EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6, Issue 2, 2020
ISSN 2694-7161
www.ejobsat.cz

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6, Issue 2
2020

Mendel University in Brno
www.ejobsat.com

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Editorial Office Address
EJOBSAT, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

Registration number MK ČR E22009
The journal is published twice a year.
First edition
Number of printed copies: 50
ISSN 2336-6494 (Print)
ISSN 2694-7161 (Online)

Number 2, 2020 was published on December 29, 2020 by Mendel University Press

CONTENTS

THE ROLE OF COMPUTER SKILLS ON THE OCCUPATION LEVEL

Brian Fabo[1,2], Martin Kahanec[3,4,5,6]

[1] *National Bank of Slovakia, Bratislava, Slovakia*
[2] *Comenius University, Bratislava, Slovakia*
[3] *Central European University, Vienna, Austria*
[4] *Central European Labour Studies Institute, Bratislava, Slovakia*
[5] *University of Economics, Bratislava, Slovakia*
[6] *Global Labor Organization, Essen, Germany*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

This paper explores the question of computer skills applicability on individual occupation level in the Netherlands using two web-based data sources: the WageIndicator online survey and job vacancies posted online. The aim of this study is to explore these innovative data sources and compare the information obtained from them with the computer skill requirements inferred from the ISCO occupation classifications. Using the WageIndicator survey, we found a very high incidence of computer use reported by the holders of nearly all office occupations and a substantial degree of computer use by the holders of skilled manual occupations. With a partial exception of the unskilled work in elementary occupations, we find that Dutch job holders are very likely to use computers even in occupations, which are not associated with any relevant tasks. We were able to confirm the robustness of our finding by benchmarking our figures against the PIAAC survey. An older version of this article has been published as a dissertation chapter (Fabo, 2017).

KEY WORDS

digital skills, job matching, vacancies, WageIndicator

JEL CODES

C81, J24

1 INTRODUCTION

Skill matching in the labor market is an important conceptual and policy issue. But to better measure skill mismatch and provide for improved labor market matching, we need to understand the interconnections between tasks, skills and occupations. Mapping these connections is however costly, when conducted by occupations experts, as it requires considerable expertise and time. Some attempts have been made to measure this mapping using new data collection techniques that utilize data originating from the Internet (Beblavý et al., 2016b; Fabo and Tijdens, 2014; Visintin et al., 2015). Each of these attempts was based on a single web-based data source, mapping either the supply side or the demand side. A pioneering, and thus far single, attempt to compare the supply and demand side using web data looked at education requirements only (Tijdens et al., 2018). In this paper, we look specifically at the highly policy relevant aspect of computer skills supply and demand across occupations utilizing two distinct data sources – the WageIndicator web survey and job vacancies posted online.

The need to measure the demand for skills and occupational skill requirements is currently quite salient due to the continued coexistence of high unemployment and growing numbers of hard-to-fill vacancies in Europe. According to Eurostat, a large proportion of Europeans, particularly young Europeans, struggle to even enter the labor market. In 2019, approximately 14% of Europeans aged 15–34 years were classified as not in education, employment, or training (NEET) and this number has dropped only slightly in spite of the robust recovery in the European labor markets in the late 2010s. The situation is particularly pronounced in Italy, where the NEET rate peaks at approximately 24%. Meanwhile, 33% of employers indicated they struggle to attract applicants with the right skills (Eurostat, 2010). According to the European Skills and Jobs (ESJ) survey, there remains a high degree of skills and qualifications mismatch (CEDEFOP, 2015).

Based on the European Skills and Jobs (ESJ) survey, in 2014, 25% of highly qualified employees were overqualified for the job. Whilst the taking of a job requiring a lower level of qualification than is held by the applicant can be a personal preference, the same survey also found that 27% of employees are stuck in the so-called 'dead end' jobs; that is in positions that do not allow the workers to develop their skills and improve their productivity (CEDEFOP, 2015). Furthermore, a large proportion of Europeans, particularly young Europeans, struggle to even enter the labor market, according to Eurostat.

Economists have long considered the role of skills central to the understanding of the matching between employees and employers in the labor market, but also from the policy perspective. Of particular importance is the human capital approach, which has become widely accepted in economics (Becker, 1962; Benhabib and Spiegel, 1994; Schultz, 1971). According to this approach, highly skilled workers are more productive and thus more valuable for employers. Unskilled workers, on the other hand, are not as productive and thus are offered only lower wages or less favorable working conditions. They may not be hired at all, if, for example, a minimum wage policy makes their employment economically unsound.

But how can a situation where workers underutilize their skills on a large scale coexist with skills shortages perceived by employers? Empirical research shows that the likely reason is a shortage of specific skills in high demand by employers (CEDEFOP, 2014). Being able to identify the demand for skills is thus crucial for informing policy makers in particular in areas such as education and training. Beblavý et al. (2016c) study IT skill demand based on a large sample of job vacancies posted online throughout 2013 for 30 common occupations in the USA. They found that (i) demand for computer skills is high across most occupations and growing as the complexity of an occupation increases, (ii) while there are many different computer skills, only a relatively small number of them is relevant for workers outside of the IT industry itself, (iii) computer skills determined on the basis of job vacancies are highly in line with the computer skills inferred from tasks defined for the individual occupations in occupational classification systems ISCO.

These findings potentially open doors for using web data to better understand the mapping between skills, tasks and occupations, and inform policy in a wide range of areas connected to skill acquisition and use. In particular, web-based data collection techniques may help us gauge information about the usefulness of skills across occupations and thus provide for enhanced policies aiming at improved skill matching in the labor market. However, the scope and usefulness of web-based data for classification of jobs and tasks pertaining thereto has yet to be tested. In this paper, we extend Beblavý et al. (2016c) in this direction by (i) benchmarking the job vacancy data against web-based survey data with respect to expert mapping of tasks to occupations and by (ii) looking at a wider scope of occupations at all skill levels.

2 THEORETICAL FRAMEWORK

Studying labor market matching is a complicated matter, because the universe of jobs and tasks that workers do, and skills they have, is large, complex and dynamic (Beblavý et al., 2016a; Fabo and Tijdens, 2014; Visintin et al., 2015). In this section, we zoom in on the current state of the art with regards to ways to systematically study the phenomenon. Jobs can be aggregated into a smaller number of clusters by abstracting from the context of work and focusing purely on the tasks performed by the workers in those jobs. These groups, called occupations, can be defined as follows: "An occupation is a bundle of job titles, clustered in such a way that survey respondents will recognize it as their job title in a valid way; an occupation identifies a set of tasks distinct from another occupation; an occupation should have at least a non-negligible number of jobholders and it should not have an extremely large share in the labor force" (Tijdens, 2010).

This aggregation of jobs is done both for practical reasons – to better organize labor (Damarin, 2006) and for research purposes – to understand skill demands on the labor market (Levenson and Zoghi, 2010). Employers and employees use occupations to characterize and refer to jobs. The skill dimension is inseparable from tasks, because a skill is "a worker's endowment of capabilities for performing various tasks" (Acemoğlu and Autor, 2011). Consequently, a salient aggregation of jobs into occupations is necessary for our understanding of the matching between employers and employees in the labor market.

Skills can be generally split into two categories: job specific and transferable. For example the O*NET classification of occupations, used widely in the US considers the "content" and "process" skills to be job specific and "social skills", "technical skills", "complex problem-solving skills", "systems skills" and "resource management skills" to be transferable. Computer skills can be placed in both categories (BGT, 2015) as they can act as the "content" of the job particularly in the IT occupations, but also can be regarded as a technical skill increasing the productivity of non-IT workers. Recently, empirical evidence has underlined the importance of the transferable dimension of computer skills across the entire labor market, including some of what was traditionally considered low skill occupations (Beblavý et al., 2016c, 2016d). Scholars along with the policy makers started recognizing exclusion from the labor market faced by workers lacking computer skills (Bührer and Hagist, 2017; CEDEFOP, 2015; Horrigan, 2016; Smith, 2015). Consequently, it appears pertinent to analyse the role of computer skills in non-IT occupations as well.

There are several approaches for determining the demand and supply for IT skills on a labor market:

The most obvious approach is connected with the occupations themselves. Because occupations are defined through tasks and tasks imply skills, it is possible to infer skills demand from the tasks associated with the individual occupations. Individual classifications of occupations, such as the International Standard Classification of Occupations (ISCO) or the American O*NET classification defines each occupation on the basis of a list of tasks assigned to it by labor market experts (Elias, 1997). The downside is that such systems are only rarely updated. For example the last update of ISCO took place in 2008, which was preceded by the 1988 update; overall there has been only four updates of the classification since the first version in 1958. Meanwhile, the nature of work is quickly changing due to numerous factors, including technological progress, outsourcing/offshoring and change in labor organization (Acemoğlu and Autor, 2011; Beblavý et al., 2016a; Cowen, 2013).

Fortunately, it is not always necessary to rely on experts. Skill and task questions have increasingly been included in surveys such as the OECD Adult Skills survey PIAAC and the web-based WageIndicator survey (Fabo and Tijdens, 2014). Importantly, the evaluation of job requirements by the job holder has been found to be largely consistent with the expert estimates, even in the case of non-representative web surveys (Tijdens et al., 2014, see the data section for our own robustness check). Such web survey-based estimates of computer skill applicability can be thus used for economic analysis (Mýtna Kureková et al., 2015; Tijdens and Visintin, 2016; Visintin et al., 2015; Drahokoupil and Fabo, 2020). In particular, the large scale, continuous web-based surveys provide up to-date-information and are usable for reliable exploratory analysis, especially if the results can be benchmarked against a representative data source on a regular basis (Fabo and Kahanec, 2018; Steinmetz et al., 2014; Tijdens and Steinmetz, 2016).

Additionally, it is possible to learn from employers. That is, because when hiring, employers tend to think of the task list the worker needs to be able to perform (Autor, 2001; Winterton, 2009). Furthermore, the Internet is increasingly becoming a place, where job vacancies are advertised and thus an important source of data on skill demand (Askitas and Zimmermann, 2015; Beblavý et al., 2016c, 2016d; Lenaerts et al., 2016; Mýtna Kureková et al., 2015; Fabo et al., 2017). Just like in the case of web surveys, there is some controversy associated with the use of these new data sources for social science research. Nonetheless, in spite of significant concerns pertaining to mainly the representativeness and the potential for generalization of results generated on the basis of this data source, there is a growing consensus among scholars that web-based data are going to be increasingly important source for the research on the labor market (Askitas and Zimmermann, 2015).

This paper scrutinizes these two approaches to mapping tasks to occupations – web survey-based and vacancy-based – to test these methodologies in terms of their precision as benchmarked to expert-based mapping. It this sheds more light on possible approaches that can help to address the need for a more precise and up-to-date efficient mapping of tasks to occupations in the rapidly developing and changing labor market.

3 METHODOLOGY AND DATA

3.1 Data Description

This paper analyses skill requirements inferred from three data sources: (i) the tasks assigned to the individual occupations by labor market experts in the ISCO2008 classification of occupations, (ii) the self-reported use of computer skills as measured by the WageIndicator survey and (iii) the share of an explicit IT requirements in vacancies posted on the Internet. Due to the availability of data, we focus on the Netherlands, as explained later in this section.

The survey-based data we obtained from the WageIndicator survey, which is a continuous, large scale, voluntary web survey covering more than 90 countries on all continents. Nonetheless, the quality of the data varies among countries. 18% of all data intake is collected in the Netherlands, where the survey originally started and where the local website hosting the survey gets approximately five million unique visitors per year. The high number of respondents to the survey in the Netherlands and the fact that Internet use is very widespread in the country result in a sample, which is quite similar to general population sample, and suitable for exploratory analysis (Fabo and Kahanec, 2018).

In the WI survey, we added a module asking employees about their use of computers at work. Specifically, we asked “When do you use a computer or tablet?” with possible answers being

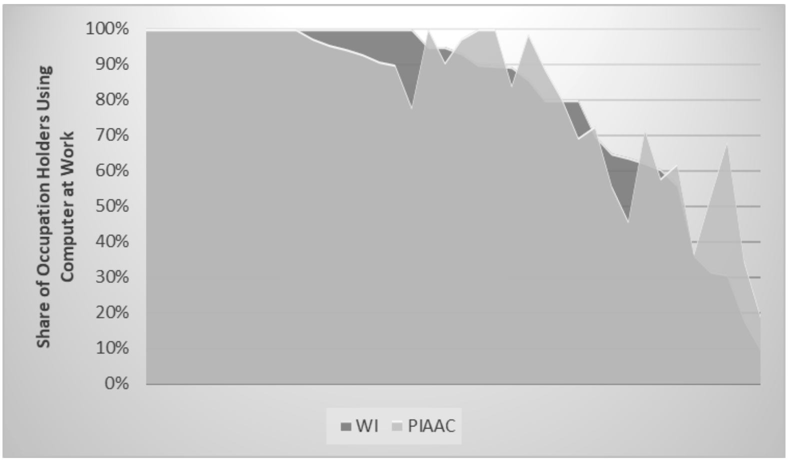

Fig. 1: Comparison of self-reported computer use per occupation between WageIndicator and PIAAC datasets

(i) Only during working hours (ii) Both during working hours and free time (iii) Only during free time (iv) Never. We have recoded these questions such that if a respondent selected either option (i) or (ii), they are considered as someone using a computer at work, otherwise we consider them a non-user.

We launched the module on the 18th of August, 2016 and collected the data until the end of the year. The respondents were not obliged to answer the questions to proceed with the questionnaire. We only used the occupations for which at least ten respondents answered the IT module. Within less than five months, we managed to cover 62 occupations (out of a total of 436 occupations defined in ISCO2008), with a total of 1644 responses, covering all aggregated occupations groups (one digit ISCO), except for military occupations and skilled workers in agriculture. As a result, our analysis is limited to common occupations and might not be representative of occupations with few holders, which is a widespread general problem in occupations research (Tijdens et al., 2014).

To get a sense of the representativeness of our data, we compare it against a representative data source, the Dutch Survey of Adult Skills (PIAAC), which is organized by OECD. One limitation is that the PIAAC data were collected in the years 2011 and 2012, so there is a 5-year asynchrony. A further limitation is that when we applied the same criterion as with the WageIndicator data and eliminated all occupations with less than 10 observations, we could only match 49 out of the 62 occupations, which are sufficiently covered in both datasets. Both datasets exhibit a largely similar picture for nearly all office jobs holders who show a very high intensity of computer use at work, while the percentage of those using computers is lower among the holders of manual jobs. The correlation between the two indicators is highly significant and very strong ($r = 0.922$, also see Fig. 1). Statistically, we can also compare the two sets of estimates using a paired t-test for two sets of estimates. This test is statistically insignificant with a t value of $t = -0.8719$ ($p = 0.806$). Where we see differences, these typically reflect a large growth of use of computers at work by some manual job holders, such as electrical mechanics, truck and taxi drivers, cooks or carpenters.

Vacancy-based data on the share of computer skills in occupations were obtained from job vacancies posted online between August and December 2016. The dataset was provided by the company Textkernel, which is the market leader in the collection, processing and analysing job vacancy data in western Europe utilizing a large number of advanced algorithms to get as representative sample of job vacancies as possible, with by far the most complete dataset being collected in the Netherlands (Zavrel, 2016). The sample covers 60 of the analysed occupations and is based on nearly 300,000 unique job vacancies posted online. In line with our previous analysis (Beblavý et al., 2016c), we calculate the share of vacancies containing at least one keyword associated with computer use within occupations defined at the most detailed (4-digit) level of ISCO occupations classification.

The Dutch labor market represents an ideal environment to explore the use of online data because it contains a very high quality online survey as well as a near-exhaustive database of online job vacancies. Furthermore, the Netherlands has a large high-skilled workforce and a high degree of computerization, which makes the market for computer skills broad enough and deep enough to get a sufficient number of observations for a large-enough number of occupations. Finally, the Dutch labor market is relatively homogenous with limited a degree of regional variations, making it particularly suitable for an analysis of occupations within the country.

3.2 Analytical Strategy

The aim of our analysis is to evaluate what the two online data sources can tell us about the intensity of computer skill use across occupations in the Dutch labor market. A key question is how we can decide which of the mappings is more accurate. After all, two things are being measured – the demand by employers And the self-reported use of skills by workers. Our strategy rests in the ISCO occupational classification itself and on the information it contains about skill use intensity across individual occupations. There are two distinct pieces of information to be identified:

First, we looked at the relevance of computer skills for the tasks associated with individual occupations in the ISCO classification. We have coded a total of 466 individual tasks such that each was either classified as clearly requiring a computer (for instance "Developing and implementing software and information system testing policies, procedures and scripts") or not necessarily requiring computer skills, but having use for them (such as "Designing and modifying curricula and preparing courses of study in accordance with requirements", which can still be performed using pen or paper, but is likely done using a text processing software by most educators) and those, that have no use for computer skills at all (such as "Maintaining discipline and good working habits in the classroom"). For coding, we included two typical uses of computer applications identified in a previous job vacancy analysis (Beblavý et al., 2016b) – general computer use, including using job-specific software and office applications such as spreadsheet and text processor.

Second, we used information on occupation complexity. The ISCO classification also associates occupations with the concept of complexity of tasks involved by sorting each civilian occupation into one of nine aggregated occupation groups plus an additional group for military occupations, which are not a subject of analysis in this paper. Eight out of those groups can be connected to degrees of complexity of tasks associated with them. The most complex, professional, occupations belong to Group 2. Groups 3–8 contain occupations with an intermediate degree of complexity, with groups 3 and 4 containing office jobs, and groups 5–8 containing manual labor. Finally, Group 9 contains elementary occupations with a low complexity of associated tasks. Group 1 is more heterogeneous than the other ones, by virtue of containing managerial occupations, which are generally associated with an intermediate to high degree of complexity (Hunter, 2009; Mýtna Kureková et al., 2013), making it impossible to exclusively assign this group to a specific complexity level.

In the analysis, we use these occupation traits determined by labor market experts when constructing the ISCO classification as a guide. More precisely, we defined three distinct groups of occupations: Those that require computer skills to perform at least one of the tasks associated with them, those that do not have any apparent use of the computer skills and finally, those that do not necessarily require the computer skills, but entail tasks, which might be performed more efficiently using a computer. Similarly, we looked at the occupation complexity to see the demand and use of computer skills across manual and office/service occupations with various levels of complexity.

We then compare the (i) measured demand for computer skills determined on the basis of actual job vacancies[1], (ii) the use of computers reported by occupation holders, and (iii) the expert-based mapping of skills to occupations based on the ISCO classification. Based on these three sources, we establish a multi-dimensional picture of the role of the IT skills in the contemporary Dutch labor market and shed light on their use and measurement options.

4 RESULTS

Having described our analytical tools and data, we are now ready to proceed with the presentation of the results. This section is organized on the basis of the ISCO classification, first looking at occupations on the basis of the applicability of computer skills inferred from tasks associated with them.

Through coding of tasks, we have identified 17 occupations, a mixture of professional and administrative occupations, which require the use of a computer for at least one task associated with them. As can be seen in Tab. 1, all or nearly all holders of these occupations use computers at work. Nonetheless, when looking at the vacancies, we see these high skill occupations (including several in IT, such as computer network professionals, which cannot be performed without computer skills) do not explicitly ask for computer skills in the majority of the cases. This likely follows from the fact that not all skill requirements are explicitly specified, because they are sometimes taken for granted. That would explain why the demand for computer skills surges for administrative occupations, such as secretaries, where those skills are not necessarily taken for granted. Nonetheless, if we accept this explanation, it strengthens the case for skepticism towards the use of job vacancies to determine the applicability of specific skills, as the scope of tendency of employers leaving out some skill requirements they take for granted is clearly a major concern (Mýtna Kureková et al., 2016).

Shifting our attention to thirteen manual occupations, both skilled and unskilled, which have no apparent use for computer skills according to tasks associated with them, a large share of workers (but certainly not all of them) in these occupations uses a computer (Tab. 2). Of particular interest are Healthcare assistants, out of whom nearly 90% use a computer, as well as Stationary plant and machine operators, Motor vehicle mechanics and repairers, which are all above 70%. This suggests a major shift of nature of those occupations as they now require skills not foreseen when constructing the 2008 update of the ISCO classification. The appearance of new tasks in existing occupations has been observed in the labor studies literature and is a common manifestation of a shifting skill demand (Barley and Tolbert, 1991; Beblavý et al., 2016a; Crosby, 2002) Turning to job vacancies, we see that demand for computer skills in these occupations is in general lower than for occupations that entail tasks associated with computer skills.

The information we can learn from job vacancies is, meanwhile, much more limited and does not match the story inferred from the WI survey. The explicit requirement for computer skills in vacancies recruiting for these jobs is negligible and does not allow us to identify the uptick in demand for computer skills in some occupations, something which is evident from the WageIndicator data.

[1] After having experimented with various computer skills definition from different sources while working with the US data (Beblavý et al., 2016b), we found the regular expression `((MS|Microsoft) Excel)|((MS|Microsoft) excel)|((MS|Microsoft) Word)|((MS|Microsoft) word)|((MS|Microsoft) Office)|((MS|Microsoft) office)|(PC)|(Computer)` the best to estimate the IT skills requirement. Luckily, the Dutch word for computer is the same as the English one, so the regular expression was usable also with the vacancies in Dutch.

Tab. 1: Applicability of computer skills for occupations requiring computer skills

Occupation	Vacancies requiring computer skills	Respondents reporting using computer in the WI survey
Secretaries (general)	25%	94%
General office clerks	17%	99%
ICT professionals	16%	100%
Accounting and bookkeeping clerks	16%	100%
Personnel clerks	15%	96%
Contact centre information clerks	14%	100%
Systems analysts	10%	100%
Advertising and marketing professionals	9%	100%
Computer network professionals	9%	100%
Industrial and production engineers	8%	100%
Accountants	8%	100%
Applications programmers	8%	100%
Software and applications developers	8%	100%
Statistical, finance and insurance clerk	8%	98%
Accounting associate professionals	7%	100%
Draughtspersons	6%	100%
Graphic and multimedia designers	4%	96%

Tab. 2: Applicability of computer skills for occupations with no apparent use for computer skills

Occupation	Vacancies requiring computer skills	Respondents reporting using computer in the WI survey
Stationary plant and machine operators	5%	76%
Health care assistants	3%	89%
Electrical mechanics and fitters	3%	56%
Freight handlers	3%	65%
Motor vehicle mechanics and repairers	2%	80%
Fitness and recreation instructors	2%	73%
Child care workers	1%	31%
Cooks	1%	61%
Waiters	1%	32%
Car, taxi and van drivers	1%	64%
Domestic cleaners and helpers	1%	10%
Carpenters and joiners	0%	18%
Heavy truck and lorry drivers	0%	36%

This leaves us with a group of 32 occupations, containing professional and administrative occupations, service and sales occupations and even some skilled manual occupations. Occupations in this group do not require computer skills, but entail tasks, which can be potentially performed more efficiently (Tab. 3). In most such occupations, the use of computers is universal or nearly universal. While there are still some occupations, such as Shop assistants, Primary school teachers or Building and related electricians, which still employ a substantial portion of workers who do not use computers at work, even in those occupations, a vast majority of workers benefit from computers. Interestingly, the requirement for computer skills in

Tab. 3: Applicability of computer skills for occupations with possible, but not necessary, use for computer skills

Occupation	Vacancies requiring computer skills	Respondents reporting using computer in the WI survey
Administrative and executive secretaries	27%	100%
Receptionists (general)	24%	90%
Production clerks	19%	100%
Buyers	14%	100%
Business services and admin. managers	11%	90%
Human resource managers	10%	100%
Product graders and testers (except food)	10%	91%
Sales and marketing managers	9%	100%
Management and organization analysts	9%	100%
Commercial sales representatives	9%	95%
Journalists	8%	100%
Stock clerks	8%	80%
University and higher education teachers	7%	100%
Managing directors and chief executives	6%	93%
Personnel and careers professionals	6%	100%
Enviro. and occup. health inspectors	6%	100%
Legal professionals	5%	100%
Employment agents and contractors	5%	100%
Nursing professionals	4%	95%
Policy administration professionals	4%	100%
Construction supervisors	4%	100%
Shopkeepers	4%	100%
Agricultural and ind. mach. mechanics	4%	80%
Vocational education teachers	3%	100%
Social work and counselling professional	3%	100%
Shop supervisors	3%	81%
Shop sales assistants	3%	62%
Building and related electricians	3%	70%
Physiotherapists	1%	100%
Primary school teachers	1%	86%
Clerical support workers	N/A	92%

posted vacancies is explicitly stated only for a few administrative positions. Consequently, on the basis of the vacancy data, we would not be able to see how widespread the use of computer skills is in these occupations.

Finally, we consider the complexity of tasks (see the discussion in the analytical strategy section) associated with the individual occupations, rather than specifically the applicability of computer skills. The results based on the WI survey and vacancy data largely match the previous results as far as division between office and manual occupations is concerned. Regardless of tasks complexity, office and service jobs require computer skills more often than manual jobs.

However, new insights emerge when looking at the details. Based on the WI data we see that nearly all holders of office/service jobs use computers, while only about 60% of skilled manual workers do. Low-skilled employment is a specific category with less than 40% of workers using

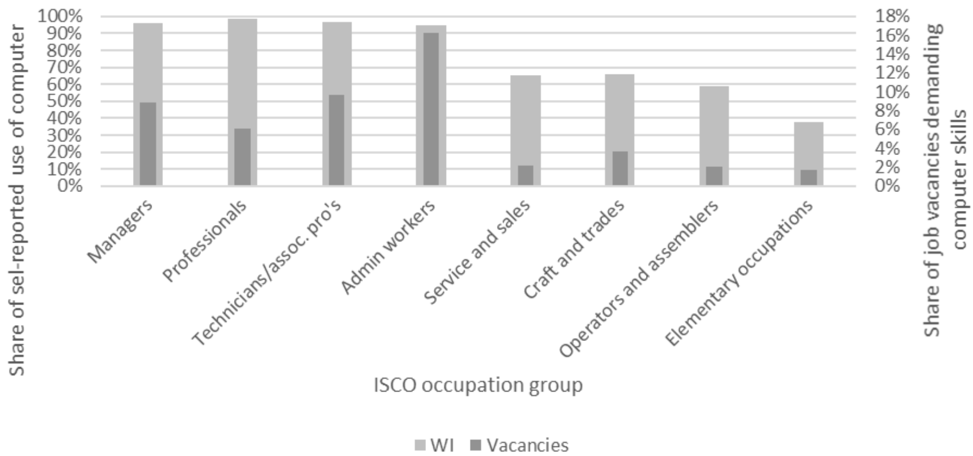

Fig. 2: Average computer skills applicability per ISCO occupation group. Notes: Inferred from self-reported rates in the WI survey (left axis) and vacancies containing IT skill requirement (right axis)

a computer at work. The detailed story told by vacancies is different from the ones inferred from the WI survey, or indeed can be expected on the basis of complexity of tasks associated with the occupations. In all categories, less than 20% of vacancies demand computer skills. Most commonly, they are required in administrative occupations (16%). Paradoxically, the lowest explicit demand for computer skills among office and service occupations is in highly complex, professional occupations (7%). The only non-office occupational group with non-negligible demand for computer skills is crafts and trades workers (4%).

5 DISCUSSION AND CONCLUSIONS

We review our findings from the three data sources in Tab. 4. What we see based on a comparison of tasks associated with individual occupations in the ISCO classification and contrasted with WI data is that while computer skills are typically necessary in particular for high-skill, professional occupations and only being rather useful for medium skill office/service occupations, nearly all workers in this category use a computer at work. Indeed, when looking at vacancy data, we see that oftentimes the explicit demand for computer skills is the highest among medium skill non-manual occupations, in particular service and sales. It appears that for high-skilled occupations, computer skills are assumed and taken for granted, by employers. Among the manual jobs holders, we see it is often the case that occupations, in which computers are rarely used, are limited to low-skill labor.

Our analysis shows the potential of web data as a source of information for scholars and professionals in need of information about the relevance of specific skills in the labor market. In particular, we show that web surveys can be used to explore computer skills applicability across detailed occupations. We also show, however, that there are limitations connected to using vacancy data. The main cause for concern is connected to the low incidence of explicit requirements for general computer skills even in some types of vacancies, where the use of such skills is clearly required to fulfil tasks associated with the respective occupations.

We show that computers have not only become nearly universal in almost any office job, but are increasingly common in manual occupations, too. Jobs not entailing any use of a computer appear increasingly concentrated in the low complexity, elementary occupations

Tab. 4: Overview of computer skills applicability based on different data sources

Skill level	Computer skills requirements inferred from tasks	Job vacancy demand for computer skills	Use of computers reported by occupation holders
High	Useful, sometimes necessary	Medium	High
Medium Office	Useful, sometimes necessary	Medium to high	High
Medium Manual	Not needed, sometimes useful	Low	Medium
Low	Not needed	Low	Low to medium

category. This represents a challenge particularly for older skilled workers lacking computer skills, who might be excluded from the labor market if not provided with appropriate IT training.

Looking at job vacancy data, we see that while for professional jobs, computer skills are not explicitly mentioned, which is likely due to expectation of employers that all candidates for some position are able to use computers, they are relatively often specified for administrative positions. Among skilled manual occupations, IT requirements appear to be often stressed for craft and trade workers. One way to interpret this finding is that in these occupations, there perhaps remains greater friction between the demand for computer skills and the supply of qualified candidates with those skills, incentivizing employers to stress the importance of computer proficiency as a requirement for employment.

Further research is needed to establish the link between explicit listing of a skill requirement and expectation of it being instrumental for labor market matching. Nonetheless, there is a potential for an interesting synergy between the two web-based data sources: An online survey can be deployed to gain a detailed account of the computerization of work across occupations, while job vacancies can tell us where there is a particular risk of mismatch between job requirements and the skill of job candidates.

Given that WI is a continuous and international survey and that job vacancies are collected in an increasingly systematic fashion, it appears to be a suitable resource to explore the progress of computerization in time and across labor markets, potentially providing for a better understanding of the labor market matching and evidence-based policy in the areas of education and training. To fulfil this potential however, substantial work is needed to further explore the apparent limitations of online job sources and how they could be potentially countered.

6 ACKNOWLEDGEMENT

The authors are thankful to Textkernel, specifically Bauke Visser and Jakub Zavrel, for providing the job vacancy data. Further, we thank Alina Poliakova for assistance with coding IT skills requirements in individual occupations on the basis of tasks associated with them.

The authors gratefully acknowledge the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program.

7 REFERENCES

Acemoğlu, D. and Autor, D. 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings. In Ashenfelter, O. and Card, D. (eds.). *Handbook of Labor Economics*, Volume 4, Part B, Chapter 12, pp. 1043–1171. DOI: 10.1016/S0169-7218(11)02410-5.

Askitas, N. and Zimmermann, K. F. 2015. The Internet as a Data Source for Advancement in Social Sciences. *International Journal of Manpower*, 36 (1), 2–12. DOI: 10.1108/IJM-02-2015-0029.

Autor, D. H. 2001. Wiring the Labor Market. *Journal of Economic Perspectives*, 15 (1), 25–40. DOI: 10.1257/jep.15.1.25.

Barley, S. R. and Tolbert, P. S. 1991. Introduction: At the Intersection of Organizations and Occupations. In Tolbert, P. S. and Barley, S. R. (eds.). *Research in the Sociology of Organizations*, Vol. 8, pp. 1–13.

Beblavý, M., Akgüç, M., Fabo, B. and Lenaerts, K. 2016a. *What Are the New Occupations and the New Skills? And How are They Measured?* InGRID Working Paper. DOI: 10.5281/zenodo.1882280.

Beblavý, M., Fabo, B. and Lenaerts, K. 2016b. *Demand for Digital Skills in the US Labor Market: The IT Skills Pyramid*. CEPS Special Report No. 154.

Beblavý, M., Fabo, B. and Lenaerts, K. 2016c. *Skills Requirements for the 30 Most-Frequently Advertised Occupations in the United States: An Analysis Based on Online Vacancy Data*. CEPS Special Report No. 11406.

Beblavý, M., Mýtna Kureková, L. and Haita, C. 2016d. The Surprisingly Exclusive Nature of Medium- and Low-skilled Jobs: Evidence from a Slovak Job Portal. *Personnel Review*, 45 (2), 255–273. DOI: 10.1108/PR-12-2014-0276.

Becker, G. S. 1962. Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, 70 (5), 9–49. DOI: 10.1086/258724.

Benhabib, J. and Spiegel, M. M. 1994. The Role of Human Capital in Economic Development Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics*, 34 (2), 143–173. DOI: 10.1016/0304-3932(94)90047-7.

BGT. 2015. *Crunched by the Numbers: Digital Skills Gap in the Workforce*. Burning Glass Technologies Report.

Bührer, C. and Hagist, C. 2017. The Effect of Digitalization on the Labor Market. In Ellermann, H., Kreutter, P. and Messner, W. (eds.). *The Palgrave Handbook of Managing Continuous Business Transformation*, pp. 115–137. DOI: 10.1057/978-1-137-60228-2_5.

CEDEFOP. 2015. *Skill Shortages and Gaps in European Enterprises: Striking a Balance Between Vocational Education and Training and the Labour Market* Cedefop Reference Series No. 102. DOI: 10.2801/042499.

CEDEFOP. 2014. *Skill Mismatch: More than Meets the Eye*. Briefing Note. DOI: 10.2801/57022.

Cowen, T. 2013. *Average Is Over: Powering America Beyond the Age of the Great Stagnation*. Dutton.

Crosby, O. 2002. New and Emerging Occupations. *Occupational Outlook Quarterly*, 46 (3), 16–25.

Damarin, A. K. 2006. Rethinking Occupational Structure: The Case of Web Site Production Work. *Work and Occupations*, 33 (4), 429–463. DOI: 10.1177/0730888406293917.

Drahokoupil, J. and Fabo, B. 2020. The Limits of Foreign-Led Growth: Demand for Skills by Foreign and Domestic Firms. *Review of International Political Economy*. DOI: 10.1080/09692290.2020.1802323.

Elias, P. 1997. *Occupational Classification (ISCO-88): Concepts, Methods, Reliability, Validity and Cross-National Comparability*. OECD Labour Market and Social Policy Occasional Paper No. 20. DOI: 10.1787/304441717388.

Eurostat. 2010. *Employers' Perception of Graduate Employability*. Analytical report. Flash Eurobarometer No. 304.

Fabo, B. 2017. *Towards an Understanding of Job Matching Using Web Data*. CentER Dissertation Series No. 528. Tilburg University.

Fabo, B., Beblavý, M. and Lenaerts, K. 2017. The Importance of Foreign Language Skills in the Labour Markets of Central and Eastern Europe: Assessment Based on Data from Online Job Portals. *Empirica*, 44 (3), 487–508. DOI: 10.1007/s10663-017-9374-6.

Fabo, B. and Kahanec, M. 2018. Can a Voluntary Web Survey be Useful Beyond Explorative Research? *International Journal of Social Research Methodology*, 21 (5), 591–601. DOI: 10.1080/13645579.2018.1454639.

Fabo, B. and Tijdens, K. 2014. *Using Web Data to Measure the Demand for Skills*. CELSI Discussion Paper No. 21.

Horrigan, J. B. 2016. *Lifelong Learning and Technology*. Pew Research Center, Internet & Technology.

Hunter, D. 2009. *ISCO-08 Draft Definitions*. ILO, Geneva.

Lenaerts, K., Beblavý, M. and Fabo, B. 2016. Prospects for Utilisation of Non-Vacancy Internet Data in Labor Market Analysis – An Overview. *IZA Journal of Labor Economics*, 5 (1), 1–18. DOI: 10.1186/s40172-016-0042-z.

Levenson, A. and Zoghi, C. 2010. Occupations, Human Capital and Skills. *Journal of Labor Research*, 31 (4), 365–386. DOI: 10.1007/s12122-010-9098-x.

Mýtna Kureková, L., Beblavý, M., Haita, C. and Thum, A.-E. 2016. Employers' Skill Preferences across Europe: Between Cognitive and Non-Cognitive Skills. *Journal of Education and Work*, 29 (6), 662–687. DOI: 10.1080/13639080.2015.1024641.

Mýtna Kureková, L., Beblavý, M. and Thum-Thysen, A. 2015. Using Online Vacancies and Web Surveys to Analyse the Labour Market: a Methodological Inquiry. *IZA Journal of Labor Economics*, 4 (18), 1–20. DOI: 10.1186/s40172-015-0034-4.

Mýtna Kureková, L., Haita, C. and Beblavý, M. 2013. *Being and Becoming Low-Skilled: A Comprehensive Approach to Studying Low-Skillness*. NEUJOBS Working Paper No. 4.3.1. DOI: 10.2139/ssrn.2402734.

Schultz, T. W. 1971. *Investment in Human Capital: The Role of Education and of Research*. New York: Free Press.

Smith, A. 2015. *Searching for Work in the Digital Era*. Pew Research Center, Internet & Technology.

Steinmetz, S., Bianchi, A., Tijdens, K. and Biffignandi, S. 2014. Improving Web Survey Quality: Potentials and Constraints of Propensity Score Adjustments. In Callegaro, M., Baker, R. P., Bethlehem, J., Göritz, A. S., Krosnick, J. A. and Lavrakas, P. J. (eds.). *Online Panel Research: A Data Quality Perspective*, Part IV, Chapter 12, pp. 273–298. John Wiley & Sons.

Tijdens, K., Beblavý, M. and Thum-Thysen, A. 2018. Skill Mismatch Comparing Educational Requirements vs Attainments by Occupation. *International Journal of Manpower*, 39 (8), 996–1009. DOI: 10.1108/IJM-10-2018-0328.

Tijdens, K., de Ruijter, E. and de Ruijter, J. 2014. Comparing Tasks of 160 Occupations across Eight European Countries. *Employee Relations*, 36 (2), 110–127. DOI: 10.1108/ER-05-2013-0046.

Tijdens, K. and Steinmetz, S. 2016. Is the Web a Promising Tool for Data Collection in Developing Countries? An analysis of the Sample Bias of 10 Web and Face-to-Face Surveys from Africa, Asia, and South America. *International Journal of Social Research Methodology*, 19 (4), 461–479. DOI: 10.1080/13645579.2015.1035875.

Tijdens, K. and Visintin, S. 2016. *What Do Workers Do? Measuring the Intensity and Market Value of Tasks in Jobs*. AIAS Working Paper No. 161.

Tijdens, K. 2010. *Measuring Occupations in Web-Surveys: the WISCO Database of Occupations*. AIAS Working Paper No. 10-86.

Visintin, S., Tijdens, K., Steinmetz, S. and de Pedraza, P. 2015. Task Implementation Heterogeneity and Wage Dispersion. *IZA Journal of Labor Economics*, 4 (20), 1–24. DOI: 10.1186/s40172-015-0036-2.

Winterton, J. 2009. Competence across Europe: Highest Common Factor or Lowest Common Denominator? *Journal of European Industrial Training*, 33 (8/9), 681–700. DOI: 10.1108/03090590910993571.

Zavrel, J. 2016. *Industry Strength Labor Market Web Mining*. Keynote speech at the Eduworks Vacancy Mining and Analysis Workshop, Amsterdam, 24 March 2016.

AUTHOR'S ADDRESS

Brian Fabo, National Bank of Slovakia, Research Department, Imricha Karvaša 1, 813 05, Bratislava, Slovakia, e-mail: brian.fabo@nbs.sk

Martin Kahanec, Central European University, School Of Public Policy, Quellenstraße 51–55, 1100 Vienna, Austria, e-mail: KahanecM@spp.ceu.edu

INSTITUTIONAL QUALITY AND INCOME INEQUALITY: EVIDENCE FROM POST-SOVIET COUNTRIES

Radek Náplava[1]

[1] *Mendel University in Brno, Czech Republic*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

This paper focuses on identifying the relationship between institutional quality and income inequality in chosen post-Soviet countries during the period 2002–2017. Using panel analysis is found a nonmonotonic relationship between institutional quality and income inequality. Increasing institutional quality is associated with growing income inequality, but only to a certain extent; from a certain level, higher institutional quality leads to a reduction in income inequality. The growing institutional quality leads to a deepening of income inequality between the richest social class compared to the poorest and middle class. Role in this process plays a particular regulatory quality, which – as it seems – favors the upper 20%.

KEY WORDS

institutional quality, income distribution, income inequality, middle class

JEL CODES

O15, P48

1 INTRODUCTION

Topic of income inequality is one of the biggest challenges in the field of labor market research, but also in the field of economic policy (Braconier et al., 2015). Disproportionate increase in income inequality can lead to a concentration of political and economic power, which can have negative consequences for economic growth and macroeconomic stability (Dabla-Norris et al., 2015). Income inequality is not only "evil", but may also be desirable as it incentives higher activity.

Empirical studies have shown that technological change and globalization have contributed not only to an increase in total wealth, but also to an increase in income inequality. In addition to technological change and globalization, attention was also paid to the influence of institutions, but this research was often reduced

NÁPLAVA, Radek. 2020. Institutional Quality and Income Inequality: Evidence from Post-Soviet Countries. *European Journal of Business Science and Technology*, 6 (2): 100–112. ISSN 2694-7161, DOI 10.11118/ejobsat.2020.007.

only to selected labor market institutions such as minimum wages, social security, unionization and employment protection (Dabla-Norris et al., 2015). On the contrary, much less attention was paid to the relationship between institutional quality and income inequality. According to Engerman and Sokoloff (2002), the initial differences in the rate of income inequality correspond to differences in institutional quality, but as point Dobre et al. (2019), there are not many studies published in this area, so the solution to the relationship is not entirely clear. This provides motivation for research in this area.

Studies examine institutional quality primarily in relation to economic performance. It seems obvious that higher institutional quality, higher quality "rules of the game", leads to higher economic performance (Acemoğlu et al., 2005; Acemoğlu, 2012; Rodrik, 2008). Institutional quality should therefore also influence the distribution of income within society (Acemoğlu and Robinson, 2002; Dobre et al., 2019). The aim of this paper is to find out how the institutional quality affects income inequality and income distribution in chosen post-Soviet countries. Selected countries are Estonia, Latvia, Lithuania, Belarus, Russia and Ukraine during the period 2002–2017. In these countries has institutional quality a significant impact on economic performance (Náplava, 2017), but the impact on income inequality and income distribution is unknown. This paper fills this gap with the help of panel regression analysis.

This paper is structured as follows. Section 2 explains the relationship between institutional quality and income inequality. Section 3 briefly describes the data and the used methods. Section 4 presents the results of the regression analysis, including robustness check, where the main explanatory variable is replaced by another indicator of institutional quality and further examines whether the relationship between institutional quality and income inequality is monotonic or nonmonotonic. In the last section 5 are discussed achieved results and this section concludes paper.

2 LITERATURE OVERVIEW

Clark and Kavanagh (1996) argue from the perspective of institutionalism that income inequality does not have to arise solely as a result of labor market developments, but that they are the result of institutions and their development. The authors state that the distribution of property rights, the distribution of costs and the distribution of power in society play a role, because it is mainly these factors that determine the ways in which redistribution occurs in the economy. The factors mentioned by the authors can be summarized under the indicator of institutional quality.

Douglass C. North (1991, p. 97) defines institutions as "Institutions are the humanly devised constraints that structure political, economic and social interaction." The worse the individual institutions play their role, the lower their perceived quality – the worse the allocation of resources in society and the more the achieved goals will differ from the expected goals (Chong and Calderón, 2000). According to Acemoğlu and Robinson (2012), institutional quality is a determinant not only for economic performance but also for the level of poverty and inequality. Conversely, poor institutional quality (especially poor political institutions and corruption) can lead to only a handful of elites having access to key resources, which can then benefit more from the country's financial development than the poor. Poor institutional quality is then reflected mainly in higher income inequality. Moreover, as pointed Chong and Gradstein (2007), in a country with poor institutional quality, there is a lack of judicial protection of the poor, thereby deepening social inequalities.

The first empirical study to explicitly capture the relationship between institutional quality and income inequality is provided by Chong and Calderón (2000). Their results imply a quadratic relationship between them. Their main results are based on observations from 70 countries. The authors use only cross-section analysis, so the dynamic relationship cannot be identified. Their article provides evidence that in poor countries, institutional quality is positively associated with income inequality, while in rich countries the relationship is negative. In other words, improvement of institutional quality should first lead to an increase in income inequality, but from a certain level it should decrease.

The dynamic relationship between institutional quality and income inequality examine Chong and Gradstein (2007, 2017), who find out the theory and empirical evidence between them. The growth of institutional quality leads to a reduction in income inequality in the long run, but the opposite relationship is stronger: the reduction of income inequality leads to an increase in institutional quality. Institutional reforms can be an effective tool for reducing inequalities if there is sufficient demand for higher redistributive policies (if political factors allow). The condition is the adaptability of the institutional environment, which is also an important factor of institutional quality. Josifidis et al. (2017) examine the effects of changes in institutional quality in 21 OECD countries between 1990–2010. Their main finding is that institutional inertia is one of the factors behind the growth of income inequality. Slow changes in the institutional environment are not able to respond to rapid technological change and the deepening of globalization. The result is insufficient redistribution and growing income inequality.

Chong and Gradstein (2017) show that political institutions, which provide support for economic policy and protect economic and political rights, have a crucial influence on the relationship between institutional quality and income inequality. In addition, they have a major impact on redistribution. Therefore, as an indicator of the quality of the institutional environment, we use the "governance matters" (GM) indicator, which evaluates the institutional environment mainly from the perspective of political institutions.[1]

3 METHODOLOGY AND DATA

To quantify the impact of institutional quality on income inequality and income distribution is employed a panel regression using an unbalanced panel of 6 countries (Estonia, Latvia, Lithuania, Belarus, Russia and Ukraine) during the period 2002–2017. All data are annual and come from the World Bank database (World Bank, 2020). Descriptive statistics of the main variables that occur in the regression analysis are given in Tab. 8 in the Annex. Pairwise correlation coefficients between institutional quality indicators and income distribution indicators are given in Tab. 9 in the Annex.

Empirical studies examining the effect of institutional quality and income inequality/income distribution have the common denominator: relatively small number of observations. Although the studies examine a large number of countries, but only primarily as a cross-section – e.g. Chong and Calderón (2000) deal 95 countries and Chong and Gradstein (2007) deal 121 countries. Other studies that use panel analysis do not usually have an annual frequency of data, but use, for example, a five-year periodicity, see Kotschy and Sunde (2017) 96 countries, Dobre et al. (2019) 28 EU countries, Josifidis et al. (2017) 21 OECD countries. Due to the annual periodicity of the data, this paper has a similar number of observations as the above study, even though it examines only 6 countries.

Institutional quality is measured as in the case of Law et al. (2014) and Brown et al. (2011) using the Worldwide Governance Indicators (WGI) variable set from Kraay et al. (2010). More specifically, the institutional quality is assessed on the basis of six composite indicators: Voice and Accountability (GM1), Political

[1] Through political institutions is distributed a political power that influences the choice of economic institutions (Acemoğlu et al., 2005).

Stability and Absence of Violence/Terrorism (GM2), Government Effectiveness (GM3), Regulatory Quality (GM4), Rule of Law (GM5), and Control of Corruption (GM6). The arithmetic mean of these components adds up the institutional quality index "governance matters" (GM), that takes values $[-2.5; 2.5]$, as well as its six components.

The model has the following form:

$$\begin{aligned} \text{gini}_{it} = \quad & \alpha + \beta_1 \text{ GM}_{it} + \beta_2 \text{ GDP}_{it} + \\ & + \beta_3 \text{ educ}_{it} + \beta_4 \text{ inv}_{it} + \\ & + \beta_5 \text{ open}_{it} + \beta_6 \text{ gov}_{it} + \epsilon_{it}, \end{aligned} \tag{1}$$

where the dependent variable (gini) for income inequality is the Gini coefficient. The main explanatory variable is institutional quality (GM). Other explanatory variables are, in addition to real GDP growth (GDP), government education expenditure (educ), net investment in government nonfinancial assets (inv), country openness measures as the volume of imports and exports divided by GDP (open) and government final consumption expenditure (gov).

Changes in income inequality are also explained with the help of individual components of the institutional quality indicator, see the following equations:

$$\begin{aligned} \text{gini}_{it} = \quad & \alpha + \beta_1 \text{ GM1}_{it} + \beta_2 \text{ GM2}_{it} + \\ & + \beta_3 \text{ GM3}_{it} + \beta_4 \text{ GM4}_{it} + \\ & + \beta_5 \text{ GM5}_{it} + \beta_6 \text{ GM6}_{it} + \epsilon_{it}, \end{aligned} \tag{2}$$

where the dependent variable (gini) is explained with Voice and Accountability (GM1), Political Stability and Absence of Violence/Terrorism (GM2), Government Effectiveness (GM3), Regulatory Quality (GM4), Rule of Law (GM5) and Control of Corruption (GM6). The arithmetic mean of the GM1–GM6 components together form the GM indicator.

In addition to income inequality, we also observe how institutional quality and other selected variables affect changes in income distribution. Specifically, we observe how institutional quality and other selected variables affect the class of the rich (5th quintile), the poor (1st quintile) and the middle class (the share of the 2nd–4th quintile in total income). The same division ("rich", "middle class", "poor") is also used by Hurley et al. (2013), Barro (1999), Atkinson and Brandolini (2013) and others; this is a standard division. The middle class is the widest and expresses the middle between the poor or at risk of poverty (lower 20%) and the upper 20% (who express "rich").

The model has the following form:

$$\begin{aligned} \text{id}_{it} = \quad & \alpha + \beta_1 \text{ GM}_{it} + \beta_2 \text{ GDP}_{it} + \\ & + \beta_3 \text{ educ}_{it} + \beta_4 \text{ inv}_{it} + \\ & + \beta_5 \text{ open}_{it} + \beta_6 \text{ gov}_{it} + \epsilon_{it}, \end{aligned} \tag{3}$$

where income distribution (id) represents either the rich (5th quintil), the poor (1st quintil) or the middle class (2nd–4th quintil). We also explain changes in income distribution (id) using the components of the institutional quality indicator, see the following equations:

$$\begin{aligned} \text{id}_{it} = \quad & \alpha + \beta_1 \text{ GM1}_{it} + \beta_2 \text{ GM2}_{it} + \\ & + \beta_3 \text{ GM3}_{it} + \beta_4 \text{ GM4}_{it} + \\ & + \beta_5 \text{ GM5}_{it} + \beta_6 \text{ GM6}_{it} + \epsilon_{it}. \end{aligned} \tag{4}$$

To provide the robust results, we estimate all models with robust standard errors clustered by country. Since we are working with an unbalanced panel, according to Brown et al. (2011) we can take into account unmeasured heterogeneity with this approach. Like Chong and Calderón (2000), in addition to pooled OLS[2], we also use two-stages least squares (2SLS) due to possible endogeneity, where the lagged explanatory variables serve as instrumental variables. After the presentation of the main results, a robustness check is presented, where the main explanatory variable changes (another institutional variable is employed).

[2]The choice of Pooled OLS is based on the results of the LM test (between OLS and RE), the test of intercepts (between OLS and FE) and the Hausman test (choice between FE and RE).

4 RESULTS

4.1 Regression Analysis

The results of the regression analysis in Tab. 1 imply a positive statistically significant relationship between the improvement of institutional quality (GM) and the growth of income inequality (gini), which is a similar conclusion as in Chong and Calderón (2000). Income inequality is also increased by other variables involved in the model, namely net investment in government nonfinancial assets, while rising government spending on education (educ) reduces income inequality, which is consistent with Acemoğlu and Robinson (2000). The growth in the share of exports and imports in GDP (open) also has a negative effect, implying that greater international trade due to deepening globalization does not necessarily mean widening inequalities (versus Dabla-Norris et al., 2015). In addition to the price of factors of production in a given country, international trade is influenced by the institutional environment, the setting of which may be biased in favor of certain groups of workers. This is confirmed in Tab. 3, where we see that international trade favors the middle class and the 1st quintile (poor), while disadvantage the 5th quintil (rich).

Tab. 2 presents the various channels through which institutional quality affects income inequality. Improving the "Regulatory quality" (GM4), which aims to develop the private sector, and increasing political stability (GM2) seems to have the greatest weight. Voice and Accountability (GM1), in other words "quality of democracy", has a negative effect, which implies that improving the quality of democratic processes would lead to a reduction in income inequality, which is consistent with Acemoğlu et al. (2015), who explain that higher quality democratic processes are usually associated with a greater tendency to redistribute and reduce income inequality in society.

Tab. 1: Institutional quality (GM) and income inequality

	Pooled OLS	2SLS
	(1)	**(2)**
Variables	**gini**	**gini**
GM	3.928***	3.609***
	(0.440)	(0.536)
GDP	0.090	0.154
	(0.082)	(0.120)
educ	−2.987**	−4.008***
	(0.744)	(0.550)
inv	0.448**	0.487**
	(0.171)	(0.194)
open	−0.077***	−0.065***
	(0.016)	(0.014)
gov	0.204	0.338
	(0.218)	(0.320)
Constant	51.680***	52.893***
	(4.809)	(5.095)
Observations	78	73
R-squared	0.810	0.770

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tab. 2: Components of institutional quality (GM) and income inequality

	Pooled OLS	2SLS
	(1)	**(2)**
Variables	**gini**	**gini**
GM1	−8.792***	−10.644***
	(1.442)	(0.592)
GM2	1.477**	1.702***
	(0.474)	(0.550)
GM3	4.540	2.724
	(2.716)	(3.096)
GM4	17.468***	22.513***
	(4.046)	(2.849)
GM5	−7.644**	−8.785***
	(2.915)	(2.383)
GM6	−6.448***	−7.582***
	(1.457)	(1.456)
Constant	27.399***	26.213***
	(0.736)	(0.525)
Observations	91	88
R-squared	0.757	0.764

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tab. 3: Institutional quality (GH) and income distribution

	Pooled OLS	2SLS	Pooled OLS	2SLS	Pooled OLS	2SLS
Variables	**(1) quintil 5**	**(2) quintil 5**	**(3) quintil 2–4**	**(4) quintil 2–4**	**(5) quintil 1**	**(6) quintil 1**
GM	2.786***	2.577***	−1.427***	−1.306***	−1.363***	−1.265***
	(0.268)	(0.396)	(0.162)	(0.260)	(0.209)	(0.185)
GDP	0.067	0.125	−0.039	−0.086*	−0.027	−0.039
	(0.056)	(0.083)	(0.032)	(0.050)	(0.024)	(0.036)
educ	−2.325***	−3.099***	1.579***	2.095***	0.750**	1.018***
	(0.541)	(0.426)	(0.322)	(0.330)	(0.245)	(0.167)
inv	0.299*	0.251*	−0.170	−0.054	−0.129	−0.191***
	(0.127)	(0.130)	(0.094)	(0.099)	(0.070)	(0.062)
open	−0.065***	−0.057***	0.048***	0.044***	0.017**	0.013***
	(0.012)	(0.011)	(0.008)	(0.010)	(0.005)	(0.004)
gov	0.138	0.210	−0.070	−0.084	−0.068	−0.128
	(0.156)	(0.240)	(0.093)	(0.172)	(0.071)	(0.095)
Constant	56.522***	58.249***	39.926***	37.967***	3.527*	3.795**
	(3.405)	(3.810)	(2.022)	(2.598)	(1.488)	(1.499)
Observations	78	73	78	73	78	73
R-squared	0.810	0.768	0.801	0.746	0.781	0.749

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Tab. 4: Components of institutional quality (GM) and income distribution

	Pooled OLS	2SLS	Pooled OLS	2SLS	Pooled OLS	2SLS
Variables	**(1) quintil 5**	**(2) quintil 5**	**(3) quintil 2–4**	**(4) quintil 2–4**	**(5) quintil 1**	**(6) quintil 1**
GM1	−6.829***	−8.169***	4.585***	5.379***	2.245***	2.815***
	(1.170)	(0.548)	(0.887)	(0.544)	(0.400)	(0.243)
GM2	0.968**	1.186***	−0.370	−0.543*	−0.606**	−0.635***
	(0.372)	(0.378)	(0.323)	(0.293)	(0.168)	(0.236)
GM3	3.885	2.691	−2.740*	−2.195	−1.102	−0.418
	(2.036)	(2.398)	(1.229)	(1.596)	(0.823)	(0.896)
GM4	13.804***	17.750***	−9.561***	−12.186***	−4.306**	−5.637***
	(3.016)	(2.207)	(1.927)	(1.587)	(1.120)	(0.659)
GM5	−6.832**	−8.208***	5.507***	6.921***	1.352	1.248
	(1.797)	(1.582)	(1.003)	(1.178)	(1.112)	(1.042)
GM6	−4.888***	−5.596***	3.051***	3.320***	1.846**	2.300***
	(1.005)	(1.002)	(0.740)	(0.719)	(0.575)	(0.587)
Constant	36.543***	35.617***	54.451***	55.065***	9.017***	9.328***
	(0.524)	(0.386)	(0.329)	(0.277)	(0.219)	(0.148)
Observations	91	88	91	88	91	88
R-squared	0.761	0.766	0.751	0.753	0.759	0.768

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Institutional quality (GM) is also statistically significant in explaining the development of income distribution (see Tab. 3). Better institutions seem to favor the richest in society (the share of the 5th quintile in total income), while they disadvantage (have a negative effect) the poorest and middle class. The current trend in income distribution in developed countries (especially in the US and Western Europe, see Acemoğlu and Autor, 2011 and Fonseca et al., 2018) is its polarization, where the share of the middle class in total income decreases, while the share of 1st and the 5th quintile grows. Our results do not indicate polarization of income distribution. Expenditure on education (educ) reduces the gap between the rich on the one hand and the middle class and the poorest on the other and the similar effect has also international trade (open).

Tab. 4 presents influence of the individual components of institutional quality on the income distribution. It seems that regulatory quality (GM4) and political stability (GM5) in particular create an environment that allows the richest (5th quintile) to get the most rich, while negatively affecting the middle class and the poorest, which is probably the reason for the positive relationship between growth of institutional quality (GM) and the growth of income inequality (gini coefficient) showed in Tab. 1.

4.2 Robustness Check

Following the same approach as Kotschy and Sunde (2017), we will use a different indicator of institutional quality to determine the robustness of the results. Instead of the Governance matters (GM), we will use the indicator of Economic Freedom index developer by Heritage Foundation "HERI" (Heritage Foundation, 2020). Similar as GM consists of six components, HERI consists of 11 components.[3] GM took values $[-2.5; 2.5]$, while HERI $[0; 100]$; in both cases, the higher the value of the coefficient, the higher the quality of the institutional environment. While the GM indicator assessed the institutional quality rather from the perspective of political institutions, HERI and its components also include the evaluation of economic institutions (especially the definition of property rights and institutions related to the markets).

The results obtained by changing the main explanatory variable to HERI are similar to those of GM. The same variables as were statistically significant for the GM model are significant now; the signs of statistically significant parameters remained the same. What is different is the value of the coefficient for the institutional variable. Here, too, the results imply that a higher institutional quality (HERI) leads to an increase in income inequality (Tab. 5) and favors the richest population over the middle class and the poorest (Tab. 6).

Tab. 5: Institutional quality (HERI) and income inequality

	Pooled OLS	2SLS
Variables	**(1) gini**	**(2) gini**
HERI	0.275***	0.254***
	(0.040)	(0.043)
GDP	0.080	0.106
	(0.078)	(0.098)
educ	−2.571**	−3.495***
	(0.794)	(0.666)
inv	0.550**	0.607***
	(0.148)	(0.216)
open	−0.078***	−0.067***
	(0.017)	(0.015)
gov	0.170	0.335
	(0.232)	(0.277)
Constant	33.681***	35.258***
	(3.925)	(3.275)
Observations	78	73
R-squared	0.824	0.794

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Our results imply a positive statistically significant relationship between institutional quality and income inequality. Given that the quality of the institutional environment has been examined above in a linear form, the relationship appears to be monotonic. However, the studies Law et al. (2014) and Chong and Calderón (2000) imply a nonmonotonic relationship between institutional quality and income inequality. Therefore, we add the square of the institutional quality indicator to the model explaining income inequality, see Tab. 7 and Fig. 1 (and Fig. 2 in the Annex). The results imply a quadratic relationship between institutional quality and income inequality. After reaching a certain level of institutional quality, income inequality seems to be declining. This implies an inverted-U curve between

[3]More specifically, HERI "Overall Score" consists from: Property Rights, Judicial Effectiveness, Government Integrity, Tax Burden, Government Spending, Fiscal Health, Business Freedom, Labor Freedom, Monetary Freedom, Trade Freedom, Investment Freedom and Financial Freedom.

Tab. 6: Institutional quality (HERI) and income distribution

	Pooled OLS	2SLS	Pooled OLS	2SLS	Pooled OLS	2SLS
Variables	**(1) quintil 5**	**(2) quintil 5**	**(3) quintil 2–4**	**(4) quintil 2–4**	**(5) quintil 1**	**(6) quintil 1**
HERI	0.196***	0.181***	−0.102***	−0.091***	−0.095***	−0.089***
	(0.026)	(0.032)	(0.014)	(0.021)	(0.017)	(0.014)
GDP	0.060	0.091	−0.035	−0.069*	−0.023	−0.023
	(0.054)	(0.067)	(0.031)	(0.041)	(0.023)	(0.028)
educ	−2.027**	−2.735***	1.419**	1.913***	0.611*	0.835***
	(0.584)	(0.516)	(0.355)	(0.379)	(0.250)	(0.192)
inv	0.372**	0.337**	−0.207**	−0.098	−0.164*	−0.233***
	(0.099)	(0.144)	(0.077)	(0.100)	(0.066)	(0.070)
open	−0.066***	−0.058***	0.049***	0.044***	0.017**	0.014***
	(0.013)	(0.012)	(0.009)	(0.011)	(0.005)	(0.004)
gov	0.112	0.208	−0.054	−0.084	−0.059	−0.126
	(0.170)	(0.211)	(0.103)	(0.157)	(0.073)	(0.082)
Constant	43.748***	45.665***	46.487***	44.333***	9.758***	9.983***
	(3.009)	(2.731)	(2.242)	(2.484)	(1.057)	(0.887)
Observations	78	73	78	73	78	73
R-squared	0.823	0.792	0.811	0.766	0.795	0.771

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

institutional quality and income inequality. We get the same result even if we use HERI variable for institutional quality. This pattern seems to be consistent.

Tab. 7: Quadratic relationship between institutional quality and income inequality

	2SLS	2SLS
Variables	**(1) gini**	**(2) gini**
GM	3.287***	
	(0.575)	
GMsquared	−4.347***	
	(0.843)	
HERI		2.274***
		(0.340)
HERIsquared		−0.017***
		(0.003)
GDP	0.164	0.117
	(0.126)	(0.108)
educ	−4.861***	−3.726***
	(0.467)	(0.511)
inv	0.508**	0.288
	(0.232)	(0.189)
open	−0.026	−0.045***
	(0.017)	(0.016)
gov	0.564	0.404
	(0.362)	(0.278)
Constant	51.559***	−25.125**
	(6.957)	(10.387)
Observations	73	73
R-squared	0.718	0.814

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

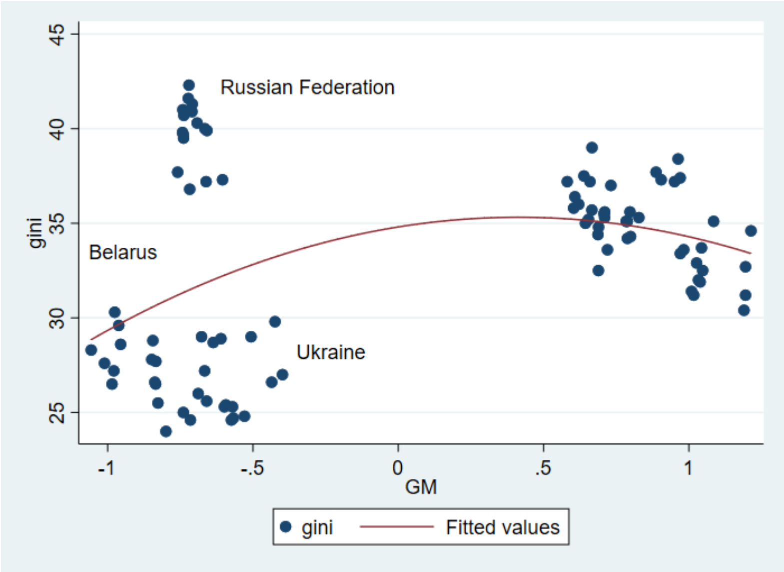

Fig. 1: Institutional quality (GM) and income inequality
Note: The x-axis represents the institutional quality (indicator GM) – the higher the number, the higher the institutional quality. The y-axis represents income inequality – the higher the number, the higher the income inequality.

5 DISCUSSION AND CONCLUSIONS

In this paper, we focused on identifying the relationship between institutional quality and income inequality (and income distribution) in selected post-Soviet countries. Technological change and globalization in particular are considered to be the main causes of growing income inequality in developed countries (Braconier et al., 2015; Dabla-Norris et al., 2015). On the other hand, the institutional environment can significantly influence the distribution of wealth in society (Acemoğlu and Robinson, 2002). The results of this paper show that the quality (and setting) of the institutional environment plays an important role in this process.

The relationship between institutional quality and income inequality seems to be non-monotonic in light of the achieved results. The growing quality of institutions leads to an increase in income inequality, but only until a certain point in time. From a certain point in time, the growth of institutional quality will start to stimulate a decline in income inequality. This relationship can be characterized as an inverted U-shape. A similar conclusion was reached by Chong and Calderón (2000), who, however, used measures from the International Country Risk Guide (ICRG) and the Business Environmental Risk Intelligence (BERI) as an indicators of institutional quality. Both indicators have a similar character to the indicators used in this article – both consist of sub-components and the average value of sub-components forms the main index that assesses institutional quality.

In connection with this result, an analogy is offered to the Kuznets curve, or to its modification by Acemoğlu and Robinson (2000). The authors confuse the original relationship between income inequality and income per capita with the relationship between income inequality and democratization, and thus offer a convincing explanation for developments within Western countries. The increase in income inequality in Western countries was usually associated with industrialization, which aroused social unrest in society. Political elites responded by expanding the right to vote (deepening democratization) to prevent deeper social unrest. Deepening democratization is associated with higher taxation and redistribution, which leads to a reduction in income inequality (Acemoğlu and Robinson, 2000). The finding of the relationship (inverted U-curve) can thus have a similar telling power.

Furthermore, the results of this paper imply, similarly to Chong and Calderón (2000) and Chong and Gradstein (2017), that the growing institutional quality in selected post-Soviet countries favors the richest social class over the poorest and middle class. Chong and Calderón (2000) argue that the growing quality of the institutional environment favors the richest in society at an early stage of institutional reform; after a while, a better institutional environment will start to generate a more equal environment. Institutional quality as a factor favoring "rich" over other social classes is consistent with Sonin (2003), who found that in Russia the rich benefit from the ability to shape economic institutions for their profit. His findings were based on observations of transformation and oligarchism during the "wild" 1990s; some trends seem to persist.

Currently, the main trend in developed countries is the polarization of income distribution (Dabla-Norris et al., 2015), which is caused by the polarization of employment, see especially Goos et al. (2009) and Acemoğlu and Autor (2011). The above results do not indicate this phenomenon. This may be due to the fact that, according to empirical evidence, employment polarization does not appear to occur in these countries, see Hurley et al. (2013) focus on the Baltic countries between 1997–2010 and Gimpelson and Kapeliushnikov (2016) examining Russia between 2000–2012.

After the collapse of the Soviet Union, the post-Soviet economies began to transform; during the period 2002–2017, the post-Soviet countries had a substantial part of the economic transformation behind them. Some of them (Baltic countries) have taken the "Western route" (democratization, growth of institutional quality and, as a result, higher economic performance), some of them seem to have returned to the idea of the Soviet Union (especially Belarus and Russian Federation). The result is a relatively low institutional quality, which, however, as in the case of Belarus and Ukraine, does not yet put upward pressure on growing income inequality. Conversely, in the case of Russia, we can observe a relatively high level of income inequality, however, due to poor institutional quality, the downward pressure through higher redistribution is not as great as in Estonia, where gini changed from 37.2 in 2003 to 30.4 in 2017; in Russia, gini changed from 37.3 (2003) to 37.2 (2017). The example of Russia confirms that income inequality is more permanent in countries with extract institutions (Dobre et al., 2019), which are institutions that favor elites (Acemoğlu and Robinson, 2012).

The potential for further research lies in the identification of channels through which institutional quality influences the development of income inequality. The results of this paper imply that the channels are mainly the "regulatory quality" and "political stability". Are the channels through which institutional quality affects income inequality in the former post-Soviet countries different from those in other (for example Central Europe) countries? There is also space for determining the right direction of the relationship. Chong and Gradstein (2007, 2017) have shown that the link income inequality – institutional quality is stronger than the link institutional quality – income inequality. Institutional quality then forms a channel through which other influences act. De Haan and Sturm (2017) report the effects of financial variables – financial development, financial liberalization or the banking crisis – all stimulate an increase in income inequality through institutional quality. Law et al. (2014), on the other hand, add that there is treshold effect of the institutional quality – from a certain level of institutional quality, financial indicators affect to a more even distribution of income. However, there is a lack of greater empirical evidence in this area, which may be a motivation for further research.

6 ACKNOWLEDGEMENT

Supported by IGA MENDELU "The routine-biased technological change hypothesis as a cause of job polarization in the Czech Republic" (PEF_DP_2020018).

7 REFERENCES

Acemoğlu, D. 2012. Introduction to Economic Growth. *Journal of Economic Theory*, 147 (2), 545–550. DOI: 10.1016/j.jet.2012.01.023.

Acemoğlu, D. and Autor, D. 2011. Skills, Tasks and Technologies: Implications for Employment and Earnings. In Ashenfelter, O. and Card, D. (eds.). *Handbook of Labor Economics*, Volume 4, Part B, Chapter 12, pp. 1043–1171. DOI: 10.1016/S0169-7218(11)02410-5.

Acemoğlu, D. and Robinson, J. A. 2000. Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *The Quarterly Journal of Economics*, 115 (4), 1167–1199. DOI: 10.1162/003355300555042.

Acemoğlu, D. and Robinson, J. A. 2002. The Political Economy of the Kuznets Curve. *Review of Development Economics*, 6 (2), 183–203. DOI: 10.1111/1467-9361.00149.

Acemoğlu, D., Johnson, S. and Robinson, J. A. 2005. Institutions as a Fundamental Cause of Long-Run Growth. In Aghion, P. and Durlauf, S. N. (eds.). *Handbook of Economic Growth*, Vol. 1, Part A, Chapter 6, pp. 385–472. DOI: 10.1016/S1574-0684(05)01006-3.

Acemoğlu, D. and Robinson, J. A. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Currency.

Acemoğlu, D., Naidu, S., Restrepo, P. and Robinson, J. A. 2015. Democracy, Redistribution, and Inequality. In Atkinson, A. B. and Bourguignon, F. (eds.). *Handbook of Income Distribution*, Vol. 2, Part IV, Chapter 21, pp. 1885–1966. DOI: 10.1016/B978-0-444-59429-7.00022-4.

Atkinson, A. B. and Brandolini, A. 2013. On the Identification of the "Middle Class". In Gornick, J. C. and Jäntti, M. (eds.). *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*, Part 2, pp. 77–100.

Barro, R. J. 1999. Determinants of Democracy. *Journal of Political Economy*, 107 (6), 158–183. DOI: 10.1086/250107.

Braconier, H., Nicoletti, G. and Westmore, B. 2015. *Policy Challenges for the Next 50 Years*. OECD Economic Policy Paper No. 9. DOI: 10.1787/2226583X.

Brown, D. S., Touchton, M. and Whitford, A. 2011. Political Polarization as a Constraint on Corruption: A Cross-National Comparison. *World Development*, 39 (9), 1516–1529. DOI: 10.1016/j.worlddev.2011.02.006.

Chong, A. and Calderón, C. 2000. Institutional Quality and Income Distribution. *Economic Development and Cultural Change*, 48 (4), 761–786. DOI: 10.1086/452476.

Chong, A. and Gradstein, M. 2007. Inequality and Institutions. *The Review of Economics and Statistics*, 89 (3), 454–465. DOI: 10.1162/rest.89.3.454.

Chong, A. and Gradstein, M. 2017. Political and Economic Inequities and the Shaping of Institutions and Redistribution. *Southern Economic Journal*, 83 (4), 952–971. DOI: 10.1002/soej.12206.

Clark, C. M. A. and Kavanagh, C. 1996. Basic Income, Inequality, and Unemployment: Rethinking the Linkage between Work and Welfare. *Journal of Economic Issues*, 30 (2), 399–406. DOI: 10.1080/00213624.1996.11505803.

Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F. and Tsounta, E. 2015. *Causes and Consequences of Income Inequality: A Global Perspective*. International Monetary Fund.

de Haan J. and Sturm, J.-E. 2017. Finance and Income Inequality: A Review and New Evidence. *European Journal of Political Economy*, 50 (C), 171–195. DOI: 10.1016/j.ejpoleco.2017.04.007.

Dobre, I., Jianu, I., Bodislav, D. A., Rădulescu, C. V. and Burlacu, S. 2019. The Implications of Institutional Specificities on the Income Inequalities Drivers in European Union. *Economic Computation and Economic Cybernetics Studies and Research*, 53 (2), 59–76. DOI: 10.24818/18423264/53.2.19.04.

Engerman, S. L. and Sokoloff, K. L. 2002. *Factor Endowments, Inequality, and Paths of Development Among New World Economics*. NBER Working Paper No. 9259. DOI: 10.3386/w9259.

Fonseca, T., Lima, F. and Pereira, S. C. 2018. Job Polarization, Technological Change and Routinization: Evidence for Portugal. *Labour Economics*, 51 (C), 317–339. DOI: 10.1016/j.labeco.2018.02.003.

Gimpelson, V. and Kapeliushnikov, R. 2016. Polarization or Upgrading? Evolution of Employment in Transitional Russia. *Russian Journal of Economics*, 2 (2), 192–218. DOI: 10.1016/j.ruje.2016.06.004.

Goos, M., Manning, A. and Salomons, A. 2009. Job Polarization in Europe. *American Economic Review*, 99 (2), 58–63. DOI: 10.1257/aer.99.2.58.

Heritage Foundation. 2020. *Economic Data and Statistics on World Economy and Economic Freedom*. [online]. Available at: `https://www.heritage.org/index/explore?view=by-region-country-year&u=637383304012865023`.

Hurley, J., Fernández-Macías, E. and Storrie, D. 2013. *Employment Polarisation and Job Quality in the Crisis: European Jobs Monitor 2013*. Dublin: Eurofound.

Josifidis, K., Supić, N. and Beker-Pucar, E. 2017. Institutional Quality and Income Inequality in the Advanced Countries. *Panoeconomicus*, 64 (2), 169–188. DOI: 10.2298/PAN1702169J.

Kraay, A., Kaufmann, D. and Mastruzzi, M. 2010. *The Worldwide Governance Indicators: Methodology and Analytical Issues*. The World Bank. DOI: 10.1596/1813-9450-5430.

Law, S. H., Tan, H. B. and Azman-Saini, W. N. W. 2014. Financial Development and Income Inequality at Different Levels of Institutional Quality. *Emerging Markets Finance and Trade*, 50 (1), 21–33. DOI: 10.2753/REE1540-496X5001S102.

Náplava, R. 2017. The Importance of Institutional Quality for Economic Performance in Post-Soviet States. In *The International Scientific Conference INPROFORUM 2017*, pp. 159–164. ISBN 978-80-7394-667-8.

North, D. C. 1991. Institutions. *Journal of Economic Perspectives*, 5 (1), 97–112. DOI: 10.1257/jep.5.1.97.

Rodrik, D. 2008. *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*. Princeton University Press.

Sonin, K. 2003. Why the Rich May Favor Poor Protection of Property Rights. *Journal of Comparative Economics*, 31 (4), 715–731. DOI: 10.1016/j.jce.2003.09.005.

World Bank. 2020. *World Bank Open Data – Indicators*. [online]. Available at: `https://data.worldbank.org/indicator`.

8 ANNEX

Tab. 8: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
gini	91	32.92747	5.133573	24	42.3
5th quintil	91	40.92637	3.971813	34.6	48.9
2nd–4th quintil	91	51.45055	2.702442	45.3	55
1st quintil	91	7.621978	1.453486	5.7	10.5
GM	96	0.0553791	0.8001192	−1.0566	1.21377
GM1	96	0.032116	0.9992681	−1.77032	1.21328
GM2	96	0.0972846	0.7510905	−1.9618	1.01276
GM3	96	0.1203478	0.786467	−1.17045	1.18493
GM4	96	0.2456136	0.9978477	−1.63846	1.69814
GM5	96	−0.015453	0.8964147	−1.28912	1.36466
GM6	96	−0.147634	0.7317866	−1.09242	1.29365
HERI	96	60.60729	12.01734	39	79.1
GDP	96	3.563034	5.479402	−14.8	12.1
educ	85	5.133266	0.8082579	3.54787	7.31364
inv	90	1.769273	1.545836	0.135798	8.12587
open	96	110.3234	33.31385	46.5181	170.428
gov	96	18.24571	1.714696	13.4298	21.3793

Note: gini = gini coefficient; 5th quintil / 2nd–4th quintil / 1st quintil = “rich” / “middle class” / “poor”; GM/HERI = institutional quality; GDP = real GDP growth; educ = government expenditure on education; inv = net investment in government nonfinancial assets; open = openness of the country; gov = government final consumption expenditure

Tab. 9: Pairwise correlation

	gini	Upper20	Middle-class	Lowest20
GM	0.3205*	0.2427*	−0.0785	−0.5170*
GM1	0.2684**	0.2021*	−0.0612	−0.4397***
GM2	0.0345	−0.0469	0.2050*	−0.2521**
GM3	0.4482***	0.3809***	−0.2303**	−0.6121***
GM4	0.4473***	0.3802***	−0.2319**	−0.6086***
GM5	0.3291***	0.2514**	−0.0879	−0.5232***
GM6	0.1997*	0.1223	0.0343	−0.3965***
HERI	0.3792*	0.3064*	−0.1505	−0.5572*

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fig. 2: Institutional quality (HERI) and income inequality
Note: The x-axis represents the institutional quality (indicator HERI) – the higher the number, the higher the institutional quality. The y-axis represents income inequality – the higher the number, the higher the income inequality.

AUTHOR'S ADDRESS

Radek Náplava, Department of Economics, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: radek.naplava@mendelu.cz

MEASURING THE PUBLIC PERCEPTION OF THE EUROPEAN INTEGRATION PROCESS: EVIDENCE FROM THE UNITED KINGDOM AND GERMANY

Daniel Pastorek[1]

[1] *Mendel University in Brno, Czech Republic*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

This paper investigates a new method of measuring public perception of European integration policies. The methodology is based on a monthly frequency of news generating a negative attitude toward the EU integration process in newspapers from 2010:1 to 2018:12. The studies concerning similar topic are mostly based on survey data, which does not allow measuring dissatisfaction with individual political acts. The newly constructed indexes show that the identified spikes correspond to all major European integration events and can capture public disagreement attitude with the implementation of individual policies on a European or national level. The results are compared with the interest of individuals with the use of Google trend data analysis. Indexes may be useful to understand the non-economic cost in policy decision making, the same as in the question of the ambivalence of Europe.

KEY WORDS

public perception, integration, European Union, Google trends, media news

JEL CODES

E70, F02

1 INTRODUCTION

In the beginning, when the European project was mainly focused on trade liberalization, public opinion was viewed as not relevant to the process. Yet, as the European Union has evolved and covering a wide range of areas, public perception is viewed more important than ever.

In terms of the European integration process, it is useful and important to distinguish between two concepts of perception. The first represents support for ideas arising from the constitutional settlement of the European Union and second, the real way how the political elite execute their policies. If these actions do not

PASTOREK, Daniel. 2020. Measuring the Public Perception of the European Integration Process: Evidence from the United Kingdom and Germany. *European Journal of Business Science and Technology*, 6 (2): 113–126. ISSN 2694-7161, DOI 10.11118/ejobsat.2020.012.

meet public expectations, Scharpf (1999) discusses that this is endemic to any regime. Even this dissatisfaction with individual political acts might be the cause for a shift in support for European integration, however such proxy to capture the public's attitude is missing.

In this paper, newly constructed media-news based indexes are presented and are evaluated as relevant proxies for capturing public disapproval with European integration policies. The contribution of these indexes compared to the use of surveys is the flexibility and ability to capture disapproval with individual political actions and may explain the change in attitude for supporting the European integration project. We can also interpret these indexes as a proxy for non-economic costs in policy-decision making. Another interesting insight occurs when we compare news coverage related to the topic in one of the major EU countries, which was the first to leave the European project, and the country which is often referred as to being the driving force of the European Union.

In addition, this study examines the public attitude and interest of individuals separately. This paper is assuming an ambiguous result in the public perception of EU integration compared to individuals' interest. Individuals search volume should be increased in events, which could directly affect an individual's welfare. Therefore, different proxies are used to measure public perceptions and individual interest. To measure individual interest, Google trend analysis is used, which shows the frequency of search terms related to the topic of the EU. This data provides powerful insight, how people react to information intermediated through the media. Results show that individuals search differs from new indexes. While peaks in individuals search can be observed only in times of European parliament election and Brexit, the newly developed indexes reliably capture peaks in main events related to European integration. The data represents a period from 2010:1 to 2018:12.

When it analyzed the relation between perception of the EU and the real economy, studies suggest numerous economic performance indicators as a significant factor in Eurosceptic opinion shaping. But there are also several authors arguing that adverse expectations or "mood" are more significant variables in perception change than macroeconomic ones.

Finally, in this context, the paper examines the effect of public perception on several macroeconomic variables using the vector autoregression (VAR) models, with the use of response impulse functions. The results suggest that reactions from increased negative perception affects variables within a delay of 2–3 months, however, response shocks are within 90% of confidence intervals bands.

The paper is structured as follows. Section 2 presents the literature review. Section 3 presents a detailed overview of used data and methods. Section 4 presents and discuss the newly constructed indexes as well as indexes from Google trend analysis data. Section 5 provides a robustness analysis. Section 6 concludes this paper. The appendix contains results of VAR models.

2 LITERATURE REVIEW

Recent studies have primarily focused on explaining so-called Euroscepticism, mostly due to recent or still ongoing disintegration pressures (sovereign debt crisis, migration crisis, Brexit, …) and threat it as the dependent variable. Far fewer studies are devoted to the issue of the effects of public perception on European integration. This perception problem is complex, with different degrees of certainty. The simple one-dimensional approach concerning if there is public support toward the EU seems to be insufficient. The lack of this approach to inherent and complex attitude problem of EU integration suggest several studies (Baute et al., 2018; Boomgaarden et al., 2011; Hobolt and Brouard, 2011; Meijers and Zaslove, 2020; Stoeckel, 2013). The disproportion between various studies explaining the determinants of public perception and less attention paid to empirical support may be due to the methods used for measuring this concept.

To capture public opinion, Eurobarometer is still the dominant data source and allows cross-national comparison over a large period, even with the issue that several questions were adjusted over time. Despite some advantages of surveys, they are limited by periodicity and generality of questions. This data is useful to refer to approval or disapproval of the integration project, and the main challenges the Union is facing, however, the questions do not bring answers to more policy-oriented, multi-faceted opinions. Eurobarometer data do not allow measuring dissatisfaction with individual political acts.

Since then, it is not possible to obtain most of the information in person, just as it is time-consuming to follow the happening from first-hand sources, the news plays an important role in the process of public perception shaping. The importance of the media's influence on perception shaping is well known and studies suggest direct and indirect effects (e.g. Gunther and Storey, 2003; Wanta et al., 2004). Wanta et al. (2004) network analysis suggests that the more negative media coverage is received, the more likely a negative opinion is created. Interestingly, positive coverage had no effect. From an economic standpoint, public economic perception and media coverage is suggested by Barnes and Hicks (2018), Hobolt et al. (2017), Linn and Kellstedt (2004) and Soroka (2006, 2014), and economy perception and influence related to vote intentions and government elections suggest Soroka et al. (2015) and Nadeau et al. (1994).

In current times of disintegration pressures, public perception is important more than ever. It may be commonly seen in the growing politicization of integration challenges among the political parties (Hooghe and Marks, 2009) and studies confirm that attitude to the issues, as a determinant of vote choices (Hobolt and de Vries, 2015; Spoon, 2012). When authorities make their policy decisions, they are a priori uncertain and motivated by both economic and non-economic (political) objectives (Pástor and Veronesi, 2013). The public perception that differs from the objectives of politicians represents a non-economic cost. However, identification of these costs remains unresolved, same as there is far less attention to research concerning how public perception shape policy decision-making (Hobolt and de Vries, 2016).

Existing research suggests, that continuously updating information about economics is costly, and people update information about economics occasionally. Mostly when the economic condition is bad, or news tends to be frequent (Doms and Morin, 2004; Lamla and Lein, 2008). Another benefit of news analysis used in this paper is, that even European integration in it's current state involves a wide range of aspects, media news can capture many publicly interesting topics. Usefulness is also confirmed by the growing studies using news-based indexes (e.g., Demir et al., 2018; Drobetz et al., 2018; Gulen and Ion, 2016).

When it analyzed the relation between perception of the EU and real economy variables, Euroscepticism is mainly explained by sensitivity to economic inequalities. Moreover, cost-benefit analysis has been shown to influence support as well, whilst calculation refers to macroeconomic indicators like growth, inflation, unemployment, inflation, budget transfers, or trade relation with other EU members. Besides the effect of performance indicators in explaining changes in domestic EU support (e.g., Gomez, 2015; Kuhn et al., 2013; Nicoli, 2017; Serricchio et al., 2013), some studying argues that developments in other EU member countries can also have an impact. If the other countries have worsened fiscal and economic conditions, this may create a negative expectation and spill-over to the domestic economy (Ioannou et al., 2015; Kang and Oh, 2020). Though, as was mention above and is well known, the perception problem is complex. There are the increasing numbers of studies, that rejecting these economic variables as a main driving force of negative perceptions. They argue that perception change is mostly due to adverse financial expectations and fear of losing cultural identity. Ritzen et al. (2014) classified these factors as "mood variables".

Kang and Oh (2020) pointing out that public perception is essential in analyzing crisis and necessity to a comprehensive analysis of diverse factors combined with "public anxiety".

In this context, beside of the main objective of this article – to better identified public disapproval with individual political acts, this study examines if captured public attitude affect suggested macroeconomic performance indicators.

3 METHODS AND DATA

The indexes are constructed for two important European countries, these countries being the United Kingdom and Germany, within the period the from January 2010 to December 2018.[1] The methodology approach is based on measuring the frequency of articles in selected newspapers. To generate disapproval attitude with European integration policies is use automated textual search following a specific combination of keywords. The choice of words is to make clear references to the European Union (1), integration (2), disapproval attitude (3). The selected article must therefore fulfill the requirement of at least one word from each category (1; 2; 3). This methodology approach follows Baker et al. (2016), who used this approach to create their economic policy uncertainty index (EPU).

The used data is extracted from digital news archives of the ProQuest database. The newspapers for the UK are Financial Times, The Times, and The Daily Telegraph. Regarding Germany, the newspapers Die Walt and Die Tageszeitung are used. The categories and combination of keywords are following: (1) "Europe" or "European" or "EU" and (2) "integration" or "enlargement" and (3) "problem" or "uncertain" or "uncertainty" or "concern". For German newspapers, their translation is: (1) "Europa" or "EU" or "europäisch" and (2) "integration" or "Erweiterung" and (3) "problem" or "unsicher" or "Unsicherheit" or "Besorgnis" or "sorge".

To aggregate new indexes, collected data is adjusted as:

$$\text{freq}_{\,i,p,t} = \frac{n_{i,p,t}}{N_{p,t}}, \tag{1}$$

where n is the number of articles meeting the condition of keyword combination i, p with $1, \ldots, p$ represent selected newspaper, where P is a total number of newspapers, t with $1, \ldots, T$ denotes time, N represents the total number of articles published in the newspaper p, at time t. Then standard deviation, σ, is computed at time T, for each i. Subsequently expression (1) standardizes as:

$$\frac{\text{freq}_{\,i,p,t}}{\text{stdev}\,(\text{freq}_{\,i,p,t})}. \tag{2}$$

The values of expression (2) are averaged together for every newspaper p, month t, and finally, normalized to have a mean of 100.

To measure individual interest Google trend analysis is used. This data provides the frequency of the search term entered into a the Googles search engine, relatively scaled on a range of 0 to 100. Each data point is divided by the number of sum searches related to the topic. The extracted data represent a similar period from 2010:M1 to 2018:12 and is obtained from both the United Kingdom and Germany. The search term represents the topic – "European Union".[2]

Many of articles showed how the performance of macroeconomic variables affects the perception of European policies. With newly developed indexes, which capture public disapproval

[1] The limitation of the database is that the articles are listed with a certain time delay. In times of data collecting and due to caution purpose year 2019 was omitted.

[2] Since individuals search is unique, the chosen topic includes all events related to the European Union, even those that the media described negatively. That make it possible to compare, whether the intensity of negative news through media correlates with the increased in the magnitude of individual search and specific events.

with individual political actions, it is possible to examine it reversely, if this attitude has affects suggested macroeconomic variables. The studies using news-based indexes tend to assign negative news with financial frictions, which result in a higher capital cost. As a result, economic activity is depressed. To investigate the link between the perception of the European integration policies and the real economy, the vector autoregression model (VAR) is estimated with the following specification (perception index, cost of borrowing, inflation, industrial production). Data are transformed using annual percentage changes of all variables with a lag estimation of three, based on Akaike information criteria. The results are reported as response impulse functions of one standard deviation of the perception index with 90% confidence bands.[3]

4 EMPIRICAL RESULTS

This section presents new indexes as a proxy to capture the public disapproval toward the European integration policies in the United Kingdom (Fig. 1) and Germany (Fig. 2). The indexes identified spikes in major events related to the European integration, within increase was expected. However, there are differences between the examined countries.

In the perspective of the UK, the highest spike was identified after the result of the referendum in June 2016, when people voted to withdraw from the European Union. Even the referendum outcome showed 52% votes in favour of leaving, the rest represent a large part of the population which was unsatisfied, and it manifested in a series of disapproval protests. Subsequently, a series of difficult negotiations took place between the UK and the EU. The referendum was preceded by a block of EU treaty with a right of veto at the end of the year 2012, which aimed to take actions addressed to the debt crisis. According to many newspapers, the EU has been described as the most divided ever in its 54-year history. The European debt crisis has been a cause for concern in Europe since 2010, but the peak of the crisis came in 2012, same results are suggested by the index. This peak is comparable to the time of Brexit, which is interestingly much larger than in the Fig. 2 for Germany.

The level of identified disapproval in the case of Germany represents a comparable level to the UK, however, no significant peak is observed, not even during the European debt crisis. This can be explained by Germany being the strongest economy in the European Union and it did not have to face a problem directly, but on the contrary, it prospered. Another reason, possibly even more important, could be a more open approach policy to finding a solution to the ongoing problem. This changed at the end of 2015, since the average values of identified disapproval have more than doubled. The reason was the migration crisis which resonated the most. Germany was an attractive opportunity for migrants, and its open policy attitude meant that the nation opened borders for many more migrants than other countries. This immigration policy of a sensitive European issue divided Germany. On one hand, a large part of the public did not agree with the admission of migrants and the other criticized the insufficient involvement of other European Union countries. This period was accompanied by a series of protests, which subsequently resulted in the German government crisis in 2018.

We can observe that the results of the indexes confidently describe the actual happening on a European or national level, related to the integration process. Individual peaks captured all the events that resonated the most in the last decade. Since the integration process is common, the spill-over effect is expected, for example Brexit. Still, we can observe that negative attitudes toward the EU differ in time, same as in the peak levels. In the UK, two peaks are identified, which preceded Brexit, and they

[3] Computation of VAR models, same as response impulse functions figures, are made in Matlab software.

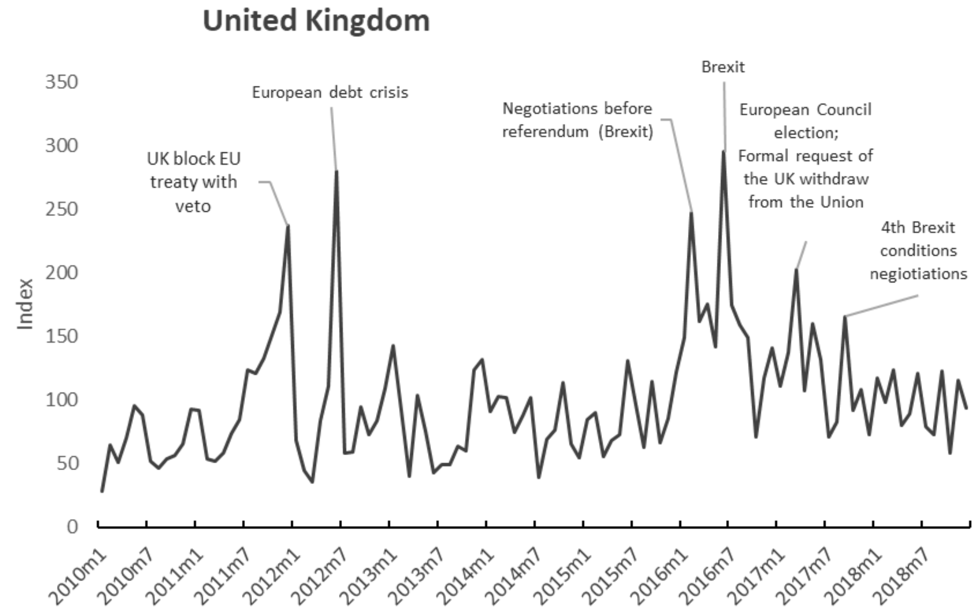

Fig. 1: Public perception index (UK)
Notes: The index represents the normalized frequency of media articles from Financial Times, The Times, The Daily Telegraph containing words (1) "Europe" or "European" or "EU" and (2) "integration" or "enlargement" and (3) "problem" or "uncertain" or "uncertainty" or "concern". The index is at the monthly frequency from 2010:M1 to 2018:12 with a mean of 100.

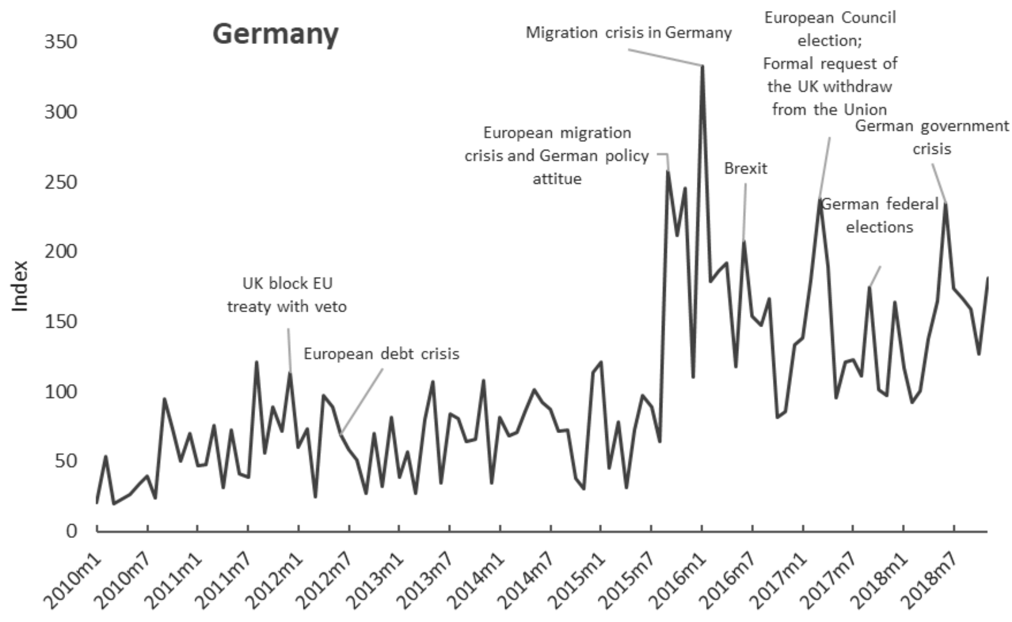

Fig. 2: Public perception index (DE)
Notes: The index represents the normalized frequency of media articles from Die Walt and Die Tageszeitung containing words (1) "Europa" or "EU" or "europäisch" and (2) "integration" or "Erweiterung" and (3) "problem" or "unsicher" or "Unsicherheit" or "Besorgnis" or "sorge". The index is at the monthly frequency from 2010:M1 to 2018:12 with a mean of 100.

describe a block of the EU treaty with the use right of veto on the issue of a joint solution to the European debt crisis and the peak of that crisis. On the contrary, in Germany, the first peak was observed on policy action to the issue of the migrant crisis. The divergence within Germany can also be seen on the long-term level increase of public disapproval. This event afterwards culminated in the German government crisis.

The description of the sequence of events above does not claim that the peaks mean a negative future progression in the EU integration process, but identification of these peaks can be a useful warning indicator of a possible shift in public perception. The shift in public perception may subsequently represent a rise or decline of non-economic cost in policy decision-making, depending on the political beliefs of the individual country's leaders toward the European Union project.

To extend empirical analysis is to investigate a possible link between the perception of the European integration policies and the real economy. The results are reported as responses of the perception index, costs of borrowing, inflation, and industrial production to one standard deviation perception shock. Separately for both countries (see Fig. 7 for UK and Fig. 8 for DE in the Annex). An increase in negative perception suggests a decrease in the costs of borrowing and a reduction in industrial production within a delay of 2–3 month. The effect on inflation is not clear. However, the shock responses to selected variables are within 90% confidence interval bands and therefore cannot be considered statistically significant.

The newly constructed indexes follow literature review and focus on negative coverage by the news and interpretation focusing on peaks, which are associated with expectation update, more likely in a negative way. The indexes are presented as a proxy for public perception, however, the effect on the individuals' interest in the topic of the European integration may not be clear. Updating information about the economy is costly and an ambiguous result in the public perception of EU integration compared to individuals' interest is expected. Individual search volume should be increased in events, which could directly affect an individual's welfare. To confirm this assumption, data obtained from Google trend analysis is presented, which shows the popularity of searches related to European integration topics in the United Kingdom (Fig. 3) and Germany (Fig. 4).

The results confirm differences between individual search and newly constructed indexes. While new indexes capture peaks in major European events, individual search volume is increased only in times of the European parliament election and Brexit, with one additional peak at the end of a period in Germany. The two most significant peaks are common for the UK and Germany. Brexit is the dominating one, naturally much larger in the UK than Germany. Nation withdrawal from the EU is an unprecedent event and represents an uncertain economic outlook for many citizens. The EU election may represent an increase in the effort to obtain additional information relevant to a direct decision of individuals. Despite this available data, the individual search seems not to be a reliable proxy in the matter of public perception concern, however, they bring interesting insights with comparison to media coverage.

5 ROBUSTNESS ANALYSIS

As mentioned above, the methodology is based on Baker et al. (2016), who they used to construct the economic-policy uncertainty index. Having a shared methodology approach and a possible similarity of some data, since the newly constructed indexes capture wide-range news concerning the European integration, the indexes will be compared together as part of the robustness analysis. First, a comparison of these two indexes is presented and afterwards, for greater detail, a table containing correlation coefficients (Tab. 1) is provided. In

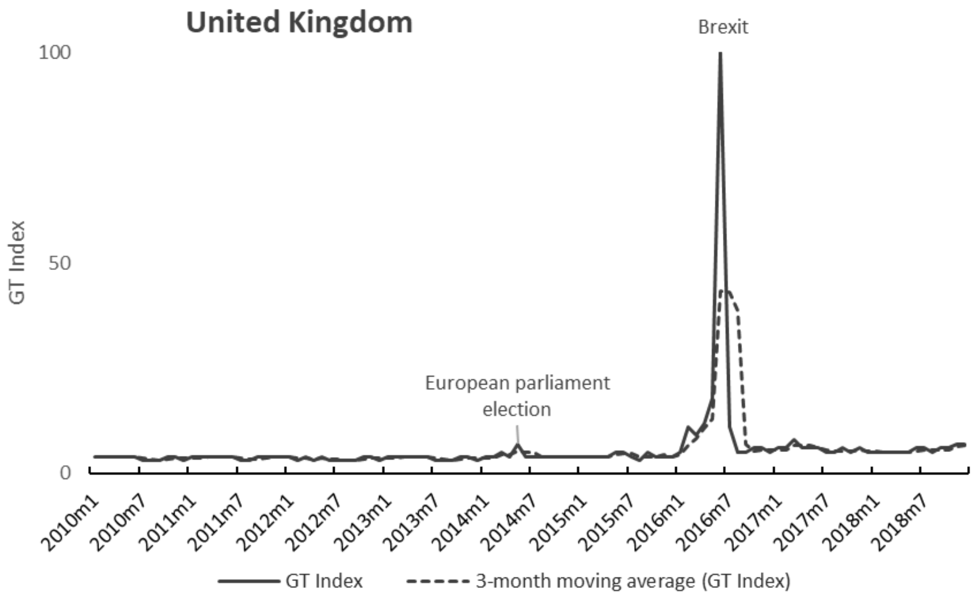

Fig. 3: Google trend analysis "European Union topic" (UK)
Notes: GT (Google Trend) index for topic "European Union" in the United Kingdom. Each data point is divided by the number of sum searches related to the topic and relatively scaled on a range of 0 to 100.

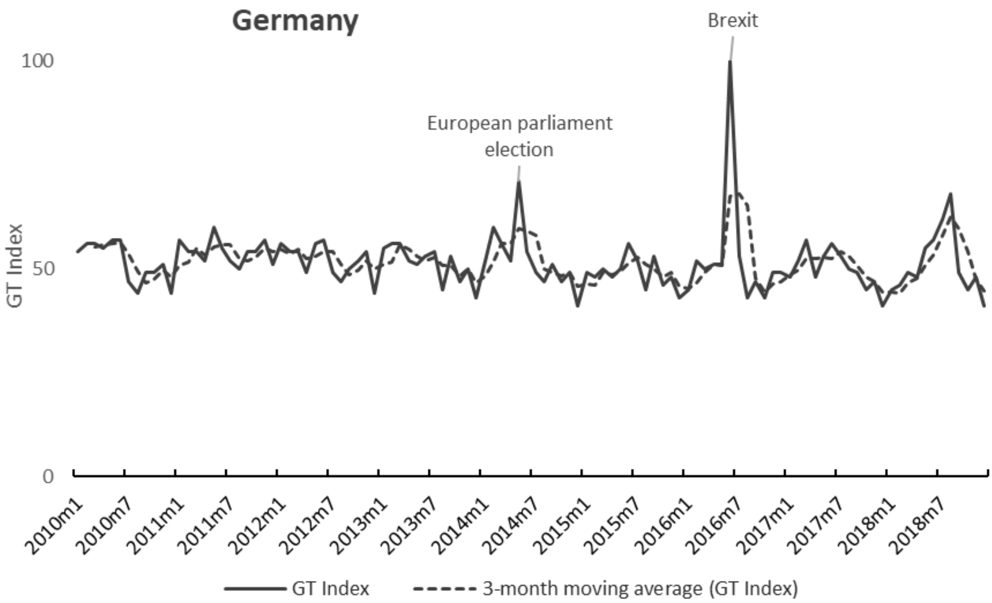

Fig. 4: Google trend analysis "European Union topic" (DE)
Notes: GT (Google Trend) index for topic "European Union" in Germany. Each data point is divided by the number of sum searches related to the topic and relatively scaled on a range of 0 to 100.

addition, the table shows the degree of intensity of correlations before and after Brexit. This event directly affects the European integration process and could represent a change in the intensity of identified rates. The results from the comparison vary at the level of individual states and are presented following as Germany (Fig. 5) and the United Kingdom (Fig. 6).

In the case of Germany, the constructed index represents a constantly lower level compared to EPU until 2015:9. At this time, the mentioned migration crisis started resonating. From this point, the level exceeds the alternative index. The United Kingdom indicates a stronger relation compared to the EPU index. Despite this similarity, it shows significant differences in events such as the UK treaty block and a peak of the sovereign debt crisis. These results are not surprising, since these events are labelled as the events that have divided the European Union the most.

Compared indexes differ in key aspects related to the focused objectives. Despite that, there are several similarities, newly constructed indexes seem to better identify negative public perception toward EU integration and thus fulfil their objectives. The following table provides more detail to the description above and shows the correlation between new indexes and EPU. Results confirm a stronger overall correlation for the UK and in the case of Germany, a stronger correlation after Brexit is seen. This unprecedented situation increases economic-policy uncertainty as much as an undisputed intervention for the European integration process. This lower correlation of the UK with index EPU after Brexit is not surprising. The economic policy uncertainty is expected to be high due to withdraw of UK and European matters are naturally becoming less essential. This may be seen in the declining trend in the level of perception index. The levels of EPU index remain higher, most likely because of still ongoing negotiations between the UK and EU.

Tab. 1: Correlation coefficients between index EPU and newly constructed indexes

	New indexes	
	UK	**DE**
EPU – Whole period		
2010:1–2018:12	0.4734*** (0.0000)	0.2930** (0.0021)
EPU – Brexit		
2010:1–2016:5	0.4504*** (0.0000)	0.1219 (0.2911)
2016:6–2018:12	0.4556* (0.0100)	0.3244* (0.0750)

Notes: Correlation coefficients between newly constructed indexes and index EPU. The table contains a correlation for the whole period and before and after Brexit. Data for the United Kingdom and Germany. Stars indicate significance level for p-value < 0.05 (*), p-value < 0.01 (**) and p-value < 0.001 (***).

6 CONCLUSIONS

This paper investigates a new method of measuring public perception toward European integration policies. The methodology is based on a monthly frequency of news generating a negative attitude toward the EU integration in newspapers. The data is presented for the United Kingdom and Germany in the period from January 2010 to December 2018. The results show that the spikes correspond to major European integration events and reliably describing European happening. The indexes are able to capture the perception of individuals to a wide range of events from economic challenges to the individual social policies of the European leaders. The newly developed indexes are compared with the trend in individuals search volume by Google trend data analysis. The most dominating peak by individual interest was identified only in times of Brexit. Thus, individual interest is unlikely to explain changes in the perception of the EU compared to the use of the new methodology. This paper finds a particular correlation between developed indexes and the widely used economic-policy uncertainty index (Baker et al., 2016) but peaks related to European integration are more

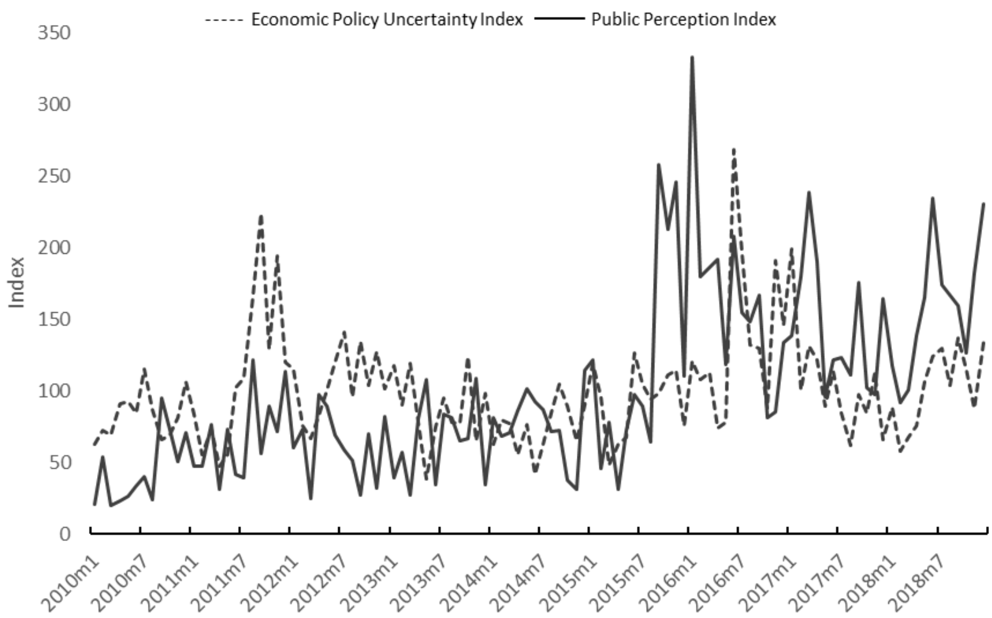

Fig. 5: Comparison with alternative index EPU (GER)
Notes: The index represents the normalized frequency of media articles from Financial Times, The Times, The Daily Telegraph containing words (1) “Europe” or “European” or “EU” and (2) “integration” or “enlargement” and (3) “problem” or “uncertain” or “uncertainty” or “concern”. The index is at the monthly frequency from 2010:M1 to 2018:12 with a mean of 100. The economic policy uncertainty index is from Baker et al. (2016).

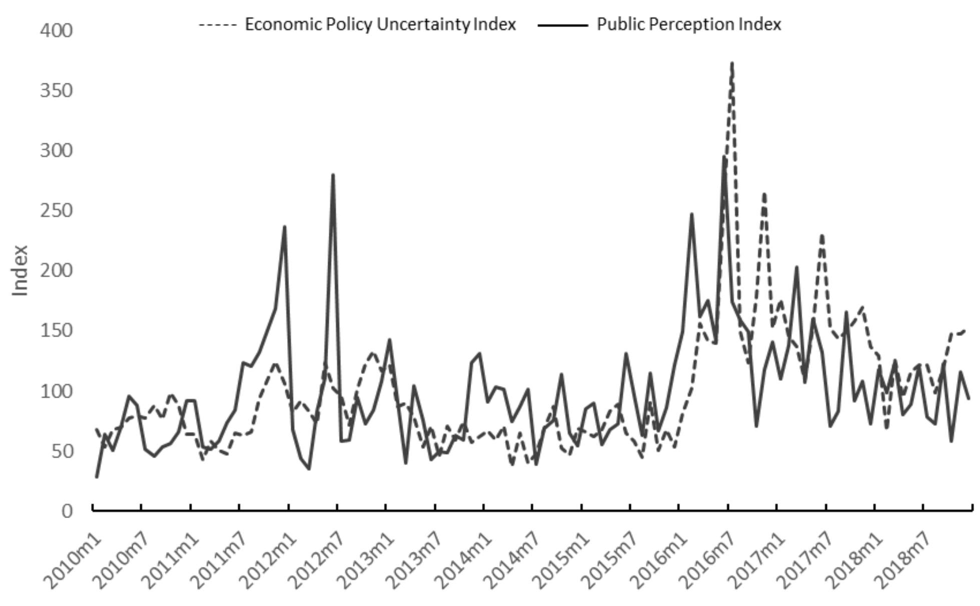

Fig. 6: Comparison with alternative index EPU (UK)
Notes: The index represents the normalized frequency of media articles from Die Walt and Die Tageszeitung containing words (1) “Europa” or “EU” or “europäisch“ and (2) “integration” or “Erweiterung” and (3) “problem” or “unsicher” or “Unsicherheit” or “Besorgnis” or “sorge”. The index is at the monthly frequency from 2010:M1 to 2018:12 with a mean of 100. The economic policy uncertainty index is from Baker et al. (2016).

pronounced in newly presented indexes. These differences are mainly in European important events (e.g. UK treaty block, migration crisis which resonated the most in Germany, negotiations before Brexit, ...). Despite well-known index EPU and its focus, newly developed index capturing all expected spikes related to the topic of European integration.

The indexes provide an alternative source for measuring public perception and the main benefit is the possibility of empirical capturing of public disagreement attitude with the implementation of individual policies pursued by authorities on a European or national level. This identification of public reaction to policies may help to alleviate upcoming disapproval peaks. When authorities make their policy decision, they are motivated by economic and non-economic costs. These non-economic costs represent the different public perception of the issue, as can be identified from the indexes. Another insight is provided within a comparison between nations, which may provide data on the question of the ambivalence of European states toward the common project.

These indexes open several paths for future research. For example, it would be interesting to extend the empirical analysis to propose different channels through public perception that could affect the real economy. However, it can also be useful in fulfiling the empirical gap in this specific topic.

7 ACKNOWLEDGEMENT

This research was funded by Internal IGA project no. PEF_TP_2020008. The Impact of Economic Crises and Financial Market Uncertainty on Economic Policy in Global Environment at Mendel University in Brno, Faculty of Business and Economics.

8 REFERENCES

Baker, S. R., Bloom, N. and Davis, S. J. 2016. Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131 (4), 1593–1636. DOI: 10.1093/qje/qjw024.

Barnes, L. and Hicks, T. 2018. Making Austerity Popular: The Media and Mass Attitudes toward Fiscal Policy. *American Journal of Political Science*, 62 (2), 340–354. DOI: 10.1111/ajps.12346.

Baute, S., Meuleman, B., Abts, K. and Swyngedouw, M. 2018. Measuring Attitudes Towards Social Europe: A Multidimensional Approach. *Social Indicators Research*, 137 (1), 353–378. DOI: 10.1007/s11205-017-1587-3.

Boomgaarden, H. G., Schuck, A. R. T. and Elenbaas, M. and de Vreese, C. H. 2011. Mapping EU Attitudes: Conceptual and Empirical Dimensions of Euroscepticism and EU Support. *European Union Politics*, 12 (2), 241–266. DOI: 10.1177/1465116510395411.

Demir, E., Gözgör, G., Lau, C. K. M. and Vigne, S. A. 2018. Does Economic Policy Uncertainty Predict the Bitcoin Returns? An Empirical Investigation. *Finance Research Letters*, 26 (C), 145–149. DOI: 10.1016/j.frl.2018.01.005.

Doms, M. E. and Morin, N. J. 2004. *Consumer Sentiment, the Economy, and the News Media*. FRBSF Working Paper No. 2004-09. DOI: 10.2139/ssrn.602763.

Drobetz, W., El Ghoul, S., Guedhami, O. and Janzen, M. 2018. Policy Uncertainty, Investment, and the Cost of Capital. *Journal of Financial Stability*, 39 (C), 28–45. DOI: 10.1016/j.jfs.2018.08.005.

Gomez, R. 2015. The Economics Strikes Back: Support for the EU during the Great Recession. *Journal of Common Market Studies*, 53 (3), 577–592. DOI: 10.1111/jcms.12183.

Gulen, H. and Ion, M. 2016. Policy Uncertainty and Corporate Investment. *The Review of Financial Studies*, 29 (3), 523–564. DOI: 10.1093/rfs/hhv050.

Gunther, A. C. and Storey, J. D. 2003. The Influence of Presumed Influence. *Journal of Communication*, 53 (2), 199–215. DOI: 10.1111/j.1460-2466.2003.tb02586.x.

Hobolt, S. B. and Brouard, S. 2011. Contesting the European Union? Why the Dutch and the French Rejected the European Constitution. *Political Research Quarterly*, 64 (2), 309–322. DOI: 10.1177/1065912909355713.

Hobolt, S. B. and de Vries, C. E. 2015. Issue Entrepreneurship and Multiparty Competition. *Comparative Political Studies*, 48 (9), 1159–1185. DOI: 10.1177/0010414015575030.

Hobolt, S. B. and de Vries, C. E. 2016. Public Support for European Integration. *Annual Review of Political Science*, 19, 413–432. DOI: 10.1146/annurev-polisci-042214-044157.

Hobolt, S. B., de Vries, C. E. and Tilley, J. 2017. Facing Up to the Facts: What Causes Economic Perceptions? *Electoral Studies*, 51, 115–122. DOI: 10.1016/j.electstud.2017.09.006.

Hooghe, L. and Marks, G. 2009. A Postfunctionalist Theory of European Integration: From Permissive Consensus to Constraining Dissensus. *British Journal of Political Science*, 39 (1), 1–23. DOI: 10.1017/S0007123408000409.

Ioannou, D., Jamet, J.-F. and Kleibl, J. 2015. *Spillovers and Euroscepticism*. ECB Working Paper No. 1815.

Kang, Y.-D. and Oh, C.-R. 2020. Spreading Euroscepticism and Its Macro-Level Determinants: Empirical Analysis of Eurobarometer Survey in 2004–2017. *Journal of Contemporary European Studies*, 28 (3), 1–18. DOI: 10.1080/14782804.2020.1733498.

Kuhn, T., van Elsas, E., Hakhverdian, A. and van der Brug, W. 2013. *An Ever Wider Gap in an Ever Closer Union. Rising Inequalities and Euroskepticism in 12 West European Democracies, 1976–2008*. GINI Discussion Paper No. 91.

Lamla, M. J. and Lein, S. M. 2008. *The Role of Media for Consumers' Inflation Expectation Formation*. KOF Working Paper No. 08-201. DOI: 10.3929/ethz-a-005640674.

Linn, S. and Kellstedt, P. M. 2004. The Political (and Economic) Origins of Consumer Confidence. *American Journal of Political Science*, 48 (4), 633–649. DOI: 10.1111/j.0092-5853.2004.00092.x.

Meijers, M. J. and Zaslove, A. 2020. Measuring Populism in Political Parties: Appraisal of a New Approach. *Comparative Political Studies*. Preprint, 36 pp. DOI: 10.1177/0010414020938081.

Nadeau, R., Niemi, R. G. and Amato, T. 1994. Expectations and Preferences in British Elections. *American Political Science Review*, 88 (2), 371–383. DOI: 10.2307/2944710.

Nicoli, F. 2017. Hard-line Euroscepticism and the Eurocrisis: Evidence from a Panel Study of 108 Elections Across Europe. *Journal of Common Market Studies*, 55 (2), 312–331. DOI: 10.1111/jcms.12463.

Pástor, L. and Veronesi, P. 2013. Political Uncertainty and Risk Premia. *Journal of Financial Economics*, 110 (3), 520–545. DOI: 10.1016/j.jfineco.2013.08.007.

Ritzen, J., Zimmermann, K. F. and Wehner, C. 2014. *Euroskepticism in the Crisis: More Mood than Economy*. IZA Discussion Paper No. 8001.

Serricchio, F., Tsakatika, M. and Quaglia, L. 2013. Euroscepticism and the Global Financial Crisis. *Journal of Common Market Studies*, 51 (1), 51–64. DOI: 10.1111/j.1468-5965.2012.02299.x.

Scharpf, F. W. 1999. *Governing in Europe: Effective and Democratic?* Oxford, UK: Oxford University Press. DOI: 10.1093/acprof:oso/9780198295457.001.0001.

Soroka, S. N., Stecula, D. A. and Wlezien, C. 2015. It's (Change in) the (Future) Economy, Stupid: Economic Indicators, the Media, and Public Opinion. *American Journal of Political Science*, 59 (2), 457–474. DOI: 10.1111/ajps.12145.

Soroka, S. N. 2006. Good News and Bad News: Asymmetric Responses to Economic Information. *Journal of Politics*, 68 (2), 372–385. DOI: 10.1111/j.1468-2508.2006.00413.x.

Soroka, S. N. 2014. *Negativity in Democratic Politics: Causes and Consequences*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781107477971.

Spoon, J.-J. 2012. How Salient is Europe? An Analysis of European Election Manifestos, 1979–2004. *European Union Politics*, 13 (4), 558–579. DOI: 10.1177/1465116512448123.

Stoeckel, F. 2013. Ambivalent or Indifferent? Reconsidering the Structure of EU Public Opinion. *European Union Politics*, 14 (1), 23–45. DOI: 10.1177/1465116512460736.

Wanta, W., Golan, G. and Lee, C. 2004. Agenda Setting and International News: Media Influence on Public Perceptions of Foreign Nations. *Journalism & Mass Communication Quarterly*, 81 (2), 364–377. DOI: 10.1177/107769900408100209.

9 ANNEX

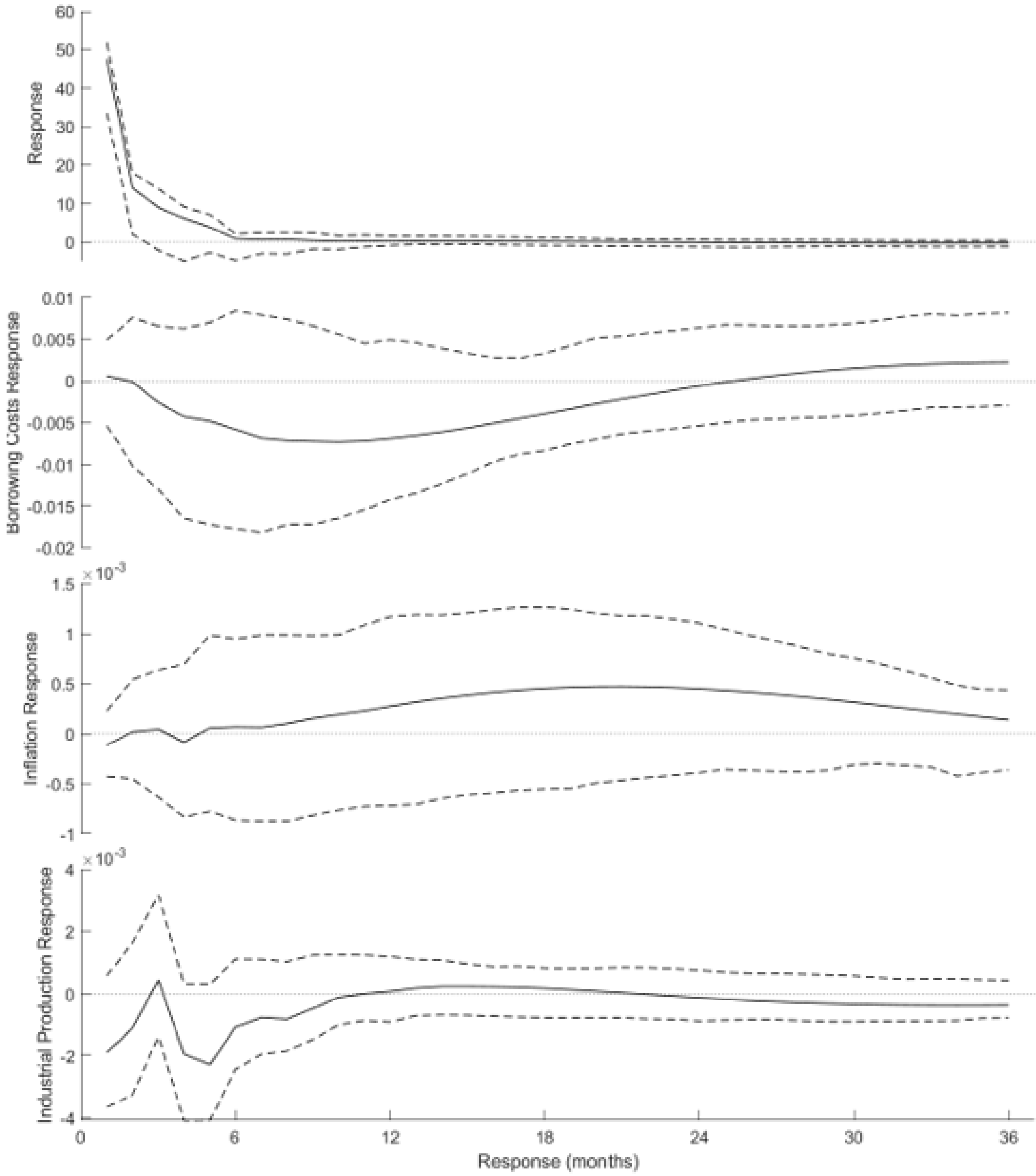

Fig. 7: Response impulse functions (UK)
Notes: Vector autoregression model (VAR) with the following specification (perception index, cost of borrowing, inflation, industrial production). Data is transformed using annual percentage changes including three lags of all variables. The results are reported as response impulse functions of one standard deviation of perception index with 90% confidence bands. Data for the United Kingdom.

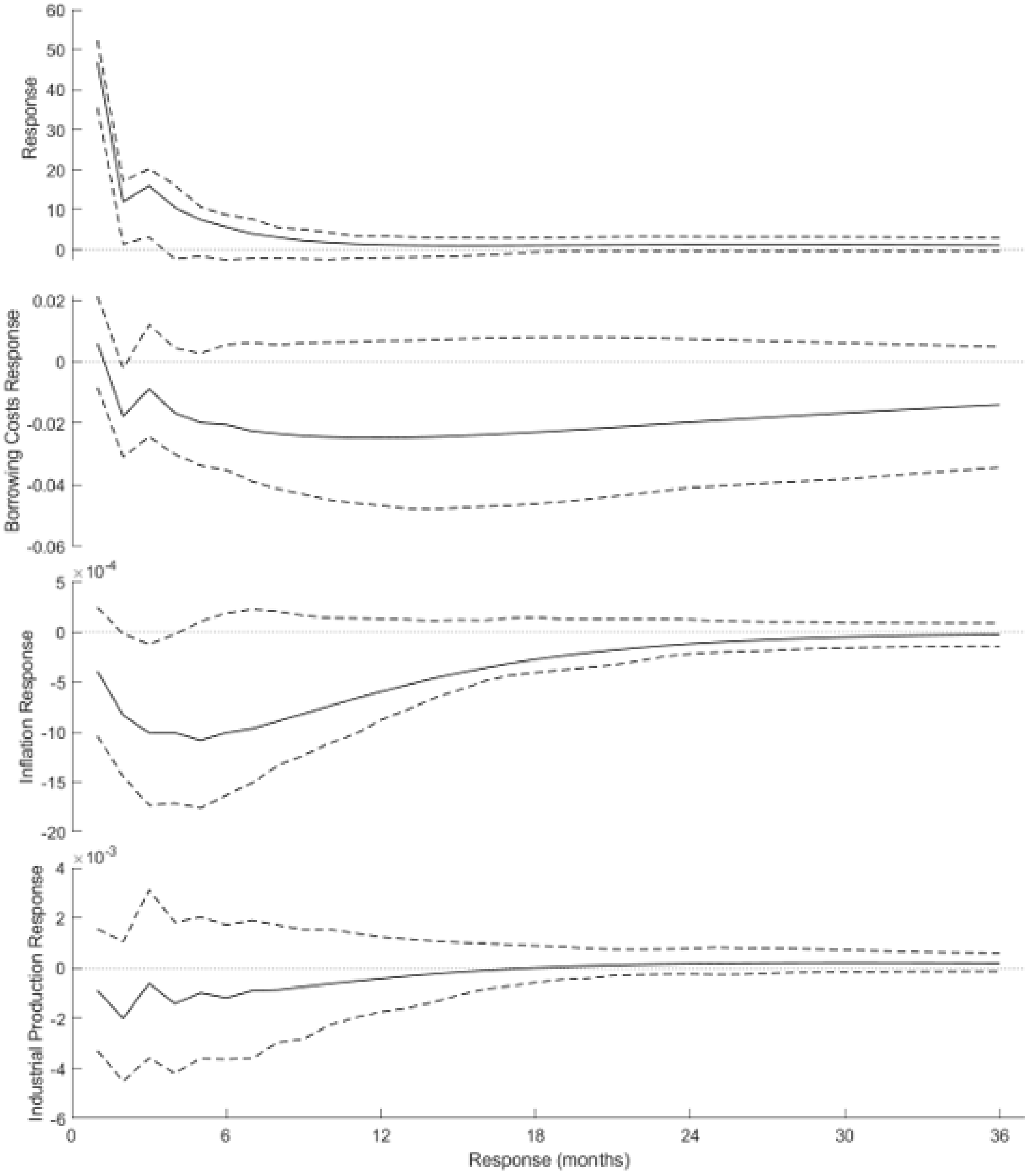

Fig. 8: Response impulse functions (DE)
Notes: Vector autoregression model (VAR) with the following specification, cost of borrowing, inflation, industrial production). Data is transformed using annual percentage changes including three lags of all variables. The results are reported as response impulse functions of one standard deviation of perception index with 90% confidence bands. Data for Germany.

AUTHOR'S ADDRESS

Daniel Pastorek, Department of Finance, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: xpastore@mendelu.cz

CONTEMPORARY OPINIONS ON THE IMPORTANCE OF ENTREPRENEURIAL COMPETENCIES

Petr Řehoř[1], Martin Pech[1], Michaela Slabová[1], Ladislav Rolínek[1]

[1] *University of South Bohemia in České Budějovice, Czech Republic*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

When starting a new business, entrepreneurs used the acquired experience, skills, and competencies. The paper aims to determine opinions on the most important competencies that entrepreneurs need to start their business. The research is based on the questionnaire to survey entrepreneurs and students of economics. We try to determine the differences in their opinions. We assume that the views of entrepreneurs already running a business near reflect the competencies' real usefulness. We used the questionnaire as an instrument for data collection. When comparing entrepreneurs and students, statistically significant differences are in most competencies. Students evaluated entrepreneurial competencies significantly less important than entrepreneurs. Entrepreneurs' opinions on competencies do not depend on enterprise size. Similarly, students' views on competencies are not related to their gender. Our study suggests a significant discrepancy between students' opinions on the importance of various competencies and reality.

KEY WORDS

entrepreneurial competencies, entrepreneurs, opinions, students, education

JEL CODES

L260, J240, A220, I250

1 INTRODUCTION

The environment in which entrepreneurs operate requires knowledge, skills, and the ability to recognize opportunities and think outside the box. Entrepreneurial competencies and expertise can play a crucial role for young people to achieve a bright professional perspective. In this context, entrepreneurship education can significantly contribute to developing their entrepreneurial competencies, attitudes, and skills. Indeed, the number of entrepreneurship

ŘEHOŘ, Petr, PECH, Martin, SLABOVÁ, Michaela, and ROLÍNEK, Ladislav. 2020. Contemporary Opinions on the Importance of Entrepreneurial Competencies. *European Journal of Business Science and Technology,* 6 (2): 127–137. ISSN 2694-7161, DOI 10.11118/ejobsat.2020.008.

programs offered increased considerably, which means there is a demand for this type of education.

Recognizing and identifying competencies is essential for educators and learning opportunities (Mitchelmore and Rowley, 2010). In 2007, the Czech Republic established lifelong learning strategies intended to improve functional and financial literacy competencies. The constant pressure to improve qualifications in the Czech Republic is a topical issue that will determine organizations' primary competitive advantage in the near future. The European Commission declares the importance of bringing Europe back to growth and creating new jobs in the Entrepreneurship 2020 Action Plan (European Commission, 2016). Given the above, there has been a growing awareness of the need to develop students' entrepreneurship in the Czech Republic. Students have to be fully involved in the classroom activities that will teach them to acquire the appropriate entrepreneurial competencies.

Furthermore, there is a significant interest of entrepreneurs and practitioners to continue developing in their career path. A review of the relevant literature showed that studies aimed at identifying core entrepreneurial competencies do not seek to compare students' and entrepreneurs' opinions on the competencies needed to start a new business. Interestingly, the differences between students' and entrepreneurs' views on the competencies can influence students in deciding which skills, competencies, and abilities to develop.

Our study aims to contribute to academia by bridging the gap between students' and entrepreneurs' opinions on the importance of entrepreneurial competencies. We believe that there is a significant difference between these groups in the perception of competencies (theory vs. practical experience). To narrow this gap, we have to identify discrepancies for universities and other institutions to change the focus and schooling methods. Our quantitative study provides empirical evidence of both higher education and the students' desired destination – the entrepreneurial world /business. We also try to offer theoretical explanations for these results that should help to reflect on this issue.

2 THEORETICAL FRAMEWORK

Entrepreneurship has been a subject of interest and attention for many in recent years. Etuk et al. (2014) indicate that SMEs' success and competitiveness depend essentially on entrepreneurial skills. Recent studies defined entrepreneurship as "the creation of new ventures, new products, and new markets" (Read and Sarasvathy, 2005); and "generating businesses using some continuous innovative methods" (Kuratko, 2005). Through an appropriate education already at the school level, students can acquire the relevant skills and mindset required for entrepreneurship (Lepuschitz et al., 2018). Langlois (2002) stated that successful entrepreneurs should be innovative, creative, and risk-taking. This view later has been reinforced in follow-up studies, like that by Wickham (2006), which identified that entrepreneurs are resourceful, seek and discover niches for market innovations, bear risk, are growth-oriented, and drive to maximize profit or investors' returns.

The skills and attributes of entrepreneurs are known as entrepreneur competencies (Asenge et al., 2018). Entrepreneurial competencies combine the personality traits, skills, abilities, and knowledge that a potential entrepreneur has and personal characteristics, significantly impacting entrepreneurial motivation (Farhangmehr et al., 2016). Entrepreneurial competencies can be considered a cognitive component of personality related to knowledge and skills and a non-cognitive component of human nature concerning attitudes (Beltrán Hernández de Galindo et al., 2019). Entrepreneurial competencies refer to knowledge, skills, and attitudes that affect the willingness and ability to perform new value creation (Lackéus and Williams, 2015). There is a difference between the terms of competence and competency. Competence is a word like the knowledge that relates to skills, standard achievements, or measured performance. The word competency is instead a noun as a skill or ability, a behavior-based term. In this paper, we prefer the term "competencies," but in the theoretical background and review of the literature, we keep the terms as stated by the authors.

According to Tittel and Terzidis (2020), there is no consistent or generally accepted source for a list of entrepreneurship competencies in literature. Many authors present competences in a narrower sense as a mix of skills, traits, or variables important for a business. Man et al. (2008) categorized entrepreneurial competencies into six to include; opportunity competencies, strategic competencies, relationship competencies, commitment competencies, and conceptual competencies. Mitchelmore and Rowley (2010) emphasize seven entrepreneurial competencies: identification of the market niche, development of products/services, creativity for the generation of new ideas, environmental analysis, and recognizing opportunities and formulating strategies. The European Commission (Bacigalupo et al., 2016) identified the sense of initiative and entrepreneurship as one of the eight essential competencies necessary for a knowledge-based society in the Entrepreneurship Competence Framework.

For the evaluation of competencies importance, in this research, we have created an appropriate set of seven competencies, each containing a few sub-skills – activities that characterize this competence:

- Opportunities perception (the ability to recognize and evaluate market opportunities)
- Communication (the ability of communication, presentation, and negotiation)
- Teamwork (the ability to cooperate, working together, and leadership)
- Risk tolerance (the ability to take with risk, responsibility and to make sacrifices)
- Creativity (the ability of creativity, initiative, and entrepreneurship)
- Problems solving (the ability to deal with problems, crises, and change)
- Intent creation (the ability to create a business plan and vision)

We independently chose these competencies based on various studies, such as Nieuwenhuizen and Swanepoel (2015). They state as the most critical competencies: the ability to recognize and evaluate opportunities in the market, the ability to develop relationships with other people in the business area, the ability to persuade and discuss with others, and the ability to make sacrifices to ensure that the business gets started. Liñán et al. (2013) emphasized creativity and entrepreneurial intent and problem-solving skills and knowledge to develop an entrepreneurial project mentioned in the comparative studies (Liñán et al., 2013).

The opinions on entrepreneurship competencies lead to several important questions, as follows. In this research, we try to determine the differences between entrepreneurs and students of economic perceptions. By comparing university students and entrepreneurs, it is possible to decide whether teaching at universities reflects the current state of affairs required by practice. Three working hypotheses, which form the subject matter, are defined.

The first hypothesis related to the question of whether students and entrepreneurs have similar opinions on the relevance of competencies to starting a new business:

H1: *The opinions on the importance of selected competencies differ among students and entrepreneurs.*

Few studies involve students and entrepreneurs to evaluate their opinions on the importance of entrepreneurs' competencies. Most studies focus on the self-evaluation of the entrepreneurial competencies of students or entrepreneurs. Alternatively, the evaluation aims to create a list of competencies needed for a business. However, Rezaei-Zadeh et al. (2017) used Irish and Iranian postgraduate students, academics, and entrepreneurs for participation in the research of interdependencies between entrepreneurial competencies that need developing in an educational context. The research results show that students rate higher importance of interpersonal skills in line with Irish entrepreneurs, but productive thinking is more critical for Iranian entrepreneurs.

To determine the causes of possible differences between groups of entrepreneurs, we further formulated the second hypothesis:

H2: *The opinions on the importance of selected competencies differ among the group of micro and small-sized enterprises.*

Several studies have explored the opinions of entrepreneurs on the competencies necessary for starting a new business. Küttim et al. (2011) examined entrepreneurs' views on a sample of 74 entrepreneurs from 4 countries (Estonia, Latvia, Finland, and Sweden). The competencies that were considered the most essential by managers were manifold like entrepreneurial knowledge, experience, personal characteristics, personal capabilities, and motivation. Schelfhout et al. (2016) identified these crucial competencies in a sample of 201 Belgian high school students: performance orientation, communication skills, taking the initiative, and planning/organizing. Morris et al. (2013) identified as significant in a sample of 20 entrepreneurs with more than 100 employees' competencies: creative problem solving, opportunity assessment, and self-efficacy. Competence in guerrilla skills placed at the end. Research of 65 Vietnamese entrepreneurs points out the importance of these competencies: spotting opportunities, ethical and sustainable thinking, valuing ideas, vision, and creativity (Devetag et al., 2020). Botha et al. (2015) analyzed the differences between an established group of entrepreneurs and a start-up group. Their study shows that functional (knowledge-related skills) competencies are more important for a regular group of entrepreneurs.

Similarly, we formulated a third hypothesis which tries to determine the causes of possible differences between groups of students:

H3: *The opinions on the importance of selected competencies differ among male and female students.*

Many studies have investigated the opinions of students on the importance of entrepreneurship competencies. However, the comparison of students' views according to gender, is usually missing. According to Ljunggren and Kolvereid (1996), women are more worried about the impact on personal life, while men are more worried about economic effects during the start-up process. Most studies on entrepreneurial competencies focus on gender differences in entrepreneurship intentions or self-assessment. Mansour (2018) shows significant gender differences in entrepreneurial preferences and perceptions to start a new business.

Kusmintarti et al. (2018) used factor analysis to determine an entrepreneur's entrepreneurial characteristics based on students' opinions. Russian students consider these competencies the most important: financial and economic literacy, motivation and perseverance, creativity, and vision. The least important is the self-awareness of efficacy and ethical and sustainable thinking (Monastyrskaya et al., 2018). Totally 128 students from India and Sri Lanka highlight these competencies: problem-solving, efficiency orientation, assertiveness, and persuasion. They consider it less significant than the competencies for success (Venkatapathy and Pretheeba, 2012).

3 METHODOLOGY AND DATA

3.1 Research Objective

The paper aims to determine whether the opinions on entrepreneurship's competencies are identical according to their importance for current entrepreneurs and students. Conceptually and empirically, the research evaluates the utility of competencies and knowledge for developing entrepreneurial skills. The paper first focuses on the importance of entrepreneurship competencies for entrepreneurs and students and their comparison. We analyze whether students are even aware of what key competencies they should have to start a new business.

Therefore, the evaluation of entrepreneurs' opinions in the real world of business moves is precious. In this way, it can provide a view that reflects the need for individual competencies in practice.

3.2 Data

The research took place in 2018–2019 and was conducted among 148 entrepreneurs of SME's and 157 students of all undergraduate and graduate programs. We used a non-probability representative sampling method in this study. The sample of entrepreneurs included participants mostly from the South Bohemian Region of the Czech Republic. Only small and micro-enterprises were selected for the research because, in the case of students – as start-up entrepreneurs in the future, we assume that they will start their business as small or micro-entrepreneurs. Entrepreneurs from micro and small enterprises we selected in a ratio of 2 : 1 for the sample. The overall response rate of 10% was for online data collection. The questionnaire classifies 148 entrepreneurs from the Czech Republic into the following categories:

- by enterprise size: micro-enterprises with less than ten employees ($n = 98$), Small enterprises with less than 50 employees ($n = 50$);
- by business category (according national occupational system and Open and accessible database of occupations): Economy, administration, human resources ($n = 6$), Banking, finance and insurance ($n = 10$), Transport and logistics ($n = 13$), Agriculture ($n = 10$), Industry and engineering ($n = 38$), Hospitality, tourism, wellness ($n = 20$), IT ($n = 6$), Business and marketing ($n = 20$), other ($n = 25$).

We interviewed 157 students of the Faculty of Economics, the University of South Bohemia in České Budějovice. The questionnaire classifies the following categories:

- by the gender of the students: 46 men and 111 women;
- by the study program: economical informatics ($n = 17$); management and economics ($n = 66$); structural politics EU for public administration ($n = 22$); accounting and financial management ($n = 55$).

3.3 Methods

In this quantitative research, we used two questionnaires to identify entrepreneurs and students of the University of South Bohemia in České Budějovice opinions. We have focused on the part of the business competencies that are important for starting a new business. We measured individual values of competencies on an ordinal scale (in our case, a 7-point Likert scale). We presented the quantitative results of the assessed opinions on the importance of competencies as the average of all respondents' given group assessments. We selected the appropriate seven competencies, each consisting of partial skills: opportunities perception, communication, teamwork, risk tolerance, creativity, problem-solving, and intent creation.

Considering that it is impossible to prove data normality (Shapiro-Wilk normality test), the non-parametric Mann-Whitney U test was used to confirm the hypotheses. We compare two independent samples with a significance level of $\alpha = 0.05$. To verify data reliability, the Cronbach alpha coefficient is the most commonly used for the Likert scale. This coefficient can take values in the range $\langle 0; 1 \rangle$, with generally acceptable values between 0.7 and 0.95. The results show Cronbach alpha 0.8488 for the sample of students ($n = 157$) and Cronbach alpha 0.8377 for entrepreneurs ($n = 148$). Both values occur within the range as mentioned above of acceptable values, and it can be concluded that the samples are consistent.

4 RESULTS

We compared the opinions on competencies as a difference in the average of all assessments of the respondents' given group. As reported in Fig. 1, the students (who are potentially going to have an entrepreneurial career) and the entrepreneurs differ regarding entrepreneurship competencies. In general, the results show that the entrepreneurs attribute greater importance to competencies at all points.

As shown in Fig. 1, students chose "Problems solving" as the most important competence for entrepreneurship (5.28), probably because they notice the changing conditions and turbulences in business in the Czech Republic. "Creativity" was identified as the second most rated competence (5.01) by students, reflecting that this competence significantly contributes to companies' competitive advantage and openness to innovation. Communication, "one of the soft skills that are very useful not just in the business area", was rated 4.99, which is the third most important competency in students' view. Surprisingly, the "Intent creation" competence, with a value of 3.57, gained the lowest importance. Students' lack of ability might cause this low rating and lacking confidence in creating entrepreneurial intent. They do not know the utility of this competence in the real environment in everyday practice.

On the contrary, entrepreneurs evaluate the following competencies the most important. With the highest value given to "Risk tolerance" (6.15), the ability to react fast and effectively in situations of varying tensions that require immediate solutions can be crucial to business survival in the market. Here it is clear that the key competence for a business's successful operation in the market is "Opportunities perception" (5.82) to find, recognize, and evaluate market opportunities, which is also essential for the functioning of companies. Close to this ranking followed "Creativity" with 5.67 points. This competence is more desired for business and other aspects of it, such as innovation, coming up with new ideas and creating a competitive advantage. "Intent creation" (5.70) means creating a business plan, vision, and mission. Keeping it in praxis or changing the business plan, flexibly is crucial to run the business properly. The results showed a surprising finding that entrepreneurs consider "Working together, teamwork and leadership" to be the least important of all (4.84).

The opinions on competencies from entrepreneurs' perspective showed that they differed from the students in assigned importance and absolute magnitude. Entrepreneurs generally give a higher value to all competencies than students, apparently because of practice experience. Another difference is the placement of competencies in terms of importance. The most significant difference between students (3.57) and entrepreneurs (5.70) is because of competency "Intent creation". This result indicates a discrepancy between practice and study. This result implies that the emphasis should be on functional knowledge, which is more than desirable in the real world of business.

We used the Mann-Whitney U test at the significance level of alpha 0.05 to test individual independent pairs of responses related to the importance of students' and entrepreneurs' opinions on competencies. The results of the working hypotheses show Tab. 1 at the 5% level of significance. Statistically significant differences between entrepreneurs and students are in most opinions on competencies. The hypothesis H1 is confirmed. These differences suggest that entrepreneurs evaluate the importance of competencies higher than students. The only exception is teamwork (p-value = 0.5423), which does not show significant differences between students and entrepreneurs.

Furthermore, we analyzed in more depth opinions on the importance of competencies for individual groups of companies according to their size and students according to gender. Indeed, from an enterprise-size perspective, the hypothesis H2 has not been confirmed. Entrepreneurs' opinions on competencies do not depend on enterprise size, for both groups are the same. We found similar results for the differences between the students according to their gender. In this case, the hypothesis H3

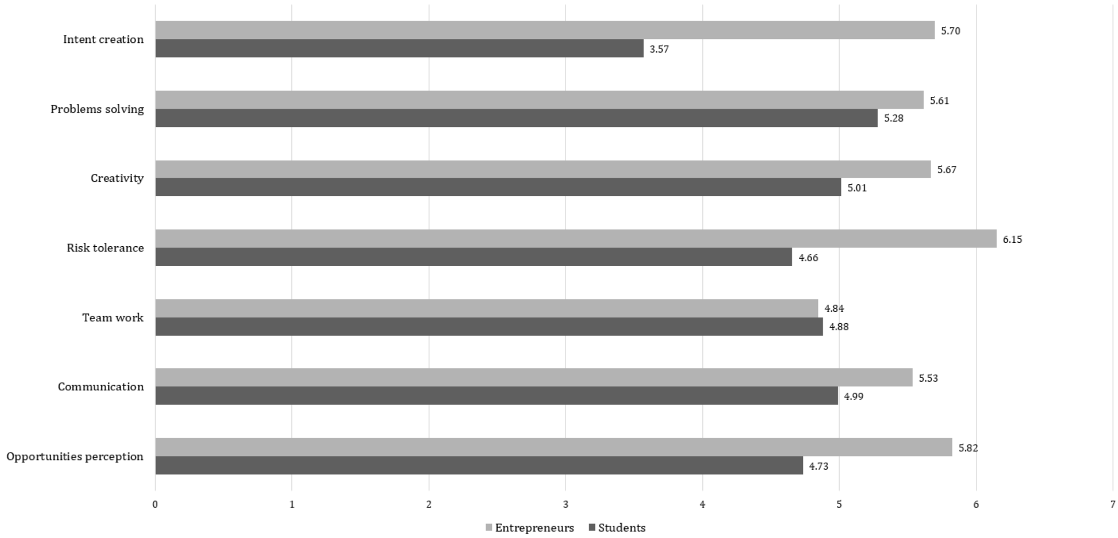

Fig. 1: Evaluation opinions on competencies between students and entrepreneurs

Tab. 1: Results of Hypotheses Testing

	H1: Students vs. Entrepreneurs		**H2: Enterprise Size (micro vs. SMEs)**		**H3: Students Gender (male vs. female)**	
	Z	p-**value**	Z	p-**value**	Z	p-**value**
Opportunity Perception	6.8073	0.0000	−1.0283	0.3038	−1.2415	0.2144
Communication	3.8050	0.0001	0.4898	0.6243	0.7838	0.4331
Team Work	0.6776	0.4980	−0.3570	0.7211	−1.4555	0.1455
Risk Tolerance	9.1490	0.0000	−1.4850	0.1375	−0.7195	0.4719
Creativity	4.4439	0.0000	0.8383	0.4019	−1.0058	0.3145
Problems Solving	2.4040	0.0162	0.1405	0.8882	−1.3780	0.1682
Intent Creation	10.1817	0.0000	−1.7399	0.0819	−1.0181	0.3086

Notes: This table presents the results of attested hypotheses (H1–H3) based on the Mann-Whitney U-test. The Z scores and p-values are included in the columns, where significant values are lower than the 5% level of significance.

was not confirmed. In general, there are no differences between the opinions on competencies from students' perspectives in terms of gender and entrepreneurs' view according to enterprise size. Groups of entrepreneurs and students' are consistent.

5 DISCUSSION

We start by identifying opinions on competencies, emphasizing specific competencies, creativity, and problem-solving. We proceed similarly to the research of (Liñán et al., 2013), who states that the general perception of the importance of entrepreneurial skills reflects the degree to which individuals believe they possess a sufficiently high level of entrepreneurial skills.

One of the most exciting survey results is comparing the students' opinions and competencies and real entrepreneurs. In particular, uncompromising information and little practical experience make students evaluate some entrepreneurial competencies significantly less critical than entrepreneurs and underestimate these competencies in the future (if they want to start a business). These differences in opinions reflect the different mentalities of the educational world and the professional world. This result is also due to the university's attitude, where not all students are allowed to study courses focused on entrepreneurship. There is little teaching of soft skills, such as leadership, communication, vision, creativity, and ability to use opportunities. However, entrepreneurs consider these competencies to be beneficial.

We found the most significant difference in the competency of "the ability to create a business plan (intent creation)" rated by the students as the least important and the third most important by the entrepreneurs. The remarkable and most different result in assessing the importance of the competence "Intent creation" confirms that students are moving away from the realistic view. Therefore, teaching in this area of business plan development and all its essentials should significantly support this skill. As the authors say (Gelard and Saleh, 2011), effective entrepreneurship education can stimulate and increase students' interest in entrepreneurship career considerations.

Another significant difference we found in the competency of "The ability to recognize and evaluate market opportunities," which the entrepreneurs rated as the second-best. In contrast, the students rated it with low value. Baum et al. (2001) also found that general competencies (organizational skills and opportunity recognition skills) have significant indirect economic growth effects. Furthermore, students' different views justify that students cannot perceive the impact of this competence on developing an enterprise in practice in terms of both competition and expansion. The training and contact with real cases are essential for students and their future professions.

Students consider "Problems solving" as the essential competency needed for entrepreneurship, however entrepreneurs' views differed. From the practice point of view, the most important competence is, above all, "Risk tolerance," "Opportunities perception." These are competencies, which are challenging to learn by studying. As also stated by Knowles et al. (2014), learning can be useful when new information is presented in real-life situations with a problem-solving approach.

In Pathak's article (2019), the emphasis is placed on preserving the effectiveness and value of entrepreneurial education and educators, which means aligning the younger generation of entrepreneurs' objectives with teaching objectives – appeal to their mindset, respond to current and future trends. The current generation of students, especially the millennials, and the generation XY, start a new business immediately after graduation. Companies hire millennials to gain practical experience before they start their own business. Thus, they take up traditional positions in companies but still focus heavily on companies that meet their preferences.

6 CONCLUSIONS

The business environment places high demands on professionals working in the organization at all management levels. The competitiveness of current entrepreneurs depends on entrepreneurs' competencies, which they should develop during professional careers and studies at universities. Students of economics expect to have a high-quality educational background, solve unexpected situations, and creative thinking for the organization's development.

The purpose of this study was to determine how current entrepreneurs and university students perceive entrepreneurial competencies. There were statistically significant differences between students and entrepreneurs in opinions on the importance of competencies. Entrepreneurs in our research identified the essential competencies "Risk tolerance, opportunities perception, intent creation." Students consider "Problems solving" as a core competence for entrepreneurship, then next "Creativity and communication." There is a gap between these two groups. It is necessary to transfer valuable information from entrepreneurs (from practice to academia) where young people are preparing for business or self-employed. The research shows the main differences that entrepreneurs generally give a higher value to all competencies than students, apparently because of practice experience.

The research results can be included in the course Entrepreneurship Support to motivate students to develop relevant competencies for successful business start-up and market survival. When teaching subjects focused on developing entrepreneurial competencies, the lecturer must consider that future students want to strengthen the competencies vital to them. At the same time, the lecturer must point out the importance of all key competencies and their interconnections and point out their development and the dynamically evolving demands of the labor market.

The contribution of this work is to propose suitable development methods, which would include such vital combinations that would also be indicators of the suitability of business graduates to run a business. Our results can improve the business schools' curriculum by defining the competencies students need to develop to succeed in entrepreneurship. Building on this finding, universities should be interested in finding effective methods of improving these competencies for students. They should effectively develop necessary competencies to establish companies and maintain the business in prosperity in traditional teachings such as seminars and lectures and new teaching forms. These include shadowing managers, implementing their projects in the classroom, using case studies.

Other practical implications of this study contribute to the importance of developing staff training, workshops, assessment centers, coworking centers, shading and coaching managers, and new effective education methods in enterprises, particularly addressing the competencies. The authors will continue future research regarding entrepreneurs and their attitudes towards entrepreneurship so that it will be possible to best tailor the content of the entrepreneurship support course. Another direction of future research may be the analysis of competencies needed for business success in the fourth industrial revolution era (Industry 4.0 competencies).

7 ACKNOWLEDGEMENT

This paper was funded by University of South Bohemia in České Budějovice [Nr. EF-150-GAJU 047/2019/S].

8 REFERENCES

Asenge, E. L., Diaka, H. S. and Soom, A. T. 2018. Entrepreneurial Mindset and Performance of Small and Medium Scale Enterprises in Makurdi Metropolis, Benue State-Nigeria. *International Journal of Innovation*, 6 (2), 124–146. DOI: 10.5585/iji.v6i2.223.

Bacigalupo, M., Kampylis, P., Punie, Y. and Van den Brande, L. 2016. *EntreComp: The Entrepreneurship Competence Framework*. EUR 27939 EN. Luxembourg: Publication Office of the European Union. DOI: 10.2791/593884.

Baum, J. R., Locke, E. A. and Smith, K. G. 2001. A Multidimensional Model of Venture Growth. *The Academy of Management Journal*, 44 (2), 292–303. DOI: 10.2307/3069456.

Botha, M., van Vuuren, J. J. and Kunene, T. 2015. An Integrated Entrepreneurial Performance Model Focusing on the Importance and Proficiency of Competencies for Start-up and Established SMEs. *South African Journal of Business Management*, 46 (3), 55–65. DOI: 10.4102/sajbm.v46i3.101.

Beltrán Hernández de Galindo, M. de J., Romero-Rodríguez, L. M. and Ramírez-Montoya, M. S. 2019. Entrepreneurship Competencies in Energy Sustainability MOOCs. *Journal of Entrepreneurship in Emerging Economies*, 11 (4), 598–616. DOI: 10.1108/JEEE-03-2019-0034.

Devetag, M. G., Zazzerini, G., Tuan, N. Q. and Hung, D. Q. 2020. Developing Entrepreneurial Competencies in Vietnam: Evidence from the Bac Ninh Province. *Journal of Entrepreneurship and Innovation in Emerging Economies*, 6 (2), 1–19. DOI: 10.1177/2393957520924987.

Etuk, R. U., Etuk, G. R. and Baghebo, M. 2014. Small and Medium Scale Enterprises (SMEs) and Nigeria's Economic Development. *Mediterranean Journal of Social Sciences*, 5 (7), 656–662. DOI: 10.5901/mjss.2014.v5n7p656.

European Commission. 2016. *The Entrepreneurship 2020 Action Plan*. [online]. Available at: `https://ec.europa.eu/growth/smes/promoting-entrepreneurship/action-plan_en`. [Accessed 2020, March 20].

Farhangmehr, M., Gonçalves, P. and Sarmento, M. 2016. Predicting Entrepreneurial Motivation among University Students: the Role of Entrepreneurship Education. *Education and Training*, 58 (7–8), 861–881. DOI: 10.1108/ET-01-2016-0019.

Gelard, P. and Saleh, K. E. 2011. Impact of Some Contextual Factors on Entrepreneurial Intention of University Students. *African Journal of Business Management*, 5 (26), 10707–10717. DOI: 10.5897/AJBM10.891.

Knowles, M. S., Holton, E. F. and Swanson, R. A. 2014. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. London: Routledge.

Kuratko, D. F. 2005. The Emergence of Entrepreneurship Education: Development, Trends, and Challenges. *Entrepreneurship Theory and Practice*, 29 (5), 577–598. DOI: 10.1111/j.1540-6520.2005.00099.x.

Kusmintarti, A., Anshori, M. A., Sulasari, A. and Ismanu, S. 2018. Student's Entrepreneur Profile: A Cluster of Student's Entrepreneurial Characteristics. *Journal of Entrepreneurship Education*, 21 (S). DOI: 10.13140/RG.2.2.15006.84804.

Küttim, M., Arvola, K. and Venesaar, U. 2011. Development of Creative Entrepreneurship: Opinion of Managers from Estonia, Latvia, Finland and Sweden. *Verslas: Teorija ir praktika [Business: Theory and Practice]*, 12 (4), 369–378. DOI: 10.3846/btp.2011.38.

Lackéus, M. and Williams, M. K. 2015. Venture Creation Programs: Bridging Entrepreneurship Education and Technology Transfer. *Education + Training*, 57 (1), 48–73. DOI: 10.1108/ET-02-2013-0013.

Langlois, R. N. 2002. *Schumpeter and the Obsolescence of the Entrepreneur*. UConn Department of Economics Working Paper No. 2002-19. DOI: 10.2139/ssrn.353280.

Ljunggren, E. and Kolvereid, L. 1996. New Business Formation: Does Gender Make a Difference? *Women in Management Review*, 11 (4), 3–12. DOI: 10.1108/09649429610122096.

Lepuschitz, W., Koppensteiner, G., Leeb-Bracher, U., Hollnsteiner, K. and Merdan, M. 2018. Educational Practices for Improvement of Entrepreneurial Skills at Secondary School Level. *International Journal of Engineering Pedagogy*, 8 (2), 101–114. DOI: 10.3991/ijep.v8i2.8141.

Liñán, F., Nabi, G. and Krueger, N. F. 2013. British and Spanish Entrepreneurial Intentions: A Comparative Study. *Revista de economía mundial*, 33, 73–107.

Man, T. W. Y., Lau, T. and Snape, E. 2008. Entrepreneurial Competencies and the Performance of Small and Medium Enterprises: An Investigation through a Framework of Competitiveness. *Journal of Small Business and Entrepreneurship*, 21 (3), 257–276. DOI: 10.1080/08276331.2008.10593424.

Mansour, I. H. F. 2018. Gender Differences in Entrepreneurial Attitude & Intentions among university Students. *International Conference on Advances in Business, and Law*, 2 (1), 193–200. DOI: 10.30585/icabml-cp.v2i1.273.

Mitchelmore, S. and Rowley, J. 2010. Entrepreneurial Competencies: A Literature Review and Development Agenda. *International Journal of Entrepreneurial Behaviour and Research*, 16 (2), 92–111. DOI: 10.1108/13552551011026995.

Monastyrskaya, T., Mikidenko, N. and Storozheva, S. 2018. Leadership and Entrepreneurial Competencies Evaluated by the Academic Community. In Strielkowski, W. and Chigisheva, O. (eds.). *Leadership for the Future Sustainable Development of Business and Education*. Springer Proceedings in Business and Economics. DOI: 10.1007/978-3-319-74216-8_35.

Morris, M. H., Webb, J. W., Fu, J. and Singhal, S. 2013. A Competency-Based Perspective on Entrepreneurship Education: Conceptual and Empirical Insights. *Journal of Small Business Management*, 51 (3), 352–369. DOI: 10.1111/jsbm.12023.

Nieuwenhuizen, C. and Swanepoel, E. 2015. Comparison of the Entrepreneurial Intent of Master's Business Students in Developing Countries: South Africa and Poland. *Acta Commercii*, 15 (1), 270. DOI: 10.4102/ac.v15i1.270.

Pathak, S. 2019. Future Trends in Entrepreneurship Education: Re-Visiting Business Curricula. *Journal of Entrepreneurship Education*, 22 (4), 1–13.

Read, S. and Sarasvathy, S. D. 2005. Knowing What to Do and Doing What You Know: Effectuation as a Form of Entrepreneurial Expertise. *The Journal of Private Equity*, 9 (1), 45–62. DOI: 10.3905/jpe.2005.605370.

Rezaei-Zadeh, M., Hogan, M., O'Reilly, J., Conningham, J. and Murphy, E. 2017. Core Entrepreneurial Competencies and Their Interdependencies: Insights from a study of Irish and Iranian Entrepreneurs, University Students and Academics. *International Entrepreneurship and Management Journal*, 13, 35–73. DOI: 10.1007/s11365-016-0390-y.

Schelfhout, W., Bruggeman, K. and De Mayer, S. 2016. Evaluation of Entrepreneurial Competence Through Scaled Behavioural Indicators: Validation of an Instrument. *Studies in Educational Evaluation*, 51, 29–41. DOI: 10.1016/j.stueduc.2016.09.001.

Tittel, A. and Terzidis, O. 2020. Entrepreneurial Competences Revised: Developing a Consolidated and Categorized List of Entrepreneurial Competences. *Entrepreneurship Education*, 3, 1–35. DOI: 10.1007/s41959-019-00021-4.

Venkatapathy, R. and Pretheeba, P. 2012. Entrepreneurial Competencies Among Management Graduates. *Gian Jyoti E-journal*, 2 (3), 135–143.

Wickham, P. A. 2006. *Strategic Entrepreneurship*. London: Financial Times/Prentice Hall.

AUTHOR'S ADDRESS

Petr Řehoř, Department of Management, Faculty of Economics, University of South Bohemia in České Budějovice, Studentská 13, 370 05 České Budějovice, Czech Republic, e-mail: rehor@ef.jcu.cz, ORCID: 0000-0003-2438-3395

Martin Pech, Department of Management, Faculty of Economics, University of South Bohemia in České Budějovice, Studentská 13, 370 05 České Budějovice, Czech Republic, e-mail: mpechac@ef.jcu.cz, ORCID: 0000-0002-0807-3613

Michaela Slabová, Department of Management, Faculty of Economics, University of South Bohemia in České Budějovice, Studentská 13, 370 05 České Budějovice, Czech Republic, e-mail: slabom02@jcu.cz, ORCID: 0000-0002-8394-1409

Ladislav Rolínek, Department of Management, Faculty of Economics, University of South Bohemia in České Budějovice, Studentská 13, 370 05 České Budějovice, Czech Republic, e-mail: rolinek@ef.jcu.cz, ORCID: 0000-0003-2587-8226

INFLUENCER IMPACT ON BRAND AWARENESS: A MIXED METHOD SURVEY IN THE GERMAN FASHION SEGMENT

Kai Dominik Renchen[1,2]

[1] *SRH Hochschule Heidelberg, Germany*
[2] *Mendel University in Brno, Czech Republic*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

This study evaluates the impact of influencer marketing on consumers. Although online influencers are long established in German B2C markets, research in the field is scarce. Marketing managers are unsure how influencer marketing strategies should be drafted and which influencers should be recruited in order to maximize their marketing efforts at the consumer level. Based on a review of previous theoretical and empirical literature, a mixed method empirical study comprising semi-structured interviews with three German fashion influencers and a consumer survey ($N = 385$) among the followers of these influencers are conducted to evaluate the impact of influencer activity on consumer followership, brand awareness and purchase intention. Results indicate that the intensity of influencer network involvement, intrinsic influencer motivation, authenticity of communication style and the real-life character of influencer posts increase the dependent variables. Hence, the empirical study has contributed to identify the most important determiners for the German B2C fashion influencer segment. Marketing managers are strongly advised to select influencers adequately and organize their marketing strategy in correspondence with their company and products.

KEY WORDS

influencer, marketing, B2C, fashion

JEL CODES

M3, M31, M37

1 INTRODUCTION

The use of the internet as an information and shopping medium by consumers has increased rapidly in recent years (Franklin, 2008). Ever more, buyers use virtual channels to find out about consumption options and products. Among others, social media are very popular.

RENCHEN, Kai Dominik. 2020. Influencer Impact on Brand Awareness: A Mixed Method Survey in the German Fashion Segment. *European Journal of Business Science and Technology*, 6 (2): 138–153. ISSN 2694-7161, DOI 10.11118/ejobsat.2020.009.

In 2017, an average US American consumer spent around 1.72 hours on social media daily. Generally, 74% of buyers are influenced by social media in their shopping behavior (Woods, 2016). In this context, influencer marketing makes use of gatekeepers in virtual social networks for advertising. Businesses denying to actively engage in these new sales strategies risk a loss of competitiveness and melt-off of established outlets (Carbonaro and Votava, 2005). However, many companies, especially small and medium-sized enterprises, are still uncertain and cautious with regard to influencer marketing in the virtual space. Only around 68% of German businesses have budgeted for influencer marketing on the Internet in 2017. "Germany has got a more reserved relationship with social media than other countries like the U.S. and the U.K." observes Davies (2017), attributing this to concerns about data security and privacy. The reasons for the uncertainty of providers in terms of influencer marketing are more far-reaching.

To date, the determinants and moderators of social media marketing effectiveness on consumer brand awareness are not fully understood in practice. In order to optimally develop virtual markets by means of influencer marketing, a comprehensive marketing concept is required that uses various online channels to address customers as comprehensively and effectively as possible (Safko, 2010). Furthermore, academic research in influencer marketing is just establishing. The use of social media marketing requires clarity about the conditions, potentials and risks as well as an understanding of determiners of effective strategies and communication methods (Rowley, 2004). German businesses' reluctance concerning influencer marketing is therefore partly due to a lack of know-how and uncertainty how consumers in fact perceive online influencers and what makes them follow those influencers.

Thus, the main objective of this paper is to close the existing research gap by investigating the impact of influencer marketing at the level of consumers and assess their quantitative effect. The conducted study identifies determinants of influencer marketing use in German companies and evaluates how influencer marketing impacts consumers' readiness to follow those influencers, their brand awareness and a as consequence, purchase intention. It is localized in the B2C (business to consumer) segment, i.e. refers to marketing directed from providers to private end consumers. In B2C marketing, psychological factors influencing the purchase decision and post purchase phase are of particular importance. The concentration on the B2C segment is thus a logical consequence of the study subject "brand awareness". It is assumed that the influencer's involvement, his or her motivation, authenticity of communication and real-life contributions are the main drivers of the postulated dependent variables. In addition, it is argued, that this relationship is moderated by influencer audience and product fit.

The approach is unique in the academic context. For the first time a comprehensive set of determinants, moderators and effects of influencer marketing is developed and empirically tested. The study gives German businesses intending to utilize social marketing channel orientation, how to make the best out of this strategy. Findings indicate that determiners of the three target parameters influencer followership, brand awareness and purchase intention show different patterns. While followership is achieved by high-influencer motivation and authenticity, brand awareness is established mainly via the real-life character and authenticity of contributions. Purchase intention is primary based on the real-life character of contributions and the fit between the influencer and his or her audience.

The structure of the paper is as follows. Section 2 reviews the literature concerning the characteristics of influencer marketing and its success factors. Section 3 introduces the method and the data applied for this research. Section 4 provides partial evidence for the postulated relationships. In section 5, results are discussed and implications for researchers and practitioners derived.

2 LITERATURE REVIEW

2.1 Influencer Marketing and Brand Awareness

The Marketing strategy of a company has to continuously adapt to new market developments, trends and requirements. Marketing anticipates consumption needs and creates new desires in consumers (Cannon et al., 2010). At the same time, consumers are not just passive targets of marketing information any longer. Instead, they participate actively in the formation of consumption trends. The internet enables consumers to share their consumption decisions and desires with peers and these readily imitate the behavior of model personalities and gatekeeper persons on the web. The internet has become a virtual environment of presentation and status demonstration and thus endows marketing with a new dynamic element (Homburg and Koschate, 2007).

The virtualization and democratization of marketing in the age of internet has increased the importance of branding in consumer marketing. According to American Marketing Association (AMA) a brand comprises "name, term, sign, symbol, or design, or a combination of them, intended to identify the goods and services of one seller or group of sellers and to differentiate them from those of competition" (Gnann, 2008). Brands create a feeling of togetherness and differentiate social groups. Brands also reduce psychological insecurity and perceived social risks (Court et al., 2009). They create a common basis of reference, communication and identification (Biel, 2001). Brands stand for experience, have got an integrating effect in social live, they are statements of personality and connect people sharing the same attitude (Schmitt, 2012). Brands thus play an important role in the integration of the individuals in social networks and even have a community-building effect. Customers reconnect to brands inner conceptions and ideals which go far beyond the product itself but comprise a whole philosophy and comprehensive context (Tan and Ming, 2003). Brand associations are partly concrete and refer to the product and its properties in more frequently abstract categories, e.g. luck, welfare or status (Bauer et al., 2002). The complex meaning of brands causes consumer convictions which are distinct from practical experience and are deeply rooted at the level of sentiment and desire (Keller, 1993).

The virtualization and growing consumer involvement with marketing has multiplied the effectiveness of brands and has endowed new dynamics to brand advertisement. Brands are created by both provider and customer (Herrmann and Schaffner, 2005). Consumers increasingly contribute to the formation of brands by their consumption behaviour and by communicating brand image through electronic word of mouth (eWOM) above all via social media. Thus, brands gain strength in a dialogue involving both parties, customers and providers (Janson, 2012). In this context, brand awareness, i.e. the conscious perception and reflection of a brand, is probably the most crucial factor in the communication process between consumer and provider (Fournier, 1998, p. 368). Shaping brand awareness is crucial in corporate brand marketing and evokes brand image, buying intention and consumer loyalty (Meyer and Schwager, 2007).

Nowadays, influencer marketing is a powerful instrument of brand communication in the age of internet and social media. Although there is no homogenous definition of influencer marketing, the term basically refers to the effect that influential communicators contribute to create awareness for a product in social media and in this way make consumption trends emerge (Evans et al., 2017; Biaudet, 2017; Baker, 2014). Influencers are consumers in a central communicative function who impact on other consumers in a targeted way to promote the consumption of certain products via social media (Keller and Fay, 2016; Dron and Mohamad, 2015). Marketing increasingly uses influencers to explicitly develop brand awareness of particular target groups from inside the group (McQuarrie et al., 2013; Dlodlo, 2014).

Influencers control and guide eWOM among their followers and are recognized as opinion leaders in their social ingroup (Dalstam et al., 2018). Influencer messages expressing positive attitudes towards brands or passion for particular brands or products, (Baker, 2014), "manipulate consumers' buying decisions" (Hu et al., 2015) and exert significant influence on prone users (Woods, 2016) Consumers' involvement with the influencer induce positive emotional, rational attitudes on the contributions and motivate consumers to follow the influencer (Riedl and von Luckwald, 2019).

Perceiving influencers trustworthy and experienced, consumers develop positive brand attitudes (Breves et al., 2019). eWOM initiated and promoted by influencers, in whom consumer trust, contributes to consumers brand knowledge (Lock, 2016). Consumers thus transfer their positive sentiments for the influencer to the brand (Evans et al., 2017). The spread of positive brand attitudes and brand involvement influences the development of a positive brand image among a majority of consumers, which again increases individual consumers brand awareness (Lock, 2016), which finally motivates consumers' purchase intention (Gadalla et al., 2019).

2.2 Determinants of Influencer Marketing Success

The question, which features of influencer marketing result in success – i.e. influencer followership, brand awareness and finally purchase intention – has been discussed to extent in previous literature, in order to enable marketing agencies to systematically select and recruit influencers, who reach an audience that buys and recommends the advertised products (Cakim, 2009).

A systematic review using an evaluation method suggested by Webster and Watson (2002), has found the following six major determiners of influencer marketing success.

The first one refers to influencer network involvement. A high level of social presence, i.e. frequent posts and intense network participation, increase influencers' fame and followers' trust in the influencer personality (De Veirman et al., 2017). Brand attitude is strengthened by famous and socially involved influencers (Jin et al., 2019). Influencers showing strong social engagement and standing, develop higher impact on consumers' brand perception and brand attitude (De Veirman et al., 2017). Youtubers' social influence contributes to perceived information credibility and consumers' involvement with product and brand (Xiao et al., 2018).

Secondly, influencer motivation plays a major role. Extrinsic as well as intrinsic factors can motivate influencers to participate in particular campaigns (Biaudet, 2017). Influencers motivated both intrinsically and extrinsically are found most effective and reliable (Hutter and Mai, 2013).

Thirdly, the motivation is followed by the influencers' communication style. Language style as well as contents of influencer communication have been found important determinants of influencer communication success with consumers. High word and image counts enhance influencers' recognition as specialists in some sectors (Chae et al., 2016). The choice of positive words connected to leisure, fun and holidays as well as the employment of positive emojis increases the reception of contributions (Jaakonmäki et al., 2017).

Fourthly, auality of influencer contributions has to be considered. High perceived argument quality contributes to influencer credibility for a sample of Chinese followers (Chae et al., 2016). High transfer of meaning by the influencer enhances Malayan consumers' brand attitude and buying intention (Lim et al., 2017). Consumers estimate eWOM reliability and high credibility information sources (Lock, 2016). According to a consumer survey, influencers' previous Instagram activity positively moderates their credibility (Breves et al., 2019).

Last but not least, the fit between the influencer and the audience as well as the product as to be discussed. The effectiveness of influencers on consumers is high when the influencer person corresponds to the target audience in character and brand identity perception (Dalstam et al., 2018). An audience perceiving influencer's congruence to own inner attitudes shows positive emotions and is ready to adapt the influencers' attitudes and accept his or her messages as a fact (Belanche et al., 2019).

In terms of the product-fit, perceived influencer credibility depends on his/her expertise with the particular product and similar products. High influencer product match indicates a positive relationship with consumer attitude towards the product and purchase intention (Lim et al., 2017). Instagram influencers disposing of high brand fit enjoy higher image, recognized expertise, trustworthiness and advertising success than their unexperienced colleagues (Breves et al., 2019). Influencers' brand experience contributes to perceptions of trustworthiness, expertise and attractiveness with followers (Eroğlu and Bayraktar Köse, 2019). Consumers recognize experienced and expert influencers as opinion leaders and follow their suggestions to interact, recommend and buy the product (Istania et al., 2019).

3 DATA AND METHODS

Theoretical model investigates influencer network involvement (IN), influencer motivation (IM), authentic communication style (IA) and quality of contributions (IR) as independent variables with a postulated influence on consumers' influencer followership, brand awareness and purchase intention. Audience (IP) and product fit (IF) act as moderators.

The model leads to the following hypotheses: There is a positive relationship between IN, IM, IA and IR and H1) consumers' influencer followership, H2) consumers' brand awareness and H3) consumers' purchase intention, positively moderated by IP and IF.

The model is investigated in the German market, as none of the evaluated studies refers to this particular market so far. Instead, the wide range of empirical scientific literature is conducted in other Northern European countries or in the USA. Nevertheless, the German influencer business, is of high interest due to its strong growth. The market volume of influencer marketing in Germany, Austria and Switzerland is expected to almost double to 990 million euros until 2020 as compared to 2017 (market volume 560 million euros).

Furthermore, previous quantitative studies are not conclusive concerning the evaluated influencer marketing strategies and success effects. Although some recent studies refer to a set of several parameters (Belanche et al., 2019; Breves et al., 2019; Eroğlu and Bayraktar Köse, 2019), none has so far evaluated the comprehensive set of possibly relevant design elements of influencer marketing found in the above review. Studies excluding one or the other relevant aspect however risk spurious results, since important side effects are neglected.

In order to test the model, an empirical survey in the German fashion influencer segment was conducted. The survey combines qualitative and quantitative research methods and evaluates two perspectives of the influencer-consumer communication process Combining a qualitative and quantitative study, the categories derived from the review are validated and then tested statistically (Cooper and Schindler, 2014).

3.1 Qualitative Interview Design

Semi-structured interviews with German fashion influencers are chosen as an approach for the qualitative section. Interviews provide in-depth and expert insights on the issue of influencer marketing, but at the same time allow the researcher to remain a less biased observer (Torbert and Taylor, 2008). The elements of the research model are addressed in the semi-structured part questions (Mayring, 2002). The model is validated by evaluating results by issue in a comparative fashion.

Influencers are a competent target group for a survey assessing the impact of influencer activity on consumers' brand awareness and purchase intention. They earn prestige and gain followers due to their personal brand commitment and the online-demonstration of this

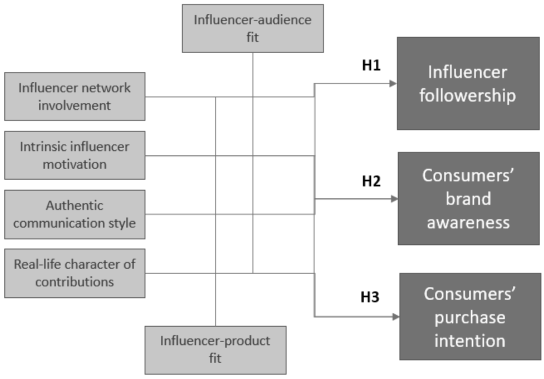

Fig. 1: Review based work model

commitment (De Veirman et al., 2017). They understand the mechanism of developing brand awareness by effectful product communication and accordingly the essence of the research model. Three female fashion influencers are recruited for the interviews, which is an acceptable number for an in-depth qualitative study (Yin, 2014). The selected influencers dispose of several 100,000 followers in the fashion segment each as of March 2020, hold central gatekeeping positions and are thus experienced enough to represent the provider perspective.

3.2 Quantitative Consumer Survey Design

A quantitative consumer survey is conducted to test the causal relationships suggested by the model using statistical hypothesis tests. It addresses consumers following the interviewed influencers. 385 participants (completed surveys) are acquired by posting a survey-link to the survey together with a positive comment below recent YouTube-contributions of the interviewed influencers. The consumer survey is anonymized and refers to the categories of the research model using several contingent research questions (items) for each category. The items scales are adopted from similar validated scales to assess followership (Kelley, 1992), brand awareness (Keller, 2001) and purchase intention (Barber et al., 2012).

The model constructs are formed from the survey items using reliability analysis. The causal relationships and hypotheses of the model are analyzed in SPSS using regression models and ANOVA tests. This methodology of causal inference is adequate for the metrically scaled data set.

3.3 Quantitative Study Representativeness

To link the quantitative study back to the interviews, the consumer survey specifically addresses the audience of the interviewed fashion influencers. To do so, a link to the recent YouTube posts of the fashion influencers containing the survey was added, asking video viewers to participate. Solely persons checking the influencers videos and reading through the YouTube comments retrieve the survey link. This method cannot avoid a further distribution of the survey link. Nevertheless, we can assume that these secondary participants are equally interested in the target influencer, since sharing happens if the secondary addressee is already involved. Altogether, 385 completed questionnaires have been collected in this way and the participants are about equally distributed between the three relevant influencers. Consequently, the survey is representative for consumers' following the interviewed influencers in the German fashion market.

This strategy limits the reach of the study to the population of the audience of these fashion influencers, however, allows to calculate the necessary sample size to reach representativeness of results. Assuming that there is no overlap between the followers of the three interviewed influencers, the total population size of their YouTube followers is 104,000. The size of the survey to reach representativeness is calculated from the size of the total population and for a certain confidence level (here 95%, resulting a z-value of 1.96) and a certain predefined error range e (here 5%), using the following rule of thumb formula for sample size (ss):

$$\mathrm{ss} = \frac{Z^2 \cdot p \cdot (1-p)}{c^2},$$

in where Z is the z-value, p stands for the percentage picking a choice (expressed as decimal and c indicates the confidence interval (also expressed as decimal).

4 EMPIRICAL RESULTS

The interview results and consumer survey are evaluated and the results are condensed to accomplish and quantify the research model. The interviewees' reflections on determiners of influencer marketing success confirm the proposed effects of influencer activity on consumers. The review has found that influencer characteristics contribute to influencer followership (H1), brand awareness (H2) and finally purchase intention (H3). In correspondence with H1, the influencers declare that businesses get into personal contact to consumers through influencers and encourage consumers to identify with particular products and brands. Followership results due to the proximity of influencers and consumers. Furthermore, followers develop brand awareness if influencers are authentic and personally identify with their model rather than just playing their role artificially. Influencers' communication policy mediates particular emotions and a special identity, which consumers assign to the product or brand. Influencer followership thus ideally contributes to brand awareness, if designed adequately (H2). In addition, followers develop own intentions to buy the product due to influencers' communication policy. Authenticity, personal engagement and the creative and innovative presentation of the product motivate consumers to adopt the product as part of their personal consumption portfolio (H3).

Hence, the three interviews confirm the relevance of the six determinants of influencer credibility derived from the review and concretize them. Participants agree that high relationship quality, intense contact and continuity are more important than a broader network. All three influencers dispose of IM originating in their personal interest in the advertised products, their creative spirit and their desire for freedom and independence on their job. They agree that the communication style should mediate honesty, correspond to personal inner attitudes, mediate true friendship, avoid exaggerate perfection and show true emotional engagement. Also content quality is of importance. Mediated contents should disclose private information, show true images, mirror daily life, provide diverting information, be plausible and evoke emotions as well as personal thoughts. Moreover, participants agree on the high importance of IP, e.g. influencers should define themselves as part or even friends of their community and feel like their followers to be perceived as authentic. IF is required according to the interviewees, as influencers have to be experienced with the products, like them personally and consider them as truly important.

A quantitative survey among followers of the interviewed influencers is conducted to test the validity of the proposed relationships. The manifest constructs measured by the questionnaire are assumed to form the latent constructs proposed in the model. To check if these constructs are unidimensional their internal consistency has been calculated using Cronbachs Alpha. Following the guidelines from Blanz (2015) all Cronbach alpha values are classified as high ($\alpha > 0.80$) and for TF and TB as very high ($\alpha > 0.90$), confirming the reliability of the latent constructs presented in Tab. 1.

Tab. 1: Construct reliability – Cronbach Alpha

Scales		Number of items	Cronbach's Alpha
IN	Influencer network involvement	3	0.87
IM	Intrinsic influencer motivation	3	0.85
IA	Authentic communication style	5	0.88
IR	Real life character of contributions	6	0.89
IF	Influencer product fit	3	0.86
IP	Influencer audience fit	3	0.88
TF	Influencer followership	6	0.90
TB	Brand awareness	6	0.91
TP	Purchase intention	4	0.87

In addition, a confirmatory factor analysis was conducted to statistically analyse the fit between the data and the coherences between manifest and latent constructs.

The confirmatory factor analysis refers separately for the dependent independent variables due to the fact, that all independent variables (IV's) have been derived from a qualitative study, while all dependent variables (DV's) have been extracted from different already validated questionnaires. Moreover, those variables have been shortened and partly adapted to fit the practical research requirements. Analysing both simultaneously could lead to constructs which are not interpretable in the conceptional realm of this research question.

Following interpretation guidelines from Hu and Bentler (1999) the results of the confirmatory factor analysis, in regard of the IV-part of the model, indicates a very good global fit, $\chi^2 = 188.83$, df = 215, $p = 0.901$, as well as a very good relative fit as shown by the Tucker-Lewis Index, TLI = 1.01, and the Root Mean Square Error, RMSEA = 0.00 [0.00, 0.01].

Regarding the global fit of the DV-part of the model, the probability value of the qui-square test is below 0.05, $\chi^2 = 2437.46$, df = 101, $p < 0.001$, indicating a rejection of the null hypothesis which states that the proposed model fits the data. Regarding the results for the relative fit, the Tucker-Lewis Index, TLI = 0.57, and the Root Mean Square Error, RMSEA = 0.25 [0.34, 0.25], furthermore indicate a poor model fit following interpretation guidelines from Hu and Bentler (1999). Therefore, a principal axis factoring analysis is conducted to explore which items may be responsible for the poor fit, to check if the hereby observed factors are interpretable and if they can be located in the conceptional realm of this research. As uncorrelated DV's are unlikely in this context and in light of the respectively poor model-fit an oblique rotation is used (Promax with a Kappa of 4), allowing for correlated constructs to obtain and evaluate a more data driven model. According to Eckey et al. (2002) as well as Bühner (2006) this usually leads to satisfactory results. Following the results from the monte carlo study conducted by MacCallum et al. (1999) the communalities in combination with the sample size will be used as an indicator if the results of a factorial analysis can be interpreted. Following the guidelines from Bühner (2006) the data can be classified as satisfactory for conducting a factorial analysis, as all communalities are above 0.50 and the sample size exceeds 300. The Kaiser-Meyer-Olkin test, KMO = 0.86, is satisfactory and does further indicate that the data is suitable for a factor analysis. The results of the Henze-Zirkler-, Mardia- and the Royston-test indicate all that no multivariate normal distribution is given. As the Bartlett test of sphericity requires such a distribution, it cannot be interpreted.

Both, the Eigenvalue criterion and the Screeplot indicate the presence of three factors which somewhat differ in their composition in contrast to the proposed model. The items TF4, TF5 and TF6 from the scale 'Influencer followership' have cross loadings on the scale 'brand awareness'. Both scales do correlate largely and significantly with each other which further indicates certain overlap of both scales, $r = 0.60$, $p < 0.001$. The same is observed for the items TB3 and TB5 from the scale 'brand awareness' regarding the scale 'purchase intention'. Both scales show strong and significant correlations as well, $r = 0.66$, $p < 0.001$. The Item TB6 from the scale 'brand awareness' does correlate distinctly higher with the scale 'purchase intention' than with 'brand awareness'. The scale 'purchase intention' and 'influencer followership' do share no cross-sectional items which fits to their weak, although significant correlation, $r = 0.23$, $p < 0.001$. The observed factorial structure did not lead to a better or at least similarly interpretable solution. Given that the three constructs have been derived from different questionnaires it appears adequate to stick to the original model as it is based on original questionnaires. As just one Item (TB6) can be clearly assigned to another scale, the current model will not be changed and used for the hypotheses tests while keeping possible distortions in mind. The constructs required for the hypothesis tests and regression models are all confirmed and no further constructs are necessary to assess the relationship.

For assessment, we carried out a multiple regression analysis. To ensure that the results of the model tests, carried out by a multiple regression analysis, can be interpreted appropriately, the necessary requirements for this procedure have been checked. The matrix scatterplot using loess smoothing indicated linear relations between all IV's. No outliers or extreme outliers have been observed. The results of the Durbin-Watson tests lie between 1.50 and 1.83 and therefor inside the interval of 1.5 and 2.5, as suggested by Brosius (2011), indicating no autocorrelations. The scatterplots, juxtaposing the standardized residuals and the predicted residuals, show no heteroscedasticity. The VIF- and TOL- values as well as the correlations seen in Tab. 2 display no presence of multicollinearity between the IV's. In conclusion all requirements are sufficiently fulfilled and the results can be interpreted without consideration of distortions.

The independent variables where standardized resulting in a mean of zero and a standard deviation of one. The usage of standardized parameters furthermore changes the unit of the beta weights from its original units to the unit of standard deviation. This eases the interpretability of the regression weights as follows: b_0 is now the intercept, and b_1 the slope for an average value of an IV.

Tab. 2: Pearson correlations of input factors

Input factors		**IM**	**IA**	**IR**	**IF**	**IP**
IN	r	−0.01	0.10	0.01	−0.06	−0.03
	p	0.928	0.052	0.926	0.259	0.620
IM	r	1	−0.02	−0.04	−0.03	0.05
	p		0.759	0.432	0.511	0.372
IA	r		1	0.01	−0.09	−0.08
	p			0.810	0.077	0.100
IR	r			1	0.05	−0.00
	p				0.315	0.985
IF	r				1	0.02
	p					0.750

Note: r is the Pearson correlation coefficient and p the level of significance.

The validated and reliability tested standardized constructs are used to test the research hypotheses.

H1 assumes that IN, IM, IA and IR impact consumers' readiness for followership positively. IF as well as IP are considered as moderating factors. Tab. 3 displays the results of the multiple regression analysis to test these assumptions of H1. All input factors IN, IM, IA and IR are significant on the one percent level. IF moderates the relations between IM, IR and influencer followership on the five percent level. In this regard H1 is accepted.

H2 assumes that IN, IM, IA and IR contribute to increase consumers' brand awareness, where IF and IP positively moderate those relationships. The results are significant for IA and IR on the one percent level. The same applies for the factor IF without being considered as a moderator. The input factors IN, IM and the moderating impact of IF and IP on all relations are not significant. In this regard H2 is accepted.

H3 assumes that the parameters IN, IM, IA and IR show a positive effect on consumers' purchase intention, positively moderated by IF and IP (compare Tab. 3). The coefficient IR is significant on the one percent level and the

Tab. 3: Results of the model analysis from multiple regression

Model	H1 Influencer Followership	H2 Brand awareness	H3 Purchase Intention
IN Influencer network involvement	0.34*** (0.02)	0.01 (0.02)	0.04* (0.02)
IM Influencer motivation	0.48*** (0.02)	−0.01 (0.02)	0.03 (0.02)
IA Authentic communication style	0.62*** (0.02)	0.55*** (0.02)	0.04* (0.02)
IR Quality of contributions	0.28*** (0.02)	0.73*** (0.02)	0.72*** (0.02)
IF Influencer audience fit	−0.01 (0.02)	0.09*** (0.02)	0.49*** (0.02)
IP Influencer product fit	−0.01 (0.02)	−0.01 (0.02)	0.13*** (0.02)
IF_IN_Interaction_term	0.01 (0.01)	−0.02 (0.02)	0.01 (0.02)
IF_IM_Interaction_term	0.04** (0.02)	0.03 (0.02)	−0.02 (0.02)
IF_IA_Interaction_term	0.01 (0.02)	0.02 (0.02)	0.03 (0.02)
IF_IR_Interaction_term	0.05** (0.02)	−0.01 (0.02)	0.06** (0.02)
IP_IN_Interaction_term	−0.02 (0.01)	−0.01 (0.02)	0.02 (0.02)
IP_IM_Interaction_term	−0.01 (0.02)	0.02 (0.02)	−0.00 (0.02)
IP_IA_Interaction_term	0.00 (0.02)	0.01 (0.02)	−0.01 (0.02)
IP_IR_Interaction_term	−0.01 (0.02)	0.00 (0.02)	0.01 (0.02)
N	384	384	384
Adjusted R^2	0.83***	0.86***	0.81***

Note: *, **, *** denote significance at the 10, 5 and 1 per cent level. All variables have been standardized before analyses. Bootstrap has been used to generate more precise standard errors. Standard errors in parentheses from the Bootstrap table.

coefficients IN and IA on the ten percent level. The coefficients IF and IP are significant on the one percent level without being considered as moderators. A moderation on the five percent level of IF has been observed for the relation between IR and purchase intention. In this regard H3 is accepted.

Based on the quantitative results the research model has been partially validated and quantified. To differentiate the results, each hypothesis is displayed in a separate part model. Fig. 2 classifies the results by relevance of input factors. Accordingly, influencer activity largely determines influencer followership. IA and perceived IM induce consumers to follow an influencer. The fit of IP or IF are insignificant in this regard. In addition, IP is moderating the relations between IM, IR and influencer followership.

The results concerning H2 show, however, that IN and IM are not relevant for consumers' brand awareness. Instead, the major determiners are IA ($\beta = 0.552$) and IR ($\beta = 0.735$), i.e. the factual output of the influencer in the form of posts. The moderator IP is of comparatively low importance (but still significant; $\beta = 0.09$), although no moderation effects have been observed.

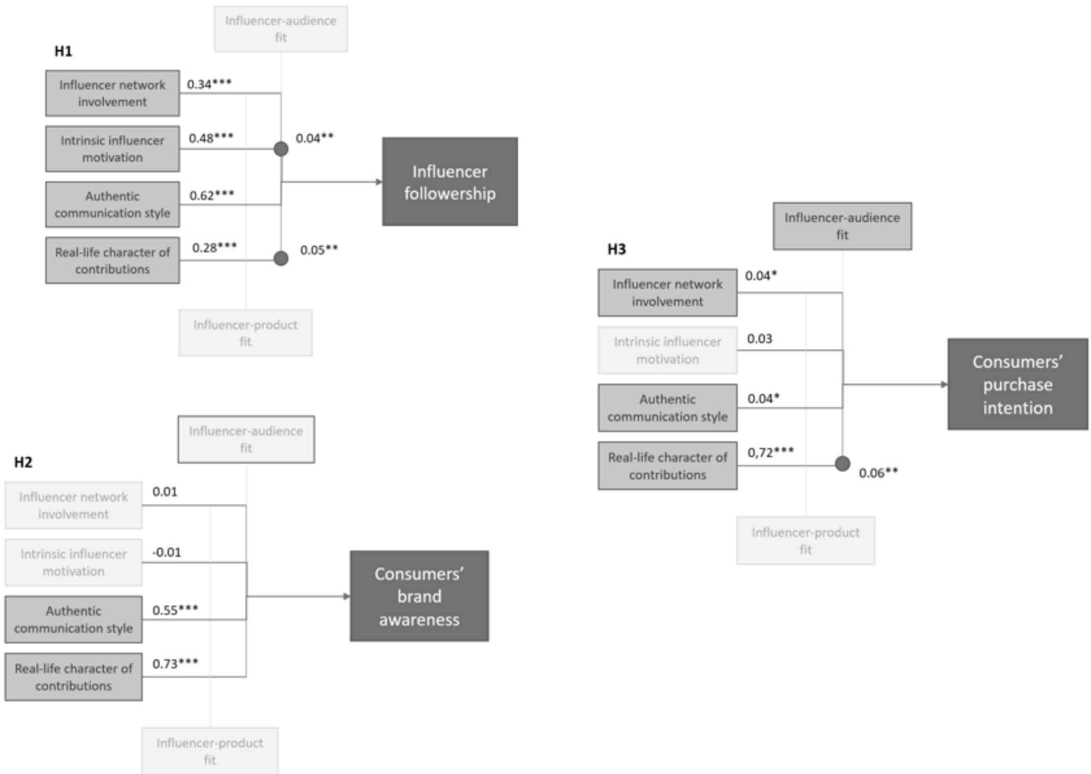

Fig. 2: Summary of empirical results

According to the results for H3, the major determiners of consumers' purchase intention are IR ($\beta = 0.72$) and the proposed moderator IP ($\beta = 0.49$), although moderation was just observed for the relation between IR and purchase intentions. IA ($\beta = 0.04$), IN ($\beta = 0.04$) and IF ($\beta = 0.13$) do also show significant impact, which however is less compared to the other factors.

These observations imply that the major determiners of the three target parameters influencer followership, brand awareness and purchase intention differ. While followership is achieved by high IM and IA, brand awareness mainly depends on IR and IR. Purchase intention mainly results from IR and IP.

One possible interpretation is that consumers' followership mainly depends on the bond influencers establish between their own personality and the audience, while consumers' product and brand related behavior rather depends on the contents of influencer messages than on the influencer herself. Yet, the target factors of the models – influencer followership, brand awareness and purchase intention – are strongly interdependent. Most consumers following the influencers equally show brand awareness and purchase intention.

These results imply that each of the determiners of influencer activity as identified from the review and influencer interviews is confirmed by the model and finally increases influencer followership, consumers' brand awareness and purchase intention. Consumers are motivated to follow due the influencers due to his or her personality (IN and IM) mainly but develop brand awareness and purchase intention due to the authenticity and real-life character of posts. All four factors interact in getting consumers involved with the brand and product.

5 DISCUSSION

The empirical study has combined a qualitative and a quantitative section, where the qualitative section focused on the supply side and the quantitative section on the demand side of the German consumer market.

Interviews with three influencers in the German fashion consumer segment have been conducted which led to a concretization of the research hypotheses. According to the experience of the influencers the major determiners of influencer effectiveness with consumers are IN, IM, IA and IR. Influencers admit that the fit IF as well as IP could moderate influencers' impact. The interviews further suggest that influencer followership, brand awareness and purchase intention are mutually interdependent objectives of influencer marketing activity.

The second part of the empirical study has been implemented in the form of a consumer survey. Specifically, consumers following the interviewed influencers have been addressed to control the influence of secondary moderators. The survey fully confirms the research model derived from the hypotheses. IN, IM, IA and IR determine consumers' influencer followership, consumers brand awareness and purchase intention. There are interrelationships between the target factors. More specifically, the survey shows that IN and IM have an effect on consumers' followership, while IA and IR impact brand awareness and purchase intention above all. The assumed moderators of IP and IF are of comparatively low relevance as compared to the dominant impact of the four determiners.

Thus, we provide a model and causal explanation of the impact of influencer activity on consumers. Our empirical study has contributed to identify the most important determiners for the German B2C fashion influencer segment. Moreover, we identified that influencers' personal engagement and the quality of their contributions interact to motivate consumers followership, brand awareness and purchase intentions.

Generally, we offer the various major contributions, which distinguish our study from earlier research in influencer marketing effectiveness.

As the review of empirical studies (chapter 2) has shown, to date no study on the impact of influencer marketing in the German B2C consumer segment has been conducted so far. To our best knowledge, this study is the first to conclusively explore the relevance of influencer on the formation of consumer attitudes and purchase intentions in German fashion marketing. Earlier studies were not conclusive with regard to the evaluated influencer marketing strategies (Lock, 2016; De Veirman et al., 2017; Belanche et al., 2019; Breves et al., 2019; Eroğlu and Bayraktar Köse, 2019). Thus, we collected all relevant determiners of influencer marketing impact in the form of a thorough review and selected the most important factors for the German B2C fashion influencer segment in a targeted way by conducting interviews with three experienced influencers.

Also, previous literature is not fully comprehensive concerning the impact of influencer marketing at the consumer level. While some assess purchase intention (Lim et al., 2017; Gadalla et al., 2019; Hughes et al., 2019; Istania et al., 2019; Riedl and von Luckwald, 2019), other focus on the impact on branding (Dalstam et al., 2018; Kucharska, 2018; Xiao et al., 2018; Belanche et al., 2019; Jin et al., 2019; De Veirman et al., 2017; Lock, 2016). Hence, we contribute by interconnecting the effects of influencer marketing with the most important elements of the purchase funnel within one single study. In addition, this study has explicitly analyzed the potential effects of moderating factors in the interviews as well as in the survey, as prior studies have vaguely referred to potential moderators of influencer activity. In accordance with previous research (Chae et al., 2016; Istania et al., 2019) we also found moderate impacts.

Finally, the study is the first to comprehensively assess all major elements that determine the effectiveness influencer marketing strategies at the consumer level. Earlier studies have been incomprehensive concerning the relevant determiners and their interaction (Xiao et al., 2018; Belanche et al., 2019; Jin et al., 2019; De

Veirman et al., 2017; Lock, 2016). Specifically, our study shows that the identified elements of influencer marketing design are independent and each factor is itself important to enhance the impact of influencer campaigns.

In sum, we postulated a model of determiners and moderators of influencer marketing effectiveness on the purchase funnel, in which we systematically extracted the key factors relevant to the German B2C fashion segment based on influencer interviews and finally confirmed the developed causal model successfully in an empirical consumer survey. The resulting model is thus validated and reliable and can guide further research in influencer marketing effectiveness in fashion and other sectors, in Germany and abroad.

The insights of the empirical study also offer valuable insights to marketing management practice. Influencers have become an important instrument in consumer marketing. So far, their impact on consumers' decision processes has found less attention in Germany than in other countries. In the age of digitalization, German consumers however follow similar trends as consumers in other industrialized countries, have become equally susceptible to social media contributions and in result, to influencers. The consumer survey illustrates that many participants are still loosely interested in the contributions of the influencers only and do not base their brand and product decisions on influencer contributions alone. Comparing these results to previous insights from other national context, influencer marketing is just about to gain ground in Germany and by now should not be applied as a standalone marketing strategy. Marketing experts in the consumer segment should certainly consider targeted influencer campaigns as part of their marketing strategy. This advance could endow companies with a first mover advantage in a just emerging marketing domain. Businesses should acquire influencers in a targeted way and choose personalities adequate for the communication of the respective product, based on their personal interest with regard to the target audience.

The study has assessed determiners of the impact of influencer activity on consumers in the German fashion segment based on a comprehensive qualitative and quantitative study that considers the supplier (influencer) and consumer perspective. The insights of the study are conclusive. The theory-founded model has been refined and adjusted to the German fashion sector in the qualitative section and has been confirmed in the course of a consumer survey of 385 participants. Still the study has got some limitations, which invite further empirical research.

To begin with, the qualitative study is based on three interviews with influencers in the German fashion sector. These are not representative for the German influencer business and not even for the German fashion influencer business as a whole since several thousands of influencers are listed online. The interviews rather represent the individual opinions of the three participants. These statements have been compared and adjusted to differentiate the research model. This strategy necessarily leads to exemplary categories, which would not necessarily occur, if other influencers were interviewed. The comprehensive review-based model design has thus been refined in a partly arbitrary way.

The second limitation refers to the consumer survey with 385 participants. All of them are followers of the previously interviewed influencers. This approach ensures the internal validity of the study but prevents external validity. The results of the survey concern the contributions of the selected three influencers only. Other influencers might make completely different statements and contributions, dispose of other followers who would probably develop different opinions on the determiners of influencer effectiveness. Hence, the results are not generalizable beyond the reach of the three influencers. The extension of the interview section to further influencers could amend on this problem but would not fully solve the issue of representativeness. Still, the selected influencers are typical for the German fashion market, which indicates that somewhat similar results should be expected for other German fashion influencer networks.

Furthermore, regression analysis was applied to assess the postulated hypotheses. This kind of analysis assess linear impacts of several determiners on a single target only. It neglects interactions between the determiners and moderators and does not directly include cross-relationships between several targets. The complementary correlation analysis used to determine interactions between the targets is possibly biased since it does not consider the mutual interdependencies between the input factors. A structural equation model could amend on this problem and should possibly be applied in follow-up studies to illustrate the full complexity of the impact of influencer strategies on consumer attitudes and behavior.

6 CONCLUSIONS

The study has shown that the choice of influencers is essential to the success of the campaign. First, influencers' personal engagement is decisive to motivate consumers to follow the influencer. The quantitative consumer survey has confirmed that the readiness to subscribe to the influencers' contributions depends significantly on IN and IM. The interviewees, however, assert that intensity of IN is more important to marketing success compared to a broad network reach. Marketing experts recruiting influencers should be aware that the influencers' contributions are decisive to influencers' impact on consumers brand and product choice, in so far that perceived IA and IR of influencers determine whether consumes develop brand awareness and form purchase intentions towards the advertised products.

Finally, the review section has shown that influencer marketing strategies should fit with the brand image the company pursues as a whole. The interviewed influencers have pointed out that self-reliance is essential to provide authentic and real-life type contributions. If businesses relegate influencers' freedom of contents and style too much, they risk that influencers loose in authenticity and in result in consumer effectiveness. Businesses intending to include influencer marketing into their strategy should balance between the development of a comprehensive marketing guideline and the admission of self-defined advertisement styles.

7 REFERENCES

Baker, S. J. 2014. *Monetizing Influencers: Diffusion of Innovations and Influencer Marketing Through Blogger Outreach*. University of Central Florida.

Barber, N., Kuo, P.-J., Bishop, M. and Goodman, R. 2012. Measuring Psychographics to Assess Purchase Intention and Willingness to Pay. *Journal of Consumer Marketing*, 29 (4), 280–292. DOI: 10.1108/07363761211237353.

Bauer, H. H., Mäder, R. and Huber, F. 2002. Markenpersönlichkeit als Determinante der Markenloyalität. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 54 (8), 687–709. DOI: 10.1007/BF03372692.

Belanche, D., Casaló, L. V., Flavián, M., Flavián, C. and Ibáñez-Sánchez, S. 2019. Is Congruity Essential for Influencer Marketing? Exploring Fashion Influencer Campaigns on Instagram. In *The 2019 WEI International Academic Conference Proceedings*, pp. 52–54.

Biaudet, S. 2017. *Influencer Marketing as a Marketing Tool: The Process of Creating an Influencer Marketing Campaign on Instagram*. International Business Thesis.

Biel, A. L. 2001. Grundlagen zum Markenwertaufbau. In Esch, F.-R. (ed.). *Moderne Markenführung*, pp. 61–90. 3rd ed. Wiesbaden: Springer.

Blanz, M. 2015. *Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen*. Stuttgart: Kohlhammer.

Breves, P. L., Liebers, N., Abt, M. and Kunze, A. 2019. The Perceived Fit between Instagram Influencers and the Endorsed Brand: How Influencer-Brand Fit Affects Source Credibility and Persuasive Effectiveness. *Journal of Advertising Research*, 59 (4), 440–454. DOI: 10.2501/JAR-2019-030.

Brosius, F. 2011. *SPSS 19*. 1st ed. Heidelberg: Verlagsgruppe Hüthig Jehle Rehm. ISBN 978-3-8266-9038-9.

Bühner, M. 2006. *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.

Cakim, I. M. 2009. *Implementing Word of Mouth Marketing: Online Strategies to Identify Influencers, Craft Stories, and Draw Customers*. New Jersey: John Wiley & Sons.

Cannon, J. P., Doney, P. M., Mullen, M. R. and Petersen, K. J. 2010. Building Long-Term Orientation in Buyer–Supplier Relationships: The Moderating Role of Culture. *Journal of Operations Management*, 28 (6), 506–521. DOI: 10.1016/j.jom.2010.02.002.

Carbonaro, S. and Votava, C. 2005. Symbole des Seins: Masse und Klasse haben als Gegensätze ausgedient. Die neue Konsumkultur sucht Qualität in jeder Lebenslage. *GDI Impuls: Wissensmagazin für Wirtschaft, Gesellschaft, Handel*, 26–30.

Chae, I., Stephen, A. T., Bart, Y. and Yao, D. 2016. Spillover Effects in Seeded Word-of-Mouth Marketing Campaigns. *Marketing Science*, 36 (1), 89–104. DOI: 10.1287/mksc.2016.1001.

Cooper, D. R. and Schindler, P. S. 2014. *Business Research Methods*. 12th ed. New York: McGraw-Hill/Irwin.

Court, D., Elzinga, D., Mulder, S. and Vetvik, O. J. 2009. The Consumer Decision Journey. *McKinsey Quarterly*, 3, 1–11.

Dalstam, M., Holmgren, D. and Nordlöf, H. 2018. *The NA-KD Truth About Influencer Marketing: Exploring Influencer Marketing Through Integrated Marketing Communication and the Influencer's Role in Strengthening a Brand*. Jönköping University, Working Paper.

Davies, J. 2017. State of Social Platform Use in Germany in 5 Charts. *Digiday UK* [online]. Available at: `https://digiday.com/marketing/state-social-platform-use-germany-5-charts/`. [Accessed 2020, July 17].

De Veirman, M., Cauberghe, V. and Hudders, L. 2017. Marketing Through Instagram Influencers: The Impact of Number of Followers and Product Divergence on Brand Attitude. *International Journal of Advertising*, 36 (5), 798–828. DOI: 10.1080/02650487.2017.1348035.

Dlodlo, N. 2014. Profiling Marketplace Change Agents (Influential) Using the Multiple Flow Communication Theory. *Mediterranean Journal of Social Sciences*, 5 (20), 705–712. DOI: 10.5901/mjss.2014.v5n20p705.

Dron, R. and Mohamad, M. 2015. *Electronic Word of Mouth for Mobile Fitness Application: An Action Case Study*. University of Salford, Working Paper. DOI: 10.13140/RG.2.1.3584.2645.

Eckey, H.-F., Kosfeld, R. and Rengers, M. 2002. *Multivariate Statistik: Grundlagen – Methoden – Beispiele*. Wiesbaden: Gabler.

Eroğlu, F. and Bayraktar Köse, E. 2019. Utilization of Online Influencers as an Experiential Marketing Tool: A Case of Instagram Micro-Celebrities. *Journal of International Social Research*, 12 (63), 1057–1067. DOI: 10.17719/jisr.2019.3297.

Evans, N. J., Phua, J., Lim, J. and Jun, H. 2017. Disclosing Instagram Influencer Advertising: The Effects of Disclosure Language on Advertising Recognition, Attitudes, and Behavioral Intent. *Journal of Interactive Advertising*, 17 (2), 148–149. DOI: 10.1080/15252019.2017.1366885.

Fournier, S. 1998. Consumers and Their Brands: Developing Relationship Theory in Consumer Research. *Journal of Consumer Research*, 24 (4), 343–373. DOI: 10.1086/209515.

Franklin, B. 2008. The Future of Newspapers. *Journalism Practice*, 2 (3), 306–317. DOI: 10.1080/17512780802280984.

Gadalla, E., Liu, R., Martin, F. and Supatchaya, N. T. 2019. Persuasive or Not?: The Effect of Social Media Influencer's Credibility on Consumer Processing and Purchase Intention. In *British Academy of Management 2019 Conference*.

Gnann, C. 2008. *Angewandte Markenforschung: Electronic Design Automation im Hochfrequenz-Markt*. Hamburg: Diplomica Verlag.

Herrmann, A. and Schaffner, D. 2005. Planung der Produkteigenschaften. In Alberts, S. and Gassmann, O. (eds.). *Handbuch Technologie- und Innovationsmanagement*, pp. 379–396. Wiesbaden: Gabler.

Homburg, C. and Koschate, N. 2007. Kundenzufriedenheit und Kundenbindung. In Sönke, A. and Herrmann, A. (eds.). *Handbuch Produktmanagement: Strategieentwicklung – Produktplanung – Organisation – Kontrolle*, pp. 844–867. 3rd ed. Wiesbaden: Gabler.

Hu, L. and Bentler, P. M. 1999. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*, 6 (1), 1–55. DOI: 10.1080/10705519909540118.

Hu, M., Milner, J. and Wu, J. 2015. Liking and Following and the Newsvendor: Operations and Marketing Policies Under Social Influence. *Management Science*, 62 (3), 867–879. DOI: 10.1287/mnsc.2015.2160.

Hughes, C., Swaminathan, V. and Brooks, G. 2019. Driving Brand Engagement Through Online Social Influencers: An Empirical Investigation of Sponsored Blogging Campaigns. *Journal of Marketing*, 83 (5), 78–96. DOI: 10.1177/0022242919854374.

Hutter, K. and Mai, R. 2013. *Effective Incentives for Buzz Marketing: How Moral Concern Moderates the Willingness to Engage as Buzz Agents*. Working Paper of ZBW, Leibniz Information Centre for Economics.

ISTANIA, F., PRATIWI, I. P., YASMINE, M. F. and ANANDA, A. S. 2019. Celebrities and Celebgrams of Cosmetics: The Mediating Effect of Opinion Leadership on the Relationship Between Instagram Profile and Consumer Behavioral Intention. *International Journal of Scientific and Technology Research*, 8 (8), 75–86.

JAAKONMÄKI, R., MÜLLER, O. and VOM BROCKE, J. 2017. The Impact of Content, Context, and Creator on User Engagement in Social Media Marketing. In *Proceedings of the 50th Hawaii International Conference on System Sciences*. DOI: 10.24251/HICSS.2017.136.

JANSON, A. 2012. *Der Kunde im Fokus: Das Konzept der Customer Journey*. Studienarbeit GRIN.

JIN, S. V., MUQADDAM, A. and RYU, E. 2019. Instafamous and Social Media Influencer Marketing. *Marketing Intelligence and Planning*, 37 (5), 567–579. DOI: 10.1108/MIP-09-2018-0375.

KELLER, K. L. 1993. Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*, 57 (1), 1–22. DOI: 10.2307/1252054.

KELLER, K. L. 2001. *Building Customer-Based Brand Equity: A Blueprint for Creating Strong Brands*. Cambridge, MA: Marketing Science Institute.

KELLER, E. and FAY, B. 2016. *How to Use Influencers to Drive a Word-of-Mouth Strategy*. Research Paper, Marketing Best Practice [online]. Available at: https://www.engagementlabs.com/wp-content/uploads/2016/05/How_to_use_influencers_to_drive_a_wordofmouth_strategy_.pdf. [Accessed 2020, July 17].

KELLEY, R. E. 1992. *The Power of Followership*. New York, NY: Doubleday Business.

KUCHARSKA, W. 2018. Personal Branding – A New Competency in the Era of the Network Economy: Corporate Brand Performance Implications. In GOLINSKA-DAWSON, P. and SPYCHAŁA, M. (eds.). *Corporate Social Responsibility in the Manufacturing and Services Sectors*, pp. 19–34. DOI: 10.1007/978-3-642-33851-9_2.

LIM, X. J., MOHD RADZOL, A. R., CHEAH, J.-H. and WONG, M. W. 2017. The Impact of Social Media Influencers on Purchase Intention and the Mediation Effect of Customer Attitude. *Asian Journal of Business Research*, 7 (2), 19–36. DOI: 10.14707/AJBR.170035.

LOCK, A. C. 2016. Impact of Brand Knowledge on Brand Trust in Private Higher Education Institutions: How do Word of Mouth Sources Intervene? *Sarjana*, 31 (2), 13–32.

MAYRING, P. 2002. *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken.* 5th ed. Basel: Beltz.

MEYER, C. and SCHWAGER, A. 2007. Understanding Customer Experience. *Harvard Business Review*, 85 (2), 116–126.

MACCALLUM, R. C., WIDAMAN, K. F., ZHANG, S. and HONG, S. 1999. Sample Size in Factor Analysis. *Psychological Methods*, 4 (1), 84–99.

MCQUARRIE, E. F., MILLER, J. and PHILLIPS, B. J. 2013. The Megaphone Effect: Taste and Audience in Fashion Blogging. *Journal of Consumer Research*, 40 (1), 136–158. DOI: 10.1086/669042.

RIEDL, J. and VON LUCKWALD, L. 2019. *Effects of Influencer Marketing on Instagram*. Open Science Publications of Access Marketing Management: A non-commercial scientific association.

ROWLEY, J. 2004. Just Another Channel? Marketing Communications in E-Business. *Marketing Intelligence and Planning*, 22 (1), 24–41. DOI: 10.1108/02634500410516896.

SAFKO, L. 2010. *The Social Media Bible: Tactics, Tools, and Strategies for Business Success*. New Jersey: John Wiley & Sons.

SCHMITT, B. 2012. The Consumer Psychology of Brands. *Journal of Consumer Psychology*, 22 (1), 7–17. DOI: 10.1016/j.jcps.2011.09.005.

TAN, T. and MING, M. 2003. Leveraging on Symbolic Values and Meanings in Branding. *Journal of Brand Management*, 10 (3), 208–218. DOI: 10.1057/palgrave.bm.2540117.

TORBERT, W. R. and TAYLOR, S. S. 2008. Action Inquiry: Interweaving Multiple Qualities of Attention for Timely Action. In REASON, P. and BRADBURY, H. (eds.). *The SAGE Handbook of Action Research: Participative Inquiry and Practice*, Part 2, Chapter 16, pp. 239–251. DOI: 10.4135/9781848607934.n24.

WEBSTER, J. and WATSON, R. T. 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26 (2), 13–23.

WOODS, S. 2016. *The Emergence of Influencer Marketing*. University of Tennessee Honors Thesis Projects.

XIAO, M., WANG, R. and CHAN-OLMSTED, S. 2018. Factors Affecting YouTube Influencer Marketing Credibility: a Heuristic-Systematic Model. *Journal of Media Business Studies*, 15 (3), 188–213. DOI: 10.1080/16522354.2018.1501146.

YIN, R. K. 2014. *Case Study Research: Design and Methods*. 5th ed. Thousand Oaks, CA: Sage.

AUTHOR'S ADDRESS

Kai Dominik Renchen, SRH Hochschule Heidelberg, Ludwig-Guttmann-Str. 6, 69123 Heidelberg, Germany; Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: kr@renchen1.de

QUALITY OF WORD VECTORS AND ITS IMPACT ON NAMED ENTITY RECOGNITION IN CZECH

František Dařena[1], Martin Süss[1]

[1] *Mendel University in Brno, Czech Republic*

EUROPEAN JOURNAL OF BUSINESS SCIENCE AND TECHNOLOGY

Volume 6 Issue 2
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

Named Entity Recognition (NER) focuses on finding named entities in text and classifying them into one of the entity types. Modern state-of-the-art NER approaches avoid using hand-crafted features and rely on feature-inferring neural network systems based on word embeddings. The paper analyzes the impact of different aspects related to word embeddings on the process and results of the named entity recognition task in Czech, which has not been investigated so far. Various aspects of word vectors preparation were experimentally examined to draw useful conclusions. The suitable settings in different steps were determined, including the used corpus, number of word vectors dimensions, used text preprocessing techniques, context window size, number of training epochs, and word vectors inferring algorithms and their specific parameters. The paper demonstrates that focusing on the process of word vectors preparation can bring a significant improvement for NER in Czech even without using additional language independent and dependent resources.

KEY WORDS

Named Entity Recognition, word embeddings, word vectors training, natural language processing, Czech language

JEL CODES

C63, C88

1 INTRODUCTION

Named Entity Recognition (NER) is one of the important subtasks of Information Extraction. It focuses on finding named entities in text and classifying them into one of the entity types. The types typically include persons, locations, organizations, temporal expressions, phone numbers, but sometimes also product names, brands, diagnoses, drug types, or pub-

DAŘENA, František, and SÜSS, Martin. 2020. Quality of Word Vectors and its Impact on Named Entity Recognition in Czech. *European Journal of Business Science and Technology,* 6 (2): 154–169. ISSN 2694-7161, DOI 10.11118/ejobsat.2020.010.

lishers (Goyal et al., 2018; Nadeau and Sekine, 2007).

Named entities can be extracted using several approaches. The *knowledge-based*, also known as rule-based approach relies on the availability of various lexicons and domain-specific knowledge (Yadav and Bethard, 2018). Knowledge-based systems can be usually easily implemented but it is difficult to define all necessary rules. The systems usually have high precision but, on the other hand, lower recall and fail on unknown cases.

Machine learning approaches strive to eliminate the problems with hand-crafted rules. The NER problem is being solved with a model automatically created by a computer. Systems using supervised learning require annotated corpora (a text with marked entities) and a learning algorithm that can automatically extract the rules for detecting entities. Systems based on unsupervised learning (no labeled data is available) require only some syntactic patterns to identify candidates for entities that can be further evaluated and disambiguated (Etzioni et al., 2005; Nadeau et al., 2006).

The crucial aspect of learning a NER model is the selection of the appropriate features. Modern state-of-the-art NER approaches avoid using hand-crafted features and rely on feature-inferring neural network systems based on word embeddings. These systems often outperform the systems using engineered features, even when they have access to domain-specific rules or lexicons (Yadav and Bethard, 2018).

A lot of research concentrates on massively used languages, like English, German, or Spanish and there have been many approaches to named entity recognition developed. For the Czech language, the situation is quite different as there is a delay in current research. There exist only a few named entity recognizers and not much attention has been devoted to the optimization of all steps of the NER procedure. A typical example is a process of preparing word vectors to be used in the NER task. The goal of the paper is thus to analyze the impact of different aspects related to word embeddings on the process and results of the named entity recognition task in Czech. The goal is not to achieve the best results and beat the current state-of-the-art approaches, which usually requires using other language-dependent resources, but to discover how different algorithms, their parameters, or the size and quality of corpora used for training can influence the result.

2 CURRENT STATE

The methods of NER often employ statistics (e.g., Conditional Random Fields – see Tkachenko and Simanovsky, 2012; Hidden Markov Models – see Zhou and Su, 2002), classification algorithms (e.g., support vector machine – see Li et al., 2005), or neural approaches (Collobert et al., 2011). In the past, classical machine learning models like SVM or logistic regression strongly relying on feature engineering were popular in NER (Goldberg, 2016). The features generally belonging to one of the three categories – document, corpus, and word-based features (Goyal et al., 2018) usually include, e.g., word length, capitalization, presence in an external list, part-of-speech, position in a sentence, the occurrence of a period or hyphen, suffixes, prefixes, or orthographic features (Zhou and Su, 2002; Tkachenko and Simanovsky, 2012).

Later, it has been found that neural models (especially deep neural models) able to learn important features directly from texts could be used also for NER. A prevalent approach is now based on neural networks with architectures such as bidirectional or convolutional LSTM (Lample et al., 2016; Chiu and Nichols, 2016; Rudra Murthy and Bhattacharyya, 2018; Chen et al., 2018). Such architectures that are suitable for processing sequential data as they have a form of memory are successfully used also in other natural language processing tasks (Mikolov et al., 2015). After a pioneering publication on word vectors training using the word2vec algorithm (Mikolov et al., 2013a), the

NER research was aimed at using word vectors also in NER in many natural languages (Nguyen et al., 2019; El Bazi and Laachfoubi, 2019; Seok et al., 2016). Word vectors (Collobert et al., 2011), which are vectors representing individual words, are able to capture the syntactic as well as semantic regularities of a language which has been found to be beneficial in many NLP tasks.

In order to learn word vectors using a neural model, texts need to be converted to a structured representation (vectors) first. The procedure can generally include several preprocessing steps like, e.g., text cleaning, white space removal, case folding, spelling errors corrections, abbreviations expanding, stemming, stop words removal, or negation handling (Dařena, 2019). For word embeddings training, some preprocessing can be applied too (Li et al., 2017; Leeuwenberg et al., 2016) which can have an impact on the context of the words, the number of unique words, and global word frequencies. Subsequently, one-hot encoded vectors (vector where only one out of its units is 1 and all others are 0) that act as the inputs and outputs of the neural models are derived (Rong, 2014). Various sets of word embeddings trained on different corpora (e.g., Wikipedia) are instantly available. Different algorithms can be also used to train their own set of embeddings, that are suitable for general use or specific task. The algorithms have various parameters that need to be set with respect to a given task. Current approaches to NER using word embeddings, however, often use the default parameters settings, and the impact of alternative settings is not evaluated.

Besides the core features derived from the text in a neural model, additional language-dependent (presence in a list of cities, countries, first and last names, days of a week, currencies, part-of-speech, singular/plural) or language-independent features (context, position, word length, fixed length prefix/suffix, presence of a hyphen) can be on the input of a NER system (Chiu and Nichols, 2016).

2.1 NER for the Czech Language

Ševčíková et al. (2007) presented the first NER system for the Czech language using decision trees analyzing handcrafted features to detect and classify entities in text. Kravalová and Žabokrtský (2009) implemented another system using SVM for classification. Král (2011) implemented a NER system for a specific purpose (searching the Czech press agency database) and demonstrated that feature selection plays a crucial role in designing a NER system. He proved that language independent features are more important than the dependent ones. Konkol and Konopík (2011) created a NER system using the Maximum Entropy algorithm which used semantic spaces that were created using the COALS method (Rohde et al., 2004). It is the first work that treated words as vectors in a multidimensional space. Another system employing Conditional Random Fields using different features and resources was presented by Konkol and Konopík (2013). In the same year, Straková et al. (2013) published another NER system for Czech using Maximum Entropy Markov Model.

The first system that employed word vectors trained using word2vec was presented by Demir and Özgür (2014). Although it used only language independent features it outperformed all existing NER systems for Czech. A better performance was later achieved by Straková et al. (2016) who use a neural network with gated recurrent units together with word vectors representing original or lemmatized words, part-of-speech tags, prefixes, suffixes, or vector representations of characters. The word vectors in both systems were trained using word2vec with the skipgram architecture. The best performance was brought by Konopík and Pražák (2018). They use a deep neural model with LSTM layers encoding character sequences and word sequences together with a wider context information obtained from Latent Dirichlet Allocation. The word sequence layer was using pretrained GloVe and fastText word vectors.

The systems using word vectors were able to improve the performance expressed by the F1-measure by a few percent. At the same time, the features did not need to be engineered manually because many useful properties and relations were encoded in the vectors. Most of the systems, however, relied on pretrained word vectors or created the vectors using default parameters of the algorithms.

2.2 Learning Word Embeddings

In machine learning, there is a general problem with choosing the right set of features for the given task (Blum and Langley, 1997). In natural language processing, there is an additional problem related to the classical representation of features derived from texts (known as bag-of-words). In this model, each word or another feature is represented by one dimension in a multidimensional space for representing the documents. Such a value does not enable sharing some information across features and is thus is independent of the others.

To solve the problem with no similarity among features, it would be possible to add other information to the existing features to better capture the context in which they appear. This, however, increases the number of dimensions in the input space and requires the combination of possible feature components to be carefully selected (Goldberg, 2016).

Some of the modern representations of texts use more dimensions to represent each word or feature. The words are embedded in a continuous multidimensional space that has typically a few hundred dimensions so we talk about word embeddings. Finding suitable values of the vector elements is based on the hypothesis stating that words in similar contexts have similar meanings (Levy and Goldberg, 2014). Because similar words (e.g., synonyms) share some information, the values of their vector elements should be similar and the vectors are located close to each other in the multidimensional space.

Popular approaches leading to generating such vectors include models using global matrix factorization like Latent Semantic Analysis (LSA) or Latent Dirichlet Allocation (LDA) and models learned by neural networks using a small context window (where word2vec is probably the most popular), see Mikolov et al. (2013a), Pennington et al. (2014). Supervised methods create embeddings that are trained towards the given goal and can capture information that is relevant for the task. They, however, require annotated data for the specific task. Unsupervised methods do not require annotated data. Their only goal is to compute embeddings that are usually learned in the task of predicting a word given its context or deciding, whether a word can belong to a context given examples of real and randomly created word-context pairs (Goldberg, 2016). Such embeddings capture general syntactic and semantic relationships and can be applied in a wide variety of tasks. When there is not enough data for domain-specific embeddings training available a model created on a general corpus can be adjusted using a smaller amount of domain-specific data (Yen et al., 2017).

Famous methods that can be used to compute word embeddings include:

- **word2vec** – a family of methods proposed by Mikolov et al. (2013a) that strongly attracted the NLP community to neural language models. The method predicts a word based on its context (the Continuous bag-of-words or CBOW approach) or the context for a word (the skipgram approach). Word2vec tried to eliminate the problems with the computational complexity of the existing neural language models. In the training phase, a neural network uses a linear activation function instead of the sigmoid function, which is typical for a multilayer perceptron, and the logarithm of the probability of predicting the word or its context is being maximized.
- **GloVe** – a model uses information about global co-occurrences of words. Word vectors are used in the task of predicting the probability with which two words co-occur. The probabilities can be calculated from a term-term matrix created from a corpus. The prediction is made by a function that takes word vectors as the input. The word

vectors are calculated in the process of word co-occurrence matrix factorization using stochastic gradient descent (Pennington et al., 2014).

- **fastText** – a model derived from word2vec, treats each word as a bag of character n-grams (which enables considering sub-word information very important for morphologically rich languages) where the vectors are associated at the n-gram level. The vector for a word is calculated as the sum of n-gram vectors. This enables, compared to word2vec and GloVe, creating vectors for words that are not in the training data (Bojanowski et al., 2017).

The skipgram technique in both word2vec and fastText algorithms can better capture semantic regularities of words. On the other hand, the CBOW approach captures syntactic regularities better (Mikolov et al., 2013b).

The methods of embeddings training require several parameters to be set – the number of vector dimensions, definition of the context (size and position), maximal number of unique words, minimal frequency of a word, number of training epochs, etc. which can significantly influence the quality of the learned vectors (Levy et al., 2015).

3 DATA AND METHODS

The quality of word vectors depends on the corpus on which they are trained. Generally, the more data is available, the better. However, the number of unique tokens, amount of errors, writing style, domain to which the texts are related etc. play a significant role also in the NER task. Here, what are an entity and its type often depends on the domain (Kulkarni et al., 2016). The number of unique tokens is especially high for morphologically rich languages where different forms of a word have an impact on the number of global occurrences as well as the number of combinations with other words. Here, normalization techniques, like stemming, lemmatization, case folding, or stop words removal can be considered (Levy et al., 2015).

3.1 Data

For the Czech language, no corpora suitable for the NER task existed before 2007 when the Czech Named Entity Corpus (CNEC) 1.0 was released (Ševčíková et al., 2007). The corpus was later extended, simplified, and transformed to a format similar to the one used by the SIGNLL Conference on Computational Natural Language Learning (CoNLL) and evolved to so-called Extended-CNEC corpus (Konkol and Konopík, 2013). The corpus in version 2.0 defining seven most commonly used entity types (numbers in addresses, geographical names, institution names, media names, artifact names, personal names, and time expressions) was used for training our NER system. This corpus has also been used by most of the other researchers so a comparison with previous research is possible.

To evaluate the impact of corpus size and quality, which are the factors influencing the quality of word vectors (Levy et al., 2015) used in the NER system, different corpora were used to learn word vectors. CWC-11 is a Czech corpus based on selected newspaper articles, blogs, and discussions on the Czech web (Spoustová and Spousta, 2012). CoNLL-2017 is a corpus released for the CoNLL conference in 2017. It contains also documents in Czech, especially from Wikipedia and other Internet sources (Zeman et al., 2018). CZES is a Czech corpus containing data from news webs from the years 1995–1998 and 2002 (Masaryk University, 2011). SYN-2015 is a part of the SYN corpus consisting of journalistic, technical, and fiction papers from the years 2010–2014 (Křen et al., 2016). EuroParl is a relatively small and specialized corpus containing texts related to the European parliament agenda in the years 1996–2011. Detailed characteristics of the data collections can be found in Tab. 1.

Tab. 1: Selected corpora containing open Czech text

Corpus	Number of words	Number of unique tokens
CoNLL-2017 (Czech part)	1.62 billion	21.5 million
CWC-2011 – articles	628 million	1.8 million
CZES	497 million	3.5 million
EuroParl (Czech part)	13 million	304 thousand
Czech Wikipedia extraction	134 million	2.6 million
SYN-2015	121 million	1.4 million
Extended CNEC 2.0	199 thousand	35 thousand

The CoNLL-2017 corpus is the largest but probably of the lowest quality, according to the number of unique tokens. The CWC-2011, CZES, and SYN-2015 corpora contain a lower number of unique tokens so the tokens should appear with higher frequencies. The difference between these corpora is mainly in their size. The EuroParl corpus contains data from a specific domain and is relatively small. Because the CNEC and EuroParl corpora are rather small, the neural network implementing NER is allowed to update the word vectors (the vectors are trainable). Although this is usually not useful, updating word vectors with respect to a specific task might be a good option when the word vectors are not good enough (Hope et al., 2017). This fine-tuning for each task can also give an extra boost to the NER system performance (Collobert et al., 2011). When the other corpora are used to train word vectors, the vectors are fixed during the NER system training.

The Extended CNEC 2.0 corpus, which is primarily used to train the NER system, was used as one of the corpora for learning word vectors. The goal was to find out whether it is beneficial to compute word vectors from the corpus that is also used to train the system for the NER task when there is only a small corpus for training a NER system (with labeled named entities) available as it is expected that different NLP tasks employ the linguistic information related with other tasks (Güngör et al., 2018).

Texts from all sources were lowercased. The reason is that some of the available texts were already in lower case so we wanted to have all of them in the same form. All non-alphanumeric characters and words with less than 5 occurrences were removed as well.

3.2 The NER System

The NER system implemented to evaluate the impact of different properties of word vectors and parameters of their learning were based on the work of Žukov-Gregorič et al. (2018) who achieved state-of-the-art results on the CoNLL NER dataset. The input to the system is a sequence of word vectors and the output is an entity type label (including a label for words that are not entities). The function mapping inputs (a sequence of word vectors) to outputs (a sequence of entity type labels) is a neural network.

The first layer of the network accepts word vectors and passes them to the hidden layer. The hidden layer uses bidirectional LSTM units. The output layer converts the signal from the hidden layer using hierarchical softmax to predict an entity type for the given input. As a stochastic gradient-based optimization algorithm, Adam (Kingma and Ba, 2014) was used to learn the weights of the network. Various hyperparameters of the network were determined experimentally, see below.

Initially, the NER system used pre-trained word vectors as input. The vectors were learned on the Czech part of the CoNLL-2017 collection with word2vec using the skipgram architecture, with context window of size 10, and word vectors having 100 dimensions (Fares et al., 2017).

The values of the hyperparameters (Feurer and Hutter, 2019) of the NER system can significantly influence the results. Because the best possible achievement in the NER task was not the main goal of the research, only an acceptably good setting was found. Initially, the hyperparameters were set to the values typical for existing research (Žukov-Gregorič et al., 2018). The values were then changed (in both directions) as long as the performance (measured by the F1-measure) of the NER system was improving.

The suitable hyperparameter values were found in the order as they appear in the list below. The suitable number of epochs was found as the average number of epochs that were needed to achieve the best result for the given combination of hyperparameters (this was 12 most of the time).

The best results were achieved with the following hyperparameter setting:

- dropout probability (the probability that a neuron will be randomly turned off): 0.7;
- learning rate (influencing how fast are neuron weights changing): 0.02;
- batch size (the number of instances used for network parameters adjustment in an iteration): 32;
- the number of hidden layer neurons: 200;
- the number of epochs (the number of passes through the training data set): 12.

With this setting, the system was able to achieve the value 0.6816 of the F1-measure on the CONLL test set without optimizing the process word vectors creation.

3.3 Changing Parameters During Word Vectors Training

There are a few aspects of word vectors training. They are evaluated in isolation in a sequence of experiments. In one phase, one parameter is investigated and its suitable value determined. The following phase works with this value and focuses on another parameter.

The corpora described in Section 3.1 were used to learn word vectors to evaluate the impact of different corpora sizes size and quality. The word2vec algorithm using the skipgram architecture, context windows of size 10, hierarchical softmax, minimal token frequency, and the number of epochs equal to 5 was used to learn vectors with 100 dimensions. The skipgram architecture is suitable for most of the NLP tasks, is often used by other authors, and has low computational complexity (Levy et al., 2015).

Another important parameter is the size of vectors. Generally, the bigger the vectors are, the more relations between words can be captured (Pennington et al., 2014). This was, however, demonstrated on the word analogy task and not on the NER task. The vector size is also related to the corpus size. The bigger the corpus is, the more words and relations can be contained there. The experiments, therefore, examined different corpus and vector sizes. Most of the previous works used word vectors with 100 to 300 values. We, therefore, examined 50, 100, 200, 300, and 400 dimensions which cover the mentioned interval as well as close values outside it.

Some of the commonly used text preprocessing techniques, namely lemmatization, case folding, stop words removal, and their combinations were applied to texts before learning word vectors (lemmatization and case folding should be then applied to training data for the NER system too). Lemmatization and case folding belonging to normalization techniques decrease the number of unique tokens and increase the global frequencies of the tokens. This might be important especially for small corpora containing texts in a morphologically rich language, like Czech.

Three algorithms, namely word2vec (using the CBOW and skipgram architectures), GloVe, and fastText (using the CBOW and skipgram architectures) were studied. In the experiments, different context window sizes (5, 10, and 15 words) at a fixed number of epochs (5) were examined. Subsequently, 1, 10, and 15 epochs (5 epochs were already included in the experiments with different context window sizes) of training using 10 words context window were applied to create word vectors.

For the best settings of word2vec and fastText algorithms, the output layer function was changed from softmax to negative sampling with 5 or 10 negative samples. In the word2vec CBOW method, summation was used together with averaging the vectors. Different n-gram sizes were studied for the fastText algorithm. Similarly to Bojanowski et al. (2017), the minimal n-gram size was 2 or 3 and the maximal size 4 or 6. In the GloVe algorithm, different exponent values in the weighting function were used.

The following list summarizes the investigated aspects of individual techniques and algorithms during word vectors learning:

- all algorithms:
  - corpus: different corpora from Tab. 1;
  - number of dimensions: 50, 100, 200, 300, 400;
  - context window size: 5, 10, 15;
  - preprocessing techniques: lemmatization, case folding, stop words removal;
  - number of training epochs: 1, 5, 10, 15;
- wor2vec and fastText:
  - architecture: skipgram or CBOW;
  - last layer function: hierarchical softmax or negative sampling (5 or 10 negative samples);
- wor2vec:
  - CBOW aggregation: sum or average;
- fastText:
  - character n-gram size: 2 to 6;
- GloVe:
  - value of exponent α in the weighting function in the cost function: 0.75, 0.5, 0.25.

The quality of word vectors can be measured in several ways. One of the popular approaches is the analogy task (Mikolov et al., 2013b). However, good results in this task do not have to automatically lead to good results in the NER task. The performance of NER systems is usually measured using precision and recall. The precision is calculated as the ratio of pieces of a text that were correctly labeled as an entity and the number of pieces of a text that were labeled as an entity. The recall is defined as the ratio of entities in the text that were labeled as entities and the total number of entities in the text. These measures are calculated for each category of entities to be identified and can be further combined to the F1-measure which is a harmonic mean of the precision and recall (Yadav and Bethard, 2018). The impact of changing different parameters during the investigation was thus measured by the F1-measure.

4 RESULTS

This section provides the results from the investigation of the impact of different aspects of word vectors learning. The aspects follow the procedure described in Section 3.3.

4.1 Corpus Characteristics

First, the suitability of different corpora for creating word vectors was evaluated. The improvement against the baseline when no word vectors were used (the input contained just word identifiers) can be found in Tab. 2.

The CoNLL-2017, CZES, and CWC-2011 corpora has brought the highest improvements of the F1-measure (more than 15%) in the NER task even without focusing on the optimization of the parameters of the algorithms used. Among these three, CoNLL-2017 has brought the least improvement despite having the nighest number of tokens. This means that not only the quantity, but also quality of the corpus is important.

Tab. 2: The values of the F1-measure obtained by the NER system when using different corpora for word embeddings training

Corpus	**F1-measure** [%]
No (baseline)	49.83
EuroParl (Czech part)	52.09
Text from Extended CNEC 2.0	52.54
Czech Wikipedia extraction	56.54
SYN-2015	61.31
CoNLL-2017 (Czech part)	64.60
CZES	65.28
CWC-2011 – articles	66.31

None from the corpora used in word vectors training was able to improve the NER outcomes so they would outperform the result achieved when training the NER system with vectors pretrained on the Czech part of the CoNLL-2017 corpus, see Section 3.2 for details (the achieved F1-measure was 0.68). This means that it makes sense to focus on the details of the algorithms of word vectors training.

The best results were achieved with the CWC-2011 corpus (a large corpus with more than 600 million words) which is used in the following experiments.

4.2 Corpus Size and Word Vectors Length

The next experiment focused on determining how corpus size and word vectors length are related and how they influence the results of NER. The outcome of this experiment is summarized in Tab. 3.

Tab. 3: The values of the F1-measure of the NER system when using different numbers of word vectors dimensions and sizes of the CWC-2011 corpus (the baseline is emphasized)

Number of dimensions	**Relative corpus size** 1%	10%	50%	100%
50	53.84	59.14	63.98	62.55
100	54.56	61.14	64.83	**66.31**
200	55.17	62.79	66.35	67.29
300	56.63	63.78	67.93	67.95
400	55.70	65.51	67.68	68.33

Most significant differences can be seen between low- and high-dimensional vectors learned on the largest corpus where smaller vectors were not sufficient for encoding all relations between words. Increasing the number of dimensions lead to improvements even for smaller corpus portions. When the dimensionality was around 300 or 400 the results stopped improving, or they even degraded. Improvements were also positively related to corpus size. While the improvements between using 1 and 50% of the corpus were around 10% in the F1-measure, the differences between using 50% of the corpus and the whole corpus were marginal. Finding a suitable amount of data from which the results stop improving had thus a positive effect on computational complexity. For Czech, corpora containing hundreds of million tokens seem to be sufficient.

In the subsequent experiments, the number of dimensions was 300 because it enabled achieving the best results. Further experiments that did not change corpus size worked with 50% of texts from CWC-2011. Both decisions did not negatively influence the performance of the NER system and had favorable computational complexity and memory demands.

4.3 Text Preprocessing Techniques

The application of normalization techniques and stop words removal lead to a decreased number of unique tokens which not only affected word vectors training but also the number of out of vocabulary words (words that are recognized in the testing phase but are unknown in the training phase) in the NER system training. The effects of the application of these techniques and their combinations can be found in Tab. 4. Based on the results, it can be noted that using normalization (here, the largest effect has lemmatization) had a positive impact especially for smaller corpora used for word vectors training. For larger text collections, especially with texts of higher quality, these techniques or their combinations were not useful. Of course, the same preprocessing techniques were applied to the texts used for the NER system training. Lowercasing is considered as a baseline since all texts were lowercased for the initial experiments (an explanation is in Section 3.1).

4.4 Algorithms and Their Parameters

Until now, only the word2vec algorithm was used to train word vectors. In the following step, other algorithms and their parameters were examined. Two parameters were relevant for all three algorithms (word2vec, GloVe, and fastText): context window size and the number of training epochs. The detailed results obtained for different parameter values can be found in Tab. 5. The table contains the values of the F1-measure for different context window sizes for a fixed number of training epochs, as well as the values of the F1-measure for different

Tab. 4: The values of the F1-measure achieved by the NER system and the number of out of vocabulary (OOV) words from 51,092 when using different portions from the CWC-2011 corpus and preprocessing techniques when creating word vectors (LC = lower case, LM = lemmatization, SW = stop words removal)

Text preprocessing technique	F1-measure [%] Relative corpus size 50%	10%	1%	Number of OOV words Relative corpus size 50%	10%	1%
No	68.70	65.94	56.96	4923	8982	21214
LC (baseline)	67.93	63.78	56.63	4173	7652	18665
LM	67.33	66.35	61.58	2744	4576	10699
SW	67.22	62.56	53.81	5367	9426	21645
LC + LM	66.81	66.22	60.80	2792	4628	10624
LC + SW	65.43	62.61	53.96	4617	8096	19107
LM + SW	68.27	64.38	58.25	2811	4662	10827
LC + LM + SW	66.85	63.23	56.39	2839	4701	10744

numbers of training epochs for a fixed context window size.

It can be seen that fastText dominated in all these experiments. Also the skipgram technique for both word2vec and fastText has brought better results than CBOW, which means that semantic similarity is more important than syntactic one (Mikolov et al., 2013b).

Tab. 5: The values of the F1-measure when using different algorithms and their parameters (context window size, number of training epochs) for word vectors training (the baseline is emphasized)

Algorithm	Context window size (training epochs = 5) 5	10	15
word2vec (CBOW)	64.58	62.69	60.53
word2vec (skipgram)	69.06	**68.70**	67.99
GloVe	63.90	64.46	64.54
fastText (CBOW)	66.01	63.91	64.39
fastText (skipgram)	72.47	71.50	72.44

Algorithm	Training epochs (window size = 10) 1	5	10	15
word2vec (CBOW)	61.46	62.69	60.65	61.18
word2vec (skipgram)	68.25	68.70	68.60	69.05
GloVe	60.43	64.46	63.04	65.00
fastText (CBOW)	64.33	63.91	63.62	63.62
fastText (skipgram)	71.72	71.50	69.36	70.48

A context window size had a larger impact on the results when the CBOW technique was used and only a negligible impact when using the skipgram technique. The number of training epochs seems to have no significant impact on the results. Five epochs of training lead to the best results in most of the experiments. The GloVe algorithm did not reach results comparable to word2vec or fastText even when changing the exponent α in the weighting function in the cost function from the default value 0.75 to 0.5 or 0.25.

Both word2vec and fastText using skipgram can use different functions of the output layer of the neural network – hierarchical softmax or negative sampling. When using 5 or 10 negative samples, no improvement was observed for fastText. On the other hand, five negative samples increased the value of the F1-measure from 69.06% to 70.37% for word2vec. When using the CBOW technique, the sum instead of the average of the output vectors improved the value of the F1-measure from 64.58% to 66.62% for word2vec. For fastText, the value of the F1-measure even decreased. In both cases, the results were worse than when using the skipgram technique. When changing the minimal and maximal n-gram sizes for fastText (minimal = 2 or 3, maximal = 4 or 6), no improvements were found compared to the default setting (minimal = 3, maximal = 6).

4.5 Best Algorithms Settings

After a series of experiments focusing on different aspects of word vectors preparation for the NER task were conducted, some recommendation for the settings have been identified:

Tab. 6: The values of the F1-measure for the NER system using recommended settings for different CWC-2011 corpus portions (the baseline is emphasized)

	No lemmatization				With lemmatization	
	Relative corpus size				Relative corpus size	
Algorithm	**100%**	**50%**	**10%**	**1%**	**10%**	**1%**
word2vec	70.66	**70.37**	66.02	53.78	66.89	59.11
GloVe	63.54	**65.00**	60.12	52.62	58.54	53.44
fastText	71.69	**72.47**	71.12	66.51	70.34	66.42

Tab. 7: Detailed NER performance measures for the best word vectors preparation settings

Entity type	**Precision [%]**	**Recall [%]**	**F1-measure [%]**	**Recognized entities**	**Number of entities**
Numbers and addresses	90.38	85.45	87.85	52	55
Geographical names	73.48	76.98	75.19	396	378
Institutions	59.83	65.74	62.65	356	324
Media	66.67	58.33	62.22	42	48
Artifacts	47.78	47.91	47.84	383	382
Persons	77.49	86.04	81.54	533	480
Temporal expressions	89.63	91.58	90.59	376	368
Total	**70.72**	**74.30**	**72.47**	**2138**	**2035**

- using the CWC-2011 corpus for training (of course, for a specific domain, other corpora might be more suitable; however, the size and quality need to be generally considered);
- training word vectors with 300 dimensions;
- using lemmatization for a small amount of text for training (i.e., 1 or 10% of the corpus);
- word2vec settings: skipgram, the context window size = 5, the number of training epochs = 5, negative sampling with 5 negative samples;
- fastText settings: skipgram, the context window size = 5, the number of training epochs = 5, hierarchical softmax as the output layer function, minimal n-gram size = 3, maximal n-gram size = 6).

The results achieved with these recommendations are summarized in Tab. 6. We can see that lemmatization makes sense in the case that just a small corpus is available for word vectors training. When more data is used, lemmatization even worsens the performance (see the columns 10% in Tab. 6.

The size of the corpus used for training had the smallest impact on the performance when using the fastText algorithm. This supports (Bojanowski et al., 2017) stating that fastText is able to learn well on smaller corpora. From a certain size, using additional texts also did not bring additional improvements in the NER task. From using 50% of the corpus, word2vec was able to provide results comparable to fastText while GloVe was about 5% behind.

Using fastText with other recommended settings of the process and fastText algorithm is therefore the best approach for preparing word vectors for the NER task for the Czech language. The NER system using word vectors prepared in this way achieved 72.47% of the F1-measure (compared to 49.83% when not using word vectors and 68.16% when using setting typically used by other authors).

The detailed performance for individual named entity categories can be found in Tab. 7. The best performance was achieved for temporal expressions (usually the names of days or months) that had very similar vector representations. On the other hand, artifact names, like units of measure, currencies, norms, or product names, were recognized with the worst performance as they cover a very wide variety of expressions. These names were often composed of more tokens and the NER system did not have to recognize all of them correctly.

Compared to word2vec, fastText was able to better recognize numbers and addresses (the F1-measure for fastText was 87.85% compared 70.91% for word2vec). It is complicated to have a separate word vector for every unique token for word2vec. On the other hand, fastText composes word vectors from character n-grams so, for example, a vector for a number will be very similar to vectors of other numbers. A similar situation is in recognizing media entities, that contain many e-mail addresses.

5 DISCUSSION

To evaluate the impact of different settings in word vectors training, the results were compared to the results of other authors that used a similar approach. This means that they used a neural model for NER together with word vectors. The results also needed to be demonstrated on the Extended CNEC 2.0 corpus while using no additional resources (e.g., gazetteers, Brown mutual information bigram clusters, regular expressions) or engineered features (e.g., lemmas, prefixes, affixes, character n-grams, orthographic features). The availability of gazetteers or well-engineered features generally improves the NER system performance so the results of the other authors, against which the comparison is made, are not the best they ever achieved. Not having exactly the same resources also would not allow a direct comparison of the results.

Demir and Özgür (2014) use a large Czech corpus containing 636 million words and 906 thousand unique tokens (the size is similar to the CWC-2011 corpus), together with word2vec using the skipgram architecture and the context window size equal to 5 for training word vectors having 200 dimensions. Straková et al. (2016) used the same algorithm applied to the large SYN corpus for word vectors training. Tab. 8 summarizes the performance of the NER system presented in this paper and the results achieved by other authors where the results were available. The first number 49.83% represents the value of the F1-measure achieved by our system when no word vectors were provided. When word vectors trained using the baseline method (see Section 4.1) were used and suitable values of the NER model hyperparameters were found the value of the F1-measure increased to 68.16%. When focusing on the optimization of different aspects of word vectors training, the value additionally increased to 72.47%.

Tab. 8: Comparison of the values of the F1-measure achieved in this paper with other research

Method	**F1-measure** [%]
This paper – no word vectors	49.83
This paper – word vectors, best hyperparameters	68.16
This paper – best setting for fastText	72.47
Demir and Özgür (2014)	64.72
Straková et al. (2016)	63.91

It is obvious that focusing on the process of word vectors preparation can bring a significant improvement to the NER system performance. This is demonstrated by the comparison to the results of other researchers that did not focus on the optimization of word vectors training. Our results are compared to the outcomes achieved without additional language independent and dependent features and other modifications of the NER algorithm.

6 CONCLUSION

The research focused on named entity recognition in Czech where the process of preparing the data for training a NER model using modern text representations has not been investigated. The main emphasis is put on the phase of preparing word vectors for training a machine learning-based NER system.

First, the NER system inspired by the state-of-art approach was created. The input to the system was a sequence of word vectors and the output was an entity type label. The function mapping inputs (a sequence of word vectors) to outputs (a sequence of entity type labels) was a neural network. The hidden layer used bidirectional LSTM units. The output layer converted the signal from the hidden layer using hierarchical softmax to predict an entity type. As a stochastic gradient-based optimization algorithm Adam was used to learn the weights of the network.

Subsequently, attention was paid to various aspects of word vectors preparation. The suitable settings in different steps were determined in extensive experiments and included the used corpus, number of word vectors dimensions, used text preprocessing techniques, context window size and number of training epochs for word vectors training and other algorithm-specific parameters. Besides suitable values for the parameters, it has been found that a sufficiently large corpus of good quality needs to be used and the number of word vectors dimensions needs to be chosen so enough relations between words can be encoded.

It was demonstrated that focusing on the process of word vectors preparation can bring a significant improvement of the NER system performance even without using additional language independent and dependent resources.

7 REFERENCES

Blum, A. L. and Langley, P. 1997. Selection of Relevant Features and Examples in Machine Learning. *Artificial Intelligence*, 97 (1–2), 245–271. DOI: 10.1016/S0004-3702(97)00063-5.

Bojanowski, P., Grave, E., Joulin, A. and Mikolov, T. 2017. Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. DOI: 10.1162/tacl_a_00051.

Chen, G., Liu, T., Zhang, D., Yu, B. and Wang, B. 2018. Complex Named Entity Recognition via Deep Multi-Task Learning from Scratch. In Zhang, M., Ng, V., Zhao, D., Li, S. and Zan, H. (eds.). *Natural Language Processing and Chinese Computing: Proceedings, Part I*, pp. 221–233. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-99495-6_19.

Chiu, J. P. C. and Nichols, E. 2016. Named Entity Recognition with Bidirectional LSTM-CNNs. *Transactions of the Association for Computational Linguistics*, 4, 357–370. DOI: 10.1162/tacl_a_00104.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P. 2011. Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research*, 12, 2493–2537.

Dařena, F. 2019. VecText: Converting Documents to Vectors. *IAENG International Journal of Computer Science*, 46 (2), 170–177.

Demir, H. and Özgür, A. 2014. Improving Named Entity Recognition for Morphologically Rich Languages Using Word Embeddings. In *2014 13th International Conference on Machine Learning and Applications*, pp. 117–122. DOI: 10.1109/ICMLA.2014.24.

El Bazi, I. and Laachfoubi, N. 2019. Arabic Named Entity Recognition Using Deep Learning Approach. *International Journal of Electrical and Computer Engineering*, 9 (3), 2025–2032. DOI: 10.11591/ijece.v9i3.pp2025-2032.

Etzioni, O., Cafarella, M., Downey, D., Popescu, A.-M., Shaked, T., Soderland, S., Weld, D. S. and Yates, A. 2005. Unsupervised Named-Entity Extraction from the Web: An Experimental Study. *Artificial Intelligence*, 165 (1), 91–134. DOI: 10.1016/j.artint.2005.03.001.

Fares, M., Kutuzov, A., Oepen, S. and Velldal, E. 2017. Word Vectors, Reuse, and Replicability: Towards a Community Repository of Large-Text Resources. In Tiedemann, J. (ed.). *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pp. 271–276.

Feurer, M. and Hutter, F. 2019. Hyperparameter Optimization. In Hutter, F., Kotthoff, L. and Vanschoren, J. (eds.). *Automated Machine Learning: Methods, Systems, Challenges*, pp. 3–33. Springer International Publishing, Cham.

Goldberg, Y. 2016. A Primer on Neural Network Models for Natural Language Processing. *Journal of Artificial Intelligence Research*, 57, 345–420.

Goyal, A., Gupta, V. and Kumar, M. 2018. Recent Named Entity Recognition and Classification Techniques: A Systematic Review. *Computer Science Review*, 29, 21–43. DOI: 10.1016/j.cosrev.2018.06.001.

Güngör, O., Üsküdarli, S. and Güngör, T. 2018. Improving Named Entity Recognition by Jointly Learning to Disambiguate Morphological Tags. In *Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics*, pp. 2082–2092.

Hope, T., Resheff, Y. S. and Lieder, I. 2017. *Learning TensorFlow: A Guide to Building Deep Learning Systems*. O'Reilly Media.

Kingma, D. P. and Ba, J. L. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference for Learning Representations*. CoRR: abs/1412.6980.

Konkol, M. and Konopík, M. 2011. Maximum Entropy Named Entity Recognition for Czech Language. In Habernal, I. and Matoušek, V. (eds.). *Text, Speech and Dialogue: Proceedings*, pp. 203–210.

Konkol, M. and Konopík, M. 2013. CRF-Based Czech Named Entity Recognizer and Consolidation of Czech NER Research. In Habernal, I. and Matoušek, V. (eds.). *Text, Speech, and Dialogue: Proceedings*, pp. 153–160.

Konopík, M. and Pražák, O. 2018. LDA in Character-LSTM-CRF Named Entity Recognition. In Sojka, P., Horák, A., Kopeček, I. and Pala, K. (eds.). *Text, Speech, and Dialogue: Proceedings*, pp. 58–66.

Král, P. 2011. Features for Named Entity Recognition in Czech Language. In Filipe, J. and Dietz, J. (eds.). *Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, Vol. 1, pp. 437–441. DOI: 10.5220/0003660104370441.

Kravalová, J. and Žabokrtský, Z. 2009. Czech Named Entity Corpus and SVM-Based Recognizer. In Li, H. and Kumaran, A. (eds). *Proceedings of the 2009 Named Entities Workshop: Shared Task on Transliteration*, pp. 194–201.

Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Kováříková, D., Jelínek, T., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondřička, P. and Zasina, A. J. 2016. SYN2015: Representative Corpus of Contemporary Written Czech. In Calzolari, N., Choukri, K., Declerck, T., Goggi, S., Grobelnik, M., Maegaard, B., Mariani, J., Mazo, H., Moreno, A., Odijk, J. and Piperidis, S. (eds.). *Proceedings of the Tenth International Conference on Language Resources and Evaluation*, pp. 2522–2528.

Kulkarni, V., Mehdad, Y. and Chevalier, T. 2016. Domain Adaptation for Named Entity Recognition in Online Media with Word Embeddings. *arXiv*. CoRR: abs/1612.00148.

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K. and Dyer, C. 2016. Neural Architectures for Named Entity Recognition. In Knight, K., Nenkova, A. and Rambow, O. (eds.). *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 260–270. DOI: 10.18653/v1/N16-1030.

Leeuwenberg, A., Vela, M., Dehdari, J. and van Genabith, J. 2016. A Minimally Supervised Approach for Synonym Extraction with Word Embeddings. *The Prague Bulletin of Mathematical Linguistics*, 105, 111–142. DOI: 10.1515/pralin-2016-0006.

Levy, O. and Goldberg, Y. 2014. Neural Word Embedding as Implicit Matrix Factorization. In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. and Weinberger, K. Q. (eds.). *Advances in Neural Information Processing Systems 27*, pp. 2177–2185.

Levy, O., Goldberg, Y. and Dagan, I. 2015. Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, 3, 211–225. DOI: 10.1162/tacl_a_00134.

Li, Q., Shah, S., Liu, X. and Nourbakhsh, A. 2017. Data Sets: Word Embeddings Learned from Tweets and General Data. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, pp. 428–436.

Li, Y., Bontcheva, K. and Cunningham, H. 2005. SVM Based Learning System for Information Extraction. In Winkler, J., Niranjan, M. and Lawrence, N. (eds.). *Deterministic and Statistical Methods in Machine Learning*, pp. 319–339. DOI: 10.1007/11559887_19.

Masaryk University, NLP Centre. 2011. *czes*. LINDAT/CLARIN Digital Library at the Institute of Formal and Applied Linguistics [online]. Available at: http://hdl.handle.net/11858/00-097C-0000-0001-CCCF-C.

Mikolov, T., Chen, K., Corrado, G. S. and Dean, J. 2013a. Efficient Estimation of Word Representations in Vector Space. In *Proceedings of the International Conference on Learning Representations*. CoRR: abs/1301.3781.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. 2013b. Distributed Representations of Words and Phrases and Their Compositionality. In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K. O. (eds.). *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Vol. 2, pp. 3111–3119.

Mikolov, T., Joulin, A., Chopra, S., Mathieu, M. and Ranzato, M. 2015. Learning Longer Memory in Recurrent Neural Networks. In *3rd International Conference on Learning Representations*. CoRR: abs/1412.7753.

Nadeau, D. and Sekine, S. 2007. A Survey of Named Entity Recognition and Classification. *Lingvisticæ Investigationes*, 30 (1), 3–26. DOI: 10.1075/li.30.1.03nad.

Nadeau, D., Turney, P. D. and Matwin, S. 2006. Unsupervised Named-Entity Recognition: Generating Gazetteers and Resolving Ambiguity. In Lamontagne, L. and Marchand, M. (eds.). *Advances in Artificial Intelligence: Proceedings*, pp. 266–277. DOI: 10.1007/11766247_23.

Nguyen, A.-D., Nguyen, K.-H. and Ngo, V.-V. 2019. Neural Sequence Labeling for Vietnamese POS Tagging and NER. In *Proceedings: 2019 IEEE-RIVF International Conference on Computing and Communication Technologies*. DOI: 10.1109/RIVF.2019.8713710.

Pennington, J., Socher, R. and Manning, C. D. 2014. GloVe: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543. DOI: 10.3115/v1/D14-1162.

Rohde, D. L. T., Gonnerman, L. M. and Plaut, D. C. 2004. An Improved Method for Deriving Word Meaning from Lexical Co-Occurrence. *Cognitive Psychology*, 7, 573–605.

Rong, X. 2014. word2vec Parameter Learning Explained. *arXiv*. CoRR: abs/1411.2738.

Rudra Murthy, V. and Bhattacharyya, P. 2018. A Deep Learning Solution to Named Entity Recognition. In Gelbukh, A. (ed.). *Computational Linguistics and Intelligent Text Processing: Revised Selected Papers*, Part I, pp. 427–438.

Seok, M., Song, H.-J., Park, C.-Y., Kim, J.-D. and Kim, Y.-S. 2016. Named Entity Recognition using Word Embedding as a Feature. *International Journal of Software Engineering and its Applications*, 10 (2), 93–104. DOI: 10.14257/ijseia.2016.10.2.08.

Spoustová, J. and Spousta, M. 2012. A High-Quality Web Corpus of Czech. In Calzolari, N., Choukri, K., Declerck, T., Doğan, M. U., Maegaard, B., Mariani, J., Moreno, A., Odijk, J. and Piperidis, S. (eds.). *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pp. 311–315.

Straková, J., Straka, M. and Hajič, J. 2013. A New State-of-the-Art Czech Named Entity Recognizer. In Habernal, I. and Matoušek, V. (eds.). *Text, Speech, and Dialogue: Proceedings*, pp.68–75.

Straková, J., Straka, M. and Hajič, J. 2016. Neural Networks for Featureless Named Entity Recognition in Czech. In Sojka, P., Horák, A., Kopeček, I. and Pala, K. (eds.). *Text, Speech, and Dialogue: Proceedings*, pp. 173–181.

Ševčíková, M., Žabokrtský, Z. and Krůza, O. 2007. Named Entities in Czech: Annotating Data and Developing NE Tagger. In Matoušek, V. and Mautner, P. (eds.). *Text, Speech and Dialogue: Proceedings*, pp. 188–195.

Tkachenko, M. and Simanovsky, A. 2012. Named Entity Recognition: Exploring Features. In *Proceedings of KONVENS 2012*, Vol. 5, pp. 118–127.

Yadav, V. and Bethard, S. 2018. A Survey on Recent Advances in Named Entity Recognition from Deep Learning Models. In *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 2145–2158.

Yen, A.-Z., Huang, H.-H. and Chen, H.-H. 2017. Fusing Domain-Specific Data with General Data for In-Domain Applications. In *Proceedings of the International Conference on Web Intelligence* pp. 566–572. DOI: 10.1145/3106426.3106473.

Zeman, D., Hajič, J., Popel, M., Potthast, M., Straka, M., Ginter, F., Nivre, J. and Petrov, S. 2018. CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pp. 1–21. DOI: 10.18653/v1/K18-2001.

ZHOU, G. and SU, J. 2002. Named Entity Recognition using an HMM-Based Chunk Tagger. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 473–480. DOI: 10.3115/1073083.1073163.

ŽUKOV-GREGORIČ, A., BACHRACH, Y. and COOPE, S. 2018. Named Entity Recognition with Parallel Recurrent Neural Networks. In GUREVYCH, I. and MIYAO, Y. (eds.). *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Vol. 2: Short Papers, pp. 69–74. DOI: 10.18653/v1/P18-2012.

AUTHOR'S ADDRESS

František Dařena, Department of Informatics, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: frantisek.darena@mendelu.cz

Martin Süss, Department of Informatics, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

CALL FOR PAPERS

Your opportunity to publish research papers.

Open access journal for researchers and specialists from around the world.

European Journal of Business Science and Technology (EJOBSAT) is an English-language, open access, double-blind refereed, multidisciplinary journal published by Mendel University in Brno, Faculty of Business and Economics.

We provide special services to the authors, namely open access, free PDFs, liberal copyright policy, special discounts on conference fees to the annual international conference Enterprise and the Competitive Environment organized by Mendel University in Brno.

The journal is distributed online via its web page or in printed form.

If you are interested in receiving the printed version of the journal, please contact `subscription@ejobsat.cz`. Please provide the volume number, issue number and shipping address. Subscription of the printed version is for free.

Subjects covered by EJOBSAT

The EJOBSAT covers the broad range areas related to empirical business sciences and empirical finance including interdisciplinary topics and newly developing areas of business, especially implementing new technology. Empirical advances in buyer behavior, organizational behavior, marketing, business decisions, processes and activities, including corporate finance, risk, investments and business financing are evaluated on a regular basis.

Abstracting & indexing

- Scopus (since 2020)
- RePEc
- DOAJ

Journal information

- ISSN 2336-6494 (Print)
- ISSN 2694-7161 (Online)
- Online submission, open access, double blind referred
- Free of charge submission and no publication fees
- EJOBSAT is published twice a year (submissions are accepted throughout the year)
- Registered members are informed about new papers
- Liberal copyright policy
- Submissions are accepted in English, in Word and TEX/LATEX

www.ejobsat.cz