above) as the areas that are most conducive to economic growth.

4. Testing Granger causality for specific countries

The table 6 lists significance levels (p-value) for hypothesis H_0 , verified for the total number of publications (divided into Documents and Citable documents) in all areas of scientific research activity (all) with reference to specific countries with lag orders 2 and 3. Lower values reflect easier rejection of hypothesis H_0 in favor of an alternative hypothesis, and they are associated with a more significant causal relationship between the past number of publications in a given country and its subsequent GDP, i.e., that the number of publications from two years earlier influences the current level of GDP to a greater degree.

The rejection frequencies (RF) for hypothesis H_0 (lack of causality) for particular groups of countries

Table 6. Significance levels (p-value) for hypothesis H_0

Test	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Albania	0.4815	0.3325	0.0158	0.0000	0.4529	0.3013	0.0205	0.0000
Algeria	0.4543	0.3028	0.0087	0.0000	0.4502	0.2984	0.0075	0.0000
Angola	0.4725	0.3226	0.0012	0.0000	0.5051	0.3589	0.0021	0.0000
Armenia	0.2154	0.0839	0.2662	0.0501	0.2486	0.1087	0.3399	0.0926
Australia	0.0010	0.0000	0.0057	0.0000	0.0007	0.0000	0.0047	0.0000
Austria	0.4351	0.2824	0.4448	0.1758	0.4061	0.2523	0.4319	0.1642
Azerbaijan	0.0063	0.0000	0.1093	0.0040	0.0077	0.0000	0.1239	0.0059
Bahamas	0.8299	0.7647	0.2392	0.0379	0.8853	0.8400	0.0836	0.0017
Bangladesh	0.0029	0.0000	0.0100	0.0000	0.0030	0.0000	0.0057	0.0000
Belarus	0.1455	0.0400	0.3553	0.1032	0.0998	0.0187	0.5857	0.3269
Belgium	0.4937	0.3460	0.5059	0.2361	0.4691	0.3189	0.5193	0.2503
Benin	0.9876	0.9825	0.1333	0.0073	0.9955	0.9936	0.1476	0.0099
Bhutan	0.0264	0.0008	0.1539	0.0112	0.0263	0.0008	0.1371	0.0080
Bolivia	0.0173	0.0003	0.0065	0.0000	0.0169	0.0003	0.0069	0.0000
Bosnia and Herzegovina	0.6256	0.5011	0.0291	0.0000	0.6334	0.5106	0.0297	0.0000
Botswana	0.5645	0.4274	0.8419	0.7045	0.5484	0.4086	0.8172	0.6627
Brazil	0.0868	0.0139	0.0016	0.0000	0.0879	0.0143	0.0020	0.0000
Brunei Darussalam	0.2576	0.1158	0.9196	0.8434	0.3771	0.2230	0.5940	0.3372
Bulgaria	0.2153	0.0838	0.5342	0.2667	0.1986	0.0722	0.5057	0.2359
Burkina Faso	0.8408	0.7793	0.6144	0.3628	0.9048	0.8668	0.4923	0.2219
Burundi	0.6766	0.5645	0.7845	0.6091	0.7108	0.6080	0.7852	0.6102
Cambodia	0.0051	0.0000	0.2228	0.0314	0.0041	0.0000	0.2450	0.0404
Cameroon	0.0821	0.0124	0.0101	0.0000	0.0748	0.0101	0.0142	0.0000
Canada	0.0018	0.0000	0.0028	0.0000	0.0031	0.0000	0.0036	0.0000
Central African Republic	0.9607	0.9446	0.9231	0.8498	0.9861	0.9804	0.7952	0.6265
Chad	0.1446	0.0395	0.0211	0.0000	0.1368	0.0354	0.0205	0.0000
Chile	0.0061	0.0000	0.0244	0.0000	0.0063	0.0000	0.0236	0.0000
China	0.0105	0.0001	0.0217	0.0000	0.0096	0.0000	0.0212	0.0000
Colombia	0.0045	0.0000	0.0170	0.0000	0.0042	0.0000	0.0132	0.0000

Table 6. Significance levels (p-value) for hypothesis H_0 (Continued)

Test	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Congo	0.6914	0.5832	0.6277	0.3801	0.5045	0.3582	0.6860	0.4597
Costa Rica	0.0240	0.0007	0.0087	0.0000	0.0339	0.0016	0.0099	0.0000
Croatia	0.9053	0.8676	0.9617	0.9236	0.8985	0.8581	0.9599	0.9200
Cyprus	0.5936	0.4621	0.0517	0.0003	0.5695	0.4334	0.0552	0.0004
Czech Republic	0.3704	0.2165	0.1049	0.0035	0.2802	0.1343	0.1043	0.0035
Denmark	0.6608	0.5446	0.5248	0.2564	0.5983	0.4678	0.3370	0.0907
Dominican Republic	0.6528	0.5346	0.6470	0.4057	0.5904	0.4583	0.6082	0.3549
Ecuador	0.7544	0.6645	0.6149	0.3635	0.6831	0.5727	0.5200	0.2511
Egypt	0.6333	0.5104	0.8447	0.7093	0.6271	0.5029	0.8468	0.7129
El Salvador	0.1801	0.0601	0.6813	0.4530	0.1701	0.0540	0.6521	0.4125
Eritrea	0.4264	0.2732	0.1156	0.0048	0.3892	0.2351	0.0838	0.0017
Estonia	0.1343	0.0341	0.3525	0.1012	0.1055	0.0210	0.3598	0.1064
Ethiopia	0.0048	0.0000	0.0190	0.0000	0.0043	0.0000	0.0154	0.0000
Fiji	0.0032	0.0000	0.0483	0.0003	0.0037	0.0000	0.0499	0.0003
Finland	0.2119	0.0814	0.2310	0.0346	0.1932	0.0686	0.2218	0.0310
France	0.2810	0.1350	0.1441	0.0092	0.2249	0.0907	0.1453	0.0095
Gabon	0.1139	0.0245	0.0329	0.0001	0.1466	0.0405	0.0845	0.0018
Gambia	0.9943	0.9919	0.6282	0.3807	0.8893	0.8454	0.2251	0.0322
Georgia	0.0032	0.0000	0.0035	0.0000	0.0031	0.0000	0.0031	0.0000
Germany	0.1397	0.0369	0.2031	0.0244	0.1418	0.0380	0.2139	0.0281
Ghana	0.9575	0.9401	0.7421	0.5426	0.9212	0.8896	0.7445	0.5462
Greece	0.2443	0.1054	0.0017	0.0000	0.1844	0.0629	0.0015	0.0000
Grenada	0.1612	0.0487	0.1875	0.0196	0.1238	0.0290	0.1831	0.0183
Guatemala	0.0074	0.0000	0.0147	0.0000	0.0101	0.0001	0.0200	0.0000
Guinea	0.4739	0.3241	0.6871	0.4613	0.4749	0.3252	0.6949	0.4724
Guinea-Bissau	0.8152	0.7449	0.8653	0.7452	0.9208	0.8889	0.8760	0.7641
Guyana	0.1445	0.0394	0.0957	0.0026	0.0844	0.0131	0.0481	0.0002
Honduras	0.0958	0.0171	0.0067	0.0000	0.1229	0.0286	0.0071	0.0000
Hong Kong	0.0042	0.0000	0.0020	0.0000	0.0057	0.0000	0.0027	0.0000

Table 6. Significance levels (p-value) for hypothesis H_0 (Continued)

Test	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Hungary	0.3506	0.1974	0.5126	0.2432	0.4107	0.2569	0.5622	0.2988
Iceland	0.7437	0.6506	0.3235	0.0820	0.7495	0.6581	0.3851	0.1253
India	0.8963	0.8551	0.1707	0.0151	0.7895	0.7108	0.0980	0.0028
Indonesia	0.0000	0.0000	0.3494	0.0991	0.0000	0.0000	0.3443	0.0956
Italy	0.1411	0.0376	0.0024	0.0000	0.1062	0.0213	0.0025	0.0000
Japan	0.2131	0.0822	0.7568	0.5653	0.2110	0.0807	0.7876	0.6141
Jordan	0.7967	0.7203	0.2561	0.0453	0.8206	0.7522	0.3858	0.1259
Kazakhstan	0.0276	0.0009	0.0368	0.0001	0.0281	0.0010	0.0313	0.0000
Kenya	0.3866	0.2325	0.0926	0.0024	0.4045	0.2506	0.1929	0.0212
Kyrgyzstan	0.1264	0.0302	0.0094	0.0000	0.1171	0.0259	0.0091	0.0000
Laos	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0003	0.0000
Latvia	0.1141	0.0246	0.1299	0.0068	0.1315	0.0327	0.1364	0.0078
Lebanon	0.0162	0.0002	0.0064	0.0000	0.0160	0.0002	0.0023	0.0000
Lesotho	0.0657	0.0076	0.1967	0.0224	0.0872	0.0140	0.2161	0.0289
Lithuania	0.0616	0.0066	0.0077	0.0000	0.0621	0.0067	0.0083	0.0000
Luxembourg	0.8931	0.8507	0.0535	0.0004	0.8873	0.8427	0.0678	0.0008
Macedonia	0.0334	0.0015	0.0141	0.0000	0.0415	0.0026	0.0195	0.0000
Madagascar	0.7443	0.6513	0.6766	0.4465	0.7991	0.7235	0.7522	0.5582
Malawi	0.6713	0.5578	0.6068	0.3532	0.6428	0.5222	0.7835	0.6076
Malaysia	0.1652	0.0510	0.0198	0.0000	0.1572	0.0464	0.0256	0.0000
Mali	0.4313	0.2783	0.0071	0.0000	0.5002	0.3533	0.0047	0.0000
Malta	0.8098	0.7378	0.0787	0.0014	0.8162	0.7463	0.0615	0.0006
Mauritania	0.0205	0.0004	0.2158	0.0288	0.0189	0.0003	0.2190	0.0299
Mauritius	0.0333	0.0015	0.0503	0.0003	0.0206	0.0004	0.0357	0.0001
Mexico	0.1487	0.0417	0.0224	0.0000	0.1592	0.0476	0.0189	0.0000
Moldova	0.0585	0.0058	0.0005	0.0000	0.0562	0.0053	0.0004	0.0000
Mongolia	0.0396	0.0023	0.0003	0.0000	0.0495	0.0040	0.0004	0.0000
Morocco	0.0157	0.0002	0.0070	0.0000	0.0082	0.0000	0.0030	0.0000
Mozambique	0.1345	0.0342	0.1283	0.0065	0.1537	0.0445	0.1281	0.0065

Table 6. Significance levels (p-value) for hypothesis H_0 (Continued)

Test	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Namibia	0.0477	0.0037	0.0938	0.0025	0.0515	0.0044	0.1096	0.0040
Nepal	0.0001	0.0000	0.0002	0.0000	0.0002	0.0000	0.0004	0.0000
Netherlands	0.2472	0.1076	0.1353	0.0077	0.2068	0.0778	0.1361	0.0078
Nicaragua	0.6506	0.5319	0.0341	0.0001	0.6130	0.4857	0.0550	0.0004
Niger	0.9509	0.9308	0.0850	0.0018	0.9349	0.9085	0.0837	0.0017
Nigeria	0.0019	0.0000	0.0021	0.0000	0.0017	0.0000	0.0018	0.0000
Norway	0.0045	0.0000	0.0050	0.0000	0.0090	0.0000	0.0075	0.0000
Pakistan	0.1993	0.0727	0.0226	0.0000	0.2100	0.0801	0.0220	0.0000
Panama	0.6655	0.5505	0.0993	0.0030	0.7233	0.6241	0.1004	0.0031
Papua New Guinea	0.5914	0.4594	0.8085	0.6483	0.6533	0.5352	0.8529	0.7236
Paraguay	0.0335	0.0015	0.1024	0.0033	0.0230	0.0006	0.0691	0.0009
Peru	0.0000	0.0000	0.0002	0.0000	0.0001	0.0000	0.0002	0.0000
Philippines	0.0000	0.0000	0.0066	0.0000	0.0000	0.0000	0.0067	0.0000
Poland	0.0021	0.0000	0.0000	0.0000	0.0017	0.0000	0.0000	0.0000
Portugal	0.9835	0.9766	0.4365	0.1683	0.9937	0.9910	0.3862	0.1262
Puerto Rico	0.0400	0.0024	0.1999	0.0234	0.0371	0.0020	0.1867	0.0193
Romania	0.5174	0.3728	0.9791	0.9578	0.5161	0.3714	0.9787	0.9570
Russian Federation	0.0357	0.0018	0.0001	0.0000	0.0435	0.0029	0.0001	0.0000
Rwanda	0.0119	0.0001	0.0293	0.0000	0.0268	0.0009	0.0548	0.0004
Saint Kitts and Nevis	0.3914	0.2373	0.4761	0.2057	0.4459	0.2938	0.3443	0.0956
Samoa	0.0702	0.0088	0.0232	0.0000	0.0835	0.0128	0.0324	0.0001
Senegal	0.4988	0.3518	0.5000	0.2299	0.4321	0.2792	0.4210	0.1547
Seychelles	0.0007	0.0000	0.0092	0.0000	0.0005	0.0000	0.0064	0.0000
Sierra Leone	0.0290	0.0011	0.0508	0.0003	0.0378	0.0021	0.0423	0.0002
Singapore	0.0009	0.0000	0.0029	0.0000	0.0010	0.0000	0.0035	0.0000
Slovakia	0.7077	0.6040	0.1328	0.0072	0.6673	0.5528	0.2027	0.0243
Slovenia	0.4949	0.3474	0.6142	0.3626	0.4830	0.3342	0.5976	0.3416
Solomon Islands	0.0114	0.0001	0.0112	0.0000	0.0235	0.0006	0.0053	0.0000
South Africa	0.0372	0.0020	0.0001	0.0000	0.0290	0.0011	0.0001	0.0000

Table 6. Significance levels (p-value) for hypothesis H_0 (Continued)

Test	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
South Korea	0.0137	0.0001	0.0128	0.0000	0.0132	0.0001	0.0195	0.0000
Spain	0.5469	0.4068	0.0728	0.0011	0.4641	0.3135	0.0266	0.0000
Sri Lanka	0.0996	0.0186	0.1342	0.0075	0.0952	0.0169	0.1263	0.0062
Sudan	0.4591	0.3080	0.3332	0.0882	0.4594	0.3083	0.3501	0.0996
Suriname	0.8657	0.8132	0.0321	0.0001	0.8491	0.7906	0.0527	0.0003
Swaziland	0.2651	0.1218	0.2927	0.0638	0.2648	0.1216	0.3397	0.0925
Sweden	0.0065	0.0000	0.0521	0.0003	0.0042	0.0000	0.0341	0.0001
Switzerland	0.0503	0.0041	0.0204	0.0000	0.0720	0.0093	0.0279	0.0000
Tajikistan	0.4754	0.3258	0.4768	0.2063	0.4781	0.3288	0.4816	0.2111
Tanzania	0.0321	0.0014	0.0277	0.0000	0.0311	0.0013	0.0265	0.0000
Thailand	0.0001	0.0000	0.0009	0.0000	0.0001	0.0000	0.0012	0.0000
Togo	0.6341	0.5114	0.0609	0.0006	0.4985	0.3514	0.0699	0.0009
Trinidad and Tobago	0.0265	0.0009	0.0133	0.0000	0.0249	0.0007	0.0246	0.0000
Tunisia	0.9833	0.9764	0.0475	0.0002	0.6559	0.5385	0.0099	0.0000
Turkey	0.0012	0.0000	0.0069	0.0000	0.0014	0.0000	0.0075	0.0000
Turkmenistan	0.4661	0.3157	0.0701	0.0009	0.4233	0.2700	0.0477	0.0002
Uganda	0.0322	0.0014	0.0665	0.0008	0.0461	0.0034	0.0655	0.0007
Ukraine	0.4539	0.3024	0.4993	0.2292	0.4920	0.3442	0.5767	0.3160
United Kingdom	0.0168	0.0003	0.0081	0.0000	0.0325	0.0014	0.0121	0.0000
United States	0.0181	0.0003	0.0048	0.0000	0.0315	0.0013	0.0104	0.0000
Uruguay	0.6434	0.5229	0.2647	0.0494	0.6357	0.5134	0.3187	0.0790
Uzbekistan	0.4317	0.2787	0.4330	0.1652	0.4196	0.2661	0.4206	0.1543
Vanuatu	0.4848	0.3362	0.7913	0.6202	0.5873	0.4545	0.8273	0.6796
Venezuela	0.7894	0.7106	0.9794	0.9584	0.7805	0.6988	0.9800	0.9595
Vietnam	0.1779	0.0587	0.2655	0.0498	0.1792	0.0596	0.2668	0.0504
Yemen	0.1335	0.0337	0.2521	0.0435	0.1136	0.0244	0.1964	0.0223
Zambia	0.0921	0.0158	0.2276	0.0332	0.1037	0.0202	0.3733	0.1163
Zimbabwe	0.1035	0.0201	0.1250	0.0060	0.1343	0.0341	0.1609	0.0127

Table 7. The rejection frequencies (RF) for hypothesis H_0 (lack of causality) for particular groups of countries and significance level 0.01

Group	RF (p-value = 0.01)								Number of countries
	Documents				Citable documents				
	Lag 2		Lag 3		Lag 2		Lag 3		
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	
All	0.174	0.382	0.236	0.611	0.181	0.368	0.243	0.611	144
G20	0.308	0.692	0.538	0.769	0.385	0.692	0.462	0.846	13
European Union	0.080	0.160	0.200	0.600	0.080	0.160	0.160	0.560	25
Very high human development	0.237	0.368	0.316	0.658	0.237	0.368	0.263	0.632	38
High human development	0.156	0.375	0.250	0.656	0.156	0.375	0.281	0.688	32
Medium human development	0.194	0.472	0.278	0.583	0.222	0.444	0.278	0.639	36
Low human development	0.108	0.297	0.108	0.568	0.108	0.270	0.162	0.514	37
Asia	0.333	0.545	0.333	0.636	0.364	0.545	0.364	0.697	33
Africa	0.070	0.302	0.163	0.581	0.093	0.279	0.186	0.535	43
North America	0.133	0.400	0.267	0.600	0.067	0.400	0.200	0.667	15
South America	0.250	0.500	0.250	0.750	0.250	0.500	0.250	0.750	12
Europe	0.146	0.317	0.244	0.634	0.146	0.317	0.220	0.610	41
Australia and Oceania	0.286	0.571	0.286	0.714	0.286	0.429	0.429	0.714	7

and significance levels 0.01 and 0.05 are listed in Tables 7 and 8, respectively.

Generally, higher values are seen than may have been expected to result from the respective significance levels (1% and 5%). The presence of relatively low values for the European Union is striking (except for two cases in test 2). This situation may be explained by a longer period of relative economic stagnation in which the development of scientific research activity has a lesser impact on GDP. The situation is different in the remaining G20 countries, where this relationship is most visible among the researched countries. When comparing continents, a relatively strong relationship between scientific research and GDP exists in Asia and South and North

America. The cause-effect relationship between scientific research activity and GDP is visibly weaker in Europe than in Australia and Oceania. This relation is especially evident in the USA, United Kingdom, Canada, Australia, and Sweden, as well as Far Eastern countries such as China (including Hong Kong), South Korea, the Philippines, Thailand, Singapore, Russia, Turkey and Poland. The table also shows the weakest relationship in the 37 countries that constitute the “low human development” group. Generally, it seems that the scientific activity is not a (significant) driver of economic development in all countries. For example, there are countries where the economy is based mainly on the mining industry or tourism, where science plays a secondary role.

Table 8. The rejection frequencies (RF) for hypothesis H_0 (lack of causality) for particular groups of countries and significance level 0.05

Group	RF (p-value = 0.05)								Number of countries
	Documents				Citable documents				
	Lag 2		Lag 3		Lag 2		Lag 3		
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	
All	0.347	0.514	0.424	0.715	0.340	0.521	0.438	0.694	144
G20	0.692	0.846	0.769	0.846	0.692	0.846	0.769	0.846	13
European Union	0.120	0.280	0.200	0.680	0.120	0.280	0.280	0.680	25
Very high human development	0.316	0.474	0.395	0.711	0.316	0.474	0.447	0.711	38
High human development	0.375	0.531	0.531	0.750	0.375	0.563	0.563	0.719	32
Medium human development	0.417	0.583	0.472	0.722	0.389	0.583	0.444	0.667	36
Low human development	0.270	0.459	0.324	0.676	0.270	0.459	0.324	0.676	37
Asia	0.545	0.636	0.515	0.818	0.545	0.667	0.545	0.758	33
Africa	0.279	0.442	0.326	0.651	0.256	0.442	0.326	0.651	43
North America	0.400	0.600	0.533	0.800	0.400	0.600	0.467	0.800	15
South America	0.500	0.667	0.583	0.833	0.500	0.667	0.583	0.750	12
Europe	0.244	0.439	0.366	0.683	0.244	0.439	0.415	0.683	41
Australia and Oceania	0.429	0.571	0.714	0.714	0.429	0.571	0.714	0.714	7

5. Testing the influence of IoSA on GDP forecasts using the example of Poland

This chapter first examines, based on the previous two tests, the level of Granger causality between the number of publications in given fields of scientific research and GDP for Poland. We chose Poland (our home country) as an example for a detailed analysis. Additionally, we chose one country due to the volume of the article. The study was carried out for lag orders 2 and 3. The table 9 lists the p-values resulting from this testing. Next, the direct influence of information contained in the IoSA vector on the quality of Polish GDP forecasts for 2011-2012 is tested using the adopted linear model (described later in this chapter).

It is clear from the results presented in the above table that the total number of publications (all) resulting from scientific research activity in Poland has a significant effect (p-value less than 0.01) in the Granger sense on Polish GDP growth. This conclusion is drawn based on both scientific research activity indicators used in the tests, as well as both tested lag orders. The same conclusions can be formulated in relation to the following fields of scientific research activities in Poland: biochemistry, genetics and molecular biology, chemistry, immunology and microbiology, medicine, and physics and astronomy. Moreover, test 2 shows that the same conclusion can be drawn for other fields of scientific research. Generalizing these results, it can be claimed that publications related to scientific research associated with highly innovative

Table 9. The result of p-values for lag orders 2 and 3

Area	p-value							
	Documents				Citable documents			
	Lag 2		Lag 3		Lag 2		Lag 3	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
All	0.0021	0	0	0	0.0017	0	0	0
Agricultural and Biological Sciences	0.5423	0.4014	0.7037	0.4853	0.5517	0.4124	0.7117	0.4969
Arts and Humanities	0.9384	0.9134	0.7852	0.6102	0.9288	0.9	0.7231	0.5139
Biochemistry, Genetics and Molecular Biology	0.0051	0	0.0002	0	0.0042	0	0.0001	0
Business, Management and Accounting	0.9409	0.9169	0.9428	0.8871	0.954	0.9352	0.9694	0.9386
Chemical Engineering	0.0558	0.0052	0.1147	0.0047	0.0499	0.004	0.1048	0.0035
Chemistry	0.0011	0	0.0004	0	0.0011	0	0.0005	0
Computer Science	0.0848	0.0132	0.2136	0.028	0.0687	0.0084	0.1822	0.0181
Decision Sciences	0.0587	0.0059	0.0435	0.0002	0.0508	0.0042	0.0371	0.0001
Dentistry	0.83	0.7648	0.058	0.0005	0.825	0.7581	0.0845	0.0018
Earth and Planetary Sciences	0.0085	0	0.028	0	0.0111	0.0001	0.0357	0.0001
Economics, Econometrics and Finance	0.7108	0.608	0.114	0.0046	0.725	0.6263	0.162	0.013
Energy	0.5472	0.4071	0.4486	0.1794	0.5147	0.3698	0.4646	0.1945
Engineering	0.1031	0.02	0.1633	0.0133	0.0879	0.0143	0.1523	0.0109
Environmental Science	0.2178	0.0856	0.0102	0	0.2203	0.0874	0.0126	0
Health Professions	0.0125	0.0001	0.0554	0.0004	0.0157	0.0002	0.0656	0.0008
Immunology and Microbiology	0.0006	0	0.0005	0	0.0004	0	0.0005	0
Materials Science	0.0463	0.0034	0.1078	0.0038	0.0498	0.004	0.1114	0.0043
Mathematics	0.2078	0.0785	0.4117	0.1468	0.2034	0.0755	0.4108	0.1461
Medicine	0.0006	0	0.0006	0	0.0004	0	0.0005	0
Multidisciplinary	0.6011	0.4712	0.3312	0.0869	0.4434	0.2911	0.4998	0.2297
Neuroscience	0.0896	0.0149	0.2501	0.0426	0.1021	0.0196	0.28	0.057
Nursing	0.0489	0.0039	0.1423	0.0089	0.096	0.0172	0.2313	0.0346
Pharmacology, Toxicology and Pharmaceutics	0.7455	0.6529	0.8929	0.7944	0.7583	0.6696	0.935	0.8723
Physics and Astronomy	0.002	0	0.0071	0	0.0019	0	0.0067	0
Psychology	0.0197	0.0004	0.019	0	0.0114	0.0001	0.0138	0
Social Sciences	0.3404	0.1878	0.1178	0.005	0.3073	0.1576	0.1089	0.004
Veterinary	0.0875	0.0141	0.0589	0.0005	0.0859	0.0136	0.0582	0.0005

Note: Scientific fields characterized by all p-values below 0.01 are emphasized in bold.

technological development have the largest impact on GDP growth in Poland (with a lag of 2 and 3 years).

The second part of the study, as stated earlier, comprises an analysis of the influence of information contained in the IoSA vector on the forecast quality of Polish GDP. In other words, an attempt is made to examine how past knowledge contained in publications (Documents, Citable documents) influences the forecast quality of future GDP. A test was carried out for the GDP forecasts for 2011-2012. Such forecasts of Polish GDP were made for these years based the model described below and subsequently checked for quality by calculating the errors, expressed as root mean square error (RMSE) and mean absolute percentage error (MAPE). The forecasts were created by first estimating the model parameters based on data from the IoSA vector for the years 1996-2010, then calculating the GDP forecast for 2011 using the estimated model parameters, and, finally, repeating the procedure for 2012 but with the model estimated based on data for 1996-2011. For both forecast years, the average forecast error was calculated as RMSA and MAPE.

An additional phenomenon was analyzed during the above test. When forecasting Polish GDP for a given year, it is possible to apply a model estimated from data pertaining solely to Poland or one estimated from data pertaining solely to other countries. One may imagine suitably formulated time series in both cases to estimate, e.g., the parameters of an autoregressive model with the appropriate lag order. It is also possible to use data pertaining to all countries except Poland to estimate the model. The first approach (estimating the model parameters using only data pertaining to Poland) corresponds to the case where all the observations are used; however, a non-zero weight (in the weighted OLS method) equal to a certain positive constant is assigned only to observations for Poland, while the remaining observations have weights equal to zero. The other approach, where a GDP model for Poland is estimated using data from 1996-2010/2011 for other countries, matches the case where all the observations are used (pertaining to Poland and other countries); however, the observation weight for Poland would assume the value of zero, and all observations relating to other countries would have a certain non-negative weight. An interesting situation arises if these weights are defined as variables. Such a case is examined below.

It is assumed that observation weights pertaining to Poland in proportion to other countries are related as Θ to $(1 - \Theta)$, where Θ denotes a number in the range $[0,1]$. Therefore, if Θ is equal to 1, only observations pertaining to Poland are used to estimate the parameters of the regression equation. On the other hand, if Θ is equal to zero, only observations from other countries are used to estimate the model. If, in turn, Θ is equal to 0.5, all the observations have the same weight, and if $\Theta = 0.75$, the weight of observations for Poland is 3 times $(0.75 / (1 - 0.75) = 3)$ greater (this corresponds to the situation in which observations pertaining to Poland would be repeated 3 times).

Two alternative model descriptions were used in this test:

$$GDP_t = c_1 + \alpha_1 GDP_{t-1} + \alpha_2 GDP_{t-2} + \dots + \alpha_p GDP_{t-p} + \beta_1 A_{t-1} + \beta_2 A_{t-2} + \dots + \beta_p A_{t-p} + \mu_t \quad (9)$$

and

$$GDP_t = c_0 + \gamma_1 GDP_{t-1} + \gamma_2 GDP_{t-2} + \dots + \gamma_p GDP_{t-p} + e_t \quad (10)$$

where A denotes Documents or Citable documents and p denotes the lag order.

The quality of the GDP forecasts for 2011 and 2012 was tested by applying random terms μ_{2012} , μ_{2011} , e_{2012} , and e_{2011} in the following models:

$$GDP_{2012} = c_{2012}^1 + \alpha_{2011}^1 GDP_{2011} + \dots + \alpha_{1996}^1 GDP_{1996} + \beta_{2011}^1 A_{2011} + \dots + \beta_{1996}^1 A_{1996} + \mu_{2012} \quad (11)$$

$$GDP_{2011} = c_{2011}^1 + \alpha_{2010}^1 GDP_{2010} + \dots + \alpha_{1996}^1 GDP_{1996} + \beta_{2010}^1 A_{2010} + \dots + \beta_{1996}^1 A_{1996} + \mu_{2011} \quad (12)$$

$$GDP_{2012} = c_{2012}^0 + \gamma_{2011}^1 GDP_{2011} + \dots + \gamma_{2010}^1 GDP_{2010} + \dots + \gamma_{1996}^1 GDP_{1996} + e_{2012} \quad (13)$$

$$GDP_{2011} = c_{2011}^0 + \gamma_{2010}^1 GDP_{2010} + \dots + \gamma_{2009}^1 GDP_{2009} + \dots + \gamma_{1996}^1 GDP_{1996} + e_{2011} \quad (14)$$

From a technical estimation point of view, information used to prepare the forecast based on the above equations can be divided into the following two types:

a) Information obtained from past relationships between past GDP and future GDP (optionally taking

into account IoSA) only for observations pertaining to Poland,

- b) Information obtained from past relationships between past GDP and future GDP (optionally taking into account IoSA) only for observations pertaining to countries other than Poland (i.e., for $144 - 1 = 143$ countries).

To eliminate the natural difference in the countries' size, GDP and IoSA data were divided by their average values for the period 1996-2010, calculated for each country. This information can be used to estimate the parameters of the tested equations with different weights, as mentioned earlier. A more accurate specification is made below on the example of estimating the error in μ_{2012} , i.e., for 2012. In the case of estimating the remaining errors (i.e., μ_{2011} , e_{2011} , e_{2012}), the procedure is analogical. Assuming

$$X_1 = \begin{bmatrix} GDP_{2010_Poland} & \dots & GDP_{2011-p_Poland} \\ GDP_{2009_Poland} & \dots & GDP_{2010-p_Poland} \\ \vdots & \vdots & \vdots \\ GDP_{1996+p-1_Poland} & \dots & GDP_{1996_Poland} \\ GDP_{2010_Country1} & \dots & GDP_{2011-p_Country1} \\ GDP_{2009_Country1} & \dots & GDP_{2010-p_Country1} \\ \vdots & \vdots & \vdots \\ GDP_{1996+p-1_Country1} & \dots & GDP_{1996_Country1} \\ GDP_{2010_Country2} & \dots & GDP_{2011-p_Country2} \\ GDP_{2009_Country2} & \dots & GDP_{2010-p_Country2} \\ \vdots & \vdots & \vdots \\ GDP_{1996+p-1_Country2} & \dots & GDP_{1996_Country2} \\ \vdots & \vdots & \vdots \\ GDP_{2010_CountryN} & \dots & GDP_{2011-p_CountryN} \\ GDP_{2009_CountryN} & \dots & GDP_{2010-p_CountryN} \\ \vdots & \vdots & \vdots \\ GDP_{1996+p-1_CountryN} & \dots & GDP_{1996_CountryN} \end{bmatrix}$$

$$X_2 = \begin{bmatrix} A_{2011-p_Poland} & \dots & A_{2011-p_Poland} \\ A_{2009_Poland} & \dots & A_{2010-p_Poland} \\ \vdots & \vdots & \vdots \\ A_{1996+p-1_Poland} & \dots & A_{1996_Poland} \\ A_{2010_Country1} & \dots & A_{2011-p_Country1} \\ A_{2009_Country1} & \dots & A_{2010-p_Country1} \\ \vdots & \vdots & \vdots \\ A_{1996+p-1_Country1} & \dots & A_{1996_Country1} \\ A_{2010_Country2} & \dots & A_{2011-p_Country2} \\ A_{2009_Country2} & \dots & A_{2010-p_Country2} \\ \vdots & \vdots & \vdots \\ A_{1996+p-1_Country2} & \dots & A_{1996_Country2} \\ \vdots & \vdots & \vdots \\ A_{2010_CountryN} & \dots & A_{2011-p_CountryN} \\ A_{2009_CountryN} & \dots & A_{2010-p_CountryN} \\ \vdots & \vdots & \vdots \\ A_{1996+p-1_CountryN} & \dots & A_{1996_CountryN} \end{bmatrix} \quad (15)$$

where country_i denotes the i-th country out of the researched countries other than Poland, $N=143$, $m=16$ (for 2012; however, $m=15$ in the forecast for 2011).

$$\Theta = \begin{bmatrix} \Theta \\ \vdots \\ \Theta \\ 1-\Theta \\ 1-\Theta \\ \vdots \\ 1-\Theta \\ 1-\Theta \\ \vdots \\ 1-\Theta \\ 1-\Theta \\ \vdots \\ 1-\Theta \\ 1-\Theta \\ \vdots \\ 1-\Theta \end{bmatrix} \quad (16)$$

May W denote the matrix with diagonal elements Θ on the main diagonal. For the discussed case, this matrix will constitute the weighting matrix, where observations pertaining to Poland are weighted with respect to observations pertaining to the remaining countries.

To shorten the notation, let us assume $X = [1, X_1, X_2]$ (17)

where 1 denotes the vector of ones related to the free term, X_1 represents GDP, and X_2 represents the IoSA vector (in the case of the model with a random term denoted as e , X_2 is an empty vector). Consequently, the WOLS (weighted OLS) estimator may be used; thus,

$$\begin{bmatrix} \tilde{\alpha} \\ \beta \end{bmatrix} = [X^T W X]^{-1} X^T W y \quad (18)$$

For the model parameters estimated in this way, the forecast errors $PK\tilde{B}_{2011}$ and $PK\tilde{B}_{2012}$ can be calculated (in other words, the realizations of random terms: μ_{2012} , μ_{2011} , e_{2012} and e_{2011}). RMSE and MAPE

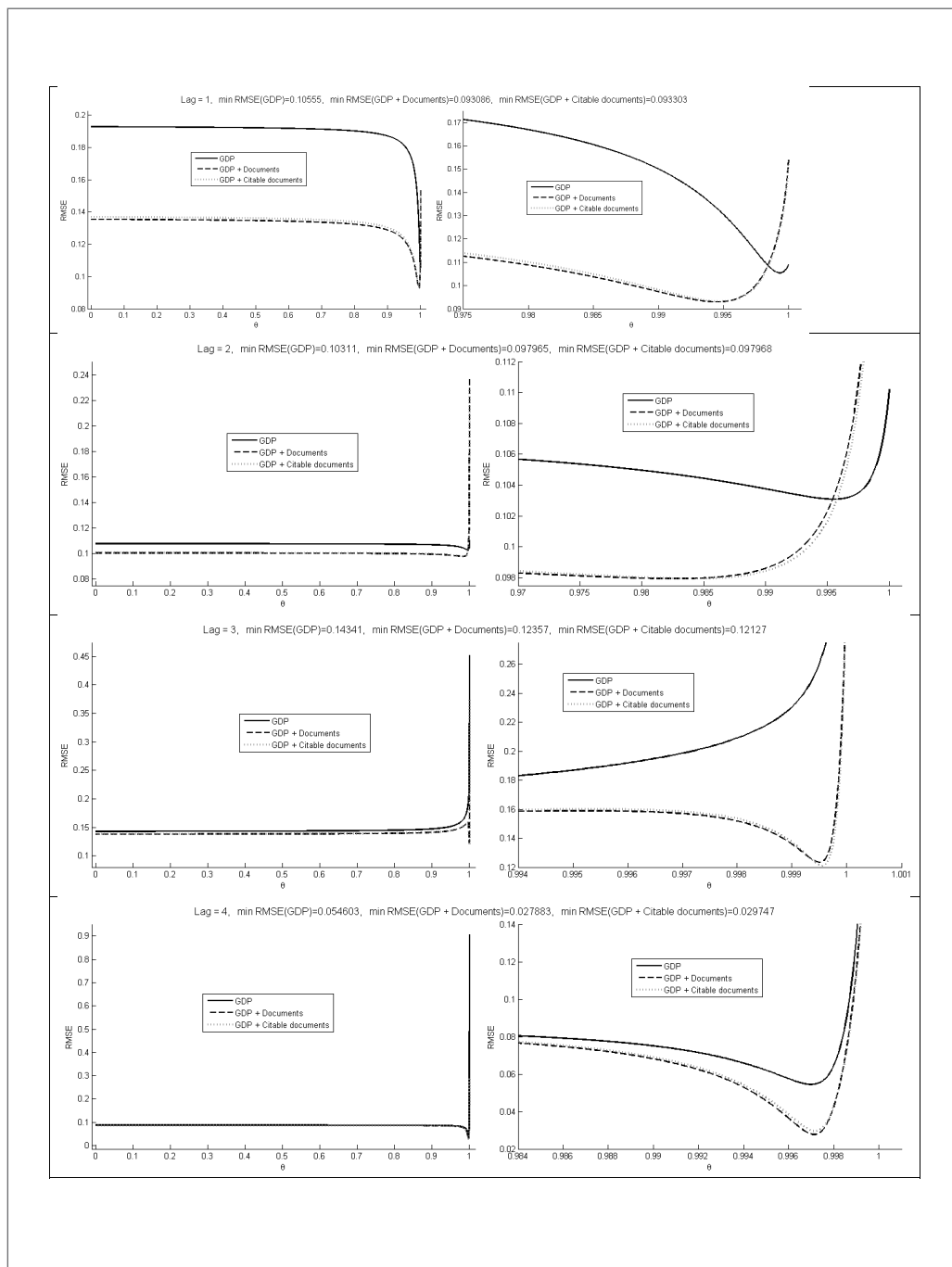


Figure 1. RMSE for "Documents" and "Citable documents" with function _ and lag=1,2,3,4

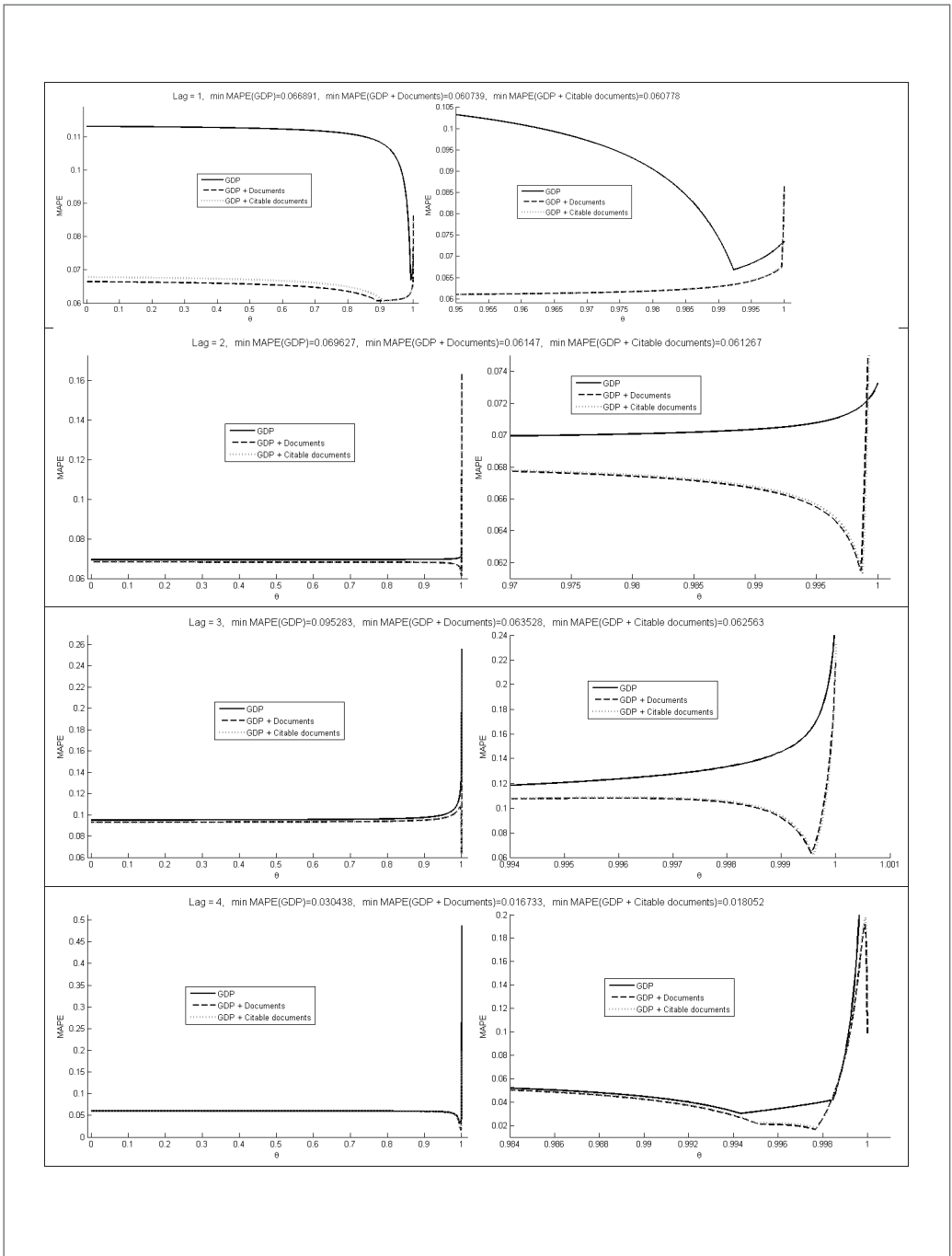


Figure 2. MAPE for "Documents" and "Citable documents" with function _ and lag=1,2,3,4

shall be used as measures of error. These are defined as follows:

$$RMSE = \sqrt{\frac{((\overline{GDP}_{2012} - GDP_{2012})^2 + (\overline{GDP}_{2011} - GDP_{2011})^2)}{2}} \quad (19)$$

$$MAPE = \frac{1}{2} \left(\left| \frac{\overline{GDP}_{2012} - GDP_{2012}}{GDP_{2012}} \right| + \left| \frac{\overline{GDP}_{2011} - GDP_{2011}}{GDP_{2011}} \right| \right) \quad (20)$$

As explained earlier, parameter Θ reflects the relationship between the share of observations for Poland and observations for other countries. Therefore, both types of errors are functions of Θ . The model error figure with and without indicators of scientific research activity (number of publications) are presented in figure 1 for various lag orders. RMSE graphs for “Documents” and “Citable documents” as a function Θ are shown for two forecast variants, i.e., with and without the adopted indicators of scientific research activity. For most cases, an interesting phenomenon is visible near the value $\Theta = 0.99$. The error curve falls to a certain minimum value, after which it rises again. Thus, information from other countries (or, more precisely, using observations pertaining to other countries in the process of estimating the parameters for the equation applied to forecast Polish GDP) improves the forecast of Polish GDP. However, this information should be balanced by an appropriately small weight. The graphs on the right side show an enlargement of the same error graph for the upper range of Θ .

Graphs representing MAPE for “Documents” and “Citable documents” with function Θ are presented in figure 2, analogically to RMSE above. The graphs on the right show an enlargement of the left error graph for the upper range of Θ .

It is worth noting that for all of the above graphs, the lowest forecast error is observed for the predictive equation variant taking into account the characteristics of scientific research activity (i.e., GDP + Documents or GDP + Citable documents). For each lag in the range of 1-4, the lowest RMSE and MAPE appear when applying a non-zero weight for countries other

than Poland. In both cases, i.e., Documents and Citable documents, the lowest RMSE appears in the model with a lag order equal to 4). In all cases, it is evident that scientific research activity indicators help better explain GDP.

Conclusions

The subject of research described in this paper was the causality, in the Granger sense, between scientific research activity and GDP. The research was carried out using annual data pertaining to 144 countries around the world and 28 distinct areas of scientific research activity. Generally, the study confirmed the main proposed hypothesis, which claims that in many countries and fields of study, a causal relationship exists between scientific research activity, measured as the number of significant publications, and GDP. This relationship is very pronounced in many countries, including the USA, United Kingdom, Canada, Australia, Russia, and China. Generally, this relationship is most visible in the G20 group of countries (which have a large potential for both scientific research activity and GDP), with the exclusion of European Union countries. The relatively less visible relationship between scientific research activity and GDP in the European Union is striking, especially in countries such as Germany, France and Italy (which may be caused by an extended period of economic stagnation). Conversely, this relationship is clearly visible on the Asian and American continents. The weakest relationship is found in countries belonging to the low human development group.

The second hypothesis, which claims a varied influence of distinct scientific research activities on GDP, proved that fields of study such as (1) biochemistry, genetics and molecular biology; (2) business, management and accounting; and (3) mathematics are most conducive to fostering economic growth.

References

- Amblard, P.O., & Michel, O. J. J. (2012). The Relation between Granger Causality and Directed Information Theory: A Review. *Entropy*, 15(1), 113–143. <https://doi.org/10.3390/e15010113>
- Ashley, R. (1988). On the Relative Worth of Recent Macroeconomic Forecasts. *International Journal of Forecasting*, 4(3), 363-376.

- Ashley, R., Granger, C. W. J., & Schmalensee, R. (1980). Advertising and Aggregate Consumption: An Analysis of Causality. *Econometrica*, 48(5), 1149-1168.
- Bodkin, R. G., Klein, L. R., & Marwah K. (1991). A History of Macroeconometric Model-Building. Northampton, UK: Edward Elgar Publishing.
- Berndt, E. (1991). The Practice of Econometrics: Classic and Contemporary. New York, NY: Addison-Wesley.
- Borensztein, E., De Gregorio, J., & Lee, J. W. (1998). How does foreign direct investment affect economic growth. *Journal of International Economics*, 45(1), 115-135.
- Can-Zhong, Y., & Qing-Wen, L. (2017). The mutual causality analysis between the stock and futures markets. *Physica A: Statistical Mechanics and its Applications*, 478, 188-204. <https://doi.org/10.1016/j.physa.2017.02.071>
- Cincera, M., & Van Pottelsberghe de la Potterie, B. (2001). International R&D spillovers: A survey. *Cahiers Economiques de Bruxelles*, 169(1), 1-20.
- Coe, D.T., Helpman, E. (1995). International R&D spillovers. *European Economic Review*, 39(5), 859-887. [https://doi.org/10.1016/0014-2921\(94\)00100-E](https://doi.org/10.1016/0014-2921(94)00100-E)
- Ding, M., Chen, Y., & Bressler, S. (2006). Granger causality: Basic theory and application to Neuroscience. In S. Schelter, M. Winterhalder, & Timmer J. (Eds.), *Handbook of Time Series Analysis* (pp. 437-460). Wienheim: Wiley. <https://doi.org/10.1002/9783527609970.ch17>
- Eckstein, O. (1984). *The Dri Model of the U.S. Manufacturing Industries*. New York, NY: McGraw-Hill.
- Engelbrecht, H. J. (2002). Human capital and international knowledge spillovers in TFP growth of a sample of developing countries: an exploration of alternative approaches. *Applied Economics*, 34(7), 832-841. <https://doi.org/10.1080/00036840110061947>
- Engle, R. F., & Granger, C. N. J. (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, 55(2), 251-276.
- Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378), 304-313.
- Geweke, J., Meese R., & Dent, W. (1983). Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence. *Journal of Econometrics*, 21(2), 161-194. [https://doi.org/10.1016/0304-4076\(83\)90012-X](https://doi.org/10.1016/0304-4076(83)90012-X)
- Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3), 424-438.
- Hamilton, J. D. (1994). *Time Series Analysis*. 1-1. New Jersey, NJ: Princeton University Press.
- Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3), 752-782.
- Lee, G. (2005). Direct versus indirect international R&D spillovers. *Information Economics and Policy*, 17(3), 334-348.
- Saggi, K. (2002). Trade, Foreign direct investment and international technology transfer: a survey. *The World Bank Observer*, 17(2), 191-235. Retrieved from <https://openknowledge.worldbank.org/handle/10986/19843>
- Sims, C. (1972). Money, income and causality. *American Economic Review*, 62(4), 540-552.
- Sims, C. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1-48.
- Solow, R. (1957). Technical change and aggregate production function. *Review of Economics and Statistics*, 39(3), 312-320.
- Solow, R. (1962). Technical progress, capital formation and economic growth. *American Economic Review*, 52(2), 76-86.
- Van Pottelsberghe de la Potterie, B., & Lichtenberg, F. R. (2001). Does foreign direct investment transfer technology across borders? *Review of Economics and Statistics*, 83(3), 490-497. <https://doi.org/10.1162/00346530152480135>
- Welfe, W. (Ed.). (2004). *Długookresowy, makroekonomiczny model W8D gospodarki polskiej [Long-term macroeconomic model W8D of the Polish economy]*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Welfe, W. (2006, July). Towards modelling knowledge-based economies. Paper presented at Second international symposium in management, engineering and informatics, Orlando, FL.
- Welfe, W. (Ed.). (2009). *Knowledge-based economies. Models and methods*. Frankfurt/Main: Peter Lang.
- Welfe, W. (2013). *Macroeconometric models*. Berlin: Springer Verlag.

- Xu, B., & Wang, J. (1999). Capital goods trade and R&D spillovers in OECD. *Canadian Journal of Economics*, 32(5), 1258-1274.
- Xu, B., & Wang, J. (2000). Trade, FDI and international technology diffusion. *Journal of Economic Integration*, 15(1), 585-601.
- Zhou, W. X., & Sornette, D. (2006). Non-parametric determinants of real-time lag structure between two time series: The optimal thermal causal path method. *Journal of Macroeconomics*, 28(1), 195-224. <https://doi.org/10.1080/14697680500383763>