

**Lukasz Arendt. University of Lodz**  
**Ewa Galecka-Burdziak. Warsaw School of Economics**  
**Fernando Núñez. University of Seville**  
**Robert Pater. University of Information Technology and Management. Rzeszow**  
**Carlos Usabiaga, Pablo de Olavide University. Seville**

**Should you have a better job with your skills? Skills demand and task-based occupations in Poland<sup>1</sup>**

**Abstract**

On-line job offers published on selected Polish job portals in 2017-2019 are used to map skills requirements in task-content groups of jobs. This approach provides new insights into analysing processes of labour market polarisation from the point of view of vacancies (unmet labour demand), which is a novel proposal, as most of the studies on RBTC hypothesis are based on employment data. With the use of sequence analysis and logistic regression methods we identified types of skills which are in high demand in the Polish labour market. Moreover, we reported on differences in skills demand between routine and non-routine jobs, pointing skills required in jobs with low probability of automation, and thus providing relatively secure employment opportunities in the short to medium run in Poland.

**Keywords:** on-line job offers, labour market polarisation, skills, logistic regression

**1. Introduction**

In this paper we investigate the polarisation hypothesis (more specifically Routinisation-Biased Technical Change hypothesis –RBTC) by analysing unmet labour demand in Poland. We examine a total number of 5,085,628 online job offers<sup>2</sup>, covering 2017-2019 years, collected within the System of On-Line Job Offers (SOJO). In the study we focus on skills' demand. The types of skills are derived from the Balance of Human Capital (BHC) study (PARP 2011), and the data match skills' requirements specified in the job offers with task-content groups proposed by Autor et al. (2003). To the best of our knowledge, we are the first ones to use such approach to analyse on-line job offers (on the skills' level) in labour

---

<sup>1</sup> This paper was prepared within the framework of the research project “The polarisation of the Polish labour market in the context of technical change” financed by the National Science Centre, Poland (contract number 2016/23/B/HS4/00334).

<sup>2</sup> Since the number of all online job offers in Poland is unobservable, we were not able to calculate the share of online job offers retrieved in SOJO out of all vacancies available in the Polish labour market.

market polarisation hypothesis. Our main goal is to map skills with particular task-content groups to show the degree of skill requirements in jobs assigned to different task-content groups. The additional goal is to test if the structure of on-line job offers in Poland is consistent with the RBTC hypothesis.

Empirical studies on RBTC hypothesis and online job offers constitute, in practice, two distinct streams of research. RBTC hypothesis is usually tested on the basis of employment data, with results revealing labour market polarisation in most of highly-developed economies, and specific polarisation patterns for CEE (Central and Eastern European) countries, including Poland. Studies dealing with (online) job offers typically focus on two aspects: technical issues related to retrieving and matching job offers with occupations and skills; and the analysis of skills requirements in different dimensions. Research approaches which try to link the RBTC hypothesis are rare (Hershbein and Kahn 2018 for the US, Colombo et al. 2018 for Italy), and to our knowledge there is no such study covering the CEE economy. In this paper, based on a dataset of on-line job offers collected from selected web portals, we link more explicitly (than in other studies) skills to task-content groups of occupations on the Polish labour market, thus contributing to the literature on RBTC. By showing the impact of skills on task-content characteristics of occupations we find which skills are connected with easily automated job, and which ones may, at least to some extent, guarantee employment security. We also characterise tasks by skills rather than occupations available in online job offers, which may be treated as a novel approach. We find that the polarisation path in the Polish labour market, from the perspective of online vacancies, is far from the standard RBTC hypothesis, revealing perceptibly too high demand for routine manual jobs. Regardless of the task-content group of jobs, mostly interpersonal, self-organisation and technical skills are required in the job offers. However, if we concentrate only on non-routine types of jobs, which are characterised by relatively low probability of automation, we find that managing and self-organisation are crucial for non-routine cognitive personal task-content group. At the same time, cognitive and technical skills are in high demand in case of for non-routine cognitive analytical jobs.

The paper is structured as follows. In Section 2 we present the literature review. Section 3 provides a synthetic description of the system of on-line job offers that retrieves data on vacancies from web portals. Section 4 describes the methodology, data and findings for the Polish labour market. Finally, Section 5 concludes.

## 2. Literature review

Acemoglu and Autor (2011) defined job task-content groups and assigned occupations to each task-content group based on their descriptions. They performed this exercise to determine whether a particular occupation is endangered by automation. In simplicity, the sophistication of the tasks workers perform determines the probability of the replacement by robots. But, in our opinion, tasks performed in jobs are more related to skills than occupations –since there might be occupations that use both replicable and unique skills– and skill’s sophistication defines whether a person can be replaced by a computer programme or not. Although groups of skills, which describe certain job, are already assigned to occupations, we argue that task-content groups should be recognised through skills, which we try to do in this paper. This is of great importance taking into account the possible technological upgrading on the labour market caused by COVID-19 pandemic.

RBTC has become a popular hypothesis explaining changes in the employment and wage structure in developed economies in recent years. The result of RBTC (as well as globalisation and offshoring; see e.g. Mandelman and Zlate 2014; Oldenski 2014) is the labour market polarisation, i.e., the change in the structure of labour demand that favours not only highly-skilled labour (as it was posited in the Skill-Biased Technical Change, SBTC theory), but also low-skilled employees, while medium-skilled labour is the biggest loser of the ICT (Information and Communication Technology)-driven technical change. In the RBTC approach the structure of the labour market is analysed in the dimension of task-content groups (not occupational groups), which are defined in line with the model presented by Autor et al. (2003), referred to as the ALM model. The ALM model distinguishes five types of tasks (task-content groups): non-routine analytical (e.g. forming and testing hypotheses, legal writing), non-routine interpersonal (e.g. persuading and selling), non-routine manual (e.g. truck driving, janitorial services), routine cognitive (e.g. record-keeping, simple calculations) and routine manual tasks (e.g. picking and sorting, repetitive assembly).

The impact of ICT (and probability of automation of these tasks) depends on the ability to provide algorithms to be performed by a computer or any other numerically controlled device. As a result, machines should potentially replace routine tasks (cognitive and manual), be complementary to non-routine analytical and interpersonal tasks, with ambiguous conclusions about non-routine manual tasks. Because routine tasks are concentrated on the middle of the skill distribution (this applies mainly to clerical and assembly line jobs), the relative share of employees in occupations requiring high and low skills, as well as the wage premium in these occupations, should grow, which leads to labour market polarisation and

growing wage inequalities. Thus, the RBTC hypothesis posits that there is a relationship between the skills distribution and the task-content distribution in the labour market, and the labour demand by skill-level forms a U-shaped curve. Because of this, some research studies on polarisation refer to skills, not to tasks.

Many research studies in the US confirm that the American labour market has polarised (see Autor et al. 2003; Autor and Dorn 2013; Cortes et al. 2017). Similar conclusions were presented for Great Britain (Goos and Manning 2007), Germany (Dustmann et al. 2009), Nordic countries (Asplund et al. 2011; Adermon and Gustavsson 2015), Western EU countries (Goos et al. 2009; Goos et al. 2014), and several OECD countries (Michels et al. 2014). The polarisation pattern was also revealed in Canada (Green and Sand 2015) and Portugal (Fonseca et al. 2018), however with some specific features diverging from the canonical RBTC model. Green and Sand (2015) argued that technical progress may explain changes in the Canadian labour market taking place in the middle and upper part of the skills distribution, while other factors (mainly the increase of labour force with tertiary education) shaped the changes in low-skilled occupations. In Portugal a decline in routine manual employment was recorded, while the decline of employment in routine cognitive task-group turned out to be modest and coupled with a higher wage premium (Fonseca et al. 2018). At the same time, Oesch (2013) argued that although in the United Kingdom, and to some extent in Switzerland, the largest decrease in employment was recorded in the middle of the skills distribution, changes in the occupational structure in these countries followed a pattern more similar to the SBTC rather than the RBTC hypothesis.

CEE countries, including Poland, have been rarely analysed from the labour market polarisation perspective. One of the few exceptions is the set of studies by the Institute for Structural Research, which focused on the evolution of the structure of tasks in the CEE labour markets. The results showed that the changes in task distribution in this region are generally in line with the trends characteristic for developed countries, with one exception –routine cognitive tasks. The intensity of these tasks grew in the CEE countries, especially since 2006. It should be emphasised that an increase in the intensity of routine cognitive tasks was recorded in Estonia, Latvia, Lithuania and Romania. A relatively stable demand for these tasks was found in the Czech Republic, Croatia, Poland and Slovakia, while in Hungary and Slovenia a downward trend emerged (Hardy et al. 2016; 2018). Employment increase instead of a decline in the middle of the skill distribution in Poland was also reported by Arendt (2018). Parteka (2018) showed that occupations with higher routine content experienced stronger downward pressure on wages. Arendt and Grabowski (2019) revealed a negative wage premium in the

case of routine jobs, however, they argued that relative wages in routine manual jobs were too high compared to predictions stemming from the standard RBTC theory. They also found a positive wage premium in the group of non-routine jobs.

Thus, we argue that labour market polarisation in Poland (like in Portugal or Canada) has a specific nature and is to a certain extent questionable. While demand for routine jobs should decrease, as RBTC hypothesis points, we have observed an increase in demand for routine cognitive jobs and relatively high wages in routine manual jobs in the Polish labour market. This phenomenon may be explained by: (i) the specific employment structure inherited from the times of the centrally-planned economy (a relatively high share of agricultural workers in total employment), as well as in the significant increase in the enrolment rate (Hardy et al. 2016); (ii) globalisation, especially the offshoring processes, which have been dynamic in Poland in recent years; or (iii) educational upgrading and the stigmatisation of vocational education in Poland (Arendt and Grabowski 2019). It seems that growing demand (and wages) for routine cognitive tasks has so far suppressed labour market polarisation in Poland (Hardy et al. 2016), but at the same time created challenges for the future. Nowadays, Poland faces relatively low level of automation in comparison to other countries –robots density in Poland in 2016 was significantly below the world average (32 and 74 robots per 10,000 employees in manufacturing industry, respectively). However, it is forecasted that robotisation boom will take place in Poland in 10-20 years' time<sup>3</sup>, and 49% of workplace activities in Poland can be potentially automated in the 2030 horizon –this is equivalent to 7.3 million of jobs (McKinsey 2018). Automation shall be beneficial for the economy (additional 15% of GDP growth between 2020 and 2030) and labour market (creation of new jobs, as a result of productivity increase, development of new technologies and society aging processes). At the same time, since around 83% of today's labour force will be still active in the labour market in 2030, and will have to adjust their skills to the requirements stemming from technical change (including changes in task-content of jobs), the scale of retraining activities will be enormous (McKinsey 2018). Moreover, because the ageing processes of the labour force are more advanced in the group of occupations characterised by a high intensity of routine tasks, in the context of displacing routine by non-routine tasks (caused by technical change) it may negatively affect the employability of older age groups in the near future (Lewandowski et al. 2017). Thus, the identification of the skills-mix required in the respective task-content groups seems to be an important exercise.

---

<sup>3</sup> <https://www.money.pl/gospodarka/wiadomosci/artykul/automatyzacja-pracy,0,0,2412288.html>

Although labour market processes are usually analysed from the point of view of realised demand (employment numbers), there are more and more attempts to study the patterns of yet unmet labour demand, mainly by analysing vacancies or job offers –with growing interest in job offers posted on-line outside the system of Public Employment Services (PES). Such approach seems highly relevant in the case of Poland, as reporting vacancies to PES by the employers has been no obligatory since February 2009<sup>4</sup>. As a result, PES database (Central Database of Job Offers (CDJO), administered by the Ministry of Family, Labour and Social Policy and available on-line at <http://oferty.praca.gov.pl/>) contains job offers which have been reported by the employers to Public Employment Services (in practice to 340 county job offices across Poland), but represents only a small fraction of all vacancies (job offers) available in the Polish labour market<sup>5</sup>. Thus, the use of on-line job offers is not only a matter of taking advantage of big data analytics, but mainly a way to complement public statistics to acquire a comprehensive picture of the unmet labour demand in the local, regional and national labour market (Dusi et al., 2015). Moreover, it has been argued that trends in the on-line job offers provide the right approximation to all vacancies available in the labour market (Hershbein and Kahn 2018).

Job offers have been intensely used and recognised to describe labour market behaviour over the business cycle. The earliest and largest analyses, named the Help-Wanted Index (HWI), have been conducted for the US economy with the use of job offers from newspapers (Abraham 1987). However, these analyses had not included skills, even though the HWI started to be calculated based on a rich repository of online job offers (Conference Board 2018). Only recently the literature on extracting skills from online job offers has been thriving, thanks to the development of webscraping methods and text analysis methods. Still, few papers have used explicit measures of skills and connect skills with job automation/routinisation.

Researchers dealing with analysis of online job offers take advantage on either one or a couple of job portals to obtain data from. Both approaches have certain properties. In case of using one specialised website/job portal, it is possible to retrieve not only data for job offers, but also for job seekers and potential matches. However, all this comes with a low coverage of the population of vacancies. This problem is solved using many web portals as a source at the same time, but such approach also has drawbacks. These drawbacks come from the fact that the mechanism of job offer posting (e.g. paid or not paid, whether they are immediately withdrawn

---

<sup>4</sup> Although there was a legal requirement to report all vacancies to PES before February 2009, in practice many companies did not comply with this regulation, as there were no legal consequences imposed.

<sup>5</sup> It has been argued that maximum 10% all job offers have been reported to PES.

if a candidate is found, whether the site also helps in the recruitment process or just provides space for advertisements, etc.) is usually not known by the researcher, and may differ among selected job portals.

Both approaches can be found in the literature dealing with the US economy. Single US website (CareerBuilder) was used by Marinescu and Wolthoff (2020), who gather data on job offers for January-March 2011, and proxied (job specific) skills by occupation/speciality, classifying them with job/occupation names according to the occupational classification (SOC) of the Bureau of Labour Statistics. Special attention is paid to low-skill and high-skill jobs to explain wage dispersion, while transversal skills were omitted. They find that occupations from job titles explain 90% of the variance in wages posted in job offers. Also these occupations explain 80% of the variance in the education and experience levels of the applicants that a vacancy attracts.

When it comes to the other approach, Burning Glass Technologies seems to be a popular tool to collect data from the US job portals and company websites (see eg. Azar et al. 2020; Deming and Kahn 2018, Hershbein and Kahn 2018) and analyses them from different points of view. Azar et al. (2020) study the geographical concentration of labour demand by occupations (at the level of 6-digit codes of the Standard Occupational Classification) with the use of data for 2016 to test for monopsony power. They reveal that more than half of labour markets are highly concentrated, and account for around 20% of employment. Both Deming and Kahn (2018) and Hershbein and Kahn (2018) reflect on skills using different skill taxonomies and time-span. Deming and Kahn (2018) classify skills from nearly 100 million electronic job postings in the United States between 2007 and 2015 to their own classification of 10 categories, restricting the study to the group of professional workers. They reveal that skills may explain wage patterns and companies performance, and occupations with high and advanced skill requirements correlate with higher productivity. Hershbein and Kahn (2018) look directly at the skill requirements in online job offers in 2007 and the period 2010-2015, introducing four proxies for skills: education, experience, cognitive skills, and computer skills. Importantly, they study upskilling (increase in education or experience, and increase in demand for the type of skills) across task-content groups over the business cycle during 2007-2015 in terms of routinisation of jobs. They find that routine-manual occupations did not face upskilling during the Great Recession, while routine-cognitive occupations did.

In Europe<sup>6</sup>, a usual practice is to collect data on online job offers from different web portals. A set of studies covering the Italian labour market was launched in the framework of the Wollybi project, which collected data on online vacancies from selected Italian portals since February 2013. The Wollybi system enabled to identify skills (divided into basic, personal and professional) and linked them with occupations (in line with ISCO-08 classification) – see Dusi et al. (2015). Colombo et al. (2018) use Wollybi data covering years 2014-2017 to analyse the probability of automation of occupations. They divide the skills into ‘soft’, ‘hard’ and ‘digital’, and match these skills with occupation according to automation probability from Frey and Osborne (2017) study, although they face problems in matching both data sources. It is revealed that basic digital skills and hard or technical skills are positively correlated with job automation probability (possibility that a particular occupation will be automated), especially the ‘low-level’ hard skills, while soft skills and advanced digital skills are negatively correlated with automation probability. More specific study was conducted by Lovaglio et al. (2018), who focused on offers addressed to ICT specialists and statisticians, collected in the third quarter of 2015 from three Italian job portals. The Wollybi tool became a predecessor of the EU-wide project launched in 2017 by the CEDEFOP, which aim is to develop a system of collecting and analysing on-line job offers covering job portals from all EU member states<sup>7</sup>. An experiment to evaluate the adequacy of this tool, called the WoLMIS system, on the basis of on-line job offers from Great Britain, Ireland, Germany, Italy and the Czech Republic was developed by Boselli et al. (2018). Importantly, WoLMIS will enable to classify a given job offer at the 4-digit code level according to the ISCO classification, as well as to assign it to the ESCO classification.

There are also examples of research studies that cover labour markets of selected EU member states. Turrell et al. (2018) studied labour market clusters in Great Britain, on a basis of 15 million online job offers collected in 2008-2016. Bhuller et al. (2020) use the National Public Employment Agency database for 2001-2016, which includes vacancies directly reported by employers to the agency, as well as from online job portals and newspapers in Norway. They analyse the matching efficiency of workers using broadband Internet access, showing that hiring and separation rates for low-skilled workers using broadband Internet access for their job search are higher than the ones corresponding to low-skilled individuals not using it. This result is interpreted as opposite to the SBTC hypothesis.

---

<sup>6</sup> The examples of research studies for other countries different from US and EU include, among others, China (see, e.g., Zhu et al. 2016, Xu et al. 2017) and India (Nomura et al. 2017).

<sup>7</sup> <https://www.cedefop.europa.eu/en/data-visualisations/skills-online-vacancies>. The European literature mostly adopts the European Commission ESCO skill classification to the analysis of online vacancies, instead of the American O\*NET classification.



### 3. The System of On-Line Job Offers (SOJO)

The data used for empirical analysis in this paper were retrieved from the SOJO, which has been developed by the Institute of Labour and Social Studies. The main rationale for constructing this system is the following: since job offers posted on recruitment web portals cover a significant part of the market, usually duplicating the offers published in other media (press, radio, television), a database containing information about on-line job offers would be the most reliable tool to analyse vacancies which were not reported to PES. SOJO collects information on vacancies from six Polish web portals: pracuj.pl, gazetapraca.pl, praca.pl, careerjet.pl, gratka.pl, olx.pl. Details of each job offer posted on these websites are collected automatically by a web crawler, which has built-in queuing and deduplication mechanisms. Mechanism for deduplication ensures that only unique job offers are downloaded (if the same job offer is posted on different web portals, the web crawler selects only one to be stored in the database). Unstructured content of all job-offers are stored in a "raw" database, which is a starting point for recoding with the use of dictionaries embedded into SOJO and other relevant data sources. SOJO provides automatic coding of the content of the job offers, assigning it to:

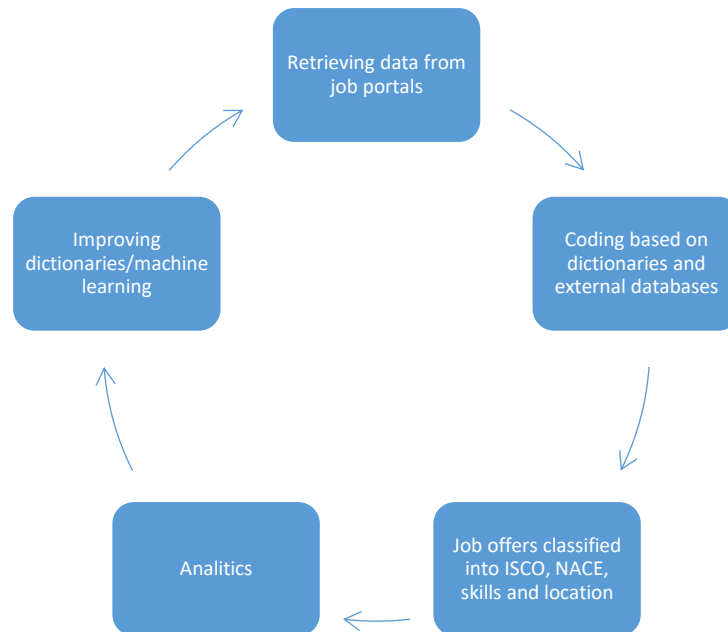
- the 6-digit occupations in accordance with the Polish Classification of Occupations and Specialities for the needs of the labour market, which is consistent with ISCO-08 classification<sup>8</sup>;
- geographical location using the TERYT database;
- skills –in accordance with the classification used in the BHC study;
- industry - according to NACE Rev. 2 classification at section level.

Dictionaries are updated on a regular basis to maximise the number of correctly classified job offers into the above described categories. However, as the quality of job offers, in terms of details describing vacancies, differs perceptibly among portals, in some cases SOJO provides information only about the occupation, while other categories –especially skills– are not available. Similarly to systems of on-line job offers developed in other countries, machine learning has also been implemented. The analytical system of SOJO includes five phases and is recursive (Figure 1).

---

<sup>8</sup> Since task-content groups correspond to ISCO occupational groups, SOJO enables to analyse polarisation processes from the point of view of vacancies.

**Figure 1. Scheme of data processing and analysis on SOJO**



Source: own elaboration based on Colombo et al. (2018).

Since SOJO contains information on occupations, not job-related tasks, specific occupations were categorised into task-content groups to enable polarisation-type of analysis. We used the tasks classification by Acemoglu and Autor (2011) based on O\*NET data and its application to Polish data proposed by Hardy et al. (2018) –in this approach occupational groups at 3-digit code level are assigned to respective five task-content groups according to the dominant task content of the jobs: non-routine cognitive analytical, non-routine cognitive personal, routine cognitive, routine manual and non-routine manual physical.

Although the most common approach to measure skills (competences) in studies on on-line job offers are O\*NET and ESCO classifications<sup>9</sup> Pater et al. (2019), who reviewed classifications of skills that can be used to measure skills demand, concluded that no international standard has been established. In case of SOJO the classification developed within the BHC study was implemented. BHC is a Polish research project conducted since 2009 by the Polish Agency for Enterprise Development and the Jagiellonian University, which provides monitoring activities with regard to demand for skills on the Polish labour market. In order to enable comparability between the demand and supply-side of the labour market along different

---

<sup>9</sup> Colombo et al. (2018) and Pater et al. (2019) used the ESCO classification, while Hershbein and Kahn (2018) and Marinescu and Wolthoff (2020) used the O\*NET classification.

periods (skills evolve with time), BHC introduced a classification of eleven “key” (general) skills –see Appendix.

#### 4. Demand for skills in on-line job offers

We performed the empirical analysis using on-line job offers published on selected Polish portals in 2017-2019, and collected within the SOJO. The total number of retrieved offers accounted for 5,085,628 units. If we make the general outlook of the data, it seems that in most cases employers looked for individuals to perform non-routine cognitive jobs (if analytical and personal task-content groups are taken together: 43%). But if we account for the structure of demand across all five task-content groups (as in line with standard ALM model), the routine manual jobs (35%) dominate (Table 1). Two conclusions emerge. Firstly, the structure of vacancies differs from the employment structure. This may be related to the recent (pre COVID-19) changes in the Polish labour market and growing problems with filling vacancies, especially in the group of low-skilled (routine) jobs. Secondly, this demand structure does not follow the standard polarisation pattern, given the too high relative importance of routine manual jobs. Job portals differ perceptibly in terms of the requirements related to the content of the job offer which may be published there on-line, and there is no uniform job offer template. In many cases the individual job offers do not contain information about demanded skills. As a result, we were able to identify and assign skills to “only” 1,893,352 offers (37% of all job offers), which constitute the basis for statistical analysis in our study. Fortunately, the structure of all on-line job offers and offers containing information about skills is almost identical (Table 1).

**Table 1. Number and frequency of job offers according to tasks-content groups**

Task-content groups	All job offers		Job offers with assigned skills	
	Frequency	Percentage	Frequency	Percentage
<b>Non-routine cognitive analytical</b>	862,762	17%	296,747	16%
<b>Non-routine cognitive personal</b>	1,312,490	26%	530,338	28%
<b>Non-routine manual physical</b>	347,549	7%	134,187	7%
<b>Routine cognitive</b>	774,041	15%	294,525	15%
<b>Routine manual</b>	1,788,786	35%	637,555	34%

Source: own computations.

For further empirical analysis, we proposed theoretical links between BHC skills and task-content groups of occupations (Table 2). Task contents and skills descriptions show that each skill may be directly related to at most three task-content groups of occupations. In this classification we assumed that skill requirements in job offers refer to higher than basic skill

level. Otherwise there are occupations, for example specialists, which require basic level of most of these skills. Also some skills at a basic level are probably needed in every job, for example availability and interpersonal skills.

**Table 2. BHC skills classified according to task-content groups of occupations (direct relationship)**

<b>Skill</b>	<b>Task group 1</b>	<b>Task group 2</b>	<b>Task group 3</b>
<b>Artistic</b>	Non-routine manual	Non-routine cognitive Analytical	Non-routine cognitive Interpersonal
<b>Availability</b>	Routine cognitive		
<b>Cognitive</b>	Routine cognitive	Non-routine cognitive Analytical	
<b>Computer</b>	Routine cognitive	Non-routine cognitive Analytical	
<b>Interpersonal</b>	Routine cognitive	Non-routine cognitive Interpersonal	
<b>Managerial</b>	Non-routine cognitive Interpersonal		
<b>Mathematical</b>	Routine cognitive	Non-routine cognitive Analytical	
<b>Office</b>	Routine cognitive	Non-routine cognitive Analytical	Non-routine cognitive Interpersonal
<b>Physical</b>	Routine manual		
<b>Self-organisation</b>	Routine cognitive	Non-routine cognitive Analytical	
<b>Technical</b>	Routine manual	Non-routine manual	

Source: own elaboration.

In the first step of our empirical approach, we computed two kinds of relative indicators. We started with analysing the relative abundancy of particular skills compared to the given task-content group. The results are presented in the upper part of Table 3 (here rows sum up to 100%). Next, we examined the relative abundancy of a given skill across task-content group. The results are presented in the lower part of Table 3 (here columns sum up to 100%). If we analyse the first part of Table 3 it seems that interpersonal, self-organisation and technical skills are the most required virtually in all types of task-content groups<sup>10</sup>. There is no doubt, that the first two types of skills are categorised as transversal, and thus are required across many occupations or task-content groups. However, if we look carefully at the definition of technical skills as proposed in the BHC study (see Appendix 1), we may come to the conclusion that technical skills may also be treated as semi-transversal –which seems to be the case in our study

<sup>10</sup> Interestingly, two types of skills from BHC list (mathematical and artistic) did not appear in our database, which may be explained by the employers' approach to posting job offers –it seems that these types of skills are treated as being so evident for certain jobs, that they are not even listed in the job offer.

on on-line job offers. In turn, if we analyse the relative importance of a particular skill across task-content groups (lower part of Table 3), it appears that non-routine cognitive personal jobs dominate in almost all types of skills (apart from office ones, and the shares are quite equal in case of interpersonal and ICT skills). If we restrict our interest just to the skills, it looks that interpersonal, self-organisation and technical ones are the most often required ones, which is of course in line with the previous findings. They appeared in 25.6, 21.5 and 24.1% of offers respectively (71.3% of all skills required). Moreover, we find that cognitive skills, availability, office, and managing skills are quite close to the threshold of 5%. The remaining skills (communication and ICT) are much less often required.

**Table 3. Relative demand for particular types of skills within a task-content group (upper part of the table) and across task-content groups (lower part of the table) (in %)**

	comm.	managing	availab.	office	Interp.	self-organ.	technical	ICT	cognitive
<b>non-routine cognitive analytical</b>	0.61	4.70	4.65	2.95	24.55	21.11	25.09	5.82	10.53
<b>non-routine cognitive personal</b>	4.73	12.38	14.05	6.23	45.66	43.74	46.22	5.69	17.09
<b>non-routine manual physical</b>	0.05	1.41	3.03	1.81	11.10	8.15	9.27	0.90	2.79
<b>routine cognitive</b>	0.59	5.13	4.75	4.29	23.26	20.44	22.06	5.51	7.59
<b>routine manual</b>	0.49	4.29	9.95	16.17	45.02	31.95	38.67	1.21	6.95
<b>non-routine cognitive analytical</b>	9.40	16.83	12.75	9.37	16.41	16.84	17.76	30.41	23.43
<b>non-routine cognitive personal</b>	73.02	44.35	38.58	19.81	30.52	34.89	32.71	29.74	38.01
<b>non-routine manual physical</b>	0.83	5.06	8.32	5.76	7.42	6.50	6.56	4.70	6.21
<b>routine cognitive</b>	9.18	18.39	13.04	13.64	15.55	16.30	15.61	28.80	16.88
<b>routine manual</b>	7.57	15.37	27.31	51.42	30.10	25.48	27.37	6.34	15.46

Source: own computations.

In order to find relations between skills and task-content occupational groups we examined what skills companies look for to fill in jobs across different task-content groups. We modelled the probability:

$$p_{t,r} = P(tcg_{t,r} = 1 | d_t, d_r, sg_{t,r}) = \exp(\mathbf{x}\boldsymbol{\beta}) / (1 + \exp(\mathbf{x}\boldsymbol{\beta})) \quad (1)$$

where  $tcg_{t,r}$  is a binary variable representing a particular task-content group of occupations  $tcg_{t,r} = 1$ , when latent variable  $tcg_{t,r}^* > 0$ , and  $\mathbf{x}\boldsymbol{\beta}$  is our functional form. In order to estimate unobserved heterogeneity across time  $\alpha_t$  and region  $\alpha_r$ , we estimated a logistic regression of a form:

$$tcg_{t,r}^* = \alpha_0 + \sum_{t=2}^4 \alpha_t d_t + \sum_{r=2}^{16} \alpha_r d_r + \sum_{s=1}^{11} \alpha_s sg_{t,r} + \epsilon_{t,r} \quad (2)$$

where  $tcg_{t,r}^*$  is a latent task-content group of occupations,  $d_t$  are time dummies encompassing years 2017-2020,  $d_r$  are regional dummies for 16 NUTS-2 regions of Poland (voivodships), and  $sg_{t,r}$  are measures of 11 skills.  $\alpha$ 's are parameters to be estimated and  $\epsilon_{t,r}$  is an error term.

In the above model, we treated each of the task-content occupation groups separately. Next, we analysed relations between each of these groups. In this case our variable of interest was categorical, but took five values, one for each task-content group. We applied a multinomial logistic model of a form:

$$p_{t,r,g} = P(tc_{t,r} = g | d_t, d_r, sg_{t,r}) = \exp(\mathbf{x}\boldsymbol{\beta}_g) / (\sum_{k=1}^5 \exp(\mathbf{x}\boldsymbol{\beta}_k)) \quad (3)$$

where  $\sum_{g=1}^5 p_{t,r,g} = 1$ . We had five task-content groups  $g$ , and  $k$  alternatives. Our functional form was analogous to (2).

Table 4 compiles the results. Firstly, it was revealed that economic cyclicality matters. After years of economic expansion and development of online vacancy market, in 2019 we observed a business cycle turning point in the vacancy market in Poland. Our data also shows that the probability of demanding cognitive task-content groups fell in response to this contraction. The demand for non-routine cognitive interpersonal tasks faced contraction even earlier, already in 2018. Changes in the demand for manual tasks occupations remained stable for non-routine tasks, and even continued to increase for routine tasks. This may be connected especially to deficient building and construction sector and transport services workers in Poland.

Secondly, there are perceptible differences in demand structure (in most cases) by the geographical dimension. Demand for routine manual jobs was visibly higher in lower-developed regions. Office skills and availability skills were most important for companies seeking workers in routine manual tasks. ICT skills and communication skills were the least expected from potential workers in these occupational groups. Contrary to routine manual, the demand for routine cognitive jobs was higher in well-developed regions. Companies often demanded strong ICT skills, but also managerial and self-organisation skills. Communication skills were the least sought-after by companies for this group of occupations. There was not visible regional pattern in the demand for non-routine manual tasks. Similarly to routine manual

tasks, availability skills were most often demanded, followed by interpersonal skills. The probability of including communication skills in job offers was very low, even lower than in both routine tasks.

**Table 4. Demand for task-content groups across time, regions and demanded skills. Results of a logistic regression**

	<b>Routine manual</b>	<b>Routine cognitive</b>	<b>Non-routine manual</b>	<b>Non-routine cognitive analytical</b>	<b>Non-routine cognitive interpersonal</b>
	Estimate Std. Error	Estimate Std. Error	Estimate Std. Error	Estimate Std. Error	Estimate Std. Error
<b>(Intercept)</b>	-0.176 *** 0.008	-1.634 *** 0.009	-2.647 *** 0.014	-1.997 *** 0.010	-1.740 *** 0.008
<b>2017</b>	0.237 *** 0.006	0.062 *** 0.007	0.262 *** 0.011	0.017 ** 0.006	-0.289 *** 0.005
<b>2018</b>	0.445 *** 0.006	-0.031 *** 0.007	0.285 *** 0.011	-0.096 *** 0.007	-0.398 *** 0.005
<b>2019</b>	0.729 *** 0.015	-0.390 *** 0.022	0.206 *** 0.027	-0.327 *** 0.023	-0.410 *** 0.017
<b>Regional dummies</b>	yes	yes	yes	yes	yes
<b>Communication</b>	-1.351 *** 0.015	-0.678 *** 0.014	-2.175 *** 0.044	-0.830 *** 0.014	1.677 *** 0.009
<b>Managing</b>	-0.586 *** 0.006	0.164 *** 0.006	-0.320 *** 0.010	-0.161 *** 0.006	0.458 *** 0.005
<b>Availability</b>	0.087 *** 0.005	-0.279 *** 0.006	0.337 *** 0.007	-0.444 *** 0.006	0.264 *** 0.004
<b>Office</b>	0.427 *** 0.005	-0.217 *** 0.007	-0.299 *** 0.010	-0.471 *** 0.007	-0.042 *** 0.006
<b>Interpersonal</b>	-0.150 *** 0.004	-0.105 *** 0.005	0.235 *** 0.008	0.038 *** 0.005	0.181 *** 0.005
<b>Self-organisation</b>	-0.383 *** 0.004	0.137 *** 0.005	-0.123 *** 0.007	-0.118 *** 0.005	0.589 *** 0.005
<b>Technical</b>	-0.282 *** 0.004	-0.175 *** 0.005	-0.225 *** 0.007	0.452 *** 0.006	0.391 *** 0.005
<b>ICT</b>	-1.649 *** 0.010	0.874 *** 0.006	-0.332 *** 0.012	0.783 *** 0.006	-0.341 *** 0.006
<b>Cognitive</b>	-0.722 *** 0.005	-0.035 *** 0.005	-0.016 * 0.008	0.526 *** 0.005	0.211 *** 0.004

Std. Errors are reported below estimates. \*\*\* significant at p=0.01.

Source: own computations.

The demand for non-routine cognitive analytical jobs was higher in some of the well-developed regions, Mazowieckie, Malopolskie and Slaskie, as well as in Podkarpackie. The latter is a low-developed region, but with growing inflow of investments from international companies, including manufacturing (aviation industry) and financial services. This might have influenced higher demand for analytical occupations. The most commonly included skills in

job offers for analytical occupations were ICT skills, then cognitive and technical skills, and finally interpersonal skills. Other skills were less important for companies, with communication skills being the least important. Occupations characterised by non-routine cognitive interpersonal tasks were demanded in most of the regions. Three well-developed regions (Lodzkie, Malopolskie and Slaskie) were exceptions. This group of occupations might be considered most versatile in the case of skills. While previous task-based groups demanded few skills, companies seeking interpersonal occupations representatives demanded a wide range of skills. Strong communication skills were the most demanded. From nine groups of BHC skills only two were generally not demanded: office and ICT skills. Other skills were included in a generic online job offer.

Finding how possessing certain skills changes the probability of being employed in a given task content group is crucial from the point of view of this study. Our reference group is the one which involves routine manual tasks, so parameter estimates are interpreted relative to that group. As it was previously stated, routine manual tasks were (in general) most demanded in the Polish labour market, so the intercept in the model is negative for every other task-based occupation (Table 5). Communication skills were rarely included in job offers for both types of manual tasks. This skill strongly increases the probability that a job offers falls to non-routine cognitive interpersonal task group. Managing skills were least demanded in routine manual tasks. They were most important in both types of cognitive tasks. Availability skill was most important for two non-routine task-groups: manual and interpersonal cognitive. Interestingly, office skills were most often included in job offers for routine manual tasks. In most cases, interpersonal skills were more often demanded than in routine manual tasks; the probability was slightly lower only for occupations rich in routine cognitive tasks. Self-organisation skills were most important in cognitive tasks, especially those non-routine. Technical skills were least demanded for occupations involving manual tasks and most demanded for those involving non-routine cognitive tasks. ICT skills increased the probability that a job offer was assigned to any type of occupations other than routine manual, especially to routine cognitive and non-routine cognitive analytical. Finally, having cognitive skills may be connected to a chance to apply for a more sophisticated job, especially one involving non-routine cognitive tasks.



**Table 5. Skills and their effect on demand for task-based occupational groups (reference group: routine manual). Results of a multinomial logistic regression**

	<b>Routine cognitive</b>	<b>Non-routine manual</b>	<b>Non-routine cognitive analytical</b>	<b>Non-routine cognitive interpersonal</b>
<b>(Intercept)</b>	-1.175 ***	-1.885 ***	-1.570 ***	-1.494 ***
	0.007	0.009	0.007	0.006
<b>Communication</b>	0.576 ***	-0.960 ***	0.471 ***	1.955 ***
	0.020	0.046	0.020	0.016
<b>Managing</b>	0.623 ***	0.163 ***	0.370 ***	0.751 ***
	0.008	0.011	0.008	0.007
<b>Availability</b>	-0.318 ***	0.244 ***	-0.445 ***	0.109 ***
	0.007	0.008	0.007	0.005
<b>Office</b>	-0.462 ***	-0.517 ***	-0.704 ***	-0.377 ***
	0.007	0.010	0.008	0.006
<b>Interpersonal</b>	-0.023 ***	0.293 ***	0.094 ***	0.193 ***
	0.006	0.008	0.006	0.005
<b>Self-organisation</b>	0.327 ***	0.102 ***	0.143 ***	0.699 ***
	0.006	0.007	0.006	0.005
<b>Technical</b>	0.058 ***	-0.041 ***	0.593 ***	0.519 ***
	0.006	0.007	0.006	0.005
<b>ICT</b>	2.138 ***	1.088 ***	2.055 ***	1.219 ***
	0.011	0.015	0.011	0.011
<b>Cognitive</b>	0.595 ***	0.536 ***	1.029 ***	0.736 ***
	0.006	0.008	0.006	0.005

Std. Errors are reported below estimates. \*\*\* significant at p=0.01.

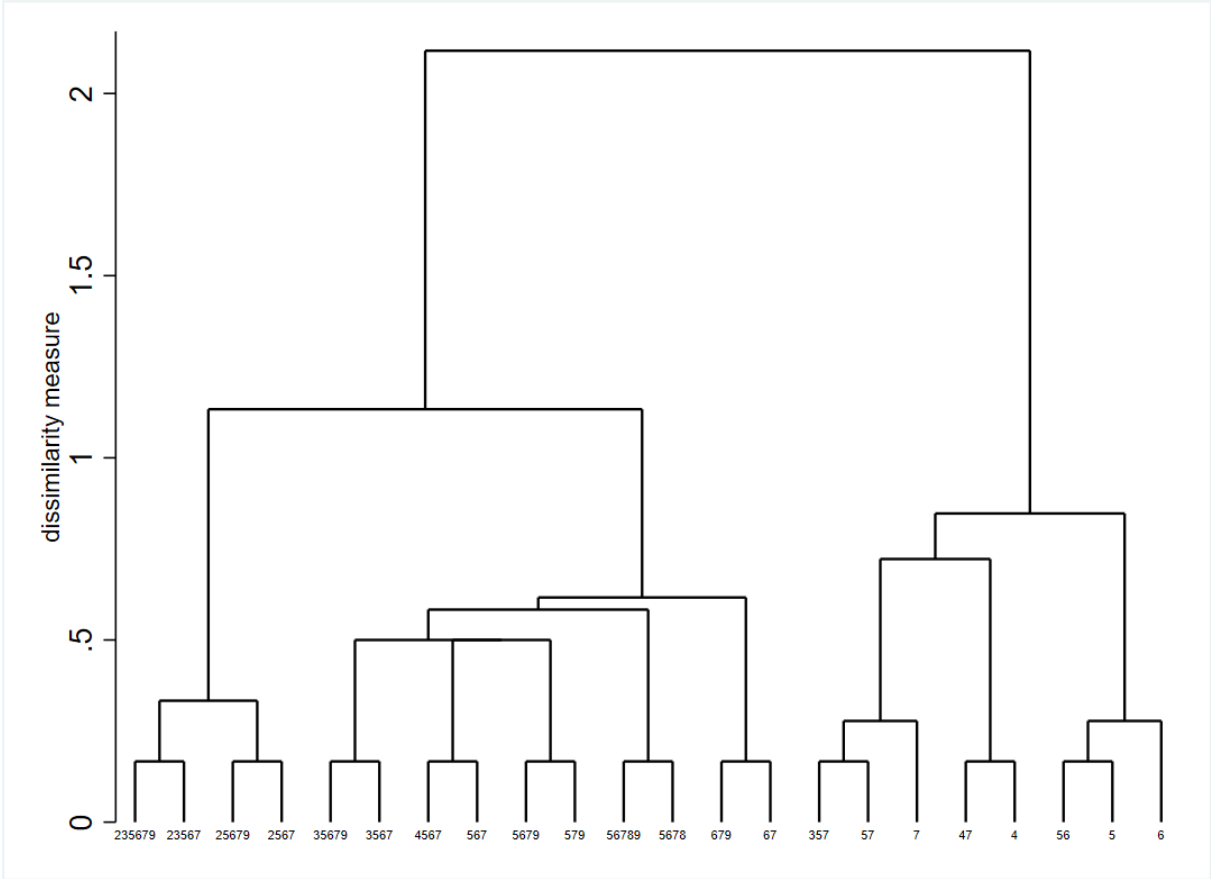
Source: own computations.

Finally, we looked at the demand for skills analysing them as a part of the job offers; we built sequences of skills. We applied optimal matching of sequences, and built the distance matrix using the Needleman-Wunsch algorithm. Then, we used the cluster analysis to generate dendrograms to show how distant particular sequences of skills in job offers were with respect to each other. The skills appeared in 365 different combinations but the frequency substantially differed. The two most common sets were: (interpersonal, self-organisation, and technical) and (interpersonal) –these two types of sequences constituted 25% of all job offers. The six most common sequences constituted 50% of all job offers, and the 65 most common sequences constituted 95% of all job offers. We examined the twenty two most common sequences of skills<sup>11</sup> that appeared in job offers. They made up 79% of all job offers, and each sequence constituted at least 1% of all sequences identified. Unsurprisingly, interpersonal, self-organisation and technical skills were the most often required ones, and none of these skills

<sup>11</sup> The results referring to all sequences are available upon request.

appeared only in one sequence studied. In 11 sequences the three types of skills were required (and accompanied by some additional ones), in subsequent 5 types of sequences any two out of three above mentioned skills were required. The relative “popularity” of each type of skill among these 22 most common sequences was the following: communication 0.00%, managing 1.42%, availability 2.54%, office 2.99%, interpersonal 24.91%, self-organisation 23.02%, technical 32.23%, ICT 1.97% and cognitive 10.93%. Based on the dissimilarity matrix we generated a dendrogram which expressed which sequences of skills appearing in the job offers were (not) more alike (Figure 2).

**Figure 2. Dendrogram for the cluster sequence analysis for the 22 most common sequences of demanded skills**



Notes: 1 – communication, 2 – managing, 3 – availability, 4 – office, 5 – interpersonal, 6 – self-organisation, 7 – technical, 8 – ICT, 9 – cognitive.

Source: own computations.

The variability in the dissimilarity matrix was quite small, but nevertheless some remarks arise. Job offers that required managing skills stood out substantially from the rest. They created a bigger group with job offers that required multiple sets of skills (at least three)<sup>12</sup>. On the other hand, most of the job offers that required predominantly one or two skills<sup>13</sup>

<sup>12</sup> The sequence of self-organisation and technical skills is an exception here.

<sup>13</sup> The sequence of availability, interpersonal and technical skills is an exception here.

(including interpersonal, self-organisation, technical, and availability) joined other job offers at the highest level of the dissimilarity measure values.

## **5. Discussion and conclusions**

It has been argued, that the timing and extent of job polarisation may differ across countries (Fonseca et al. 2018). While most of the studies on RBTC hypothesis are based on employment data, in our study we focused on testing the polarisation hypothesis from the perspective of unmet labour demand (job offers). Such an approach is scarce in the literature. The results of our study indicated a rather unusual pattern of polarisation in the Polish vacancies market, with almost an equal division between offers for non-routine and routine jobs. This conclusion is consistent with previous studies for Poland, though dealing with the RBTC from the point of view of the employment structure. They pointed out to specific pattern of labour market polarisation with increasing demand for routine cognitive jobs and relatively high wages in routine manual jobs. Our results might be explained by the following drivers: large migration outflow of low-skilled individuals after Polish accession to the European Union; changes in the skills composition of labour –increasing number of tertiary educated individuals and decreasing number of vocationally educated workers; and the neglect of vocational education by the government in 1990-2000s. Perceptible wage increase for low-skilled occupations, especially in building and construction industries, has been a side-effect of these changes.

By mapping skills versus task-content groups we shifted from characterising tasks by occupations to link job-related tasks with skills requirements. This approach enables to identify which skills are related to jobs susceptible of automation, and which ones may, at least to some extent, provide employment security.

The quantitative conclusion points out to the importance of communication, managing and self-organisation skills in case of non-routine cognitive personal jobs and ICT, cognitive and technical skills in case of non-routine cognitive analytical jobs –both groups characterised by relatively low probability of automation. Importantly, the demand for non-routine cognitive jobs is, in general, concentrated on the Polish regions which are more developed –this may pose growing challenges to regional labour markets in less developed regions stemming from the potential acceleration of polarisation processes in the future.

The qualitative conclusions reveal that in most on-line job offers available in the Polish labour market, regardless of the task content group, employers require interpersonal, self-organisation and technical skills. If job offers are most often restricted to one or two skills required (among the three mentioned above) they considerably stand out. Similarly, those jobs

that require managing skills seem to be distinguished among the others. Another set of job offers are predominantly those which require multiple skills.

Diversified labour demand in terms of required skills poses a challenge for shaping labour supply. Occupations are more and more often characterised in terms of executed tasks, what makes a challenge to map skills with these tasks. Forecasts of labour demand indicate possible routes of development, signalling a change in the relative importance of particular qualifications. That is why it seems especially important to follow the labour demand from a detailed perspective, both regionally and over time. This type of analysis constitutes a useful tool for the diagnosis and projection of the labour market development, which we aim to extend in future research.

## **6. References**

- Abraham, K.G. (1987), Help-wanted advertising, job vacancies, and unemployment, *Brookings Papers on Economic Activity*, 18 (1), 207-248.
- Acemoglu, D., Autor D. (2011), Skills, tasks and technologies: Implications for employment and earnings, in *Handbook of Labor Economics*, 4B, Card. D, Ashenfelter, O (Eds.), Elsevier, Amsterdam, 1043-1171.
- Adermon, A., Gustavsson, M. (2015), Job polarization and task-biased technological change: Evidence from Sweden, 1975-2005, *Scandinavian Journal of Economics*, 117 (3), 878-917.
- Arendt, Ł. (2018), Is Polish labour market heading towards polarisation?, *Olsztyn Economic Journal*, 13 (3), 309-322.
- Arendt, Ł., Grabowski, W. (2019), Technical change and wage premium shifts among task-content groups in Poland, *Economic Research-Ekonomska Istraživanja*, 32 (1), 3392-3410.
- Asplund, R., Barth, E., Lundborg, P., Nilsen, K.M. (2011), Polarization of the Nordic labour markets, *Finish Economic Papers*, 24 (2), 87-110.
- Autor, D., Dorn, D. (2013), The growth of low-skill service jobs and the polarization of the US labor market, *American Economic Review*, 103 (5), 1553-1597.
- Autor, D., Levy, F., Murnane, R.J. (2003), The skill content of recent technological change: An empirical exploration, *Quarterly Journal of Economics*, 118 (4), 1279-1333.
- Azar, J., Marinescu, I., Steinbaum, M., Taska, B. (2020), Concentration in US labor markets: Evidence from online vacancy data, *Labour Economics*, 66, 101886.
- Bhuller, M., Kostøl, A.R., Vigtel, T.C. (2020), How broadband Internet affects labor market matching, IZA Discussion Paper No. 12895.

Boselli, R., Cesarini, M., Marrara, S., Mercurio, F., Mezzanzanica, M., Pasi, G., Viviani, M. (2018), WoLMIS: A labor market intelligence system for classifying web job vacancies, *Journal of Intelligent Information Systems*, 51 (3), 477-502.

Colombo, E., Mercurio, F., Mezzanzanica, M. (2018), Applying machine learning tools on web vacancies for labour market and skill analysis, *mimeo*.

Conference Board (2018), The Conference Board Help Wanted OnLine (HWOL), Technical Note.

Cortes, G.M., Jaimovich, N., Siu, H.E. (2017), Disappearing routine jobs: Who, how, and why?, *Journal of Monetary Economics*, 91 (C), 69-87.

Deming, D., Kahn, L.B. (2018), Skill requirements across firms and labor markets: Evidence from job postings for professionals, *Journal of Labor Economics*, 36 (S1), S337-S369.

Dusi, S., Mercurio, F., Mezzanzanica, M. (2015), Big data as a new data source for improved regional and local labour market monitoring, in *Big Data and the Complexity of Labour Market Policies. New Approaches in Regional and Local Labour Market Monitoring for Reducing Skills Mismatches*, Larsen, C., Rand, S., Schmid, A., Mezzanzanica, M., Dusi S. (Eds.), Rainer Hampp Verlag, Munich, 31-44.

Dustmann, C., Ludsteck, J., Schonberg, U. (2009), Revisiting the German wage structure, *Quarterly Journal of Economics*, 124 (2), 843-881.

Fonseca, T., Lima, F., Pereira, S.C. (2018), Job polarization, technological change and routinization: Evidence for Portugal, *Labour Economics*, 51 (C), 317-339.

Frey C.B., Osborne M.A. (2017), The future of employment: How susceptible are jobs to computerisation?, *Technological Forecasting and Social Change*, 114 (C), 254-280.

Goos, M., Manning, A. (2007), Lousy and lovely jobs: The rising polarization of work in Britain, *Review of Economics and Statistics*, 89 (1), 118-133.

Goos, M., Manning, A., Salomons, A. (2009), Job polarization in Europe, *American Economic Review*, 99 (2), 58-63.

Goos, M., Manning, A., Salomons, A. (2014), Explaining job polarization: Routine-biased technological change and offshoring, *American Economic Review*, 104 (8), 2509-2526.

Green, D., Sand, B. (2015), Has the Canadian labour market polarized?, *Canadian Journal of Economics*, 48 (2), 612-646.

Hardy, W., Keister, R., Lewandowski, P. (2016), Technology or upskilling? Trends in the task composition of jobs in Central and Eastern Europe, IBS Working Paper No. 01/2016.

Hardy W., Keister R., Lewandowski P. (2018), Educational upgrading, structural change and the task composition of jobs in Europe, *Economics of Transition*, 26 (2), 201-231.

Hershbein, B., Kahn, L.B. (2018), Do recessions accelerate routine-biased technological change? Evidence from vacancy posting, *American Economic Review*, 108 (7), 1737-1772.

Lewandowski, P., Keister, R., Hardy, W., Górka S. (2017), Routine and ageing? The intergenerational divide in the deroutinisation of jobs in Europe, IZA Discussion Paper No. 10732.

Lovaglio, P.G., Cesarini M., Mercurio F., Mezzanzanica, M. (2018), Skills in demand for ICT and statistical occupations: Evidence from web-based job vacancies, *Statistical Analysis and Data Mining*, 11 (2), 78-91.

Mandelman, F., Zlate, A. (2014), Offshoring, low-skilled migration, and labor market polarization, Federal Reserve Bank of Atlanta, Working Paper No. 2014-28.

Marinescu, I., Wolthoff, R. (2020), Opening the black box of the matching function: The power of words, *Journal of Labor Economics*, 38 (2), 535-568.

McKinsey & Company (2018), Shoulder to shoulder with robots. Tapping the potential of automation in Poland, Report.

Michels, G., Natraj, A., Reenen, J.V. (2014), Has ICT polarized skill demand? Evidence from eleven countries over 25 years, *Review of Economics and Statistics*, 96 (1), 60-77.

Nomura, S., Imaizumi, S., Areias, A.C., Yamauchi, F. (2017), Toward labor market policy 2.0. The potential for using online job-portal big data to inform labor market policies in India, World Bank Policy Research Working Paper No. 7966.

Oesch, D. (2013), Occupational Change in Europe. How Technology and Education Transform the Job Structure, Oxford University Press, Oxford.

Oldenski, L. (2014), Offshoring and the polarization of the U.S. labor market, *Industrial and Labor Relations Review*, 67 (3), 734-761.

PARP (2011), Bilans Kapitału Ludzkiego w Polsce, Raport podsumowujący pierwszą edycję badań realizowaną w 2010 roku, Polska Agencja Rozwoju Przedsiębiorczości, Warszawa.

Parteka, A. (2018), Import intensity of production, tasks and wages: Micro-level evidence for Poland, *Entrepreneurial Business and Economics Review*, 6 (2), 71-89.

Pater, R., Szkoła J., Kozak M. (2019), A method for measuring detailed demand for workers' competences, *Economics: The Open-Access, Open-Assessment E-Journal*, 13 (2019-27), 1-29.

Turrell, A., Thurgood, J., Copple, D., Djumalieva, J., Speigner, B. (2018), Using online job vacancies to understand the UK labour market from the bottom-up, Bank of England Staff Working Paper No. 742.

Xu, H., Gu, C., Zhou, H., Kou, S., Zhang, J. (2017), JCTC: A large job posting corpus for text classification, *mimeo*.

Zhu, C., Zhu, H., Xiong, H., Ding, P., Xie, F. (2016), Recruitment market trend analysis with sequential latent variable models, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'16*, ACM, New York, 383-392.

## Appendix

**Table. Classes of skills used in the BHC study (and SOJO)**

Skills	Behavioural dimension	Behavioural sub-dimensions
Cognitive	Retrieval and analysis of information; drawing conclusions	Quick summarising large amounts of text
		Logical thinking, analysis of facts
		Constant learning of new things
Mathematical	Performing calculations	Making simple calculations
		Performing advanced mathematical calculations
Computer	Computer and Internet use	Basic knowledge of office software
		Knowledge of specialised software, ability to develop software or creating websites
		Internet use: searching websites, e-mail operations
Artistic	Artistic and creative abilities	-
Physical	Physical fitness	-
Technical	Technical imagination and the use of devices	Use of devices
		Ability to repair the devices
Self-organisation	self-organisation of work and taking initiative (planning and implementation of tasks on time, achieving goals)	Making decision independently
		Entrepreneurship and taking initiative
		Creativity (being innovative, coming up with new solutions)
		Resistance to stress
		Implementation of planned activities on time
Interpersonal	communication with other people - colleagues, clients or subordinates	Team-working
		Ease in entering into communication with co-workers and customers
		Being communicative and ability to communicate thoughts clearly
		Solving conflicts between people
Office	Organising and performing office work	-
Managerial	Managerial skill and work organisation of others	Assigning tasks to the employees
		Work coordination
		Providing discipline at work
Availability	Availability	Readiness for frequent business trips
		Flexible working time

Source: PARP (2011, p. 32).