

Lenaerts, Karolien; Beblavý, Miroslav; Fabo, Brian

Article

Prospects for utilisation of non-vacancy Internet data in labour market analysis: An overview

IZA Journal of Labor Economics

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Lenaerts, Karolien; Beblavý, Miroslav; Fabo, Brian (2016) : Prospects for utilisation of non-vacancy Internet data in labour market analysis: An overview, IZA Journal of Labor Economics, ISSN 2193-8997, Springer, Heidelberg, Vol. 5, Iss. 1, pp. 1-18, <https://doi.org/10.1186/s40172-016-0042-z>

This Version is available at:

<https://hdl.handle.net/10419/152364>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



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

ORIGINAL ARTICLE

Open Access



Prospects for utilisation of non-vacancy Internet data in labour market analysis—an overview

Karolien Lenaerts^{1*} , Miroslav Beblavý¹ and Brian Fabo^{1,2}

* Correspondence:

karolien.lenaerts@ceps.eu

¹Centre for European Policy Studies (CEPS, Brussels), CEPS, 1 Place du Congrès, 1000 Brussels, Belgium
Full list of author information is available at the end of the article

Abstract

Along with the advancement of the Internet in the last decade, researchers have increasingly identified the web as a research platform and a data source, pointing out its value for labour market analysis. This article presents a review of online data sources for this field. Specifically, the article introduces web-based research, focusing on the potential of relatively new data sources such as Google Trends, social networks (LinkedIn, Facebook and Twitter) and Glassdoor (surveys). For these data sources, a review is done and recent empirical applications are listed. Web-based data can further our understanding of the dynamics of the labour market.

JEL codes: E4, J2

Keywords: Labour market, Web-based research, Data sources, Google Trends, Social networking sites, Glassdoor

1 Introduction

The last few decades have been characterised by an astonishing advancement of the Internet, which has substantially transformed many aspects of life. As a result, the number of Internet users has skyrocketed. A recent report by the International Telecommunications Union shows that at the end of 2014, almost three billion people had access to the Internet (ITU 2014). Globally, close to 44 % of households have Internet access at home.

The economic and societal changes that result from these developments are considerable and have attracted the attention from academics and policymakers. Along with the advancement of the Internet, researchers have increasingly shown interest in the worldwide web, not only as a research subject but also as a potential data source. This interest has not been limited to a single field but stretches out across many different domains. In the discipline of economics, labour economics has been identified as a field for which web-based data are particularly promising. In their seminal papers, Kuhn and Skuterud (2004) and Askitas and Zimmermann (2009, 2015) argue that web data could be very valuable to do research on the labour market. The Internet has transformed the labour market in many ways. Autor (2001) identifies three dimensions through which the web has affected the labour market: how firms and workers search for each other (e.g. employers look for “passive” candidates, on-the-job search), how

labour services are delivered (e.g. skills required online) and how local labour demand is shaped (e.g. impact of e-commerce). Technological change indeed has greatly affected job search, application, selection and matching (Carnevale et al. 2014; Kuhn 2014; Kuhn and Mansour 2014). A recent survey suggests that 80 % of the HR teams use social media when hiring new employees (Konetic 2014).

In this article, we aim to provide an overview of novel web-based data sources and their potential as sources for labour market research. In this way, we aspire to contribute to the literature on the link between the Internet and the labour market. There already is a large and quickly growing body of literature on the use of data from job portals for labour market analysis. Indeed, in this very issue of *IZA JOLE*, there is a paper by Kureková et al. (2015) on the methodological issues involved in using vacancies from job portals (and also, to a lesser degree, the WageIndicator survey). For this reason, we decided to survey other sources of data. While these sources have not been ignored by researchers, their utilisation has been lower, and there is also a lack of survey literature bringing together all strands. We therefore describe alternative data sources covering the supply, demand or both sides of the labour market and explain why we are convinced of the potential of these data as excellent platforms and sources for labour market research.

The remainder of this paper is organised as follows. In Section 2, a brief overview of web-based research is presented. The section first discusses how the Internet became more prominent as a research platform and data source. To this end, the section includes an analysis of early web-based work. As a second step, a series of relatively new web-based data sources is described. These sources comprise data extracted from online job boards and intermediaries, Google Trends, social networks and surveys. Sections 3, 4 and 5 present more details on the new data sources. In each of these sections, first, the data source is described and then empirical applications that make use of the source are surveyed. We consider not only articles published in scientific journals but also prospective work and other contributions. We examine how the data source is, or could be, used for labour market analysis. Section 3 introduces Google Trends. In Section 4, we focus on three social networks: LinkedIn, Facebook and Twitter. Section 5 deals with web-based surveys and Glassdoor in particular. Section 6 of the article summarises our findings and outlines some paths for future research.

2 Using the Internet as a data source for labour market analysis

With the growth of the Internet and its increasing coverage across countries and socio-demographic groups, the web itself has become an interesting research subject and a tool for data compilation. Hooley et al. (2012) therefore distinguish between studies that cover the Internet and studies that make use of the Internet to conduct research but note that these two domains actually are strongly connected. The earliest studies were mostly of the first type, with research that focused on the social dimension of the web (Freeman 1984; Finholt and Sproull 1990). Shortly after these first studies, work that used the Internet to do analyses emerged (Kiesler and Sproull 1986; Kehoe and Pitkow 1996; Foster 1994). As the field expanded, new approaches and data collection methods were developed, which often were strongly embedded in the existing methodological framework and enriched with insights from technological progress.

Online research can take many forms. In their book, Hooley et al. (2012) concentrate on four types: surveys, interviews and focus groups, ethnographies, and experiments. *Surveys* were among the first research activities performed online. In fact, the first recorded email survey was done in 1986 (Kiesler and Sproull 1986) and the first recorded web survey in 1994 (Kehoe and Pitkow 1996). Compared to traditional paper-and-pencil methods, online surveys have the advantage to be flexible, fast, cheap and easy to set up. Data can be collected from a larger and more diverse sample, which has a positive impact on data accuracy. At the same time, the anonymity of the respondents is ensured. Web surveys are also easier to analyse than traditional surveys. Disadvantages of online surveys include sample bias, measurement error, non-response and dropout, and other technical and ethical issues. For the USA, there are two Internet panel surveys that we wanted to bring to the reader's attention: RAND's American Life Panel (ALP) (6000 participants) (<https://alpdata.rand.org/index.php?page=main>) and the Understanding America Study (UAS) panel (2500 participants, CESR, University of Southern California, http://static.usc.edu/sites/files/UAS_Brochure.pdf). Both panels are representative for the US population of ages 18 and up. *Online interviews and focus groups* have developed more slowly. This research mainly concerns asynchronous email interviews, although limited work does consider synchronous interviews and focus groups (O'Connor et al. 2008; Foster 1994). Online interviews and focus groups are more flexible and cost- and time-effective. They do, however, require technical competence of the participants, shift the power balance in their favour and constrain the researcher from observing any nonverbal communication. *Online ethnographers* examine how humans live and interact online; research commonly deals with social interactions on online communities, networks, gaming, discussion groups, bulletin boards and social media (Hookway 2008; Thelwall 2008). Lastly, *online experiments* have been used extensively beyond the boundaries of the social sciences. Examples are Krantz and Dalal (2000), Musch and Reips (2000), Horton et al. (2011), Pallais (2014), Pallais and Sands (Pallais and Sands: Why the referential treatment? Evidence from field experiments on referrals, forthcoming) and Horton (Horton: The effects of algorithmic labor market recommendations: evidence from a field experiment, forthcoming). Pallais (2014) performs an experiment on oDesk—an online marketplace—to test the hypothesis that young workers have a higher probability of being unemployed than older workers because of barriers to labour market entry. oDesk is also used by Pallais and Sands (Pallais and Sands: Why the referential treatment? Evidence from field experiments on referrals, forthcoming), to examine why referred workers have a higher chance to be hired, and by Horton (Horton: The effects of algorithmic labor market recommendations: evidence from a field experiment, forthcoming), to investigate the role of recommendations (which appear to raise the probability of hiring). Horton et al. (2011) perform a set of experiments on Amazon Mechanical Turk (MTurk). More specifically, the authors replicate three classic experiments online and prove that such experiments are valid and beneficial to researchers. Besides oDesk and MTurk, there is another platform on which researchers can do online experiments: TESS (Time-sharing Experiments for the Social Sciences, <http://www.tessexperiments.org/>). Researchers can submit proposals for experiments, which are peer-reviewed. When a proposal is approved, TESS does the experiment free of charge on a representative sample of US-based adults. Horton et al. (2011)

demonstrate that online experiments are more flexible, faster, cheaper and easier to conduct than real-life experiments, and they allow for a broader scope. The sample of participants that one can reach also is more diverse and larger. These features improve the quality of the study.

In recent years, the analysis of labour demand, supply and matching became increasingly based on online data (Askatas and Zimmermann 2009, 2015; Kuhn 2014; Kuhn and Mansour 2014). Internet data can contribute to our understanding of the labour market by filling the gaps that currently exist in the literature (Kureková et al. 2015). A first web data source that could be particularly useful in this case is a dataset extracted from *online job portals*. Many job boards are not limited to vacancies but also collect CVs and resumes and offer wage comparisons, employer evaluations and career advice. We do not cover portals in detail in this survey, but refer to Kureková et al. (2015). However, we do present some examples of recent contributions in this field. Marinescu (2015) uses CareerBuilder to investigate geographical mismatch in the USA, while Marinescu and Rathelot (2015) rely on this portal to study unemployment insurance. With a sample of vacancies from Burning Glass, Hershbein and Kahn (2015) provide evidence for upskilling: employers require more in areas with higher unemployment rates. Using a similar approach and dataset, Sasser Modestino et al. (2014) examine changes in employers during the Great Recession (2007–2012) and the subsequent recovery (2010–2012). In “bad” labour markets, employers are more demanding, in terms of both education and experience. Kudlyak et al. (2012) rely on matched applicant-vacancy data from SnagAJob to assess how job seekers’ behaviour changes during their job search. In another contribution based on SnagAJob, Faberman and Kudlyak (2014) report that job seekers’ search effort declines with search duration. Brenčić (2014) looks into information acquisition through portals and resume databases. Agrawal and Tambe (2014) use online resumes to track workers’ career paths, focusing on workers previously employed in firms acquired through leveraged buyouts.

Note that one could make the distinction between online job portals and other *online labour market intermediaries* like oDesk, Amazon Mechanical Turk (MTurk), CoContest and TaskRabbit. Most work on these market intermediaries has focused on MTurk, an online marketplace through which employers offer tasks that require human intelligence (i.e. that computers are unable to do). Horton (2011), for example, examines the fairness of MTurk employers. Buhrmester et al. (2011) evaluate MTurk’s potential as a data source in the field of psychology and the social sciences. Maselli and Fabo (2015) use CoContest to evaluate the potential income designers can earn via the platform. Other studies cover oDesk (Ghani et al. 2012; Pallais 2014). Research seems to be more limited for other online market intermediaries.

Another potential data source is data obtained from *online web surveys*. Surveys such as Glassdoor and WageIndicator collect data on salaries, working conditions or company reviews and often reach a large number of participants. Online surveys have proven their value as research platforms in the past and should therefore be considered for labour market analysis too. However, participation is voluntary; information on specific questions may be missing or only a part of the population is involved. This can result in low response rates (Benfield and Szlemko 2006). Other issues are the (computer) literacy of participants and the “rapport” with the respondents. In this article, we focus on Glassdoor. We do not present more details on WageIndicator as

this source was already discussed, to some extent, by Kureková et al. (2015). Other interesting online sources are *Google Trends* and *social networking sites*. These sources are sometimes overlooked but do hold valuable information. Google Trends serves as an excellent tool to detect new trends. In addition, job opportunities are often posted on companies' Facebook pages or Twitter profiles, while job candidates can create a LinkedIn profile to connect with potential employers. Social networking sites further can reduce search frictions.

Although several of the *advantages and limitations of traditional and web-based data sources* have already been pointed out above, we devote more attention to this topic in this paragraph. Labour market data are typically initiated and made available by governments or international organisations. These data generally are considered to be more accurate, better structured and more complete than data from other sources. Traditional sources such as the CPS or Census further have the advantage to be based on a randomly selected sample of the population (representative). A disadvantage of traditional sources, however, is that statistics usually are distributed with a lag. Some databases are not updated regularly or instead frequently revised. Data are often assembled from administrative sources or surveys, which could result in a small sample size or data unavailability for sectors or regions with a limited coverage (Wright 2012). For many less-developed countries, labour market data are simply absent or of a low quality (Shapiro 2014).

In contrast, online data are available in real time (no lags, no revisions) and thus are an excellent tool to detect current trends and dynamics. Web data allow researchers to fill gaps where traditional data sources are absent or weak (e.g. due to low coverage or limited quality) (Autor, 2001; Shapiro 2014). Online, researchers can assemble large, diverse and potentially more representative datasets. As an increasing part of the population is active online, sampling may become unnecessary in the future (Askatas and Zimmermann 2015). Other advantages are that data collection and analysis are easy, fast, flexible and relatively inexpensive and that logistical issues associated with traditional sources can be avoided (e.g. tedious data entry prone to errors) (Benfield and Szlemko 2006). Another advantage of web data is that they facilitate research on self-employment, which is a key driver for entrepreneurship and job creation and may become even more relevant in the labour market of the future. This is also highlighted in a recent OECD report: the web is a catalyst for business innovation, across all sectors of the economy, but it is not easy to study these dynamics (OECD 2014). At the same time, self-employment is difficult to measure on the basis of traditional data sources (because data are lagged and the definition of self-employment differs across data providers; Fairlie and Robb 2009). Many datasets cover information on either the owner or the business, but not on both. Web data can be a solution, as users can indicate in their profiles whether they are self-employed and information on the business can be found online. Another advantage is that online portals and social networks facilitate on-the-job search. Stevenson (2009) reports that the 77 % of the people that use the web for job search are employed. These employed job seekers are more likely to leave their current job, transition from employment to employment and have better negotiation positions vis-à-vis their employer. Stevenson (2006) already concluded that the Internet leads to better matches for the employed (e.g. higher wage growth when changing jobs). Kuhn and Skuterud (2004) assess which unemployed workers look for a

job online and whether they became reemployed more quickly. Online job search seems ineffective in reducing unemployment durations. However, an alternative explanation is negative selection on unobservable variables (e.g. poor networks).

Web data, however, are also characterised by some limitations (Benfield and Szlemko 2006; Shapiro 2014; Carnevale et al. 2014). There are ethical and technical issues (e.g. privacy, familiarity with a computer). Furthermore, online data and vacancy data in particular are incomplete by their very nature, and vacancies only represent a small part of labour demand (Wright 2012). Vacancy data may be biased towards specific sectors, regions or applicants. For example, Carnevale et al. (2014) find a bias towards high-skilled, white-collar workers and STEM positions in the vacancies published online. Autor (2001) further points to adverse selection of job applicants (applying for a job is cheap and easy; therefore, job seekers apply for many jobs, for which they could be over- or underqualified). His paper also has a segment on geography and inequality. Although job search is cheaper, markets are integrated and labour services can be delivered online; this is not necessarily beneficial to all groups. Selection is an important issue for web data (see Carnevale et al. 2014; Kearney and Levine 2015). Websites or online platforms could attract specific users, which affects data representativeness. For this reason, we describe the demographics of the social networks below. Kureková et al. (2015) examine the representativeness of vacancy data. These data are not missing at random; this results from sampling. The quality, consistency, accuracy and volatility of web data should be examined (Carnevale et al. 2014).

In this paper, we concentrate on three web-based data sources: Google Trends, social networks and surveys. Through these sources, researchers can identify new trends, find job offers (title, description and requirements, labour demand) and detail user profiles (describing education and skills, experience, labour supply). In each case, we first describe the source (using information available on the websites). Then, we discuss which information can be collected from the source, with a focus on labour-related features. We continue with an overview of applications that either cover the source itself or use it for research on other topics. Because the field is rapidly developing, we do not limit our analysis to articles published in academic journals but also consider works-in-progress and other contributions.

3 Google Trends

Google Trends (<https://www.google.com/trends/>) was launched by Google in 2006. The service is based on Google Search and analyses a percentage of these searches. In particular, Google Trends allows users to check how often search terms, or combinations thereof, are entered relative to the total number of searches done (by region, across time). When multiple search terms are entered, their relative popularity is compared. More precisely, on the web page, there is a search button which allows users to type in their search term of interest. Google Trends displays the interest in this search term over time (on the rise or declining) and by region (global, national or regional level) and related searches (spilt out into topics and queries). Regional searches are possible because a user's location can be identified via his/her IP address. Google Trends provides a normalised volume of queries for a specified term which is presented in a graph. Spikes in the graph are associated with news headlines, when possible. However, to protect the privacy of its users, Google does not publish results when there are

insufficient observations. Trends data exclude searches made by very few people, duplicate searches and special characters. Searches and search outcomes can be manipulated by Google, and one has to keep in mind that it is a company that develops content, sells advertisements and promotes its sub-brands (e.g. Yahoo Finance). This could particularly affect small- and medium-sized firms, which see their search ranks worsen and lose significant amounts of traffic. Moreover, organisations and large companies are able to manipulate search results too, to maximise traffic and exposure.

Google Trends further has lists of “hot searches” and “hot topics”. The former tracks the searches that are rising the fastest at the moment, while the latter captures trending terms in the news and on social media. Google Trends also features top stories that can be filtered by region and topic (e.g. business or health). In 2008, Google launched Google Insights, an extension to Google Trends, which allowed users to track words and phrases entered into search boxes, analyse results and structure data. The tool was integrated into Google Trends in 2012. All data can be downloaded in the .csv format. Because Google Trends enables users to verify which (combinations of) search terms are on the rise, the platform provides us with more insight into the type of positions that job seekers are looking for, the types of skills that are high on demand, the industries that are booming and many other aspects. Google Trends presents information on labour demand and supply.

Nevertheless, there are some caveats to Google Trends. Because only a sample of searches is used and searches for which there are insufficient observations are excluded, Google Trends data could be affected by *sample bias* (in small samples, only random draws with enough observations are shown) (Kearney and Levine 2015). A second issue is *sampling variability* (problematic for standard error calculations when data are treated as fixed rather than random variables). To address these issues, the authors repeat their searches on Google Trends several times and calculate the average of the indices (to reduce the sampling variability). Another shortcoming of Google Trends data is that demographic information is not available. As temporal and geographic variations are sources of variation that labour economists typically rely on, the above issues are important to account for.

3.1 Empirical applications

Google Trends serves as a data source in a large number of contributions. One of the most well-known applications is *Google Flu Trends*. In an influential article published in *Nature*, Ginsberg et al. (2009) explain how Google Trends can be used to improve the early detection of seasonal influenza by monitoring search engines like Google. This approach seems to work well because of the high correlation between the percentage of doctor visits and the relative frequency of specific queries on Google. The authors can predict weekly influenza activity in the USA (with a time lag of one day). Other studies have used Google Trends to examine health-related topics as well (e.g. papers that extend or improve Ginsberg’s method or focus on other diseases).

The strand of literature that relies on Google Trends for *forecasting and now-casting* is extensive too. Choi and Varian (2012) show that Google Trends is a useful tool for predicting the “present” (in the form of subsequent data releases, i.e. the short-term future) due to the correlation between queries and economic indicators. They illustrate this result with the examples of travel, retail sales, home sales and automotive sales. Carriere-Swallow and Labbé (2013) work on a related topic, focusing on automobile

purchases in Chile. Preis et al. (2013) relate Google queries to stock market dynamics and show that losses are often preceded by a growing volume of specific stock market search terms. In a recent publication, Chen et al. (2015) evaluate to what extent Google search queries can be used to “now-cast” business cycle turning points during 2007–2008. Schmidt and Vossen (2012) use Google Trends to account for special events in economic forecasting. In another paper, Preis et al. (2012) link queries, and whether they refer to the future or past, to countries’ economic success. Constant and Zimmermann (2008) use Google Trends to measure economic and political activities, while Askitas and Zimmermann (2009) and Choi and Varian (2009) use it to predict unemployment.

Other studies use Google Trends for *behavioural analysis*. In a series of articles, Stephens-Davidowitz uses Google Trends to explore topics such as racism, religion, prejudice and health.¹ Rode and Shukla (2014) use a Google search query to examine racial differences in labour market outcomes in the USA. The authors provide evidence for racial prejudice: in metropolitan areas with more racially charged searches, black-white gaps in annual income, hourly wage and annual hours worked are wider. This result appears to be somewhat stronger for less educated workers. Another relevant paper is Kearney and Levine (2015), who combine data from Google Trends, Twitter and two other sources to examine how media images affect adolescents’ attitudes and outcomes for the case of MTV’s reality TV show “16 and Pregnant”. Interestingly, the TV series appeared to raise the amount of Google searches and tweets on birth control and abortion. Moreover, the show is associated with a 5.7 % reduction in teen births in the 18 months after its introduction. Kearney and Levine (2015) do point to potential endogeneity: the interest in “16 and Pregnant” is likely higher in areas where the teen birth rate is higher or where it is rising or falling more slowly. While the former can be tackled via geography-fixed effects, the latter is addressed with an instrumental variables (IV) strategy in which ratings are instrumented with ratings of any MTV show broadcasted during the same time in the previous period.

4 Social networking sites

In the last few years, many studies have appeared that concern social networking websites. Social networks commonly have large user bases comprising individuals, firms and other organisations. User profiles often contain detailed information about current employment, experience, educational attainment and other qualifications (labour supply). Information about individuals’ behaviour and preferences can easily be obtained from these sites. In addition, firms and organisations often have profiles on these networks as well, through which they can interact with current employees and interested job applicants and share vacancies (labour demand). Information is often publicly available. A vast majority of companies use social networks to look for candidates. Social networks can reduce search frictions. Acquisti and Fong (2015a, b) investigate how information available on job applicants’ profiles affects their interview invitation rates. A third of employers searched online for information on the candidates. Results also suggest that employers in the Republican parts of the USA have a significant bias against Muslim candidates and in favour of Christian applicants.

4.1 LinkedIn

Of the social networks discussed in this article, LinkedIn (www.linkedin.com) is the most obvious candidate to serve as a data source for labour market analysis because of

its focus on professional networks. By enabling users to set up profiles, connect with other users and find or list job openings, LinkedIn aims to “connect the world’s professionals to make them more productive and successful” and to “transform the ways companies hire, market and sell”. LinkedIn was founded in 2002 and became available online in the spring of 2003. About 4500 users signed up during the first month. Since then, LinkedIn has developed into the largest global online professional network, connecting over 364 million users (individuals and organisations) in over 200 countries and territories. In the first quarter of 2015, over 75 % of LinkedIn’s new users were not US-based. LinkedIn currently supports 24 languages. In Europe, LinkedIn has more than 89 million users. Two new users sign up every second. In the USA, 28 % of the adult Internet users use LinkedIn (Duggan et al. 2014). The website is particularly popular among college graduates, higher-income households and the employed. LinkedIn is the only platform where people aged 30–64 are more likely to be users than those aged 18–29 (Duggan et al. 2014).

Because LinkedIn holds profiles of companies, job seekers and recruiters, it is an interesting platform to analyse labour market dynamics. Firms can use LinkedIn to set up a “Company Page” on which they can post job opportunities or create dedicated “Career Pages” for this purpose. LinkedIn users can go through company pages to find job opportunities, use the general search options that LinkedIn offers, or connect with recruiters. Since 2011, users can even apply for jobs directly by using their profile as a resume when they click on the “Apply with LinkedIn” button on job listing pages. Another feature is the option to establish or become a member of an interest group (e.g. the group “Java Developers”). User profiles of employees and job seekers further comprise valuable information on their education level and skills. In the fall of 2012, LinkedIn added a feature through which users can comment on each other’s profiles and endorse each other’s skills. From this, it is clear that LinkedIn reduces search frictions, as job seekers (employed or unemployed) can easily find positions in their organisation or sector of interest, and employers can easily browse through a large set of profiles to find their ideal candidate (also passive candidates are available). LinkedIn therefore is a good starting point for labour market analysis. For the 94 % of job candidates (two thirds of recruiters), LinkedIn is the most important social network for job hunting (candidate sourcing) (Right Management 2015).

4.1.1 Empirical applications

Currently, the majority of the work on LinkedIn covers the platform itself. For example, there are several studies that examine how LinkedIn can be used in selection, recruitment or other business processes (Caers and Castelyns 2011; Bonsòn and Bednárová 2013; Rangel 2014; Zide et al. 2014; taking the perspective of job seekers and employers). Jarrow et al. (2011) discuss LinkedIn’s stock price. On the other hand, there are only a few contributions in which LinkedIn serves as a data source. An interesting example is State et al. (2014), who examine migration to the USA among professional workers of different education levels with a database of geo-located career histories from LinkedIn. Boucher and Renault (2015) use a dataset compiled by hiQ Labs, which comprises many job titles and LinkedIn profile summaries, to construct a job classification. Gee (2015) takes vacancies published on LinkedIn to do an online experiment that covers 2.3 million job seekers. She demonstrates that reporting the number of previous

applications increases the likelihood of application, especially among female job seekers. Tambe (2014) examines how labour market factors shape early returns to investment in big data technologies such as Hadoop and Apache Pig on the basis of LinkedIn. Other studies, of which several are related to the analysis of jobs and skills, can be found on the LinkedIn website (<http://data.linkedin.com/publications>).

A final set of applications worth mentioning are embedded in LinkedIn's "Economic Graph" challenge. This challenge was launched in 2012 and sets out to create an "economic graph" within a decade (i.e. to digitally map the world economy). For this challenge, teams were invited to propose how they would use LinkedIn data to research a range of topics related to the job market. For the challenge, 11 teams were selected (see <http://economicgraphchallenge.linkedin.com/>).

4.2 Facebook

Facebook (www.facebook.com) is a well-known online social network that was launched in 2004. On Facebook, users can set up a profile on which they can post messages, photos and videos; update their status; and use other features. User profiles can be public or private. On their profile, users can share their employment status or occupation, education level, family situation, skills, interests and hobbies, and other information (labour supply). Users can connect with others by becoming "friends", in which case they receive notifications when a friend updates his/her profile (via the "news feed"), and are able to send messages or chat. Since 2004, users have the possibility to create or become a member of a (private) Facebook group. As of March 31, 2015, Facebook had 1.44 billion monthly active users. The average number of daily active users during March 2015 was 936 million. About 83 % of the daily active users do not reside in the USA or Canada. This number reveals that Facebook has an extremely large global user base. The website has a much larger network of users than any of the other social networks discussed. Duggan et al. (2014) find that Facebook is the most popular social network: it is used by 71 % of online American adults. Women are particularly likely to use Facebook compared to men (66 % of men; 77 % of women have a profile).

Companies, recruiters and other organisations can also create a Facebook profile. This possibility was introduced in 2007 and is known as "Facebook Pages". Facebook Pages are public profiles held by celebrities, businesses, organisations and brands. On their profile, firms can present themselves to users, interact with them, introduce new products and post job vacancies (labour demand). On this note, there are also many job portals that have their own Facebook page through which they look for new talent, share job opportunities and offer career advice (labour demand, matching). Some examples are Indeed and Monster. In 2012, Facebook launched a job board, the "Social Jobs Partnership Application", which is the result of a collaboration with the Department of Labor, the National Association of Colleges and Employers, the Direct Employers Association, and the National Association of State Workforce Agencies. The introduction of the job board was motivated by the outcome of a survey performed by the National Association of Colleges and Employers (NACE), which targeted 530 employers and recruiters in the spring of 2012.² This survey revealed that (i) 50 % of the employers used Facebook in the hiring process, (ii) almost 90 % of the companies claimed that recruiting via Facebook is more cost-effective than through other channels

and (iii) especially networking and referrals are key features to find new employees (e.g. engaging in a network with a candidate who “liked” the company’s Facebook page). Via the application, recruiters can post job opportunities which can be sorted by location, industry and skills. When the application was first launched, it combined offers from BranchOut, Direct Employers Association, Jobvite, Monster and Work4 Labs. The goal of the project is to support finding and sharing jobs via Facebook. A final option is to exploit Facebook Graph Search (Headworth 2014). Note that Facebook lowers search frictions (users can easily connect with an employer of interest or join a group; employers can browse through profiles and discover interesting candidates more easily via groups).

4.2.1 Empirical applications

There are many studies on the topic of Facebook, but at first glance, only a few studies exist that use Facebook data to analyse labour market dynamics. Overall, the literature is extensive and covers many fields. Some examples are health-related issues, network analysis and education. Wilson et al. (2012) examined the research on Facebook in the social sciences. Their analysis is based on 412 articles published up until the first of January in 2012. Wilson et al. (2012) classified these articles into five categories that reflect five research questions: (i) Who is using Facebook, and what are users doing while on Facebook? (ii) Why do people use Facebook? (iii) How are people presenting themselves on Facebook? (iv) How is Facebook affecting relationships among groups and individuals? and (v) Why are people disclosing personal information on Facebook despite potential risks? The bulk of the articles addressed the fourth question. These questions are also particularly interesting to get more insight into labour market dynamics, recruitment and selection. Some work has been done on these issues, from the perspective of both job seekers and that of employers and recruiters. Research confirms that Facebook is a popular tool to screen applicants (Karl et al. 2010a, b; Kluemper and Rosen 2009). This, however, implies that employers and recruiters can also evaluate job applicants on other criteria, such as their gender or race or “inappropriate” material on their profile (Kluemper and Rosen 2009; Bohnert and Ross 2010). Kelkar and Kulkarni (2013) discuss the usefulness of Facebook from the labour supply side. While pointing out Facebook’s possible advantages for job seekers, Kelkar and Kulkarni (2013) do find that only a small number of members actually use Facebook to look for jobs and for networking. The authors criticise the Social Jobs Application, which tends to yield inconsistent and ineffective results and does not appear to add much to what is already available on job portals. One can conclude that Facebook has the potential to serve as a source for recruitment and selection, but this has not yet been fully exploited. Two other examples of work on Facebook are Gee et al. (2015a) (on the link between job transmission and Facebook) and Gee et al. (2015b) (on weak ties). Facebook appears to be an interesting and valuable data source that has been used for labour market research. Wilson et al. (2012), however, do note that data crawling techniques are becoming less effective because of stricter privacy settings.

4.3 Twitter

Twitter (www.twitter.com) is a micro-blogging website through which users can send and read short messages (of no more than 140 characters) called “tweets”. Whereas

these messages can be read by anyone, only registered users can send tweets. Users are “connected” to each other when they follow or are followed by other users. Moreover, messages sent out by one user can be re-tweeted by others. Tweets can cover any topic and can be grouped by topic or via hash tags. Twitter also tracks “trending topics” (global and regional, via an algorithm that accounts for the location and interests of users). Twitter was launched in 2006 and has grown substantially ever since. About 500 million tweets are sent each day, most of which are accessible for public view as tweets are publicly visible by default. As of March 31, 2015, Twitter has 302 million monthly active users. Twitter supports 33 languages. 77 % of Twitter’s accounts are held outside the USA, which illustrate the website’s global outreach. Twitter’s mission is to “give everyone the power to create and share ideas and information instantly, without barriers”. This idea is put into practice through following and followers, re-tweeting and the public nature of the service.

Twitter can be used by job seekers, employers or companies, and recruiters and is therefore a useful tool in the analysis of the labour market. Job seekers can use Twitter to get more information on companies, discover job openings and find out more about the qualifications required by following these firms, and they can interact with current employees (labour supply). Moreover, job seekers can also find vacancies via the general “search” function. Companies, on the other hand, cannot only turn to Twitter for marketing purposes and sales; they may also use the website to tweet about vacancies and look for employees (labour demand). Such tweets often consist of a job title, a brief description of the position and a link to a webpage with more information. Companies can also use Twitter to strengthen their profile and spread information to clients and (potential) employees. A further option that companies have is to work with third parties, for example “Tweet My Jobs”, to share their vacancies (Schawbel 2012). Recruiters can also rely on Twitter to find talent, through their own account, by becoming a member of groups or by following other users. Another feature that is particularly useful in this regard is the option to embed a web link in the company’s or recruiter’s Twitter profile page, which directs job seekers to their website. Furthermore, many job portals, such as CareerBuilder, Indeed, Simply Hired and Monster, have their own Twitter accounts through which they share job listings and offer job search and career advice. As Twitter messages are fairly short, employers generally will not use Twitter as their main recruiting tool, but rather as a part of a whole recruitment strategy (Larsen 2011). Duggan et al. (2014) suggest that 23 % of the US adults online use Twitter. The site is particularly popular among those younger than 50 and the college educated.

4.3.1 Empirical applications

Even though Twitter was only introduced in July 2006, there already exists an extensive literature on the micro-blogging website. Academics and other researchers from many different fields have taken an interest in Twitter, which resulted in a high number of studies on a variety of topics (including but not limited to the field of computer and information sciences, physics, linguistics and economics). This interest is motivated by the scale of the database (the large amount of users and tweets) and its time dimension. In a recent article, Williams et al. (2013) focus on Twitter-based research to identify and classify the types of studies that are being done. From their review of the literature,

they conclude that generally the following four elements are covered: the message, user, technology and concept. Other elements that are considered in some papers are the domain (e.g. education, health, business and security), data and research method. For a sample of 575 papers on Twitter published between 2007 and 2011, the authors show that most research deals with the content of Tweets followed by work on the users (together they represent 80 % of the papers). Some more specific examples of recent work on Twitter are the paper by Achrekar et al. (2011) on the prediction of flu trends, by Murth (2015) on elections and by Yu and Wang (2015) on sentiments in tweets during the World Cup of 2014. Despite the large number of topics covered, research on the relation between Twitter and labour market dynamics and the use of Twitter data to identify or classify occupations and skills is limited. Kearney and Levine (2015) use Google Trends and Twitter but find that the latter is harder to access than the former. Historical data cannot be accessed nor is there information on the frequency of tweets on the site. A library of past tweets can be obtained, but this library is difficult to manipulate due to its size and format. Data can be obtained through third-party vendors. Other limitations are that geographical information is difficult to obtain (no IP address) and that information on demographics is unavailable. Only a few papers seem to use Twitter for the behavioural analysis.

In February 2014, Twitter launched a pilot project called “Twitter Data Grants”. The project consisted of a call for proposals for research institutions to collaborate with Twitter staff and obtain historical and public data (see <https://blog.twitter.com/2014/twitter-datagrants-selections>). Six teams were selected. Before the Data Grant programme was launched, free access to data was limited to the last seven days. The difficulty in accessing Twitter data also spurred several papers on the topic. Kwak et al. (2010), for example, describe crawling the platform to examine its topological characteristics. Twitter, in contrast to other platforms, is mainly relevant for labour demand as user profiles are fairly limited. One option, however, is that job seekers can spread their resume via Twitter in the hope to attract companies’ and recruiters’ attention.

5 Online web surveys—Glassdoor

Web-based surveys, such as WageIndicator and Glassdoor, have already been introduced in the first sections of this article. In this section, we present Glassdoor (www.glassdoor.com), a popular career community website that was first launched in 2008. Although the website operates as a job board, Glassdoor goes beyond traditional job portals as the company also targets employers, recruiters, career centres and libraries. The company is US-based and managed to develop into one of the largest websites there, and its user base is rapidly expanding towards a more global audience. In fact, Glassdoor currently has over 30 million members in over 190 countries worldwide. The Glassdoor website is organised into *six categories*. The first four categories are oriented towards job seekers and employees (who are invited to post their resume), while the last two target firms and recruitment agencies. The first out of the six categories is “*jobs*”. Similarly to the traditional portals, Glassdoor functions as a job board on which millions of vacancies are listed. Job seekers can browse through these job openings to discover which firms are hiring and which positions are available. On the website, job seekers have the opportunity to look for positions by location, title, occupation or keywords. Glassdoor also presents them with a list of

popular and related searches. The second category is “*companies*”. Job seekers can look for firms in a specific location and are redirected to a detailed company page when they select a firm. On a company page, job seekers can upload or read company reviews, find CEO approval ratings, discover the salaries that the firm offers for specific roles, ask questions to current or former employees, find office photos and other information, read interview tips and so on.³ One of the innovative features of Glassdoor is that all this information is provided by firms’ former or current employees. The website covers over 400,000 firms worldwide. To write a company review, users need to have an account. Users are only allowed to write one review per firm worked at per year. Reviews are published anonymously. Firms have no information on the identity of the employee that posted the review and cannot manipulate or remove reviews. Before a review is published, it has to be approved by Glassdoor. Reviews that do not meet the guidelines are not published (e.g. reviews should be balanced, cannot disclose trade secrets). Employers can respond to reviews and flag reviews that do not meet the guidelines, are inappropriate or fraudulent. The third category on Glassdoor is “*salaries*”, which can be viewed for specific positions. The fourth category targeted towards job seekers is “*interview*”. Job seekers can find information on the interview style, level of difficulty and sample questions. Since much of the information provided on the Glassdoor is provided by former and current employees, the site clearly has a “web survey” dimension.

Glassdoor also aims to attract employers. The two remaining categories are “*employers*” and “*post a job*”. On the website, employers can set up an “Enhanced Employer profile” on which they can share their history and discuss their activities, promote vacancies or post a link to their Facebook account. Glassdoor offers employers recruiting and branding solutions via “*Glassdoor for Employers*”. Employer branding implies that employers can track the candidates that are looking into their company and the reputation that their firm has on the website. Job advertising involves listing a single or all available jobs, targeting job seekers and using analytics to improve matching. Glassdoor has over 2000 clients and partners for which they do employer branding promotion, job advertising (especially to candidates who may not have been aware of the position) or both. Glassdoor also offers solution for companies in specific sectors, such as tech, telecom and SaaS companies; banking and finance companies; and consulting firms. Finally, Glassdoor also reaches out to *career centres and libraries*, which can offer unlimited access to the website without having to post information.

Glassdoor is an excellent source for the analysis of labour market dynamics as the website collects a very large amount of real-time information on several dimensions. The database comprises millions of vacancies, company reviews, CEO approval rates, and salary and benefit reports. For employers, the website provides analytics of search behaviour, targeted job advertising and employer branding solutions. According to Glassdoor, users of all ages and backgrounds are on the website. Moreover, the average company rating is 3.4 (on a 1–5 scale). Of the employees, 70 % indicate “OK” or “satisfied” when asked about their employer. A survey of over 4600 Americans revealed that 2201 of them use Glassdoor (Osterhaus 2014). About 50 % consult Glassdoor at the start of their job search to identify top employers. Osterhaus (2014) reports that especially job seekers aged 55–64 and those earning between \$25,000 and \$49,999 annually are active on the site. Most respondents live in urban or suburban areas. Glassdoor

seems to attract users of all ages and income levels. As Glassdoor engages more users and companies, representativeness may increase further.

5.1 Empirical applications

There are numerous studies that use data extracted from Glassdoor to study the labour market. In many cases, data on wages and other information posted on Glassdoor are used to complement a more extensive analysis. An example of this is a study on the opportunities of females in IT by Thiele (2014), who uses Glassdoor to find the median wages for several IT professions (including software trainers, system programmers and network technicians). Massimino et al. (2015) use Glassdoor to find employee satisfaction rates. Other publications refer to Glassdoor as a source of information on interview techniques, skill requirements and other useful tips. Kaplan (2014), for example, refers to Glassdoor in her article on how to prepare for job interviews. Another example is Lauby (2013), who considers Glassdoor's growing popularity as an indication of the rising importance of employer and career branding. Chandra (2012) uses Glassdoor data on the work-life balance ratings across firms to compare Eastern and Western perspectives on work-life balance. American and European companies rank higher than Indian companies as they pay substantial attention to this issue.

Glassdoor also has its own research team of economists and data scientists: Glassdoor Economic Research (<http://www.glassdoor.com/research/>). Because Glassdoor gathers an enormous amount of real-time data on different labour market aspects, the website is a unique and rich data source. These data have been difficult to collect in the past, especially on such a large scale. Recent studies published on the site deal with hiring times, jobs affected by the introduction of a minimum wage, and the link between salary and employee satisfaction. The site will offer downloadable data in the future and supports "Job Tools". The latter are two interactive map-based tools that can be used by job seekers nationwide. The first tool is the "Job Explorer", through which job seekers can find where their skills are in demand. The second tool is the "Apprenticeship Finder" that can be used to explore apprenticeship and career opportunities. With the second tool, job seekers can view on a map in which states there are apprenticeship opportunities (high to low). For example, in the state of California, 5704 apprenticeship jobs were available at the end of July 2015, most of which in the areas around San Francisco, Sacramento, San Diego, Los Angeles and Silicon Valley. With the first tool, the job explorer, job seekers can select a job category or type in a keyword to view on a map where these skills or jobs are concentrated. For California, this resulted in 94 "programming skills" jobs and 44,556 "programmer" jobs. The tool indicates "top cities" for the jobs and lists "other jobs you should consider" (for "programmer", these jobs are senior programmer, consultant, software engineer, programmer analysts).

6 Conclusions

This study comprises an overview of web-based data sources that can be used in the analysis of the labour market. Our work is motivated by the substantial advancement of the Internet during the last decade and the growing body of literature it is accompanied by. Several studies have pointed to the value of Internet data for the field of labour economics (Askatas and Zimmermann 2009, 2015). Most work in the field, however,

relies on vacancy or CV data extracted from online job boards or online labour market intermediaries. Nevertheless, there are many other web-based data sources that could be valuable for labour market analysis. These sources have not been overlooked completely, but their utilisation is more limited. In this study, we have therefore explored the potential of Google Trends, social networks and Glassdoor.

From our survey, it is also clear that there are many possible avenues for future research. Some first examples are the use of web data for research on self-employment and on-the-job search. Research could also focus in more depth on the demographics of the users of the different platforms, from the perspective of both the job seekers and the companies or recruiters. Although social networks are used by employers to share vacancies, little work seems to exist on the nature of these vacancies (e.g. in which sectors, types of jobs). In addition, only few studies seem to examine the effectiveness of the different platforms for job search, selection and matching. The future of Internet-based labour market research therefore seems to be bright.

Endnotes

¹Many of his articles are published as newspaper columns and can be accessed via <http://sethsd.com/>.

²The survey is available at <http://naceweb.org/uploadedFiles/NACEWeb/Connections/social-jobs-partnership-executive-summary.pdf>.

³Note that the website also offers reviews on interviews and benefits, but because the guidelines are very similar to those for the company reviews, we only focus on those in our discussion.

Competing interests

The IZA Journal of Labor Economics is committed to the IZA Guiding Principles of Research Integrity. The authors declare that they have observed these principles.

Acknowledgements

The authors would like to thank Ilaria Maselli and Elisa Martellucci for their insightful comments. Brian Fabo acknowledges the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program.
Responsible editor: Klaus Zimmermann

Author details

¹Centre for European Policy Studies (CEPS, Brussels), CEPS, 1 Place du Congrès, 1000 Brussels, Belgium. ²Central European University (CEU Budapest), CEU, 9 Nador u., 1051 Budapest, Hungary.

Received: 1 September 2015 Accepted: 7 January 2016

Published online: 18 January 2016

References

- Achrekar, Harshvardhan, Gandhe, Avinash, Lazarus, Ross, Su-Hsin Y, Benyuan L (2011) Predicting Flu Trends using Twitter data. In: Computer Communications Workshops (INFOCOM WKSHPS), IEEE, pp. 702-707
- Acquisti A, Fong C (2015) "An experiment in hiring discrimination via online social networks", Working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2031979
- Acquisti A, Fong C (2015) "Online information and mechanisms of discrimination", Working paper
- Agrawal AK, Tambe P (2014) "Private equity and workers' career paths: the role of technological change" Available at SSRN: <http://ssrn.com/abstract=2286802>
- Askitas N, Zimmermann KF (2009) Google econometrics and unemployment forecasting. *Appl Econ Q* 55(2):107–120
- Askitas N, Zimmermann KF (2015) The Internet as a data source for advancement in social sciences. *Int J Manpow* 36(1):2–12
- Autor DH (2001) Wiring the labor market. *J Econ Perspect* 15(1):25–40. doi:10.1257/jep.15.1.25
- Benfield JA, Szlemko WJ (2006) Internet-based data collection: promises and realities. *J Res Pract* 2(2). Article D1. Retrieved [date of access] from <http://jrp.icaap.org/index.php/jrp/article/view/30/51>
- Bohnert D, Ross WH (2010) The influence of social networking web sites on the evaluation of job candidates. *Cybertechnol Behav Soc Netw* 13:1–7
- Bonsón E, Bednárová M (2013) Corporate LinkedIn practices of Eurozone companies. *Online Inf Rev* 37(6):969–984
- Boucher E, Renault C (2015) Job classification based on LinkedIn summaries—CS 224D, Stanford University, Available at http://www.deccocel.com/docs/linkedin_project.pdf

- Brenčić V (2014) Search online: evidence from acquisition of information on online job boards and resume banks. *J Econ Psychol* 42:112–125
- Buhrmester M, Kwang T, Gosling SD (2011) Amazon's Mechanical Turk—a new source of inexpensive, yet high-quality, data? *Perspect Psychol Sci* 6(1):3–5
- Caers R, Castelyns V (2011) LinkedIn and Facebook in Belgium: the influences and biases of social network sites in recruitment and selection procedures. *Soc Sci Comput Rev* 29(4):437–448
- Carnevale AP, Jayasundera T, Repnikov D (2014) Understanding online job ads data: a technical report, Georgetown University, McCourt School on Public Policy, Center on Education and the Workforce, April 2014, 28 pages.
- Carrière-Swallow Y, Labbé F (2013) Nowcasting with Google Trends in an Emerging Market *J Forecast* 32(4):289–298
- Chandra V (2012) Work-life balance: eastern and western perspectives. *Int J Hum Resour Manag* 23(5):1040–1056
- Chen T, So EPK, Wu L, Yan IKM (2015) The 2007–2008 US Recession: what did the real-time Google Trends Data tell the United States? *Contemp Econ Policy* 33(2):395–403
- Choi H, Hal V (2009) Predicting initial claims for unemployment insurance using Google Trends, Technical report, Google, Available at: <https://static.googleusercontent.com/media/research.google.com/en//archive/papers/initialclaimsUS.pdf>
- Choi H, Varian H (2012) Predicting the present with Google Trends. *Econ Rec* 88(s1):2–9
- Constant A, Zimmermann KF (2008) Im Angesicht der Krise: US-Präsidentenwahlen in transnationaler Sicht. *DIW Wochenbericht* 44:688–701
- Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M (2014) Social media update 2014, PEW Research Center, <http://www.pewinternet.org/2015/01/09/social-media-update-2014/>
- Faberman RJ, Kudlyak M (2014) "The intensity of job search and search duration", Federal Reserve Bank of Richmond, Working Paper, No. 14–12
- Fairlie RW, Robb A (2009) Entrepreneurship, self-employment and business data: an introduction to several large, nationally-representative datasets., IZA discussion papers 4052
- Finholt T, Sproull LS (1990) Electronic groups at work. *Organ Sci* 1(1):41–64
- Foster G (1994) Fishing with the Net for research data. *Br J Educ Technol* 25(2):91–97
- Freeman LC (1984) The impact of computer-based communication on the social structure of an emerging social scientific speciality. *Soc Networks* 6:201–21
- Gee LK (2015) The more you know: information effects on job application rates by gender in a large field experiment, Tufts University Working Paper, http://laurakgee.weebly.com/uploads/2/2/4/8/22485690/2015_06_25_gee_moreyouknow.pdf
- Gee LK, Jones J, Burke M (2015) Social networks and labor markets: how strong ties relate to job transmission using Facebook's social network. Working paper. <http://laurakgee.weebly.com/research.html>
- Gee LK, Jones J, Fariss CJ, Burke M, Fowler JH (2015) The paradox of weak ties in 55 countries. Working paper. <http://laurakgee.weebly.com/research.html>
- Ghani E, Kerr WR, Stanton CT (2012) Diasporas and outsourcing: evidence from oDesk and India, NBER. Working paper 18474
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L (2009) Detecting influenza epidemics using search engine query data. *Nature* 457:1012–1014
- Headworth A (2014) Recruiters: why are you not using Facebook Graph Search? Blog post. Available at: <http://sironaconsulting.com/2014/09/recruiters-facebook-search/>
- Hershbein B, Kahn L (2015) Is college the new high school? Evidence from vacancy postings. Available at: http://works.bepress.com/brad_hershbein/6
- Hookway N (2008) Entering the blogosphere: some strategies for using blogs in social research. *Qual Res* 8(1):91–113
- Hooley T, Marriott J, Wellens J (2012) What is online research? Using the Internet for social science research. London: Bloomsbury Academic 2012.
- Horton JJ (2011) The condition of the Turkling class: are online employers fair and honest? *Econ Lett* 111(1):10–12.
- Horton JJ, Rand D, Zeckhauser R (2011) The online laboratory: conducting experiments in a real labor market. *Exp Econ* 14(3):399–425
- International Telecommunications Union (ITU) (2014) Measuring the information society report 2014
- Jarrow R, Kchia Y, Protter P (2011) Is there a bubble in LinkedIn's stock price? *J Portf Manag* 38(1):125
- Kaplan K (2014) Interviews: en garde. *Nature* 506:397–399
- Karl K, Peluchette JV, Schlaegel C (2010a) A cross-cultural examination of student attitudes and gender differences in Facebook profiles content. *Int J Virtual Communities Soc Netw* 2:11–31
- Karl K, Peluchette JV, Schlaegel C (2010b) Who's posting Facebook faux pas? A cross-cultural examination of personality differences. *Int J Sel Assess* 18:174–186
- Kearney M, Levine P (2015) Media influences and social outcomes: the effect of MTV's "16 and Pregnant" on teen childbearing. *A Econ Rev* 105(12): 3597–3632
- Kehoe CM, Pitkow JE (1996) Surveying the territory: GUV's five WWW user surveys. *World Wide Web J [Online]* 1(3):77–84
- Kelkar A, Kulkarni S (2013) Value of Facebook for job search: languishing present to a lucrative future, information society (i-Society). International Conference, 24–26 June 2013, 222 – 226
- Kiesler S, Sproull LS (1986) Response effects in the electronic survey. *Public Opin Q* 50(3):402–13
- Kluemper DH, Rosen PA (2009) Future employment selection methods: evaluating social networking web sites. *J Manag Psychol* 24:567–580
- Konetic (2014) 80 % of HR teams use social media for sourcing and hiring new staff. <http://www.recruitmentgrapevine.com/article/2014-09-16-80-of-hr-teams-use-social-media-for-sourcing-and-hiring-new-staff#.VCQXMfIdUf0>
- Krantz JH, Dalal RS (2000) Validity of web-based psychological research. In: Birnbaum MH (ed) *Psychological experiments on the Internet*. Academic Press, San Diego, pp 35–60
- Kudlyak M, Lkhagvasuren D, Sysuyev R (2012) "Systematic job search: new evidence from individual job application data". Federal Reserve Bank of Richmond, Working Paper, No. 12-03R (Revised September 2014)

- Kuhn P (2014) The Internet as a labor matchmaker. *IZA World of Labor* No. 18
- Kuhn P, Mansour H (2014) Is Internet job search still ineffective? *Econ J* 124(158):1213–1233
- Kuhn P, Skuterud M (2004) Internet job search and unemployment durations. *Am Econ Rev* 94(1):218–232
- Kureková L, Miroslav B, Thum-Thysen (2015) Using Internet data to analyse the labour market. A methodological enquiry. *IZA Discussion Paper* No. 8555, 29 pages
- Kwak H, Lee C, Park H, Moon S (2010) What is Twitter, a social network or a news media? *WWW Proceedings of the 19th international conference on world wide web*, ACM New York
- Larsen M (2011) How the big boys do Twitter recruitment, blog post at the recruiter.com. Available at: <https://www.recruiter.com/i/twitter-recruitment/?PageSpeed=noscript>
- Lauby S (2013) The new social workplace. *Hum Resour Manag Int Dig* 21(3):3–5
- Marinescu I (2015) "The general equilibrium impacts of unemployment insurance: evidence from a large online job board". Working paper. http://www.marinescu.eu/Marinescu_UL_2014.pdf
- Marinescu I, Rathelot R (2015) "Mismatch unemployment and the geography of job search". Working paper. http://www.marinescu.eu/MarinescuRathelot_geomismatch.pdf
- Maselli I, Fabo (2015) Digital workers by design? An example from the on-demand economy. CEPS Working Document No. 414/October 2015. <https://www.ceps.eu/publications/digitalworkers-design-example-demand-economy>
- Massimino PM, Kopelman RE, Joseph ML (2015) Explaining hospital performance via the Cube One framework. *J Organ Eff People Perform* 2(1):73–90
- Murth D (2015) Twitter and elections: are tweets, predictive, reactive, or a form of buzz? *Inf Commun Soc* 18(7):816–831
- Musch J, Reips U-D (2000) A brief history of web experimenting. In: Birnbaum MH (ed) *Psychological experiments on the Internet*. Academic Press, San Diego, pp 61–87
- O'Connor H, Madge C, Shaw R, Wellens J (2008) Internet-based interviewing. In: Fielding N, Lee RM, Blank G (eds) *The Handbook of Online Research Methods*. Sage Publications, Thousand Oaks
- OECD (2014) "Skills and jobs in the Internet economy". OECD Digital Economy papers, No. 242, OECD publishing, Paris
- Osterhaus E (2014) How job seekers use Glassdoor reviews. Survey by Software Advice. Available at: <http://new-talent-times.softwareadvice.com/how-job-seekers-use-glassdoor-0114/>
- Pallais A (2014) Inefficient hiring in entry-level labor markets. *Am Econ Rev* 104(11):3565–3599
- Preis T, Moat HS, Stanley HE, Bishop SR (2012) Quantifying the advantage of looking forward. *Sci Rep* 2:350
- Preis T, Moat HS, Stanley HE (2013) Quantifying trading behavior in financial markets using Google Trends, *Nature, Scientific Reports* 3. Article number: 1684
- Rangel L (2014) Writing a LinkedIn profile to get found by recruiters. *Career Plann Adult Dev J* 30(2)
- Right management (2015). Available at: <http://www.kent.ac.uk/careers/jobs/social-networking.htm>
- Rode A, Shukla A (2014) Prejudicial attitudes and labor market outcomes. Royal Economic Society. Annual Conference, 7–9 April 2014, Manchester. Available at <http://webmeets.com/files/papers/res/2014/686/RodeShukla-RevisedOct2013.pdf>
- State B, Rodriguez M, Helbing D, Zagheni E (2014) Migration of professional to the US—evidence from LinkedIn data, published in *Social Informatics (SOCINFO 2014)*, edited by Aiello and McFarland. *Lect Notes Comput Sci* 8851:531–543
- Sasser Modestino A, Shoag D, Balance J (2014) "Upskilling: do employers demand greater skill when skilled workers are plentiful?" Federal Reserve Bank of Boston. Working Paper No. 14–17
- Schawbel D (2012) How recruiters use social networks to make hiring decisions now. *Time*. Business Section. Published on July 9, 2012. Available at: <http://business.time.com/2012/07/09/how-recruiters-use-social-networks-to-make-hiring-decisions-now/>
- Schmidt T, Vossen S (2012) Using Internet data to account for special events in economic forecasting, *Ruhr Economic Papers*, no. 382
- Shapiro H (2014) "Use of real time labour market data". Presentation during CEPS Expert workshop 2 20/10/2014 of the InGRID Research Project
- Stevenson B (2006) "The impact of the Internet on workers flows". Available at: http://users.nber.org/~bstevens/papers/Stevenson_Internet.pdf
- Stevenson B (2009) "The Internet and job search" studies of labor market intermediation. David Autor (eds), Chicago, University of Chicago Press.
- Tambe P (2014) Big data investment, skills, and firm value. *Management Science*. 60(6): 1452–1469
- Thelwall M (2008) Social networks, gender, and friending: an analysis of MySpace member profiles. *J Am Soc Inf Sci Technol* 59(8):1321–30
- Thiele L (2014) Opportunities for women in IT, Available at: <http://cdn.aimforbrilliance.org/pdf/Opportunities%20for%20Women%20in%20IT.pdf>
- Williams SA, Terras M, Warwick C (2013) What people study when they study Twitter: classifying Twitter related academic papers. *J Doc* 69(3)
- Wilson RE, Gosling SD, Graham LT (2012) A review of Facebook research in the social sciences. *Perspect Psychol Sci* 7(3):203–220
- Wright J (2012) "Making a key distinction: real-time LMI & traditional labor market data". Blog post on *Economic Modeling*, February 7th, 2012. <http://www.economicmodeling.com/2012/02/07/making-a-key-distinction-real-timelmi-traditional-labor-market-data/>
- Yu Y, Wang X (2015) World Cup 2014 in the Twitter World: a big data analysis of sentiments in US sports fans' tweets. *Comput Hum Behav* 48:392–400
- Zide J, Elman B, Shahani-Denning C (2014) LinkedIn and recruitment: how profiles differ across occupations. *Employ Relat* 36(5):586–604