The Impact of Workers’ Skills on the Quality of Work in Crowdsourcing Environments: A laboratory study

Panagiotis Tsaktsiras

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in Applied Economics & Data Analysis

School of Business Administration
Department of Economics

Master of Science in “Applied Economics and Data Analysis”

September 2017
Three-member Dissertation Committee

**Research Supervisor:** Manolis Tzagarakis  
Assistant Professor

**Committee Member:** Nicholas Giannakopoulos  
Associate Professor

**Committee Member:** Dimitrios Tzelepis  
Assistant Professor

The present dissertation entitled

«The impact of workers’ skills on the quality of work in crowdsourcing environments: A laboratory study»

was submitted by **Panagiotis Tsaktsiras, Sid 1017042**, in partial fulfillment of the requirements for the degree of Master of Science in «Applied Economics & Data Analysis» at the University of Patras and was approved by the Dissertation Committee Members.
I would like to dedicate my dissertation to my wife Helen, for all of her love and support.
Acknowledgments

I would like to thank my supervisor, Manolis Tzagarakis for his guidance and support. You have given an example of excellence as an instructor.

I would like to thank PhD candidate, Vaggelis Mourelatos for all of his guidance through this process; your comments and discussion were absolutely invaluable.

I would also like to thank my dissertation committee members.

I would especially like to thank my amazing family for the love, support, and constant encouragement I have gotten over the years. In particular, I would like to thank my wife Helen, my parents and my brother John.

I would also like to thank all my fellow graduates.

Last but not least, I would like to thank my beloved grandparents Panagiotis and Helen.
Summary

Crowdsourcing is a type of participative online activity in which an individual proposes to a group of individuals of varying knowledge, heterogeneity, and number, the voluntary undertaking of a task.

Even though the application of crowdsourcing services has been increased dramatically over the years, one traditional problem that exist is the relatively low quality of labor received by users.

This is mainly because more and more workers - mostly from developing countries - guided through their money incentives, take on tasks that fail to carry out or poorly process them.

In this dissertation, we aim to examine which characteristics of workers, and especially which skills, affect (and to what degree) the quality of work in a crowdsourcing environment.

Keywords: Crowdsourcing, Online Task, Cognitive Skills, Personality Traits, Quality of Work
Περίληψη

Με τον όρο «πληθοπορισμός» (crowdsourcing) ορίζεται η διαδικασία ανάθεσης μίας εργασίας σε ένα εξωτερικό, ανώνυμο πλήθος ατόμων μέσω ανοικτής πρόσκλησης.

Αν και η χρήση υπηρεσιών πληθοπορισμού έχει αυξηθεί ραγδαία τα τελευταία χρόνια, ένα παραδοσιακό πρόβλημα που ακόμη αντιμετωπίζουν τέτοιου είδους υπηρεσίες είναι η σχετικά χαμηλή ποιότητα της εργασίας που λαμβάνουν οι χρήστες.

Αυτό οφείλεται στο γεγονός ότι αρκετοί εργάτες - κυρίως από αναπτυσσόμενες χώρες - έχοντας ως κίνητρο το χρηματικό ποσό, δέχονται εργασίες τις οποίες είτε δεν μπορούν να ολοκληρώσουν είτε τις εκτελούν πλημμελώς.

Στόχος της διπλωματικής αυτής εργασίας είναι να μελετήσει ποια χαρακτηριστικά εργατών, και ειδικότερα ποιες δεξιότητες, έχουν επίδραση -και ποιες- στην ποιότητα εργασίας που υποβάλλεται σε περιβάλλοντα πληθοπορισμού.

Λέξεις κλειδιά: Πληθοπορισμός, Διαδίκτυος Εργασία, Γνωστικές Δεξιότητες, Προσωπικά Χαρακτηριστικά, Ποιότητα της εργασίας
Contents

1 Introduction

2 Literature Review
  2.1 Crowdsourcing
  2.2 Quality of Work
  2.3 The Big Five Model

3 Data
  3.1 Objective Tasks
  3.2 Experimental Design
  3.3 Descriptive Statistics
    3.3.1 Demographics
    3.3.2 Cognitive Skills
    3.3.3 Non-cognitive Skills
    3.3.4 Social Economic
    3.3.5 Work Effort

4 Empirical Model

5 Estimation Results
  5.1 OLS regression
  5.2 Discussions

6 Conclusions

References
List of Figures
List of Tables
Chapter 1

Introduction

Crowdsourcing is a type of participative online activity in which an individual suggests to a group of individuals of different knowledge, heterogeneity, and number, the voluntary undertaking of a task.

The term of this phenomenon was initially appeared in 2005 by Jeff Howe and Mark Robinson, editors at Wired, in order to delineate how businesses and organizations were using the Internet to "assign work to the crowd". Howe was the first to issue a definition of crowdsourcing.

Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed in co-operation), but is also often undertaken by unique individuals. The basic prerequisite is the use of open calls and the large network of potential workers (Howe 2006).

Through the existing literature it is well known that online labor markets have some great advantages over the laboratory experiments and the traditional field experiments. In online labor markets, it is possible for workers from around the world to participate in tasks that can complete from a distance. In these markets, it is also possible to employ
a huge number of diverse workers ready to engage in such tasks, with much less knowledge of experiments and the way they work, compared to the exact same workers in laboratories.

**Figure 1.1:** Where does crowdsourcing stands in Open Source

While advantages above surely are some significant features, the main advantage of online labor markets is that they are fully controlled, from individual payments to being able to exclude users with invalid accounts within the market and prevent workers from communicating with other workers in the same market. Today, it is often used for specific types of projects such as translation services, microtasks, image tagging and transcription. (Estellés-Arolas et al. 2012).

Crowdsourcing has received in recent years the interest of researchers in various fields that aim to analyze, comprehend, assess and even improve this new form of labor and finally find strategies and frameworks in order to increase the quality of the work being done in high levels (Howe 2008).

An overview of the general principles of crowdsourcing aimed towards achieving high quality of work is given by existing literature (Yuen et al. 2011). Since online labor market and crowdsourcing are something
relatively new, the quality of workers output is currently under thorough investigation.

This is the main reason why many studies try to recommend a variety of models and metrics to evaluate and secure the quality of work in such contexts, as best possible. Based on the existing models, there are two specific approaches to achieve quality results; an approach focusing on the standards and the planning of the submitted task and an approach structured by the profile of individual workers.

The purpose of this dissertation is to examine which characteristics of workers, and especially which skills, affect (and to what degree) the quality of work in a crowdsourcing environment. Towards this, we conducted an experiment.

In this study we focus on the first technique by conducting an experiment in laboratory controlled environment and by using regression analysis we try to estimate several factors affecting the quality of workers’ work.

This dissertation is organized as follows. In chapter 2, we included the literature review about Crowdsourcing and the determinants of quality work in online tasks, based on online workers skills. In chapter 3, we present the data processing and the descriptive analysis of the dataset. Chapter 4 includes our empirical analysis followed by the estimation results in chapter 5 and the conclusion of our research in chapter 6.
Chapter 2

Literature Review

2.1 Crowdsourcing

Businesses, nonprofit organizations, and government agencies regularly integrate the creative energies of online communities into day-to-day operations, and many organizations have been built entirely from these arrangements. This deliberate blend of bottom-up, open, creative process with top-down organizational goals is called crowdsourcing (Brabham 2013).

Online labor markets like crowdsourcing environments are an increasingly method for individuals and organizations to find and hire workers with specific skills, in order to finish a vast variety of tasks in many aspects of modern life, like marketing, accounting, finance, programming, office work and many more.

Crowdsourcing is also very useful for businesses to gather market information, find solutions to problems or finish some of their tasks. The process of crowdsourcing is often used to subdivide tedious work by combining the efforts of numerous self-identified volunteers or part-time workers, where each contributor of their own initiative adds a small portion to the greater result (Howe 2006).
2.1 Crowdsourcing

As it seems, online communities are prolific fountains of novelty and innovation. As a result, there are more and more researches on how crowdsourcing is working and why over the years. Although there has been a significant growth in empirical research about crowdsourcing, scholars are still write about this phenomenon free of these findings. This is mostly because there is a big variety of differing definitions and interpretations of crowdsourcing. It is not easy to exploit what empirical researchers have find out about crowdsourcing.

Crowdsourcing as a definition, consists of three specific parts: an open call, the crowd and a task.

First, an open call is an invitation to participate in a crowdsourcing project (Estellés Arolas & González-Ladrón-de-Guevara, 2012). Therefore, it is obvious that everybody is invited to follow that invitation and that no restriction or preselection criteria limit participation in the task (Aitamurto, Leiponen & Tee, 2011).

Second, the crowd which is characterized as a large group of people (Estellés Arolas & González-Ladrón-de-Guevara, 2012) from the majority of academic literature. Also, it is widely agreed that there should be heterogeneity in crowd’s characteristics, like skills and knowledge (Selzer & Mahmoudi, 2013).

Third, there is a vast variety of tasks that workers should complete. It can range from simple copywriting, copy editing or other micro tasks to the evolution of a new product. No matter the task that the crowd is demanded to find a solution, it is essential to have a clear objective (Estellés-Arolas & González-Ladrón-de-Guevara, 2012).
2.1 Crowdsourcing

There is a mutual benefit from crowdsourcing for both individual-worker and individual-requester. Through the crowdsourcing process the company gets access to ideas, innovations, information and external knowledge, which it uses to generate value (Estellés-Arolas & González-Ladrón-de-Guevara, 2012).

It is essential to understand that crowdsourcing as a phenomenon is not one-dimensional. According to (Howe, 2006), crowdsourcing consists of four specific strategies, which are:

1. Crowd Wisdom – uses crowd’s knowledge for specific tasks
2. Crowd Creation – uses crowd’s creativity to develop innovative concepts
3. Crowd Voting – uses crowd’s decisions to organize huge pieces of information
4. Crowd Funding – uses crowd’s shared wallet for new projects or product

**Figure 2.1:** Architecture of crowdsourcing
2.2 Quality of Work

In the literature, work quality has been related to crowd demographics (Ross, Irani, & Silberman, 2010; Sheng, Provost, & Ipeirotis, 2008), contributors’ gender, profession and age (Downs, Holbrook, Sheng, & Cranor, 2010), and other worker characteristics (Kazai, Kamps, & Milic-Frayling, 2011). So it is obvious that the expected quality of work is promptly affected by crowd features.

It is therefore very important for individuals, businesses and policymakers to investigate and understand which skills or abilities can determine high quality of results in the labor market.

On one hand, cognitive ability is the single most important determinant of labor market outcomes (Herrnstein and Murray, 1994). On the other hand, a different side supports that noncognitive skills such as persistence, motivation, emotional stability and social skills are equally or more important than cognitive skills (Heckman, Stixrud and Urzua 2006).

Cognitive skills involve conscious intellectual effort, such as thinking, reasoning, or remembering. In order to measure these skills workers require to demonstrate their capabilities in many areas.

Noncognitive or “soft skills” are related to motivation, integrity, and interpersonal interaction. They may also involve intellect, but more indirectly and less consciously than cognitive skills. Soft skills are associated with an individual’s personality, temperament, and attitudes. For virtually all jobs, a worker needs the soft skills associated with working well with other people and functioning effectively in a work environment.

There is not clear evidence in the existing literature in favor of either side. A large part of the literature validates that the measures of the cognitive skills (like IQ) are robust predictors of the outcome of the la-
2.3 The Big Five Model

The Big Five Framework of personality traits (Costa & McCrae, 1992), also known as Five Factor Model (FFM), has burst as a widely approved model for understanding the robust relationship between the personality of an individual and assorted academic behaviors (Poropat, 2009). It consists of five factors which are:

1. Openness to experience
2. Conscientiousness
3. Extraversion
4. Agreeableness
5. Neuroticism

Labor market and therefore, the quality of work. The estimated effect of noncognitive skills in quality of work varies significantly in the literature. However, it is not an easy task to compare these two sides, since the importance of noncognitive skills is difficult to be taken into account, due to shortage of legitimate measures.

Most researches in economics measure personality traits through self-reported questionnaires. Compared to IQ tests, such measures are less reliable and less precise (Borghans et al., 2008b). Furthermore, the evaluation of cognitive and noncognitive skill is probable to diverge from one sector to another.

More recently, economists have started to focus on the importance of non-cognitive skills in determining earnings (Heckman and Kautz, 2012). Recent studies have linked job performance and wages to the so-called 'Big Five' personality traits: openness, conscientiousness, extraversion, agreeableness and neuroticism (Heckman et al., 2006).
2.3 The Big Five Model

Openness is reflected in a strong intellectual curiosity and a preference for novelty and variety. Conscientiousness is exemplified by being disciplined, organized, and achievement-oriented. Extraversion is displayed through a higher degree of sociability, assertiveness, and talkativeness. Agreeableness refers to being helpful, cooperative, and sympathetic towards others. Finally, neuroticism refers to degree of emotional stability, impulse control, and anxiety.
Chapter 3

Data

3.1 Objective Tasks

The aim of the research mentioned in this study is to jointly assess the actual responses to the tasks as well as to investigate the impact of the following groups of variables on the quality of the results provided by the workers.

In respect to the cognitive skills of each individual, we sort workers in accordance with their computer skills and english language skills (Campbell et al., 2001).

Additionally, we capture the non-cognitive skills of each individual by the Big Five Personality Test variables (Borghans et al., 2008). Finally, the work effort done by the workers is approached by the time of completion of the task of the workers and the number of the submitted keywords, before the final submission (Heckman et al., 2012).
3.2 Experimental Design

In order to research better the main factors that affect the quality of results of the workers’ work, we firstly describe the structure of the experiment. The experiment was pertained to the task of image tagging, a very common and widespread microtask in crowdsourcing online platforms.

In particular, we created a session to the laboratory of Department of Economics in the University of Patras. Workers were freshmen students of the Department and their motivation to deliver distinguished results was the enhancement of their grade in a specific semester’s class.

Subsequently, we asked the workers to observe very carefully some pictures of random content and then to move on to image tagging. Image tagging is the process of accurately identifying image characteristics as well as identifying the category under which an image may fall. Workers had to provide us a number of keywords for every image of the experiment.

For the purpose of estimating and evaluating the result from each individual specific, we had sent the image database of the task -the total of the pictures that were shown to the subjects in the experiment- to professional photographers in order to send us all the potential keywords that a single picture can include.

The experiment begun with participants completing a demographic questionnaire. The demographic questionnaire was consisted of specific queries about their gender, their age, their country and city of birth, their country of residence -which was the same for all subjects- and their city of residence and their ethnicity.
3.2 Experimental Design

Figure 3.1: A typical online labor market workflow

Afterwards, workers were asked to take a self-evaluation test regarding their computer skills, their level of English language competence and their education level. Eventually, they filled out a questionnaire referring to their personality. This test had the form of a five-level Likert scale, which is a psychometric scale commonly involved in scholar research.

Ultimately, workers were asked to proceed to the execution of the experiment. Every individual had to fill in keywords for six specific images and then proceed to submitting the final results. Each subject had the possibility to run this experimental task once. In that way, we avoided having workers that had run the task more than once and therefore acquiring experience and familiarization with the concept of the task.

In order to get results for the individual quality of performance, we compare participant workers’ outcomes with professional photographers’ results (golden standard).
3.3 Descriptive Statistics

To begin with, it is worth noting that the average quality of the results (performance) was close to 73%, with a maximum value close to 97% and the minimum close to 46%, which may indicate the existence of extreme observations (outliers) in our dataset.

To proceed with our analysis we obtained 98 answers from our workers based on our laboratory experiment. Therefore, we derived our data from the experiment containing information about various variables in relation to the workers participating to the task. Afterwards, we sorted the variables into large skill groups:

- **Demographics** (age, female)

- **Cognitive skills** (computer & english competency)

- **Non-cognitive skills** (extraversion, agreeableness, consciousness, emotion stability & openness)

- **Work effort** (sum of total tagged keywords & time of task completion)

- **Social economic** (poor status & good health)

This was essential in order to set up our data to proceed with our further analysis. Table 1, presents our variables, which we used in our analysis. Quality of performance is our dependent variable.
Table 3.1. Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of performance</td>
<td>.731</td>
<td>.1159</td>
<td>.459</td>
<td>.969</td>
</tr>
<tr>
<td>Age</td>
<td>18.6</td>
<td>1.76</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>Female</td>
<td>.510</td>
<td>.502</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Computer competence</td>
<td>2.591</td>
<td>1.208</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>English competence</td>
<td>3.326</td>
<td>1.043</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Sum of totals</td>
<td>33.29</td>
<td>5.908</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>Completion time</td>
<td>13.44</td>
<td>3.774</td>
<td>6</td>
<td>28</td>
</tr>
<tr>
<td>Poor status</td>
<td>2.265</td>
<td>.650</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Good health</td>
<td>.612</td>
<td>.489</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Extraversion</td>
<td>27.52</td>
<td>4.367</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>34.35</td>
<td>4.865</td>
<td>22</td>
<td>45</td>
</tr>
<tr>
<td>Consciousness</td>
<td>31.33</td>
<td>5.050</td>
<td>20</td>
<td>43</td>
</tr>
<tr>
<td>Emotion stability</td>
<td>23.80</td>
<td>4.862</td>
<td>13</td>
<td>35</td>
</tr>
<tr>
<td>Openness</td>
<td>35.14</td>
<td>4.588</td>
<td>25</td>
<td>45</td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations. 
Notes: Number of observations equals to 98. Age is measured in years. Task completion time is measured in minutes. Poor status is a categorical variable and good health a binary variable.

3.3.1 Demographics

When it comes to independent variables, we observe that the mean of the age of the workers was extremely close to the minimum value (18.6), suggesting that the majority of the participants were freshmen students of the Department of Economics. Also, almost half of the participants were female (51%).

3.3.2 Cognitive Skills

Regarding the cognitive skills group, figures 3.2 and 3.3 shows in our experiment that computer competence levels and english competence levels differ among participant workers. Furthermore, it is important to mention that, since participants were freshmen students of the Department of Economics is Patras University, educational level is not essential for our research.
Most of the participants underlined that their computer level was "basic" (category 2), while the majority of them marked their English competence as "Intermediate" (category 3) and "Advanced" (category 4).

**Figure 3.2:** Computer competence histogram

**Figure 3.3:** English competence histogram
3.3 Descriptive Statistics

As a result, in order to capture better the effect of a worker’s computer and english competence level on his task performance, we generate a new dummy variable comlevD, (value 1 : includes workers with at least professional computer competence) and a dummy variable en-glevD, (value 1: includes workers with at least advanced english competence).

In table 3.2 we observe that only 11 out of 98 participant workers have at least professional computer competence, with a mean score of almost 90% (std. dev. .053) and minimum score of 80% and maximum score of 97%.

The rest of participant workers (which is the majority of them, 87 out of 98) have a mean score of quality of performance of 71% (std. dev. 105) with minimum score of 46% and maximum score of 91%.

<table>
<thead>
<tr>
<th></th>
<th>Quality of Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Obs</td>
</tr>
<tr>
<td>Low Computer Competence</td>
<td>87</td>
</tr>
<tr>
<td>High Computer Competence</td>
<td>11</td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations.

It is clear that workers with high computer competence attribute better quality in their performance in average (89%) as compared with the workers with lower computer competence (71%).

In table 3.3 we observe that participant workers are splited on almost two groups regarding their english competence level. 43 of workers state that they have at least advanced english competence level with a mean value of 78% (std. dev. .1) and minimum score of 46% and maximum score of almost 97% among them.
On the contrary, the rest of the workers (55 out of 98) have a lower level of english competence with a mean score of 70% (std. dev. .110) in their quality of performance and a minimum value of 47% and a maximum value of 91%.

<table>
<thead>
<tr>
<th></th>
<th>Quality of Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low English Competence</td>
<td>Obs 55 Mean .690 Std. Dev. .110 Min .470 Max .911</td>
</tr>
<tr>
<td>High English Competence</td>
<td>Obs 43 Mean .784 Std. Dev. .100 Min .459 Max .969</td>
</tr>
</tbody>
</table>

Table 3.3. *English level binary variable decomposition*

It is again clear that higher English level means better quality of performance in average terms (78.4% for high English level against 69% for low English level workers).

### 3.3.3 Non-cognitive Skills

Regarding non-cognitive skills, for our analysis we used the Big Five Personality Test. This test consists of fifty personality traits that you must rate on how true they are about you, on a five point scale where 1=Disagree, 3=Neutral and 5=Agree (Goldberg et al. 1992).

For our crowdsourcing environment experiment, we examined the workers under the variables of extraversion (a worker’s degree of sociability, assertiveness, and talkativeness), agreeableness (a worker’s degree of being helpful, cooperative, and sympathetic towards others), conscientiousness (a worker’s degree of being disciplined, organized, and achievement-oriented), emotional stability versus neuroticism (a worker’s degree of emotional stability, impulse control, and anxiety) and openness to experience (a worker’s degree of intellectual curiosity and a degree of being inventive and creative) (John et al. 1999).

Figure 3.4 shows the distribution of responses to the Big Five Per-
For extraversion, we observe a mean value of 27.52 (std. dev. 4.367) which indicates a slight inclination to talkativeness, while for openness we observe a mean value of 35.14 (std. dev. 4.588) which expound no specific tendency to this trait.

In respect to agreeableness trait, we observe a mean value of 34.35 (std. dev. 4.865) which indicates that participant workers tend to be cooperative and helpful, while for consciousness trait, we observed a lower mean score of 31.33 (std. dev. 5.050) suggesting that we really should not expect a highly detailed job from workers in our task.

Finally, for emotion stability over neuroticism, we observe a mean value of 23.8 (std. dev. 4.862) which suggests an anxious nature of the workers.

**Figure 3.4:** Workers' big five personality test distribution
3.3.4 Social Economic

Social economic factors are elements of lifestyle and measures of economic viability and social status. They directly affect the social privilege and levels of economic independence. Factors such as health condition and wealth status are concluded in this research as to how each one influences participant workers behavior and circumstances.

So, regarding social economic variables and more specifically about poor status, in table 3.4 we observe that the majority of participant workers (51%) are poor or have been poor sometime in their life, while only 38% of them never have been poor in their life. Also, 11 workers declare themselves as poor, right now.

<table>
<thead>
<tr>
<th>Poor Status</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Now</td>
<td>11</td>
<td>11.22</td>
<td>11.22</td>
</tr>
<tr>
<td>All the times or sometimes</td>
<td>50</td>
<td>51.02</td>
<td>62.24</td>
</tr>
<tr>
<td>Never</td>
<td>37</td>
<td>37.76</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>98</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations.

Regarding the quality of performance, in table 3.5 we observe that there is a positive correlation between poor status and quality of performance in our task. More specifically, workers who have never been poor in their life have a mean score of 76.2%. Workers who have been poor all the time or sometimes in their life have a mean score of 72% in their quality of performance, while workers who are poor the day they took over the task have a mean score of 68% in their quality of work.
3.3 Descriptive Statistics

Table 3.5. Poor status and quality of performance

<table>
<thead>
<tr>
<th>Quality of Performance</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Now</td>
<td>11</td>
<td>.681</td>
<td>.158</td>
<td>.470</td>
<td>.896</td>
</tr>
<tr>
<td>Sometimes or all the times</td>
<td>50</td>
<td>.719</td>
<td>.103</td>
<td>.459</td>
<td>.906</td>
</tr>
<tr>
<td>Never</td>
<td>37</td>
<td>.762</td>
<td>.112</td>
<td>.5</td>
<td>.969</td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations.

Furthermore, in table 3.6 we perceive that almost 62% of the participant workers have good health status while 38% has not.

Table 3.6. Good health variable tabulation

<table>
<thead>
<tr>
<th>Good Health</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>38</td>
<td>38.78</td>
<td>38.78</td>
</tr>
<tr>
<td>Yes</td>
<td>60</td>
<td>61.22</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>98</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations.

Regarding the quality of performance, in table 3.7 we observe that there is a positive correlation between health status and quality of performance in our task. More specifically, workers with good health status have an average of 76% in their quality, while workers with bad health status have a lower average quality of performance (69%).

Table 3.7. Health status and quality of performance

<table>
<thead>
<tr>
<th>Quality of Performance</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good Health</td>
<td>60</td>
<td>.759</td>
<td>.105</td>
<td>.459</td>
<td>.969</td>
</tr>
<tr>
<td>Bad Health</td>
<td>38</td>
<td>.688</td>
<td>.119</td>
<td>.470</td>
<td>.911</td>
</tr>
</tbody>
</table>

Source: Dataset with results drawn from laboratory experiment. Authors calculations.
### 3.3.5 Work Effort

Finally, when it comes to work effort variables and more specifically the sum of tags submitted by the workers, we observe in figure 3.5 that there is a negative correlation between the quality of performance of the participant workers and the total sum of tagged keywords for the task.

In other words, the more tags a worker submit for the pictures of the task, the more his quality of performance decline.

![Figure 3.5: Sum of tags correlated with quality of performance](image-url)
Finally, we don’t discern a significant correlation between the completion time of the task and the quality of performance of the participant workers, even though a negative tendency exists.

**Figure 3.6:** Task completion time correlated with quality of performance
In order to investigate the impact of the observed variables on the quality of performance of workers in crowdsourcing environments, an econometric model has been developed that aims to predict the quality of work based on skill groups as well as environmental control variables.

In particular, we use an OLS linear regression model, which was applied by correcting for heteroskedasticity and by using all the groups of variables mentioned before. This model estimates how the above variables affect the quality of workers’ performance, in average. That is, it can answer the question: "In order to achieve quality results in crowdsourcing tasks, are demographics and personal skills essential?"

In linear regression, each regression coefficient estimates the change in the mean response per unit increase in X when all other predictors are held constant. In other words, a regression coefficient represents the increase in the response variable generated by a unit increment in the predictor variable associated with the coefficient.

In our research we contain a selection of "skills" and "environmental" controls for the observed performance of participant workers in crowdsourcing environment task. The general form of the econometric specifications of our OLS model is the following:
\[ Q_i = \alpha + \beta D_i + \gamma C_i + \delta NC_i + \epsilon WE_i + \zeta SE_i + e_i \] 

where \( Q_i \) is the quality of performance attainment indicator of the \( i^{th} \) worker in logarithmic values, \( D_i \) is a vector of demographics outcomes (age and gender) of the \( i^{th} \) worker, \( C_i \) is a vector of cognitive skills outcomes (computer competence, English levels) of the \( i^{th} \) worker, \( NC_i \) is a vector of non-cognitive skills outcomes (Big Five Personality Test variables) of the \( i^{th} \) worker, \( WE_i \) is a vector which includes the work effort variables (completion time and sum of tagged keywords) of the \( i^{th} \) worker, \( SE_i \) is a vector of social economics variables (health condition and poor status) of the \( i^{th} \) worker and \( e_i \) contains all the unobserved effects (i.e regression’s disturbance term).
Chapter 5

Estimation Results

If we apply model (4.1) to our data set, we obtain the following OLS linear regression model in table (5.1). It is essential to report that for the dependent variable we use the logarithmic value for normalization and easier interpretation.

5.1 OLS regression

The above results suggest that there are differences in the distribution of the quality of workers’ performance in relation to the skills and demographic characteristics of the workers. In particular, the first column in Table 5.1 shows the coefficients for the OLS linear regression model. The results show that more than half of the independent variables are statistically significant.

For example, having higher levels of English competency and computer competency significantly increased a worker’s quality of performance (at the 1% and 5% level of significance respectively and the independent variable is positively relate to the dependent variable).

More specifically, workers with at least professional English competency scored 8.7 points of percentage of quality higher on average than those with lower level of English competency and those with good com-
petency at computers scored 9.6 percentage points of quality higher on average than those with poor.

Moreover the regression coefficient for female workers was positive and for age was negative but not significantly different from zero, suggesting that females did not provide better results in crowdsourcing environment tasks than males and being younger or older doesn't really matter in crowdsourcing environment tasks.

Among the personality traits, we find that extraversion and emotion stability has a worthy of reference effect on quality of results at 1% level of statistical significance.

Regarding the effect of a worker’s work effort, our results revealed a strong effort-performance relation. In other words, the sum of tagged keywords of a worker affects significantly at 1% level of significance, his performance on the task, with the right hand variable being negatively related to the left hand variable.

Furthermore, the findings of OLS regression also showed a strong correlation between a worker’s wealth status and his quality of results, at 1% level of statistical significance.

More specifically, workers who have been poor their whole life or sometimes in their life scored about 14 percentage points of quality higher than those who are poor the day they took over the task. Also, workers who have never been poor in their life scored about 12 percentage points of quality higher than those who are poor right now (at 5% level of statistical significance).

Last but not least, our data analysis reveals a noticeable relationship (at the 10% level of significance) between the quality of results and the general health status of a worker, with the independent variable being positively related to the dependent variable.

Workers with high level of health status scored about 5 percentage
points of quality higher than the ones with low. It is a good indication that individuals with better health can provide better results in overall.
**Table 5.1. The determinants of performance on a crowdsourcing task**

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Mean</th>
<th>[1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.153</td>
<td>(0.105)</td>
</tr>
</tbody>
</table>

Demographics

| Age | -0.006 | (0.004) |
| Female | 0.013 | (0.022) |

Cognitive Skills

| High computer competence | 0.096** | (0.037) |
| High english competence | 0.087*** | (0.027) |

Personality Traits

| Extraversion | 0.085*** | (0.028) |
| Agreeableness | 0.015 | (0.025) |
| Consciousness | 0.034 | (0.027) |
| Emotion stability | -0.125*** | (0.028) |
| Openness | -0.031 | (0.024) |

Work Effort Variables

| Sum of tags | -0.007*** | (0.002) |
| Task completion time | 0.034 | (0.023) |

Social Economic Variables

| Good health status | 0.048* | (0.027) |
| Poor status: Sometimes or all the times poor | 0.140*** | (0.043) |
| Poor status: Never | 0.118** | (0.047) |

| Observations | 98 |
| R-squared | 0.607 |

Source: Dataset with results drawn from laboratory experiment. Authors calculations.

Notes: Dependent variable: Quality of performance. In parentheses heteroskedasticity corrected standard errors. Poor status "Now" is the reference category of the categorical variable of the wealth condition. Statistical significance: *** p<0.01, ** p<0.05, * p<0.1
5.2 Discussions

The skills and attributes that characterize a person who participates in online labor markets like crowdsourcing, can have a significant impact on the performance of his work. Cognitive or non-cognitive skills affect individuals when it comes to decisions concerning the implementation of a crowdsourcing task.

It has been demonstrated by the analysis that high levels of English competence and computing as well as high values of extraversion and stability of emotions play a significant role in the quality of performance of workers participating in crowdsourcing environment tasks.

It has also been found a strong positive relationship between a worker’s wealth status and quality of performance. In addition, the worker’s variable of work effort has had a significant negative impact on the performance in the crowdsourcing task.

Finally, our research reveals a significant impact of a worker’s health status on his quality of performance on online labor markets.
Chapter 6

Conclusions

Skills leap out as a crucial factor in achieving high quality work in crowdsourcing tasks. In addition, measuring and evaluating these skills is also a critical aspect. This dissertation presents an initial attempt to understand the role of cognitive and non-cognitive skills, work effort and demographic characteristics and their effect in quality of performance.

This research generate some interesting results. The analysis shows that different characteristics and skills contribute to the quality of performance in online labor market tasks, at a different level. For example, this research shows that high levels of English competence and computing affect significantly the quality of performance of workers participating in crowdsourcing environment tasks.

But the same applies to non-cognitive skills also, like high values of extroversion and stability of emotions that affect workers’ quality of performance in an important way.

In addition, regression results also provide some valuable data for the different relationships of explanatory variables with the quality of workers’ performance. Specifically, some variables, such as health and wealth status and the sum of tagged keywords, have a great impact on the quality of the results.
These results add to the ensemble of research that aims to explain the performance of workers in online labor markets like crowdsourcing.

Online activity is increasing, largely influenced by the ongoing development and evolution of internet technologies. This will greatly force many businesses and organizations to turn to the online communities and seek solutions in various tasks from individual workers.

Comprehension of the composition and the nature of crowdsourcing and its perspectives is essential in order to understand, improve and even create a favorable environment seeing that the already maturated market for the crowd labor remains almost entirely unstable and unregulated.
References


Ross, J., Irani, L., Silberman, M., Zaldivar, A., & Tomlinson, B. (2010, April). *Who are the crowdworkers?: shifting demographics in mechanical turk*. In CHI’10 extended abstracts on Human factors in computing systems (pp. 2863-2872). ACM.
