





© 2018 Mourelatos et al. This article follows the  Open Access policy of CC BY NC under Creative Commons attribution license v 4.0.  Revistia

Submitted: 04/10/2018 - Accepted: 04/11/2018 - Published: 28/12/2018

A Study on the Evolution of Crowdsourcing Websites

Evangelos Mourelatos¹

Nikos Frarakis¹

Manolis Tzagarakis¹

¹Department of Economics, University of Patras, Greece

Email: vagmour@upatras.gr

DOI: 10.26417/ejser.v11i1.p29-40

Abstract

Today hundreds of millions of Internet users are using crowdsourcing platforms either to work to an online job as workers, or as a model of problem solving and production as requesters. This growing workforce makes it necessary from the perspective of the online platforms, to fully understand the factors which contribute on this emerging and innovative “online labor”. We present a study comparing how crowdsourcing platforms have evolved over two time periods. The study explores in addition the effect of several factors of such websites on their traffic data and attention seeking. The results show that several website’s characteristic are strong indicators of its attempting to get and optimize potential workers’ or requesters’ attention.

Keywords: Crowdsourcing, Online labor, Websites Review, Alexa Rankings, Regression Models

JEL Classification: A19, C14, J49, O10

Introduction

Since Jeff Howe introduced the term "Crowdsourcing" in 2006 for the first time in history [3], crowdsourcing has become a pivotal part of today’s Internet focus where everything is designed to take advantage of the networked world. Thousands of

workers every day categorize images; write articles or take part to several kinds of online tasks. Crowdsourcing, as a term, is a strategic model to attract an interested, motivated crowd of individuals capable of providing solutions superior in quality and quantity to those that even traditional forms of business can [1]. Today the term “Crowdsourcing” is used today as equivalent to online labor. With the growth of online platforms with crowdsourcing and crowdfunding services like Amazon Mechanical Turk and Kickstarter respectively, a huge work force and a large knowledge base can be easily accessed and utilized to tackle problems requiring human intelligence. Users have only recently been recognized as an alternative source of new product ideas. Whereas some have attributed great potential to outsourcing idea generation to the “crowd” of users (“crowdsourcing”), others have clearly been more skeptical [2]. Given the success of the crowdsourcing paradigm, we expect that this field will continue to grow rapidly. It’s not outsourcing, it’s crowdsourcing [3].

The continuous development of these online platforms providing crowdsourcing and crowdfunding services, in general, impose the need for continuous and responsible destination management in order to achieve and maintain an appropriate level of sustainability and competitiveness of this new way of online labor. In our study, we use Alexa rankings data as a measure of a website’s attention seeking indicating its level of sustainability and competitiveness. We determine the factors that affect and to which degree the ‘combination’ of the estimated average monthly unique visitors and the estimated number of page views incurred in these online platforms based in our two time-period study [4].

Related Work

In recent years, many studies have elaborated proposals for defining and measuring the quality of work in crowdsourcing environments [5], while others focused on the level of engagement and motivation of the participants called “workers” in these online tasks [6]. Due to these issues, crowdsourcing has attracted the interest of researchers from various fields. Many research efforts have as a focal point the development of several mechanisms in order to make possible the quality control and cheat detection [7], while others aimed at answering important questions in the field (e.g. What tasks are most paid? or When are the users of my platform active?) by analyzing the anatomy of a crowdsourcing platform [8]. Nevertheless, an online platform is in fact a type of online business and all businesses aim at improving their conversion rate in order to have sustainability and competitiveness. Yet, it is unclear, which factors enact on a crowdsourcing website’s effort to attract targeted traffic, so as to achieve financial success.

This paper aims at investigating how crowdsourcing platform characteristics correlate with their traffic and by extension to their ranking as reported by Alexa. The overall goal is to investigate how crowdsourcing platforms change over time in order to provide their services to a greater user base. This study builds and extends earlier research that aimed to provide a first systematic review of such kind of sites [9].

Data Analysis

Data source

Data regarding the traffic of the studied crowdsourcing websites were collected from alexa.com in values of global ranking, while the data regarding their provided services and mode of operation have been collected by visiting each site included in the survey. The study considered the Alexa top-ranked one hundred crowdsourcing platforms for two periods; January of 2014 to January of 2015 (which from now on will be referred to as year 2014) and January of 2016 to January of 2017 (which from now on will be referred to as year 2016). Alexa was chosen as the preferred site of web traffic data because, related to Webometrics i.e. the process of measuring various aspects of websites that include their popularity and usage patterns, Alexa has been shown to outperform other similar services such as Google Trends for Websites and Compete (Vaughan and Yang 2013). We gathered the data the exact time period in order not to have traffic seasonality issues, so that we can make safe conclusions through the comparison. Last but not least, we didn't include in our study the period of January of 2015 to January of 2016, because we wanted our comparison study to reflect all the websites' changes occurred since 2014. For that reason, we needed an unobserved time period in order the websites' change process to be completed.

Methodology

We selected a number of popular websites offering crowdsourcing services for review. The same methodology was used to select the crowdsourcing sites for both time periods (2014 and 2016) and was based on a two scaled selection methodology. Initially, we used the most three popular search engines – Google, Bing and Yahoo!Search - to search for crowdsourcing online platforms that stakeholders are likely to encounter. For the 2014 dataset, all searches were performed from 1 January to 30 April of 2014 and for the 2016 dataset searches were conducted from 1 August to 30 November using the same set of keywords (“crowdsourcing”, “crowdfunding”, “online platforms”). The final list of crowdsourcing sites was compiled by selecting sites that met specific criteria that included:

Language: All crowdsourcing websites reviewed had to present their services in English. This facilitated the work of assessing services provided and comprehending their use.

Presentation of the type of service provided i.e. type of tasks accepted: Websites had to explicitly report all the necessary information related to the types of tasks it accepts in order to make their review possible.

All information needed for completing the review had to be offered. Many websites may not disclose all information required and such websites were excluded from our analysis.

Based on the above criteria a final list of crowdsourcing sites was compiled and for each site its Alexa ranking were retrieved. For each site the Alexa rankings for the years 2014 (Alexa 2014) and 2016 (Alexa 2016) were collected resulting in the top 100 crowdsourcing websites as ranked by Alexa. The selected websites were assessed against a number of criteria, which aimed to capture the site's characteristics. In particular, these criteria cover technical as well as operational features. Below we present the criteria in greater detail:

Type of service provided

The term "type of service provided" refers to the type of tasks the site specializes in. Services provided by websites were grouped into the following ten categories [10]:

- a. Microworks/Simple tasks, which are considered the smallest unit of work in a virtual assembly line, e.g. categorisation, tagging, Web research, transcription, etc.
- b. Crowdfunding, which is the collection of finance from backers (the crowd) to fund an initiative (project). Crowdfunding has its origins in the concept of crowdsourcing, which is the broader concept of an individual reaching a goal by receiving and leveraging small contributions from many parties. Crowdfunding is the application of this concept to collect funds through small contributions by many parties in order to finance a particular project or venture
- c. Mobile crowdsourcing services, which is concerned with issues related to applications for mobile phones.
- d. Content Generation services, in which content is generated by the crowd. This method is becoming increasingly popular because it offers an alternative to content creation and content curation.
- e. Data Entry services, which are projects using many different *modi operandi*, e.g. Excel, Word, electronic data processing, typing, coding and clerical assignments.
- f. High knowledge intensity services, which are specialised services in specific fields such as health, law, insurance, consultancies, data management, market research and cloud applications.
- g. Program developing services, which focus on the design and implementation of software.
- h. Web and graphic design services, which use the crowd contribution in the creation of Web and artistic projects.
- i. Translation services, which target content translation tasks from one language into another.
- j. Product reviews and testing, in which reviews and tests are requested.

Quality & Reliability

This variable is used to capture the methods the website employs to ensure the quality of work provided by workers. It also includes methods that the platform employs to prevent cheating and ensure workers' reliability [11].

Region

Indicates the region the platform is operating in [12]. Websites were classified in one of the following regions: North America, Europe, Australia and Asia.

Online Imprint

This variable reflects the strategies a platform uses for digital marketing and includes three categories; social networks, video streaming-sharing communities and blogs/forums [13].

Descriptive Statistics

The study essentially consists of two databases. In particular, we examined the top 100 crowdsourcing websites of 2014 and 2016 respectively, based on the aforementioned criteria. It is noteworthy to point out that 74 websites in the 2014 list were absent in the 2016 database because either they closed; or they changed their field of service. In the reviewed websites 57.47% remained in the top 100 list in the reported time period while 42.53% of these for several reasons appear to exit the top100 listings. With respect to the latter, 19% of them have ceased their operation, 7% have changed its business scope and 5% have merged with other crowdsourcing websites. Table 1 gives an overview of the variables collected for each reviewed crowdsourcing website. It shows also summary statistics for each variable for the set of reviewed websites.

As it is shown from the following table 1 most of our variables are binary (i.e termed 0 and 1). Thus, for example regarding the crowdfunding websites, are coded with "crf" and the mean shows that 33.5% of our sample is coded with "1", meaning that on average in both periods of our study one out of three online platforms provided crowdfunding services.

Table 1. Summary statistics

Variables	Code_Name	Obs	Mean	Min	Max	Std. Dev.
Alexa Ranking of 2016 (Dependent)	lnalexa_16	174	10.726	6.516	13.816	2.070
Alexa Ranking of 2014	lnalexa_15	174	10.747	6.261	13.678	1.948
Microworking websites	mwk	174	.175	0	1	.381
Crowdfunding websites	crf	174	.335	0	1	.473
Mobile Crowdsourcing sites	mcw	174	.050	0	1	.218
Content Generation websites	cntg	174	.085	0	1	.280
Data Entry websites	dte	174	.065	0	1	.247
High Tech Services websites	hts	174	.200	0	1	.401
Product Development websites	pdvp	174	.060	0	1	.238
Design Services websites	dsns	174	.130	0	1	.337
Translation Services websites	trs	174	.045	0	1	.208
Products' Review websites	rtp	174	.060	0	1	.238
No Quality Assurance	qr1	174	.300	0	1	.460
Reviews & Ratings	qr2	174	.355	0	1	.480
Workers Profile	qr3	174	.465	0	1	.500
Skills & Practice Tests	qr4	174	.140	0	1	.348
Spamming Tools	qr5	174	.270	0	1	.445
Websites from North America	NorthAmerica	174	.685	0	1	.466
Websites from Europe	Europe	174	.215	0	1	.412
Websites from Australia	Australia	174	.030	0	1	.171
Websites from Asia	Asia	174	.070	0	1	.256
No Imprint	NoImpnt	174	.040	0	1	.196
Social Communities	Soc_Com	174	.125	0	1	.332
Video Communities	Vid_Com	174	.065	0	1	.247
Blogs or Forums	BlogsForums	174	.770	0	1	.422

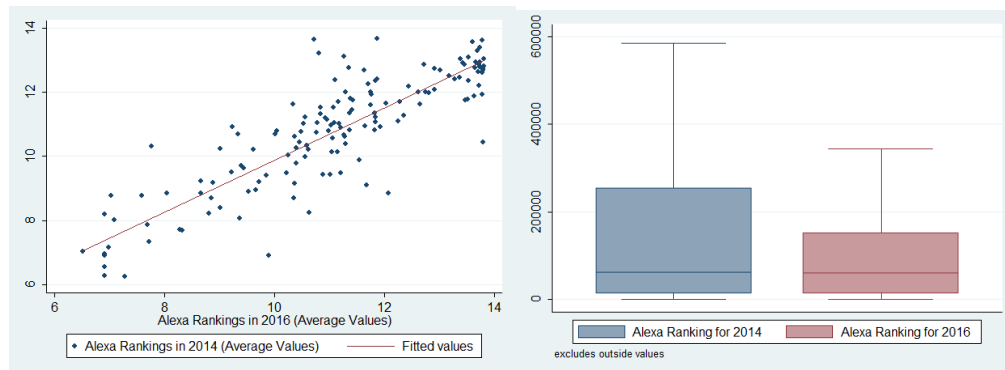
Source: Dataset with results drawn from alexa.com. Author's calculations.

Note: Alexa rankings are on logarithmic scale.

Furthermore, our study contains a total 174 unique crowdsourcing websites in both years of research. In order to investigate what changes have occurred in the 2014 and 2016 rankings drawing additionally some first conclusions in the evolution and preference of such kind of platforms we grouped the variables in six major categories: Alexa rankings, websites' quality assurance mechanisms, Type of services provided, Region of origin and Online imprint) and compared the how these groupings changed over the two year period (2014-2016).

More specifically, with regard to Alexa rankings, we crawled from alexa.com the average Alexa ranking values of 2014 and 2016 of each crowdsourcing platform. The results showed a positive relationship between them (figure 1), a statistically significant difference on the average values at 1% (table 3) and similar medians but different distribution characteristics (figure 1, box plot). This means that, a crowdsourcing online platform that was in high-ranked places in 2014 preserved its traffic over time and remained in similar ranking place, while in general we can observe that the values of alexa rankings in 2016 has a higher concentration around the mean and less extreme values (figure 1, Box Plot).

Figure 1 shows the relationship and the distribution of the values of alexa ranking over time. In the left graph y-axis measures the alexa ranking of our sample in 2014 values and x-axis in 2016 respectively. In the right graph y-axis has as a unit of measurement values of alexa ranking.

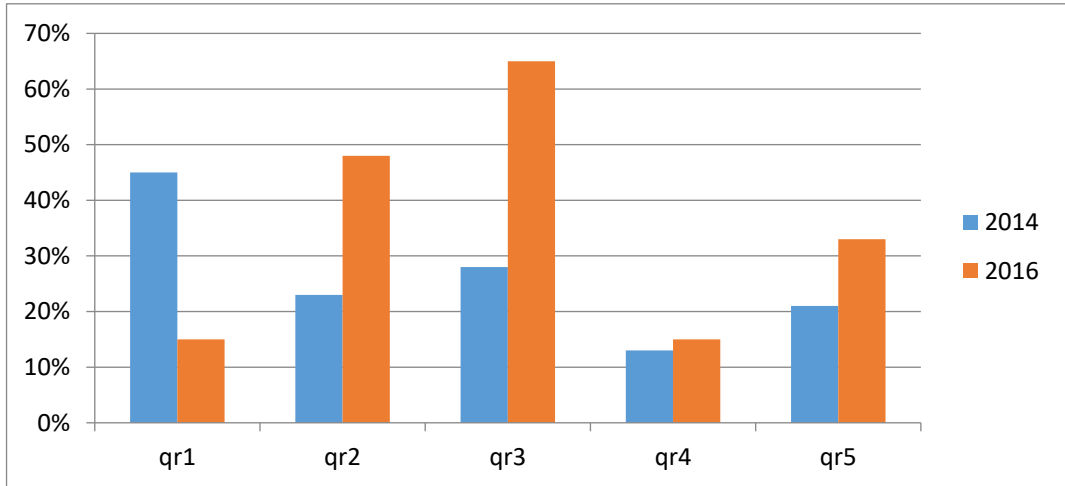


In addition, with respect to the quality mechanisms that the crowdsourcing online platforms offer as a tool for controlling the quality of responses and detecting cheats, the analysis revealed some interesting results between the two reviewed periods (table 2 and figure 2). First, the percentage of websites that don't provide any cheat detection mechanism has declined from 45% to 15% in the studied time period, while the percentages of all these type of quality control methods being provided by the crowdsourcing websites have been steadily rising, with the percentages' difference being statistically significance in most cases (table 2).

It is known that, commercial crowdsourcing applications suffer from workers, who try to submit invalid or low quality work in order to maximize their received payment while reducing their own effort. Our study revealed that, despite the success of many crowdsourcing platforms, it is only the last 2 years that they realise the processes that lead to high quality output by espousing several quality control mechanisms. It is known that, the strategies used to assure high quality depend largely on the tasks and outputs of crowdsourcing projects and for that reason many researchers have already introduced several types of quality control. In our research we categorise them into four main groups. "Reviews & Ratings", meaning whether or not a crowdsourcing platform gives the opportunity to the requesters to find the best worker for his task-needs, based on the reviews and ratings of the worker's job activity. "Workers profile", meaning whether or not a crowdsourcing platform provides on requesters, a profile for each worker containing demographic characteristics. "Skills & Practice Tests", meaning whether or not a crowdsourcing website gives the opportunity to the requesters to pre-select workers and through specific tests to understand which workers are most suitable to their crowdsourcing job. "Spamming Tools", meaning whether or not an online crowdsourcing platform has embodied in their

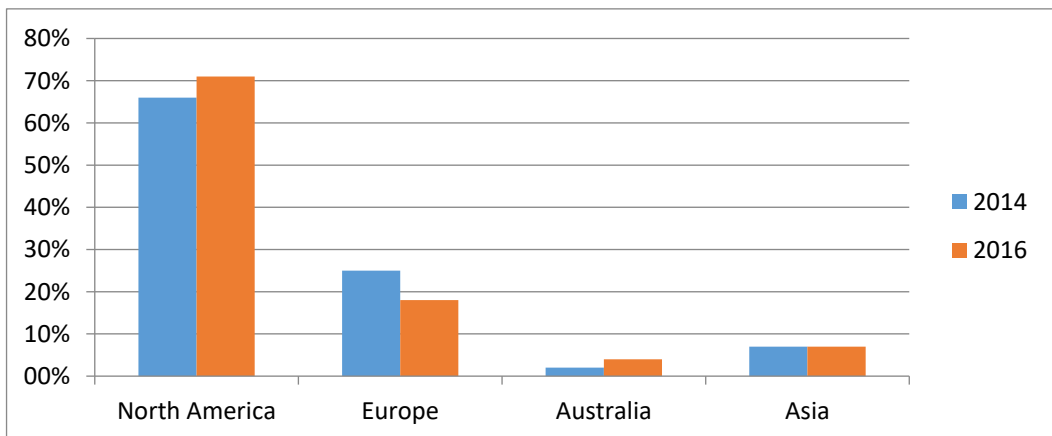
crowdsourcing process automated methods to aggregate a user's contributions in a way that promotes high quality output [14].

Figure 2 shows the percentages of "quality assurance" mechanisms over time



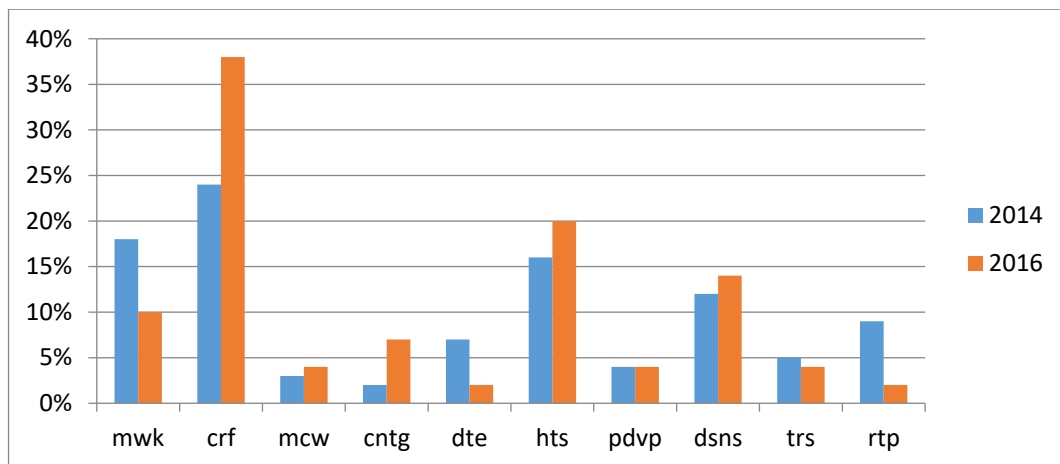
Moreover, concerning the websites' region of origin, we grouped both databases in four major groups, considering the origin country, North America, Europe, Australia and Asia. Figure 2 shows that from 2014 to 2016, a small increase of the percentage of these websites in North America (66% vs. 71%), in Australia (2% vs. 4%) was observed, while in Asia the percentage remained stable (7%) while in Europe a decrease was observed (25% vs. 18%). Our results showed that, crowdsourcing despite its web-based aspect; its process (i.e. its key activities, incentives of participation and structure), is strongly related to several geographical characteristics such as the platforms' geographical origins [15].

Figure 3 shows the percentages of websites' region of origin over time.



Concerning the websites' type of services provided, the review pointed to some noticeable outcomes. Figure 4 presents the major changes that occurred over the reported time period.

Figure 2 Comparison of websites' type of services provided over time



In particular, the number of websites supporting crowdfunding has increased significantly in the reported time period (24% vs. 38%). Similarly, the percentages of websites providing mobile crowdsourcing (3% vs. 4%), content generation (2% vs. 7%), high-tech services (16% vs. 20%) and web & graphic design tasks (12% vs. 14%) have increased in 2016 when compared to 2014. On the other hand, the percentage of websites offering microtask services (18% vs. 10%), data entry services (7% vs. 2%), translation (5% vs. 4%) and review & testing products (9% vs. 2%) show a decrease when compared to 2014. Websites with programming services remained stable (4%) over the reviewed time period.

Table 2 also shows the websites' online imprint over time. The term "online imprint" is used to denote the strategy that a website follows with respect to their digital marketing. In 2016 the percentage of websites using social networks such as Facebook and LinkedIn as well as video sharing sites such as Youtube are increased. A minority of the reviewed websites (2%) does not employ any digital marketing strategy in 2016.

**Table 2. Independent Samples t-tests for equality of means of websites' Alexa Rankings,
Quality Assurance Mechanisms & Online Imprint, collected in 2014 and 2016 respectively**

	Year of Research		Difference [2]-[1] [3]	t-test
	2014 [1]	2016 [2]		t [4]
Alexa Rankings	124.711	201.716	77.005	4.36***
<i>Websites' Quality Assurance</i>				
No Quality Assurance	.45	.15	-.30	4.87***
Reviews & Ratings	.23	.48	.25	3.81***
Workers' Profile	.28	.65	.37	5.62***
Skills & Practice Tests	.13	.15	.02	0.41
Spamming Detector Tools	.21	.33	.12	1.97*
<i>Websites' Online Imprint</i>				
No Social Network	.11	.02	-.09	2.61***
Facebook	.83	.88	.05	1.01
Pinterest	.09	.21	.12	2.40**
Youtube/Vimeo	.19	.49	.30	4.70***
Linkedin	.41	.52	.11	1.56
Twitter	.84	.89	.05	1.03
GooglePlus	.40	.39	.01	0.14
Blog or Forum	.71	.83	.12	2.09**

Source: Dataset with results drawn from Alexa.com. Author's calculations.

Note: Statistical significance: *** 1%, ** 5% and * 10%.

In order to have a better picture of the evolution of crowdsourcing platforms over time, we conducted a cross-tabulation analysis. Table 3 presents the relationship between each websites' quality assurance mechanisms and where they are hosted. It is noticeable that in 2016, all quality assurance mechanisms except skills and practise tests (qr4) are increasingly being adopted by crowdsourcing online platforms located in North America, Europe and Australia. On the other hand, crowdsourcing websites located in Asia emphasize more on providing information about their workers through worker profiles (qr3 is seen in 85.7% of the websites in 2016). This differentiation in preference of quality control mechanisms in Asia located crowdsourcing websites may be explained by the fact, that crowdsourcing in Asia is at an early stage of adoption and crowdsourcing in Asia is just warming up [16]. Yet, the general increase of the adoption of quality assurance mechanisms by the crowdsourcing websites over time indicates that while crowdsourcing platforms are maturing, online platforms realize the importance of "quality control" and "cheat detection" mechanisms [17].

Table 3. The relationship of a website’s region and its quality assurance mechanisms over time.

Region \ Years	2014				2016			
	qr2	qr3	qr4	qr5	qr2	qr3	qr4	qr5
North America	37.1	25.7	20	40	42.3	59.2	16.9	35.2
Europe	37.5	62.5	25	25	61.1	72.2	11.1	33.3
Australia	0	100	50	0	75	100	0	25
Asia	75	25	0	0	57.1	85.7	14.3	14.3

Source: Dataset with results drawn from Alexa.com. Author’s calculations.

Note: A website may have more than one quality control mechanism.

Unit is the percentage of websites that meet the current criteria

Concerning the relationship between the type of services being provided by crowdsourcing websites and their location and quality control mechanisms, our analysis show some significant results (table 4 & 5). In 2016, a major increase of crowdfunding websites (crf) in all regions is observed while websites focusing on microtask (mwk) services are severely limited. Moreover, it can be observed that, over time, the percentage of websites providing mobile crowdsourcing tasks in Asia and Australia remain non-existent, while in Europe and North America appear with very low percentage. This is rather counterintuitive considering the use of smartphones. With respect to other type of services, changes are observed which are summarised in table 4.

Table 4. The relationship of website region and its provided type of services over time.

Region	Type of Services (%)										
	mwk	crf	mew	cntg	dte	hts	pdvp	dsns	trs	rtp	
2014	North America	28.6	28.6	4.3	12.9	17.1	17.1	7.1	11.4	4.3	11.4
	Europe	26.7	26.7	6.7	20	13.3	26.7	0	6.7	6.7	13.3
	Australia	0	0	0	0	50	0	50	100	0	0
	Asia	75	0	0	0	0	0	75	75	25	0
2016	North America	7	36.6	5.6	7	2.8	22.5	4.2	12.7	2.8	2.8
	Europe	22.2	44.4	0	11.1	0	16.7	0	5.6	5.6	0
	Australia	0	25	0	0	0	0	25	75	0	0
	Asia	14.3	42.9	0	0	0	14.3	0	14.3	14.3	0

Source: Dataset with results drawn from Alexa.com. Author’s calculations.

Note: A website may provide more than one type of service. Unit is the percentage of websites that meet the current criteria.

Regarding the websites adoption of quality control mechanisms and their provided services, table 5 shows that in 2016 websites have a more specific orientation in quality mechanisms depending on the type of the provided services. More specifically, in 2016 websites focusing on microtasks (mwk), crowdfunding (crf), software development (pdvp) and web & graphic design tasks (dsns) offer only basic tools of quality control and cheat detection, namely via publishing the worker’s history and

job success. The review also indicates that, in 2016 compared to 2014, websites focusing on mobile crowdsourcing have stopped providing trivial quality assurance mechanisms at all and have moved on using other forms. This may indicate that such kinds of mechanisms are ineffective in such tasks.

Table 5. The relationship of Website quality assurance mechanisms and its provided type of services over time

Quality Mechanims	Type of Services (%)										
	mwk	crf	mcw	cntg	dte	hts	pdvp	dsns	trs	rtp	
2014	qr2	42.1	16.7	50	55.6	50	33.3	71.4	71.4	0	33.3
	qr3	26.3	33.3	25	22.2	25	66.7	14.3	42.9	33.3	16.7
	qr4	31.6	0	50	33.3	41.7	11.1	0	0	33.3	33.3
	qr5	47.4	50	50	55.6	58.3	0	28.6	14.3	33.3	50
2016	qr2	60	42.1	0	28.6	50	55	25	64.3	50	50
	qr3	70	78.9	0	28.6	50	55	75	85.7	25	50
	qr4	50	0	0	14.3	0	20	25	21.4	50	0
	qr5	60	36.8	0	42.9	0	35	25	21.4	25	0

Source: Dataset with results drawn from Alexa.com. Author’s calculations.

Note: A website may provide more than one type of service and have more than one quality control mechanism.

Unit is the percentage of websites that meet the current criteria.

The analysis of the crowdsourcing websites reveals that these underwent many changes in a rather short period. This may indicate that the field of crowdsourcing is still evolving trying to find ways to properly offer and assess such kind of services.

Empirical Evaluation

Overview

In this section we present a way to model the Alexa rankings of online crowdsourcing platforms, using the characteristics of websites which have been discussed previously. The aim of this model is to investigate the determinants of a website’s Alexa traffic rankings and to what extent affects its attention seeking. The model might provide some useful insights related to the popularity of websites and suggest ways for improvement. Towards this, a OLS linear regression model is used, which applied to each ranking separately. Nevertheless, we also wanted a more comprehensive picture of predictors’ effect on the response variable, in order to investigate, if their effect exists and to what extent, on low and high values of the alexa ranking list and be compared to their effect on median. For this reason, we also used for our estimation a quantile regression. This allows comparing how some quantiles (25%,50%,75%) of the rankings, may be more affected by our determinant-factors, than others.

Empirical Model

Based on our methodology, the econometric specification in our study is of the following general form:

$$AR_{16i} = \alpha + \beta AR_{14i} + \gamma QA_i + \delta R_i + \varepsilon S_i + \zeta OI_i + u_i \quad (1)$$

Where AR_{16i} is the logarithmic values of the aforementioned alexa ranking in reversed values (dependent variable) for the i^{th} online platform for the year 2016. Factor AR_{14i} in the model is the vector containing the logarithmic values of rankings in reversed values of the i^{th} online platform for the year 2014. QA_i is a vector which includes the quality assurance mechanisms of the i^{th} online platform (No quality assurance mechanisms, Reviews & Ratings, Workers' Profile, Skills & Practice Tests, Spamming Detector Tools), R_i is also a vector for region of the i^{th} online platform (North America, Europe, Australia, Asia) and S_i is a vector of the services being provided by the i^{th} online platform (Microtasks, Crowdfunding, Tasks on Mobile, Content Generation, Data Entry, High Tech, Program Developing, Graphic Design, Translation, Reviews & Testing Products). Last but not least, OI_i is a vector of the online imprint of the i^{th} online platform (Social Networks, Video Streaming-Sharing, Blogs and Forums). u_i is the disturbance term.

Estimation Results

The OLS linear regression model and the coefficient for the 25th, 50th and 75th quantiles are shown in table 6. In the next section, each result of each model is discussed in more detail.

OLS Regression

The results suggest that there are remarked differences across the distribution of websites' Alexa rankings with respect to the independent/explanatory variables.

Recall, our dependent variable (i.e. crowdsourcing websites' alexa rankings in year 2016) and crowdsourcing websites' alexa rankings in year 2014 are in reversed values based on the equation; reversed values = maximum value + minimum value - initial values. We did so, in order to have a better interpretation with the regression outputs. In particular, now, higher values of our dependent variable reflect higher levels of online website traffic.

Thus, the first column in table 6 shows the coefficients for the OLS linear regression model. The results show that several independent variables are statistically significant. For example, a website's Alexa ranking of previous years (i.e. 2014) is a strong indicator of its future traffic (at the 1% level of significance and the right hand variable is positively relate to the left hand variable). Concerning the group variables of "quality assurance mechanisms", the reference group was websites with that did not disclose any quality control and cheat detection mechanisms. The results show that websites providing mechanisms for spam detection and reviews & ratings for

their workers have a high Alexa ranking compared with the websites which belong to the reference group ($r=1.330$, $p.value = .098$ and $r= .450$, $p.value = .030$ respectively). Furthermore, having as a reference group websites from North America, the model indicates at the 5% level of significance that websites originated from Asia have an approximately 60% lower average traffic compared with crowdsourcing websites from North America. The OLS coefficients for the various types of services offered show some interesting results. Having as a reference group online platforms providing microworks and simple tasks (i.e. tasks that are performed in parallel by large, paid crowds in a short time), we find that the effect of websites' providing service vary on their traffic and therefore in their position in the Alexa rankings.

Table 6. The determinants of the 2016 Alexa Rankings for the reviewed crowdsourcing online platforms.

	OLS	Quantile Regression		
	Mean	0.25	0.50	0.75
Independent variables	[1]	[2]	[3]	[4]
Constant	1.602*** (.551)	1.110*** (.142)	.872* (.567)	1.841*** (.532)
Alexa ranking (year 2014)	.920*** (.045)	.969*** (.036)	.999*** (.044)	.924*** (.042)
<i>Websites' Quality Assurance</i>				
Reviews & Ratings	1.330* (.810)	.808** (.423)	1.540 (2.912)	2.179* (1.262)
Workers' Profile	-.275 (.200)	.218** (.097)	-.223 (.224)	.390* (.223)
Skills & Practice Tests	-.175 (.213)	.161 (.213)	-.008 (.281)	.790** (.352)
Spamming Detector Tools	.450** (.206)	.390*** (.112)	.220 (.260)	.477** (.210)
<i>Websites' Region</i>				
Europe	.079 (.166)	.152*** (.019)	.194 (.179)	-.149 (.183)
Australia	.282 (.278)	.737* (.409)	.256 (.246)	-.469 (.519)
Asia	-.599** (.276)	-.346 (.299)	-.446 (.450)	-.705* (.401)
<i>Websites' Type of Services/Tasks</i>				
Crowdfunding Tasks	-.781*** (.175)	-.522** (.274)	-.626*** (.207)	-.635*** (.221)
Tasks on Mobile Applications	-1.415*** (.296)	-1.408** (.598)	-1.385*** (.268)	-1.651*** (.335)
Content Generation Tasks	.351 (.302)	1.012 (.798)	.465 (.490)	.267 (.014)
Data Entry Tasks	-.093 (.449)	-.156 (1.478)	-.281 (.799)	.304 (.404)
High Tech Tasks	-.531** (.225)	-.896** (.283)	-.600* (.333)	-.165 (.244)
Programm Developing Tasks	.400 (.638)	.545 (.703)	-.212 (5.020)	-.596 (1.548)
Graphic Design Tasks	-.195 (.254)	-.213 (.254)	-.067 (.239)	.029 (.361)
Translation Tasks	-.137 (.197)	.025*** (.198)	-.088 (.206)	.262 (.399)

Reviews & Testing Products Tasks	.288 (.389)	.185* (.105)	-.068 (.387)	.985 (.758)
<i>Websites' Online Imprint</i>				
Social Networks	.267 (.327)	.535** (.270)	.107 (.455)	.549* (.410)
Video Streaming-Sharing	.611* (.327)	.080 (.455)	.441* (.303)	.679 (.794)
Blogs and Forums	.048 (.230)	-.506 (.234)	-.199 (.316)	.297 (.276)
R-squared	.861	.684	.653	.670
Observations	174	174	174	174

Source: Dataset with results drawn from Alexa.com. Author's calculations.

Notes: Dependent variable: Websites' Alexa Rankings for year 2016. In parentheses heteroskedasticity corrected standard errors.

For Websites' Quality Assurance categorical variable the reference group is websites with no quality assurance mechanisms, for Websites' Region categorical variable the reference group is websites located in North America, for Websites' Type of Services/Tasks categorical variable the reference group is websites providing Microtasks/Simple Tasks Services and for Websites' Online Imprint categorical variable the reference group is websites with no online imprint.

Statistical significance: *** 1%, ** 5% and * 10%.

For example, a crowdfunding online platform has an approximately 78% lower traffic than a microtask website, with the result being strongly significant at the 1% level of significance. Similarly, websites having tasks on mobile or tasks requiring high-tech knowledge have lower traffic than the reference group (at the 1% and 5% level of significance respectively). Last but not least, regarding the impact of a website's digital marketing strategy (online imprint) on its traffic, the study revealed a weak effect of video sharing communities. More specifically, crowdsourcing online platforms that have a strong profile in video sharing communities such as Youtube and Vimeo, have 61.1 % more traffic than the ones with no profile on such networks.

Quantile Regression

The quantile regression makes possible to statistically examine the extent to which a website's quality assurance mechanisms, types of services provided, location and online imprint differ when considering the distribution of their traffic. The column 2, 3 & 4 in table 6 presents the regression for the 25th, 50th and 75th quantiles respectively. Recall that quantile regressions were estimated to determine if the impact of the above independent variables varied for crowdsourcing websites at different points in the distribution of their average Alexa rankings in 2016. The results of the quantile regressions show primarily that a website's traffic in previous years is a strong indicator for its traffic in the near future, regardless its position (low or high) in the Alexa list. Moreover, concerning the websites' quality assurance mechanisms, the quantile regressions not only confirm the OLS results, but also reveal that among

low-ranked websites (at the 25th quantile) another mechanism called “workers’ profile” have a strong effect on their traffic (at 5 % level of significance and the right hand variable is positively relate to the left hand variable). Similarly, among high-ranked websites (at the 75th quantile), the opportunity of a requester of a crowdsourcing task to conduct skills and practice tests in order to decide which workers to hire, has a notable effect on the online platform’s general traffic at 5 % level of significance. The quantile regressions also reveal that among low-ranked websites, a website that originated in Europe has 15.2% more traffic than websites originated in North America at 1% level of significance ($r = .152$ and $p \text{ value} = .000$), while among high-ranked websites a website that originated in Asia has 70.5% less traffic than websites originated in North America at 10% level of significance ($r = -.705$ and $p \text{ value} = .081$). Additional evidence reveals that the effect of a website’s type of service differs among the distribution of the alexa rankings. For example, among websites in low traffic positions, a website that provides translation tasks or reviews and testing products achieves more traffic when compared to websites offering microtasks, while websites with crowdfunding, high-tech services of tasks on mobile have in general less traffic than the reference group. Finally, with regard to the websites’ online imprint, quantile regression does not confirm the OLS concerning the effect of a strong website’s profile on video sharing communities on Alexa ranking, because at the 25th quantile and at 75th quantile this effect is substituted from a strong website’s profile on social networks such as Facebook, Twitter, LinkedIn etc. (at 5% and 10% level of significance and the right hand variable is positively relate to the left hand variable respectively).

Discussion

The tools that a crowdsourcing website offers to the potentially requesters, the region where it is located, the type of tasks it accepts and the strategies for efficient digital marketing that applies, can have important effects on its recognition and popularity. The abovementioned characteristics of a crowdsourcing platform affect the decision of a requester, concerning the participation and accomplishment of a crowdsourcing job in this platform. Obviously, a high-attracted crowdsourcing online platform has major indexes of traffic and as an extent a high position in alexa ranking list. As a consequence, highly-ranked websites may have more crowdsourcing sessions when compared to low-ranked sites, giving them the opportunity to respond efficiently to their competitive operating environment.

Conclusions

The purpose of this research is twofold. First, it aimed at investigating how crowdsourcing websites have evolved in terms of their characteristics and services offered in a period of two years (2014-2016).

Second, it attempted to devise a theoretical framework to investigate the effect of the websites’ characteristics on their ranking as these are reported by Alexa. The

theoretical evaluation's purpose is to investigate and understand how the individual characteristics of crowdsourcing websites influence their traffic and popularity. Give the dynamics of the field in recent years, such results may offer insights in improving their services and operation.

References

- [1] Brabham, C. Daren "Crowdsourcing as a model for problem solving an introduction and cases." *Convergence: the international journal of research into new media technologies* 14.1 (2008): 75-90.
- [2] Poetz, K. Marion, and Martin Schreier. "The value of crowdsourcing: can users really compete with professionals in generating new product ideas?." *Journal of Product Innovation Management* 29.2 (2012): 245-256.
- [3] Jeff Howe, "The rise of crowdsourcing." *Wired magazine* 14.6 (2006): 1-4.
- [4] Vaughan, Liwen, and Rongbin Yang. "Web traffic and organization performance measures: Relationships and data sources examined." *Journal of informetrics* 7.3 (2013): 699-711.
- [5] Evangelos Mourelatos and Manolis Tzagarakis. "Worker's Cognitive Abilities and Personality Traits as Predictors of Effective Task Performance in Crowdsourcing Tasks." *PQS 2016 5th ISCA/DEGA Workshop on Perceptual Quality of Systems*. 2016.
- [6] Kaufmann, Nicolas, Thimo Schulze, and Daniel Veit. "More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk." *AMCIS*. Vol. 11. 2011.
- [7] Donmez, Pinar, Jaime G. Carbonell, and Jeff Schneider. "Efficiently learning the accuracy of labeling sources for selective sampling." *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009.
- [8] Matthias Hirth, Tobias Hoßfeld, and Phuoc Tran-Gia. "Anatomy of a crowdsourcing platform-using the example of microworkers. com." *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2011 Fifth International Conference on*. IEEE, 2011.
- [9] Mourelatos, E., Tzagarakis, M. and Dimara E. "A Review of Online Crowdsourcing Platforms", *South-Eastern Europe Journal of Economics*, 14 (1), 2016.
- [10] Eric Schenk and Claude Guittard. "Towards a characterization of crowdsourcing practices." *Journal of Innovation Economics & Management* 1 (2011): 93-107.

- [11] Wang, Jing, Panagiotis G. Ipeirotis, and Foster Provost. "Managing crowdsourcing workers." The 2011 winter conference on business intelligence. 2011.
- [12] Ross, Joel, et al. "Who are the crowdworkers?: shifting demographics in mechanical turk." CHI'10 extended abstracts on Human factors in computing systems. ACM, 2010.
- [13] Rosemary Thackeray, B. L Neiger, C. L. Hanson & J. F. McKenzie, "Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media." Health promotion practice 9.4 (2008): 338-343.
- [14] Hansen, D. L., Schone, P. J., Corey, D., Reid, M., & Gehring, J. (2013, February). Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at familysearch indexing. In Proceedings of the 2013 conference on Computer supported cooperative work (pp. 649-660). ACM.
- [15] Hossain, M. (2012, May). Crowdsourcing: Activities, incentives and users' motivations to participate. In Innovation Management and Technology Research (ICIMTR), 2012 International Conference on (pp. 501-506). IEEE.
- [16] Alec. Lynch, "Crowdsourcing is booming in asia." TechCrunch, 8Dec (2012).
- [17] Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H. R., Bertino, E., & Dustdar, S. Quality control in crowdsourcing systems. IEEE Internet Comput, 17(2), (2013) :76-81.
- [18] Vaughan, L and Yang, R. (2013). Web Traffic and Organization Performance Measures: Relationships and Data Sources Examined. Journal of Informetrics 7(3):699-711.