

Use of OJA and big data in research

Nikos Askitas, IZA – Institute of Labor Economics

April 15, 2021

CEDEFOP, Getting the future right | Towards smarter and people-centred skills intelligence

Online job advertisements: Gaining real-time insights into skill trends

Thessaloniki/Zoom.

Table of contents

1. Introduction
2. Reflections on OJA and Big Data

Introduction

- I am a **mathematician** by training, persuasion and denomination, a **technologist** (as a result of economic disruption, serendipity and because it is fun), a data scientist working as an **economist** at the IZA – Institute of Labor Economics.
- I spent most of my life as a social scientist losing sleep (and hair) over the pros and cons of online/big **data**.
- *The Internet as a data source for Social Science* is a **research theme** at my RDC.
- Together with Peter Kuhn we run a **workshop** titled *Matching Workers and Jobs Online - New Developments and Opportunities for Social Science and Practice*.
- Together with GESIS I run the **EDDI** - European DDI Conference.
- I invented **JoSuA** a web based (code-to-data) solution which enables data custodians to open even the most sensitive data for research by roaming researchers (RDC of **IAB** - Research Institute of the German Fed. Empl. Agency, RDC of **BIBB** - Federal Institute for Vocational Edu. and Training, RDC of **IQB** - Institute for Edu. Quality Improvement).
- At the IDSC we are all about data: do research with data and on data.

Besides getting informed about CEDEFOP's work and enjoying discussing data with experts I want to briefly reflect on three questions I was kindly given in the hope that my remarks will somewhat enhance this session which I think as a process of cross pollination of each other. So the best that I can achieve today is to infect you with an idea which you will go away and incubate.

- Q1: How OJA and big data boost labour market research?
- Q2: What are the key challenges when working with such data?
- Q3: What needs to be done to support research activities?

Reflections on OJA and Big Data

- Markets (marriage, labor, information, transportation etc) are either born or are moving online: most transactions in most markets are computer mediated and take place on the internet.
- Why? ICT is ideally suited to solve the core market problem: matching of supply and demand.
- Moreover: you can easily run experiments and RCT's to improve matching and you can easily record, rewind and replay the market in order to better understand it. Also: online markets are non-invasive.
- Hence there is no need to boast about boosting: If you want to understand markets you need the Internet and you need to solve the problems it comes with.
- For Labor Economics: OJA data is a *sine qua non*.

- Day laborer hotspots (yes street corner offline matching still exists: 5th century BC in ancient Athens Colonos in Agora, Cologne, Home Depot parking lots in the USA etc)
- Temporary work agencies (reduce risk of bad matching in a rapidly changing economy)
- Platform Economy and Gig Work (digital version of Day Laborer Hotspots, driven by both labor demand and labor supply)
- Public and Private Job Boards (main source of OJA data, driven by labor demand)
- Social Media and Offline networks (first is observable, second is not but we know exists).
- Apprenticeships (well developed in Germany): demand informs supply.

- A job is “simply” a bundle (set, array, etc) of tasks.
- A worker is a bundle of skills.
- Studying the properties (duration, mutual satisfaction, quality of fit etc) of matched bundles is our core problem.
- In historical periods where technology is relatively static, production is “streamlined” and job bundles are relatively fixed hence: you have time to teach and produce your worker bundles.
- In times of technological, climatic, demographic (and recently epidemiological) disruptions the need for **timely skills intelligence** becomes acute and its production becomes critical for society.

Q1:How OJA and big data boost labour market research? 1/3

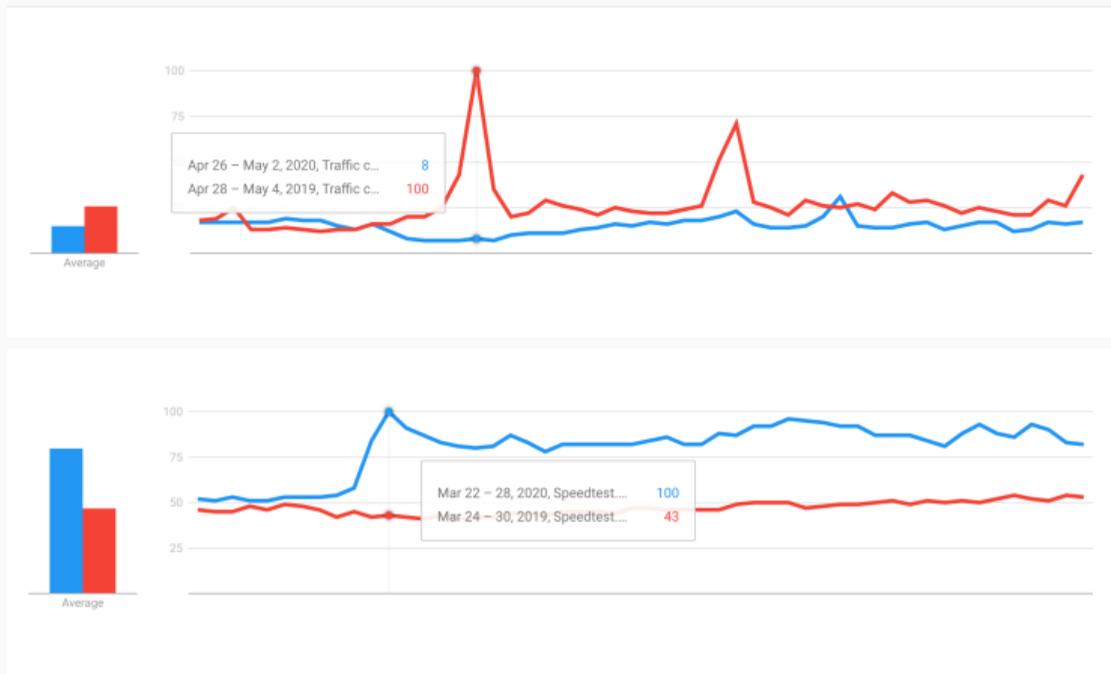


Figure 1: Traffic Jam (top) and speedtest.net in global Google searches in 2019 (red) & 2020.

- In ongoing work with K. Tatsiramos, using BGT data, we can describe how businesses in the UK were responding to the oncoming BREXIT after the referendum. To mention but one result: in regions with larger trade exposure to the EU we observed larger shifts from design jobs to maintenance jobs in the industrial sector as apparently firms were postponing investment until further due.
- *Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study* Askitas, Tatsiramos, Verheyden 2021. Combines data from the Oxford COVID-19 Government Response Tracker and the Google Mobility Reports. We show how policies impacted COVID cases and population whereabouts and provided a behaviorally founded estimate of the incremental effect of each lockdown policy conditional on other concurrent policies.

There is a multifaceted LM research ongoing using the transaction data from various Internet based matching facilitators and that data takes place both in Academia as well as within the matching facilitators themselves. In the US it is more common for an academic labor economist to work with data of a matching facilitator in a kind of win-win cooperation: this type of research produces cutting edge labor economics publications while at the same time it helps the matching facilitator better understand and hence improve the matching performance of their own online service. In Europe this is unfortunately not as common.

- Lack of optimal Data Access.
- Many “technical” multi-disciplinary issues: e.g.:
 - 1. misrepresentation of the worker bundle or the job bundle due to e.g. strategic behavior or cognitive and other limitations,
 - 2. Non-standard or different taxonomies used.
 - 3. Omitted but self-evidently required skills or tasks to be fulfilled (e.g. MS Excel).
 - 4. Homeseme, homophone (e.g. Administrator)
- Inertia on behalf of researchers, lack of understanding on behalf of the policy maker (although this is slowly changing as suggested by evidence from the IZA Discussion Paper series, my workshop with Peter Kuhn, CEDEFOP etc).
- In Europe cultural problems: Working as a research economist within or for a Matching Facilitator.
- Skills Gap in Graduate School

- Policy makers need to know (and act like it) that good policy making needs good and timely research and good research needs good and timely data and they both cost money. Furthermore just like you need researchers to extract information from data you need data to extract information from.
- Research is a data refinery producing (hopefully) actionable intelligence: Properly documented relevant data must be available timely, affordably and easily.
- Foster experimentation: while there is no denying the merits of a solid taxonomy, the world often changes like a mutating organism and not like a geometric puzzle. Things appear in “magical” unexpected ways (which we can always “explain” after the fact...). Think of what we could have done in terms of data quality before the Internet in order to see it coming...

- Research and Policy makers form a “knowledge market” of their own. We need a better alignment of the supply (research) and demand (policy making) in it. In research we often have answers to questions policy makers don't ask or questions they don't know that they need answers to...
- Job Market Facilitators should standardize their data models and taxonomies much like the car industry did in the early days of the automobile. CEDEFOP and Eurostat might help move this along...
- All stakeholders involved (researchers, policy makers, official statistics) need to be more vigilant to get a head start on new developments. It is not enough to come after the fact and "clean up" you need to have skin in the game persistently.

Big data starts small and grows in baby steps

Thank you for your attention

EDDI: <https://www.eddi-conferences.eu/>

JoSuA: <https://josua.iza.org/>

CfP: http://conference.iza.org/conference_files/DATA_2021/call_for_papers

Profile: <https://www.iza.org/home/askitas>

IZA DPs <https://www.iza.org/publications/dp>