
Building the Online Labour Index: A Tool for Policy and Research

Otto Kässi

Oxford Internet Institute,
University of Oxford
1 St Giles, OX1 3JS
United Kingdom
otto.kassi@oii.ox.ac.uk
@ottokassi

Vili Lehdonvirta

Oxford Internet Institute,
University of Oxford
1 St Giles, OX1 3JS
United Kingdom
vili.lehdonvirta@oii.ox.ac.uk
@ViliLe

Author Keywords

labour markets; employment index; online freelancing;
gig economy; sharing economy; public policy

ACM Classification Keywords

Economics; Measurement; Electronic commerce

Introduction

Labour markets are undergoing a digital transformation today. Yet the quantitative impact of this shift is still largely invisible to public policy makers and labour

Abstract

We present our research agenda for building the *Online Labour Index (OLI)*, the world's first indicator of online labour market utilisation. We argue that a new aggregate index on online labour markets is needed for public policy and research purposes, and discuss the data collection efforts that this entails.

market researchers, because conventional labour market indicators published by statistical agencies and labour market organisations are ill suited to measuring work that is transacted solely via online channels [1].

We are working on creating the *Online Labour Index (OLI)*, a set of aggregate measures on online labour utilisation. This paper explains why such a measure is needed and discusses some of the methodological choices and opportunities involved. Comments on our plans ideas for collaboration are extremely welcome!

Motivation

Finding reliable information on the total volume of the online labour markets is difficult due to lack of publicly available data. Various disparate sources and estimates nevertheless suggest that the transaction volume of online labour markets is already substantial and is projected to grow considerably in the next few years.

For example, a World Bank report [2] estimates that the total gross services revenue of the three biggest online labour market platforms exceeded \$1 billion in 2014. It also reports yearly growth projections of over 20 percent. The same report estimates that there are over 100 million workers employed either part or full time via online labour market platforms, though most likely this reflects the number of registered accounts

rather than currently active workers. It is possible that online labour markets are starting to obtain national economic significance especially in low-income countries, whose citizens are overrepresented among online workers [3]. Online labour market activity rates should thus be of keen interest to public policy makers and labour market researchers.

Existing labour market indicators

A standard ILO measure of employment rates used by statistical agencies counts as employed anyone gainfully employed for at least one hour in a week. This measure fails to capture any incremental effects of online work – if someone already has a job and does a second job online, their efforts are not captured in employment statistics. Further, it is not clear to what extent online workers choose to report their earnings to tax agencies, especially if the earnings are small. This might be an especially relevant concern for the large share of online workers living in developing countries, where the informal economy dominates. Finally, even when online earnings are duly reported, the existing statistical categories do not allow such earnings to be distinguished from those earned from the domestic labour market. New measures are thus clearly needed.

There is some prior work on using online market data to measure labour market activity. The Monster Employment Index was a measure of employer online recruitment activity, calculated monthly from vacancies posted across a wide variety of corporate career sites and job boards. The index was discontinued in 2012. A project focusing specifically on an online labour market

is Panos Ipeirotis' MTurk Tracker¹, which tracks task availability and completion over time on Amazon Mechanical Turk. These examples show that it is possible to track online labour market utilisation, and that it is possible to construct a composite index from multiple disparate online data sources.

Online labour market platforms

Which platforms should be included in the index is obviously a key question. For practical reasons data collection will have to be limited to a subset of the platforms. Platform size is one obvious selection criteria, and another consideration is geographic and linguistic coverage.

A key question will be industry/segment coverage; some platforms are more generalist, whereas others cater to very specific types of work only. For now, we are primarily focusing on platforms that focus on remotely delivered labour as opposed to localized services such as transport, though a sub-index for such services could be added later.

In constructing measures of online labour market utilisation, it is useful to note that platforms offering online work feature three main types of mechanisms for matching workers with employers:

1. *Supply side mechanisms*, where workers post their "virtual resumes", including requested wages. Employers can then "bid" for the workers' time by contacting them. If a worker is contacted, they typically enter into a negotiation phase, where the employer and the worker agree on the details of the project. For

¹ <http://www.mturk-tracker.com/> (see [4])

example, Upwork.com and Freelancer.com feature a supply side mechanism.

2. *Demand side mechanisms*, where employers post details of tasks or projects into the platform, and workers bid on them by posting their resume and wage requests. A negotiation phase typically follows. For instance, a Chinese online labour market Witmart.com only features a demand side mechanism. Upwork.com, and Freelancer.com, on the other hand, feature both demand and supply side mechanisms.

3. *Spot markets*, where the pricing and task content are fixed and there is no negotiation phase; the first worker to apply gets the task automatically. "Microwork" platforms such as Mechanical Turk are the most prominent examples of spot markets.

The main difference between supply and demand side mechanisms and spot markets is that spot markets have taken the commodification of labour to extreme, whereas the supply and demand mechanisms rely

Measuring online labour utilisation

Traditional offline labour markets are measured by surveying workers and establishments on a regular basis. Because the surveys are often conducted via telephone or (non-electronic) mail, national statistical agencies have put considerable effort into developing sampling schemes that provide reliable estimates with small sample sizes. In contrast, data collection from online platforms is highly scalable. Many platforms provide a general purpose developer API that can be utilised for data collection. If an API is not available, some types of relevant data can also be scraped from the platforms' web user interface.

Due consideration must be given to legal and ethical questions, but in general this type of data collection has been considered acceptable in social science research as long as the platform is not burdened with excessive requests and user privacy is respected.

The most straightforward piece of data to collect is often the *total size* of platform measured either by total number of projects or registered worker accounts. In many cases, platforms publish such this data on their front pages². If such a number is not available, the total size and its changes can be approximated by conventional web analytics measures.

If one wants a more nuanced view of online labour markets, segmented by for example geographical location or project type, a dataset with individual worker or project level data is needed.

Labour supply can be studied by taking snapshots of the registered labour force on a platform. This allows us to calculate the total labour supply in the platform and to disaggregate it either geographically or by skills of workers. Comparing the hours worked of a worker between two points in time allows us to calculate the *online workforce utilisation rate* on each platform. This measure is the online labour market analogue to rate of employment in traditional labour markets.

Turning to labour demand – in contexts of both demand side mechanisms and spot markets – we can track it in a similar fashion to supply. Here, we periodically take

² For example, on 2015-12-09, Freelancer.com reported that it had "17 443 533 registered users and 9 082 390 available projects".

snapshots of the open projects posted on the platform, and see how large a share of them have been completed between the two snapshots. The projects can be disaggregated again geographically (by employer's location), or by skills required to fill the project.

Sampling and aggregation

In practice, both the supply and demand side indicators can be also calculated by using representative samples of workers and projects respectively. Though developing representative samples is not without its difficulties, in practice this may be the best approach for most platforms. This will reduce storage and data processing requirements substantially, and require much more modest data access than full snapshots.

Finally, measures obtained from different platforms must be combined into an aggregate online labour index, or a set of related indices describing workforce size, utilisation rate, wages, and so on. Because of differences in data collection methods and underlying assumptions, it may not be possible to produce aggregates of absolute numbers; instead, the likely product is a set of indices that track changes over time in relation to an arbitrary starting value – a familiar tool to policy analysts and labour market researchers.

A practical example

Finally, to fix ideas, we present an example of the collection of OLI for one particular online labour market platform.

This platform features both, a supply and a demand side mechanism. Since the two are, essentially, two

sides of the same issue, we concentrate only on demand.

We have an regularly updated estimate for the total number of workers on the platform. In addition, we have collected a worker panel of roughly 30 000 registered users. The users' hours worked are automatically collected from the platform in the beginning of each month.

The former of the metrics accounts for the size of the worker base in the platform, and the latter captures the rate of utilisation within it. We calculate the OLI sub-index for this platform by simply multiplying the two metrics with each other. Finally, we normalise the metric so that the first period of measurement equals 100.

Conclusions

The full impact of online labour markets to national economies remains poorly understood. A key reason for this is the lack of economic indicators that would cover online labour markets. We are working on Online Labour Index (OLI), a set of aggregate indices tracking online labour market utilization and related variables.

References

1. Sundararajan, A. (2014). Peer-to-Peer Businesses and the Sharing (Collaborative) Economy. Written testimony for the Committee on Small Business of the U.S. House of Representatives hearing on 15 Jan. <http://1.usa.gov/1rig5yx>
2. World Bank Group (2015): The Global Opportunity in Online Outsourcing.

<http://pubdocs.worldbank.org/pubdocs/publicdoc/2015/6/212201433273511482/Global-OO-Study.pdf>

3. Ajay Agrawal, John Horton, Nicola Lacetera, Elizabeth Lyons. 2013. Digitization and the Contract Labor Market: A Research Agenda. NBER Working Paper 19525.
<http://www.nber.org/papers/w19525>
4. Panos Ipeirotis, 2010. Analyzing the Amazon Mechanical Turk Marketplace. ACM XRDS, Vol 17, Issue 2, Winter 2010, pp. 16-21.