WHICH SKILLS ARE THE MOST PRIZED? ANALYSING MONETARY VALUE OF GEOGRAPHERS’ SKILLS ON THE LABOUR MARKET IN SIX EUROPEAN COUNTRIES

Danuta Piróg, Adam Hibszer

1 Institute of Law, Economics and Administration, University of the National Education Commission, Krakow, Poland
2 Institute of Social and Economic Geography and Spatial Management, Faculty of Natural Sciences, University of Silesia in Katowice, Sosnowiec, Poland

Abstract: The objective of the study is to identify those skills that are actually needed by the labour market and allow university graduates to achieve the highest remuneration. To achieve this objective, the authors monitored, for 18 months, online job postings from six countries addressed to geography graduates. Online job postings are the most up-to-date and reliable source of data about the salaries that employers are willing to offer for specific skills or skillsets. A total of 17,397 advertisements were collected, out of which 7,407 included information about the offered salary. Applying text mining and regression tree (classification and regression tree [CART]) analyses, the authors identified skills that significantly differentiate annual salaries. The group of competences associated with higher earnings includes highly specialised geographic information system (GIS), statistical and geological skills. Lower salaries were linked to some general skills such as communicating in a native language as well as some specialised skills, but only to those related to teaching and conducting scientific research.

Keywords: demand, job postings, monetary value, regression tree (CART), remuneration, skills, text mining

Introduction

For over a decade, theoretical and empirical studies on the functioning of the labour market and university-to-work transition have been strongly skill-oriented (García-Aracil, Van Der Velden 2008; Storen, Wiers-Jenssen 2016; Giotis 2018; Krafft 2018; Naess 2020; Liwiński, Pastore 2021). Skills are key to selecting candidates for a given post and determining salary levels. They polarise jobs into those that are better paid that require non-routine tasks, more sophisticated skills that are more difficult to acquire, and those that are paid less and rely on recurrent tasks and skills that are relatively easy to master (Robst 2007; McGuinness, Sloane 2011; Wolbers 2013; Montt 2017; Ortiz-Gervasi, McGuinness 2018). Therefore, more often, young people make university choices based on economic calculation. Finding a good job, i.e. related to one’s field of
study and well paid, has become a leading criterion for making education-related decisions and one of the primary motivations for enrolment to degree programmes (Vrontis et al. 2007; Piróg 2018). Young people consider their university education to be an investment in their future, and like with any investment, they expect notable returns. The greater the investment in higher education, the more serious the concern for investing in the right type of education. This relationship is particularly evident in the education choices made in a knowledge-based economy. In this type of economy, education is considered a springboard to social and material advancement. Consequently, in a knowledge-based economy, a large part of the population views the entire education process (from kindergarten to university) as a long-term, resource-intensive investment in one’s social position (Nicolescu, Paun 2009; Mateos-Romero, del Mar Salinas-Jiménez 2017; Šafránková, Šikýř 2017; Ortiz-Gervasi, McGuinness 2018; Naess 2020; Tzanakou et al. 2020).

Therefore, for many young people, it has become crucial to choose a degree programme that will ensure a successful university-to-work transition after graduation (Teichert et al. 2022). There is also growing pressure for universities to equip graduates with competences that match and satisfy employers’ needs (Teichler 2018). It is an important component of the mission of higher education institutions (HEIs) which, besides scientific, research and education purposes, should serve the public, i.e. its stakeholders (Sheng-Ju, Jing-Wen 2016). Equipping young people with the expected so-called hot skills is defined as the responsiveness of HEIs to the career needs of their graduates (Bardhan et al. 2013). Identifying the skillset currently in demand on the labour market is a milestone for a smooth transition to work. This raises a question: Which skills should be developed? It is not only universities that opt for a market-oriented education model that seek an answer to this question. It is also an issue for those HEIs that want to duly perform the role entrusted in universities, as well as attract further populations of young people (Jurše, Tominc 2008; Bridgstock 2009; Wilton 2012; Farmaki 2018). Finding an answer to the question of which skills are actually, and not merely declaratively, sought-after by employers underpins a more in-depth analysis of the contemporary labour market, i.e. the traits and factors that impact the market dynamics as well as predict further directions of change (Kircher 2020; e.g. Agasisti et al. 2021).

In line with neoliberal orientation, it is the financial value of intellectual capital acquired during the course of study that is considered to be one of the most important premises on which people make a decision whether investing in higher education is economically viable. Consequently, it is considered important for research to establish not only the list of skills in demand but also their monetary value. The need to associate skills with monetary value has been justified in the literature using three arguments. First, changes in the size and structure of demand for specific skills and their value perfectly reflect the changes in the mechanisms operating on the labour market (Storen, Wiers-Jenssen 2016; e.g. Ortiz-Gervasi, McGuinness 2018). Second, earnings are an important premise for reflection on the value of human capital, largely collected during university education. Such data can fill a knowledge gap about the financial value of skills acquired or developed during higher education, the completion of which often involves a significant financial investment in one’s future position on the labour market (Jimeno et al. 2016). Third, evidence of earnings constitutes an important database of information for university candidates and HEIs. Owing to such information, young people with market-oriented education goals can make more informed decisions about higher education (Ayllón, Ramos 2019). On the other hand, HEIs can utilise the data to revamp those curricula that are aimed at training employable graduates and ensure that the ‘investment’ in higher education yields the highest return possible (Belfield, Bailey 2019). Knowing which skills are worth most, HEIs can also inform their students what areas require more self-study if financial aspects of employment are important to them (Brunello, Wruuck 2021). Data on the possible earnings of graduates also constitute the most valuable promotional content for HEIs. Nowadays, prospective salaries after graduation, which are linked with acquired skills, are a more important factor for university candidates than the mere chance of finding employment. Salary levels are one of the key push and pull factors
for students who consider a given degree programme or university (García-Aracil, Van Der Velden 2008; Bukowska, Łukasiewicz 2017). This factor also has a significant impact on the degree completion rate (Bardhan et al. 2013).

A clear research gap and, at the same time, the greatest practical potential can be seen in studies not linked with the estimated or declared value of skills, but with the actual monetary value of specialised skills, i.e. the skills acquired at specific degree programmes. Specialised skills are crucial in finding employment corresponding to the discipline of study, getting high-paying jobs and determining the levels of skill premium (Robst 2007; Reinhold, Thomsen 2017). The need for such studies is justified by the fact that available analyses are based on data regarding the entire population of university graduates and, thus, bring about results that are too general. Consequently, such studies can only give a general idea about the tendencies in changing demand for knowledge and skills on the labour market but do not specifically examine the structure of the demand. Subsequently, there is a need for more specific and detailed studies in order to collect ‘much detailed knowledge’ about these skills and give all stakeholders the basis for more targeted actions. Systematic research involving specialists from a single degree programme would allow for the revision of curricula on an ongoing basis and, thus, a quicker response of HEIs to employers’ needs (Mok, Wu 2016). For jobseekers to attain the highest salary levels, it is crucial to first know which specialised skills are needed by the employers and will be adequately remunerated and then to go about acquiring them (Deming, Kahn 2017; Kracke et al. 2018; Krafft 2018; Belfield, Bailey 2019).

**Theoretical framework**

Once we know that there is a need to establish a structure of the skills desired on the labour market and the price tags associated with them, another question emerges: Where can reliable data be obtained? (Heckman, Kautz 2012; Suleman 2018). The literature review has shown that hard and reliable sources of this kind of data are online job postings. Since they are exhaustive, they are a source of highly valuable, current, free of charge, non-reactive and non-declarative data on the actual directory of skills needed by employers (Smith, Ali 2014; Faberman, Kudlyak 2016; Anabo, Albizuri 2017; Gürtzgen et al. 2021). They are easy to access, relatively standardised and widely available on the Internet for almost all industries, occupations and a wide variety of countries. Such postings form a directory of data on the required qualifications, competences, foreign language skills and others. Paired with information about the remuneration offered to employees with the given skillset, job postings provide information about the actual monetary value of skills. Consequently, they are the best database about the market value of employees’ skills (Bennett 2002; Marinescu, Wolthoff 2020; Deming, Kahn 2017). In addition, compared to point-in-time snapshots provided by survey-based labour market data, which rely on random sampling, these data provide knowledge about changing demands over time and space and, thus, can underpin labour market forecasts (e.g. Clyde 2002; Kopf et al. 2012; Carnevale et al. 2014; Kim, Angnakoon 2016; Karakatsanis et al. 2017; McLaughlin et al. 2018). Nowadays, nearly 80–90% of openings that require at least a bachelor’s degree are posted online (Carnevale et al. 2014; Kureková et al. 2015). This makes online job postings the best source of knowledge on employers’ demand for certain skills for university degree holders (Di Meglio et al. 2007; Ledermüller 2011; Kureková et al. 2015; Beblavý et al. 2016; Dunbar et al. 2016; Kim, Angnakoon 2016; e.g. Anabo, Albizuri 2017; Boselli et al. 2017). For these reasons, utilising online job postings for academic investigations into labour market dynamics has been on the rise (Brown, Souto-Otero 2020).

Recently, it has been highlighted that this data source not only has a great potential and provides high-quality data but also deserves a multidimensional examination, instead of simply ‘scratching the surface’. An in-depth investigation of online job postings can lead to valuable and unique conclusions, both from academic and practical points of view. Such studies require exploration, applying novel procedures and advanced textual data analysis methods of large corpora. The methodology needs to be verified in terms of its efficiency to achieve its research goals (Kircher 2020).
The topic of demand for skills acquired in the course of higher education has attracted the growing attention of academics in recent years. However, this question has so far been explored mainly through surveys conducted among different groups of stakeholders. Research projects have dealt with employers’ expectations regarding the skills of their potential employees (e.g. Andrews, Higson 2008; König et al. 2016; Sarkar et al. 2016). Suleman (2016) investigated employers’ satisfaction with graduates’ skills. Graduates were asked to assess the usefulness of skills/competences acquired during their degree programmes in their job-seeking efforts (Nagarajan, Edwards 2015) or in their actual jobs (García-Aracil, Van Der Velden 2008; Cattani, Pedrini 2021; Liwiński, Pastore 2021). Some studies looked into how useful vocational skills were in relation to characteristics of specific occupations (Girsberger et al. 2018). Interviews were held with employers and academic lecturers regarding their opinion and thoughts about competences that are expected on the labour market from people with higher education (e.g. König et al. 2016). All these studies provide data that allow for making predictions about which skills can prove useful in finding employment. However, to a large extent, they are merely hypothetical and wishful declarations of employers and guesses of academics regarding the most likely expected skills rather than an actual expression of demand for certain competences. Thus, such responses cannot be the basis for implementing conscious market-oriented corrective actions at a given HEI or specific degree programme. Neither are they a hard premise for establishing which skills should be mastered by the graduates individually if they want to find or change a job (Sarkar et al. 2016).

Such a premise would be the value of skills expressed in monetary terms – something which has been analysed much less frequently. To date, researchers have made mainly retrospective judgements based on surveys. Through such studies it has been established that over one-third of differences between HEI graduates’ salaries can be accounted for by differences in their competences. It has also been demonstrated which participative and methodological skills are financially rewarded and which have no impact on remuneration (García-Aracil, Van Der Velden 2008). Some research projects have examined the relationship between salaries and specific skillsets or individual skills. It was found that cognitive and social skills are predictors of salary differences, i.e. the higher the salary, the higher the expected level of cognitive and social skills (Deming, Kahn 2017). Chiu and Chuang (2016) have concluded that their teamwork skills have no impact on salary levels. It was established that employers reward soft and hard skills differently. Soft skills translate to higher earnings only in jobs related to the field of study and only when one such skill is specifically listed by employers as being useful in the given position. Hard skills also bring a high wage premium when they are reported as very useful but not completely relevant (Liwiński, Pastore 2021). Researchers also indirectly investigated the dependence between skills and remuneration by looking into the impact of achievements – specifically final grades on university diplomas – on salaries. It has been established that good grades have a positive impact on salaries of graduates with a bachelor’s degree and a negligible impact when it comes to master’s graduates (Rudakov, Roshchin 2019).

Finally, a very small body of research refers to the relationship between remuneration and the skills of graduates of specific degree courses. Such studies were conducted solely for degree programmes such as marketing, general management, finance, human resource management (Bennett 2002) and science, technology, engineering and maths (STEM) (Altonji et al. 2012; Croce, Ghignoni 2020). To the best of our knowledge and conducted literature reviews, so far no research has been based on job postings to identify which skills actually result in better pay and which – even if declared as useful by employers – are associated with a lower price tag. At the same time, due to the premises outlined above, it is necessary to academically investigate this topic. It is vital to establish which skills result in the best possible match between university graduates’ competences and the current demand of employers who are looking for specific specialists. Such knowledge gives graduates the best chance for the highest skill premium, which results from their investment in a specialised university degree course (Kircher 2020).

Our article is an attempt to fill the research gap regarding the monetary value of the skills really needed by employers, i.e. those listed in actual
job postings addressed to specialists, i.e. candidates with a higher education diploma. We have tried to assign a value not to individual skills but to specific skill sets that translate into the highest remuneration for geography graduates. Besides expanding our knowledge, the study results could be useful to two groups of stakeholders: specialists responsible for making geography degree curricula more market-oriented and students and graduates who would like to acquire and develop skills that can improve their chances of finding a well-paid job.

Research materials and methods

Owing to the identified need to analyse demand for specialised skills expressed in monetary terms and examine the potential of job postings as a data source, the main research objectives were:
- to identify the monetary value of individual skills actually expected by the labour market from specialists with a geography degree;
- to separate a subgroup of skills that are crucial in applying for top-paying jobs, i.e. skills that yield the highest monetary skill premium.

A complementary research objective results from the postulates outlined in the literature review regarding the need to implement advanced analysis methods for examining large textual datasets in the field of the labour market. Consequently, a complementary research objective is to verify whether regression tree methodology can be used to achieve the primary research objective.

The research is underpinned by the rational choice theory framework that assumes that an individual, when making a decision, represents the attitude of *homo oeconomicus*. This means that in their decision-making process, including the decision to invest in skill acquisition, an individual is driven by pragmatism (e.g. Lindenberg 1992; Olssen, Peters 2005). People are rational beings and are driven by logic. They make informed decisions that will ensure the highest possible profit (Breen, Goldthorpe 1997; Jaeger 2007). A pragmatic attitude in the decision-making process translates into performing a cost-benefit analysis of material and non-material gains and then making a choice that will maximise the former (Eriksson 2011). Therefore, *homo oeconomicus*, being a rational consumer, considers university education or ongoing training to be ‘an investment’ with a view of obtaining maximum benefits. In line with this theory, people considering accumulating more skills capital calculate first and foremost whether it will bring them material gain, usually associated with the salary level. They also expect non-material gains such as higher social status, professional recognition and intellectual stimulation.

In line with the paradigm of perfect rationality, decisions should be made based on exhaustive information, e.g. about real gains from an investment (Sordyl 2009). Not knowing the monetary value of skills on the labour market, employees do not have reliable premises on which to base their decision-making processes. If such information is permanently missing or incomplete, then both HEIs and their former and recent graduates are not able to respond accurately, i.e. obtain the extra skills in demand. The greater the mismatch between labour market needs and graduates’ skills, the greater the likelihood of search frictions, i.e. impediments to a match. They result in vertical or horizontal mismatches and trigger a lower earning risk because employers need to bear the costs of employee adaptation to the tasks assigned (Mavromaras et al. 2013). Employees who, as a result of bad educational decisions, are not aligned with market needs end up earning less than those whose skills are a closer match, thus incurring a wage penalty in the form of lower remuneration (Robst 2007; McGuinness, Sloane 2011; Wolbers 2013). Therefore, determining the monetary value of skills that are really needed by employers can provide the necessary knowledge to make more informed decisions about investing in one’s skills capital, which should result in skill premia on the labour market.

The data collection process was as follows. To collect data on the real demand of employers for skills and knowledge of graduates, we used job adverts posted online over a period of 18 months, i.e. from the 1st of February 2019 to the 31st of July 2020. Before we took the decision to collect the data over a period of 18 months in several countries, we studied the relevant literature on the value and reliability of job adverts as a data source for academic research.

We chose GEES (Geography, Earth and Environmental Sciences) because this academic
Discipline combines the scientific and cognitive aspects with elements of a practical nature. It has always been an important part of geography’s mission to educate graduates who are well-prepared to the needs of the labour market and can make a smooth university-to-work transition.

To make sure the data sample justifies drawing general conclusions and allows for detecting possible regularities on an interregional level, we collected job postings from websites in the following six countries: Austria, Germany, Great Britain, Ireland, Poland, and Switzerland. In three countries, the adverts were collected in German, two in English, and one in Polish. The following keywords and their German and Polish equivalents were used in the search: Geography, Geology, Geographic Information System (GIS), Earth Sciences.

Having verified in the literature review that job adverts are a reliable data source providing information about the real demand, we made sure to collect job postings from the leading job search websites with a proven track record. Therefore, a total of 116 commercial and institutional websites with job postings were selected for the study (see Appendix 1). The postings were downloaded on the first working day in a given week. In most cases, we collected the adverts automatically (web-scrapping) using Octoparse software. Because the data were gathered partially automatically, the collected job postings were then checked to eliminate duplicate records, incomplete downloads, or adverts that were collected by mistake and were not addressed specifically to people with a geography degree. We eliminated duplicate records in two stages. The first stage was fully automatic, where adverts with identical content or ID numbers specific for a given website were identified and duplicates removed from the database. Second, the automatic process was verified through queries to the database (identical advert title, location, employer, remuneration). In this way, we were able to identify adverts with a high degree of similarity, which were then compared and, subsequently, either removed or kept in the database. This was the second method of data deduplication. Finally, we entered 17,397 adverts with information about salaries, out of which 3,553 included information about remuneration in British pounds (GBP), 3,470 in Polish zlotys (PLN), and 384 in euros (EUR). We standardised the data for each currency by converting the numbers into annual salaries. Next, further analyses were conducted for each of the three currencies.

The extraction of specific data from job adverts to fulfill a defined research objective is a challenging task and requires the application of special tools (Colombo et al. 2019). In our case, to extract words related to skills from such an extensive body of text, we first used standard Statistica text mining to extract word occurrences in specific adverts. It was a binary word occurrence analysis, i.e. 0 denoted no occurrence of the word in the advert and 1 meant that the word occurred in the advert at least once. As a result, we compiled a list of 20,709 words for GBP, 3,176 words for PLN, and 5,291 words for EUR. It is a highly valued and commonly applied method for advanced statistical quantitative analysis of texts and documents (Rybchak, Basystiuk 2017). Next, from a list of words ordered from the most to the least frequently occurring, we manually selected terms broadly related to skill areas. For this procedure, we used 49 geographic and general skill areas in professional geography, adopted by Solem et al. (2008). As a result, we extracted a list with the information about salaries offered for each of the postings and the corresponding word occurrence statistics.

A matrix was created where each row signified an individual advert and each column represented a variable with words concerning the required specific (individual) skills and key variable, i.e. remuneration.

Having extracted the individual skills, in line with the objective of the study, we tried to indicate the skills that are key when applying for best-paying jobs, i.e. which provide the highest monetary skill premia. For this reason, we did not use the ESCO (European Skills, Competences, Qualifications and Occupations) multilingual classification of skills, competences and occupations, which is, however, highly useful to group skills similar in terms of content and characteristics into classes (De Smedt 2015; Colombo et al. 2019; Chiarello et al. 2021), but in our study, the matrix was the input data for regression tree
analysis, which belongs to the group of decision trees (classification and regression tree [CART]).

These are graphically displayed models of observations divided into a specific number of subsets. The objective of this kind of analyses is to obtain subsets (words) that are as homogenous as possible from the point of view of the value of the dependent variable (Breiman et al. 1984; Martin et al. 2018).

The CART is branded by the fact that it constructs binary trees. It means that each internal node has exactly two outgoing edges, which generate the decision tree with variable branches per node. This is “a binary recursive partitioning procedure, capable of processing continuous and nominal attributes both as targets and predictors” (Almunirawi, Maghari 2016: 757). It is used both for classification of objects (dataset) and predicting the dependent variable or, specifically, the relationship between independent variables and the dependent variable. Classification trees are used for data classification/segmentation and are created when the dependent variable is qualitative. On the other hand, regression trees are used to predict the value of a dependent variable of a continuous type, i.e. a variable that can take on any value within a finite or infinite range. In our research, the dependent continuous variable is a specific amount of money that is offered by the employer for a given skill area and independent variables are a set of skill areas.

A considerable advantage of using regression trees is the fact that in the case where the response variable does not have classes, a regression model is fitted to each of the independent variables, isolating these variables as nodes where their

<table>
<thead>
<tr>
<th>word/sentence</th>
<th>F-value</th>
<th>p-value</th>
<th>word/sentence</th>
<th>F-value</th>
<th>p-value</th>
<th>word/sentence</th>
<th>F-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mother tongue</td>
<td>43.21917</td>
<td>0.000000</td>
<td>geological</td>
<td>100.2249</td>
<td>0.000000</td>
<td>gis</td>
<td>159.2739</td>
<td>0.000000</td>
</tr>
<tr>
<td>teaching</td>
<td>19.85193</td>
<td>0.000011</td>
<td>study</td>
<td>60.5423</td>
<td>0.000000</td>
<td>python</td>
<td>116.1116</td>
<td>0.000000</td>
</tr>
<tr>
<td>foreign language</td>
<td>13.66830</td>
<td>0.000250</td>
<td>pedagogical</td>
<td>49.5815</td>
<td>0.000000</td>
<td>javascript</td>
<td>75.1077</td>
<td>0.000000</td>
</tr>
<tr>
<td>interdisciplinary</td>
<td>12.11479</td>
<td>0.000558</td>
<td>land</td>
<td>44.3487</td>
<td>0.000000</td>
<td>implement</td>
<td>54.3710</td>
<td>0.000000</td>
</tr>
<tr>
<td>computer</td>
<td>8.83863</td>
<td>0.003136</td>
<td>draft</td>
<td>41.5861</td>
<td>0.000000</td>
<td>software</td>
<td>53.7989</td>
<td>0.000000</td>
</tr>
<tr>
<td>implement</td>
<td>6.06630</td>
<td>0.014219</td>
<td>advanced</td>
<td>34.7165</td>
<td>0.000000</td>
<td>teaching</td>
<td>45.4052</td>
<td>0.000000</td>
</tr>
<tr>
<td>statistical</td>
<td>5.77829</td>
<td>0.016701</td>
<td>field</td>
<td>33.6697</td>
<td>0.000000</td>
<td>geospatial</td>
<td>44.6137</td>
<td>0.000000</td>
</tr>
<tr>
<td>professional</td>
<td>5.12849</td>
<td>0.024096</td>
<td>hydrogeological</td>
<td>33.6245</td>
<td>0.000000</td>
<td>technological</td>
<td>40.1315</td>
<td>0.000000</td>
</tr>
<tr>
<td>publication</td>
<td>4.72761</td>
<td>0.030295</td>
<td>hydrography</td>
<td>32.4548</td>
<td>0.000000</td>
<td>explaining</td>
<td>37.2314</td>
<td>0.000000</td>
</tr>
<tr>
<td>processing</td>
<td>4.42281</td>
<td>0.036115</td>
<td>documentation</td>
<td>30.7242</td>
<td>0.000000</td>
<td>geoenvironmental</td>
<td>35.9375</td>
<td>0.000000</td>
</tr>
<tr>
<td>investigative</td>
<td>4.17468</td>
<td>0.041718</td>
<td>graphics</td>
<td>29.1576</td>
<td>0.000000</td>
<td>solution</td>
<td>33.9021</td>
<td>0.000000</td>
</tr>
<tr>
<td>international</td>
<td>3.66011</td>
<td>0.056477</td>
<td>executing</td>
<td>27.1883</td>
<td>0.000000</td>
<td>define</td>
<td>32.1312</td>
<td>0.000000</td>
</tr>
<tr>
<td>graphic</td>
<td>3.64130</td>
<td>0.057112</td>
<td>English</td>
<td>25.5585</td>
<td>0.000000</td>
<td>reliability</td>
<td>32.0249</td>
<td>0.000000</td>
</tr>
<tr>
<td>precise</td>
<td>3.50102</td>
<td>0.062095</td>
<td>construction</td>
<td>25.2428</td>
<td>0.000001</td>
<td>investigative</td>
<td>31.7571</td>
<td>0.000000</td>
</tr>
<tr>
<td>relations</td>
<td>3.49585</td>
<td>0.062288</td>
<td>flexible</td>
<td>24.8398</td>
<td>0.000001</td>
<td>project</td>
<td>31.4953</td>
<td>0.000000</td>
</tr>
<tr>
<td>negotiating</td>
<td>3.12685</td>
<td>0.077810</td>
<td>oligophenopadagogy</td>
<td>23.7844</td>
<td>0.000001</td>
<td>data</td>
<td>31.0935</td>
<td>0.000000</td>
</tr>
<tr>
<td>coping</td>
<td>2.72856</td>
<td>0.099390</td>
<td>drilling (adjective)</td>
<td>22.4606</td>
<td>0.000002</td>
<td>geological</td>
<td>30.9687</td>
<td>0.000000</td>
</tr>
<tr>
<td>teaching</td>
<td>2.56951</td>
<td>0.109768</td>
<td>drilling (noun)</td>
<td>21.4213</td>
<td>0.000004</td>
<td>technical</td>
<td>30.9504</td>
<td>0.000000</td>
</tr>
<tr>
<td>university</td>
<td>2.48751</td>
<td>0.115581</td>
<td>office</td>
<td>18.9589</td>
<td>0.000014</td>
<td>classroom</td>
<td>30.8648</td>
<td>0.000000</td>
</tr>
<tr>
<td>independence</td>
<td>2.35215</td>
<td>0.125938</td>
<td>map</td>
<td>16.7032</td>
<td>0.000045</td>
<td>academical</td>
<td>30.0520</td>
<td>0.000000</td>
</tr>
<tr>
<td>speech</td>
<td>2.34226</td>
<td>0.126735</td>
<td>managing</td>
<td>16.2222</td>
<td>0.000058</td>
<td>ensure</td>
<td>30.0199</td>
<td>0.000000</td>
</tr>
<tr>
<td>innovative</td>
<td>2.29678</td>
<td>0.130469</td>
<td>training</td>
<td>14.9204</td>
<td>0.000114</td>
<td>manage</td>
<td>29.9769</td>
<td>0.000000</td>
</tr>
<tr>
<td>qualification</td>
<td>2.26085</td>
<td>0.133507</td>
<td>water</td>
<td>14.4314</td>
<td>0.000148</td>
<td>geotechnical</td>
<td>29.2295</td>
<td>0.000000</td>
</tr>
<tr>
<td>practical</td>
<td>2.13958</td>
<td>0.144364</td>
<td>geotechnician</td>
<td>14.1588</td>
<td>0.000171</td>
<td>cooperation</td>
<td>28.6344</td>
<td>0.000000</td>
</tr>
<tr>
<td>utilise</td>
<td>2.12955</td>
<td>0.145305</td>
<td>autodoc</td>
<td>13.4466</td>
<td>0.000249</td>
<td>enterprise</td>
<td>27.6629</td>
<td>0.000000</td>
</tr>
<tr>
<td>expert opinions</td>
<td>2.01128</td>
<td>0.156949</td>
<td>designing</td>
<td>12.7017</td>
<td>0.000370</td>
<td>lead</td>
<td>26.5837</td>
<td>0.000000</td>
</tr>
<tr>
<td>drafting</td>
<td>2.00620</td>
<td>0.157472</td>
<td>individual</td>
<td>12.2152</td>
<td>0.000480</td>
<td>progress</td>
<td>26.1714</td>
<td>0.000000</td>
</tr>
<tr>
<td>management</td>
<td>1.92907</td>
<td>0.165669</td>
<td>law</td>
<td>12.0555</td>
<td>0.000523</td>
<td>improve</td>
<td>26.0998</td>
<td>0.000000</td>
</tr>
<tr>
<td>solving</td>
<td>1.87776</td>
<td>0.171393</td>
<td>laboratory</td>
<td>11.6474</td>
<td>0.000650</td>
<td>identify</td>
<td>25.1227</td>
<td>0.000001</td>
</tr>
<tr>
<td>applying</td>
<td>1.74394</td>
<td>0.187431</td>
<td>resocialisation</td>
<td>11.4585</td>
<td>0.000720</td>
<td>practical</td>
<td>25.1004</td>
<td>0.000001</td>
</tr>
</tbody>
</table>

Table 1. The best predictors for a dependent variable (remuneration) by currency.

Source: own study.
inclusion decreases error. Therefore, regression trees are an alternative to classic multiple regression when it comes to examining, i.e. detecting and presenting a graphic visualisation of relationships between a set of independent variables and a metric/numeric dependent variable (Loh 2011; Almunirawi, Maghari 2016; Breiman et al. 1984; Martín et al. 2018). Furthermore, regression trees not only allow for detection of single words (skill descriptions) but also relationships between them, thus showing sequences of skills that lower or increase average salaries in the posted adverts. It is a tool used for the mathematical valuation and graphic representation of these skill sets (not necessarily similar in terms of their characteristics) which yield the highest monetary skill premium, and this was the secondary objective of the study. A standard procedure of a regression tree model is to indicate the most predictive variables (see Table 1). It was assumed that for each group, there will be 30 variables, i.e. skills or terms that describe skill areas (Solem et al. 2008), with the highest importance.

Next, using the selected variables, we created models of regression trees (CART). At the same time, we selected the so-called stop parameters in order to get the best structure of the tree. We divided the sample into learning and testing groups at the levels of 70% and 30%, respectively (see Fig. 1). For each currency, we generated a tree with the optimum structure of leaves and branches.

**Results**

As demonstrated by the regression tree structure (see Fig. 2), the skills of geography graduates which primarily increase salary levels on the British market are specialised skills in the field of GISs. On an annual basis, these skills increased remuneration by over 25%, from nearly 36,000 GBP to just over 46,000 GBP per annum. Subsequent capital components that had a significant positive impact on salaries were skills related to advanced processing of geographic information using Python and JavaScript programming languages. The combination of the three skills, namely, GIS, Python and JavaScript, translate into earnings that were higher by just over one-third. At the same time, they are associated with the highest average annual salaries, amounting to nearly 75,000 GBP per annum.

Subsequent word sequences that refer to skills required by employers and result in higher earnings demonstrate the demand for GIS skills that can be applied in a geographical environment. These skills include identifying processes and phenomena, spatial planning, implementing actions for regional development, etc. They are represented by such word co-occurrence patterns as GIS and identify; GIS and enterprise; GIS and geospatial; GIS and implement. Each of these word sequences increased salaries by over one-fifth – by between 8,000 and 11,000 GBP per annum. The most financially rewarding was the combination of the words GIS and enterprise.

In the postings where the word GIS did not appear, the combination of words associated with higher earnings included such skills or skill areas as solution and reliability; solution and implement as well as practical and lead. They resulted in a skill premium amounting to approximately 15%. Moreover, we identified single skills/skill areas that are rewarded with higher earnings, i.e. lead and cooperation.

Another group of skills that are highly correlated with the field of study and in demand, but
are associated with lower salaries are explaining (remuneration lower by one-sixth); research (investigative) and geo-environmental skills (in both cases, salaries were lower by one-tenth) as well as teaching. The latter skill was associated with the most significant decrease in the offered salary, i.e. by one-fourth: from 46,000 to 34,500 GBP per annum.

For job postings published in Poland, we did not detect word sequences describing skill sets that would impact the levels of salaries offered. However, the regression tree clearly demonstrates which single skills increase or decrease the level of remuneration. The first skill associated with a higher salary was related to geology, i.e. conducting geological field studies. This skill was rewarded with salaries that were higher by 75%: from approximately 20,000 to 36,000 PLN per annum. Next on the list were skills from the field of oligophrenopedagogy, which translated into salaries that were higher by almost 50%. English speaking skills bring about a skill premium amounting to 38%. The tree then indicates that flexibility is the skill that increases earnings significantly, i.e. by as much as 180%. Given the above, we analysed in depth the set of job postings classified in the tree under the ‘flexibility’ leaf. It turned out that all of them referred to managerial positions in a variety of sectors (education, information systems, environmental management). However, a small number of postings (N = 8), classified under this leaf in the tree and its positioning, suggest excluding this competence from the group of skills that actually determine the most financially rewarding jobs.

The only skill identified to lower the offered salaries on the Polish market are pedagogical skills related to the teaching profession. These skills lower the earnings by over 25% (see Fig. 3).

Among the group of job postings with earnings in euros, the regression tree model, like in the case of Polish postings, did not show any word sequences describing skills that determine salary levels. However, it did show which

---

Fig. 2. Regression tree for GBP.
Source: own work.
Fig. 3. Regression tree for PLN.
Source: own work.

Fig. 4. Regression tree for EURO.
Source: own work.
individual resources of educational capital have an impact on remuneration. First, the model distinguished skills related to lower earnings. These include communication skills in the mother tongue and in a foreign language. Both these skills significantly lowered the average salaries offered. Mother tongue communication skills translated into remuneration that was lower by nearly two-thirds: from nearly 35,000 to approximately 10,000 EUR per annum. Owing to this significant reduction, we examined job postings classified under this leaf in the regression tree. We found that the adverts were addressed mainly to fresh geography graduates and included positions related to selling travel products or equipment, devices and technologies used in the GEES sector. The candidates were offered a lower base salary and actual compensation depended on employees’ efficiency because it included sales bonuses. On the other hand, foreign language skills were associated with a drop in salaries by more than 50%. Another examples of skills from the ‘cheaper’ group turned out to be international and processing. They resulted in salaries lower by approximately one-seventh.

The first competence related with a skill premium were statistical skills. They resulted in salaries that were higher by over one-third: from nearly 35,000 to 47,700 EUR per annum. Statistical skills were also associated with the highest annual salaries, amounting to 47,728 EUR per annum. Another skill that increased the offered remuneration was the ability to implement, which resulted in salaries higher by more than 50% (see Fig. 4).

Discussion and conclusions

In the course of the study, we managed to identify a group of geography graduates’ skills/skill areas that are really in demand on the labour market along with their monetary value. In job postings with salaries in GBP, employers were willing to pay the most, i.e. by about one-sixth to one-third more, for skills related to modern and specialised GIS tools used in spatial planning and environmental management. On the other hand, the set of skills associated will lower earnings (by about 17–25%) included abilities related to teaching geography and conducting environmental studies. In PLN adverts, the most prized skills were those related to geological research, teaching children with special needs, ability to communicate in English and flexibility. They translate into a significant, yet very diverse, skill premium ranging from 38% to 73%. Among the competencies associated with lower salaries (by 28%) were, like in the case of British postings, pedagogical skills. In the group of adverts in EUR, the most rewarded were abilities to utilise mathematical knowledge for environmental, social and market statistical analyses as well as implementation skills. They were associated with a skill premium ranging from 36% to 50%. On the other hand, lower salaries were offered for communication skills, collaboration in international teams and data processing. Communication skills translated into an all-time low remuneration – average salaries decreased by as much as 72%.

Comparative analyses of regression trees for individual currencies highlighted both similarities and differences. They referred to both the structure of single skills or the combinations of skills associated with either a skill premium or lower salaries and the degree of impact on remuneration levels.

A common denominator for all trees was that the group of most valued skills, associated with the highest skill premium, was dominated or entirely populated by specialised skills. Some specialised skills, usually those related to teaching and environmental studies, are also present in the group that lowers remuneration, but then they are usually accompanied by some general skills. The finding that the higher monetary value of specialised skills of geography experts often refers to the use of new tools and technologies corresponds to the conclusions reached by other researchers. Studies demonstrated that hard, highly specialised skills that distinguish specific employees are a gateway to higher earnings (e.g. Corominas et al. 2010; Naess 2020). The lack of such skills or their insufficient occupational specificity is the reason for difficulties in university-to-work transition for graduates of the so-called soft degree programmes such as humanities or social sciences (Klein 2010). Research found that humanities graduates earn 40% less than engineers, 36% less than social science graduates and 26% less than experts in natural sciences (Grave, Goerlitz 2012). However, the question is why do specialised skills vary so significantly
in their monetary value? The reasons are twofold. On the one hand, it is linked to how costly they are to acquire. In other words, these specialised skills that increase salaries are more expensive to learn. On the other hand, it may be because there is a certain set of skills that are taken for granted in the case of graduates of specific degree programmes and employers are not prepared to additionally reward them (Suleman 2018).

However, there are differences in how significant the impact of these skills on salaries in currency groups is. The highest skill premium for certain skills (or skillsets) are given to employees earning in PLN, then in EUR and finally in GBP. In turn, the greatest impact of the identified skills on decreasing remuneration was found in adverts in EUR, then PLN and finally GBP. Geographical variations in the size of skill premia and penalties (with the former being more pronounced) have also been found in other studies. Wage penalties amount to approximately 4-10% (Diem, Wolter 2014). Skill premia range between 25% and 15% (Liwiński, Pastore 2021). Our study confirms these trends, and in the sample of geography experts, they are even more prominent.

We argue that the adopted research procedure is well suited to recognising the monetary value of skills that are really needed on the labour market. Job postings provide current and unique source data about the demand for specific skills and the price offered for them by employers. Utilising big data collected on the Internet to real-time labour market analyses responds to calls for innovation in data sources and methods used in higher education and labour market studies (Burrows, Savage 2014). Text mining techniques and regression trees are conducive to detecting and highlighting words or word sequences that describe skills for which employers are willing to pay specific amounts. They provide hard quantitative underpinning for identifying a group of skills that are key to top-paying jobs that can be further analysed using qualitative methods. The success of the adopted methodology in reaching the research objective allows us to recommend it to other researchers. The presented method of data collection and analysis with text mining techniques and regression trees is universal and can be applied to research on any degree programme in any region. To our knowledge, the procedure adopted in this study has not been used so far in research on the market monetary value of skills capital of HEI graduates.

Given the long duration of data collection (18 months), the variety of sources and the international scale, our results should have evened out seasonal variations in the job market (Bennett 2002). Therefore, the results allow us to formulate detailed conclusions about the regularities in the field of geography and provide a useful yardstick for other degree programmes. Our research results can also serve a variety of stakeholders to implement their goals more efficiently. For instance, for market-oriented university candidates, the results can – in line with the rational choice theory – help make decisions about which degree programme to pursue. Given the current, highly competitive situation on the labour market, HEI graduates need to have a comprehensive and matching repertoire of knowledge and skills for successful university-to-work transition (Cattani, Pedrini 2021).

For recent graduates looking for their first jobs, as well as experienced professionals, the conclusions can be a benchmark regarding which skills they should particularly invest in if salary levels are their key criteria. The results can also help the youth acquire some of the skills required in the labour market before they leave the education system (Scarpetta, Sonnet 2012). Finally, they can be a useful tool for HEIs that train GEES specialists to review their curricula and strengthen these competences that translate into the highest skill premium (Mok, Wu 2016). There is no contradiction between teaching skills that are sought after on the market and the concept of university education based on co-teaching and problem-based learning (Naylor, Veron 2021).

**Limitations and future research**

The presented research procedure has two key limitations. First, even though job postings are clearly a high-quality data source about the demand for skills, it needs to be emphasised that not all of them include information about remuneration (Brenčič 2012; Marinescu, Wolthoff 2020). This was also the case in our study where some employers did not publish salary information, especially in the groups of adverts in EUR. In such cases, regardless of the motivation and
determination of the researcher, the dataset is more limited which, in turn, hinders them from formulating general conclusions. Second, job postings vary in how much detail they provide on the skills expected from candidates. There is a regularity that the higher the salary offered, the more detailed the expectations (Bennett 2002). This observation was corroborated in our study. Limited detailed information on skills for lower-paid jobs is not a hindrance if we want to identify the group of the most expensive skills. However, it can restrict analyses that focus on the skills associated with a wage penalty.

There are several main topics for future research in this field. They are worth academic investigation because of the high academic and practical value of the studies on demand for skills and their monetary value. It is to be expected that the competition on the labour market will intensify and exacerbate differences between HEI graduates of different degree programmes and their chances to obtain better paid jobs. Therefore, the question of demand for skills and their associated value needs to be further explored in relation to graduates of different degree programmes, both on national and international levels. Such investigations will make it possible to draw comparisons between different degree programmes and recognise changing trends on the labour market (Reimer et al. 2008; Kracke et al. 2018; Krafft 2018). The results could be used to revise curricula for specific degree programmes (Mok, Wu 2016). A second important aspect, challenging from a research point of view, is to recognise the decision-making process with regard to selecting and hiring a specific candidate vis-à-vis the job requirements published in the posting. The main question is to what extent were employers’ expectations met regarding job requirements and how many requirements candidates had to meet (if not all) in order to be considered for the job and then actually hired (Brown, Souto-Otero 2020). It would be important to identify – both for top-paying jobs and those associated with a wage penalty – the skills that were absolutely crucial and those that were considered complementary. A comparative study on the flexibility of employers regarding candidates’ skills could be a valuable contribution to the body of research on shortage and surplus occupations.

Acknowledgments

The study was financed from budget funds of the National Science Centre, Poland as research project No 2018/29/B/HS4/00847.

Author’s contribution

DP: conceptualisation; investigation; funding acquisition; writing – original draft; methodology; validation; visualisation; software; formal analysis; project administration; data curation; supervision; resources; AH: investigation; resources, project administration.

References


Storen A.L., Wiers-Jenssen J., 2016. Transition from higher education to work: Are master graduates increasingly...


Appendix 1. List of websites used for the job ads data collection (in an alphabetical order)

- at.indeed.com
- at.linkedin.com
- bip.amu.edu.pl
- bip.edu.bydgoszcz.pl
- bip.imgw.pl
- bip.krakow.rdos.gov.pl
- bip.lodzkie.pl
- bip.malopolska.pl
- bip.mos.gov.pl
- bip.slaskie.pl
- bip.uj.edu.pl
- bip.umwp.wroclapodlasia.pl
- bip.umww.pl
- bip.uni.lodz.pl
- bip.uni.wroc.pl
- bip.up.krakow.pl
- bip.usz.edu.pl
- bip.warmia.mazury.pl
- bydgoszcz.lento.pl
- ch.linkedin.com
- de.indeed.com
- de.linkedin.com
- gazetapraca.pl
- ie.indeed.com
- ig.up.krakow.pl
- jaslo.lento.pl
- ko.poznan.pl
- ko.rzeszow.pl
- ko-gorzow.edu.pl
- kuratorium.bydgoszcz.uw.gov.pl
- kuratorium.kielce.pl
- kuratorium.krakow.pl
- mbopn.kuratorium.waw.pl
- mota-engl.elevato.net
- nabory.kprm.gov.pl
- oferty.kuratorium.opole.pl
- oferty.praca.gov.pl
- ogloszenia.kuratorium.wroclaw.pl
- ogloszenia.trojmiasto.pl
- o-gorzow.edu.pl
- oig.ug.edu.pl
- piaszczno.lento.pl
- pl.indeed.com
- pl.jobsora.com
- pl.joboole.org
- poczta.wp.pl
- pracadawcy.pracuj.pl
- publicjobs.ie
- uml.lodz.pl
- uratorium.bydgoszcz.uw.gov.pl
- us.edu.pl
- wupkrakow.praca.gov.pl
- ww.jobs.ch
- www.absolvent.pl
- www.aplikuj.pl
- www.arbeiten.de
- www.bip.podkarpackie.pl
- www.careerjet.at
- www.careerjet.pl
- www.careerteachers.co.uk
- www.eani.org.uk
- www.educationposts.ie
- www.goldenline.pl
- www.gospodarkamorska.pl
- www.gowork.pl
- www.gumtree.pl
- www.gup.gdansk.pl
- www.ie.indeed.com
- www.ie.linkedin.com
- www.igf.edu.pl
- www.igipz.pan.pl
- www.indeed.ch
- www.indeed.de
- www.inned.de
- www.infopraca.pl
- www.ing.pan.pl
- www.inig.pl
- www.innmed.de
- www-jobagent.ch
- www.jobcentreonline.com
- www.jober.pl
WHICH SKILLS ARE THE MOST PRIZED? ANALYSING MONETARY VALUE OF GEOGRAPHERS’ SKILLS ON THE LABOUR...