Mourelatos, Evangelos; Giannakopoulos, Nicholas; Tzagarakis, Manolis

Working Paper

Personality Traits and Performance in Online Labour Markets

GLO Discussion Paper, No. 338

Provided in Cooperation with:
Global Labor Organization (GLO)

Personality Traits and Performance in Online Labour Markets

Evangelos Mourelatos, Nicholas Giannakopoulos, Manolis Tzagarakis

Department of Economics, University of Patras

March 2019

Abstract

In this paper we investigate the impact of non-cognitive skills on the quality of task-specific outcomes by conducting a quasi-experiment on a well-known online crowdsourcing platform. We show that a worker's performance varies with personality traits, gender, human capital, crowdsourcing experience and work effort. Regarding the effects of non-cognitive skills, we find that workers' performance in online microtasks is positively related to extraversion and agreeableness. The positive impact of extroverts is also revealed when performance is adjusted for task completion time. These findings provide implications regarding the integration of selection mechanisms in online labour matching platforms aiming in uncovering micro-workers soft skills to improve performance and consequently the allocation of resources in online microtasks.

Keywords: Crowdsourcing, online labour, quality of work, cognitive abilities, personality traits, workers
JEL Codes: O33, J40, J24

1. Introduction

In this study we investigate worker performance in online microtasks by conducting a quasi-experiment on a crowdsourcing platform. The expansion of Web 2.0 has transformed the functioning of the traditional labour market and contributed in the creation of the online labour market (Autor, 2001; Horton et al., 2011; Bloom et al., 2014; Agrawal et al., 2015; Acemoglu & Restrepo, 2018). In this context, crowdsourcing constitutes a mechanism that optimally reallocates resources through labour matching and performance (Pallais, 2014; Pallais & Sands, 2016; Dube et. al., 2018). Specifically, online marketplaces[1] allow requesters (i.e., individuals and/or organizations) to provide wage offers to the motivated crowd of individuals (i.e., micro-workers) who are interested and capable to deliver labour related tasks (Chen & Horton, 2016). Preliminary evidence suggests that micro-worker performance in online microtasks depends on demographic, human capital and income-related factors (Ipeirotis, 2010) and it could be empirically tested within a task-specific framework which uncovers effectively worker unobserved attributes regarding individual performance (Autor & Handel, 2013). While a plethora of studies in typical labour markets highlight the role of cognitive and personality traits in explaining individual performance (Borghans et al. 2008; Cobb-Clark & Schurer, 2012) analogous evidence from online labour markets is rare (Kazai, et. al., 2012). In this paper, we conduct a quasi-experiment as an attempt to empirically investigate how the tasks performed on online markets are influenced by personality traits.

The empirical investigation of individual performance relies on data drawn from several sources, i.e., surveys, administrative datasets and lab experiments. More recently, behavioural studies with experimental data drawn from the online laboratories are becoming widespread and considered reliable in analysing several socio-economic outcomes (Arechar et. al. 2018;

[1] Online labor markets are organized in platforms where the creators of these markets provide the environment for individual-specific payments, screen out users who do not have valid accounts and prevent workers from communicating with each other. Some of the most popular platforms are Mechanical Turk, oDesk, Freelancer, Crowdflower, MobileWorks, ManPower, microWorkers.

Brañas-Garza et. al. 2018). The advantages of using information from online lab experiments relate to the precision in measuring performance of experimental subjects, to assessing how certain personality traits are directly relevant to performance and to move out noise resulting from unobserved factors such as workplace environment and peer effects.

In this paper, we conduct an online experiment using the microWorkers platform where the performance of micro-workers is based on their correct answers regarding the listening of a music sample with lyrics. We isolate cognitive skills from personality traits, and we focus on the role of these traits since they capture cross-sectional differences in performance indicators. Following Ipeirotis (2010), Kazai et al. (2012) and Kokkodis & Ipeirotis (2018) we allow worker performance to vary across several observed attributes (e.g., skills, education, work history, certifications) and latent characteristics (e.g., expertise and ability). Regarding personality traits we adopt the 44-item Big Five Inventory Survey (John & Srivastava, 1999; McCrae & Costa, 1999) which provides measures for each personality trait i.e., Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism (OCEAN[2], hereafter). The role of personality traits is well established in standard models of individual behaviour regarding task performance and the adopted OCEAN taxonomy captures individual-specific differences in the ways of thinking, feeling, and behaving (Filiz-Ozbay, et al., 2018).

The relevant literature is quite extensive, and it based on empirical studies that utilize data collected from laboratory experiments. Depending on the outcome, several facets of the OCEAN taxonomy have been found to correlate with task-specific measures of performance (Tett et al., 1991; Barrick & Mount, 1991; Salgado, 1997; Witt et. al., 2002; Rothman & Coetzer, 2003; Schmitt, 2007; Ng & Feldman, 2010; Dohmen & Falk, 2011; Müller & Schwieren, 2012; Larkin & Leider, 2012; Gow et al., 2016), productivity in experimental tasks

[2] Openness refers to the tendency to be creative and unconventional, Conscientiousness to the tendency to be organized and disciplined, Extraversion to the tendency to be sociable and active, Agreeableness to the tendency to be trusting and modest and Neuroticism to the tendency to experience negative emotions.

(Dohmen, & Falk, 2011; Cubel et. al., 2016), worker earnings (Nyhus & Pons, 2005; Mueller & Plug, 2006; Borghans et al. 2008; Heineck & Anger, 2010; Cobb-Clark & Schurer, 2012; Fletcher, 2013; Gensowski, 2018), recruiting outcomes (Baert & Decuypere, 2014), employment decisions (Fletcher, 2013; Gensowski, 2018), educational outcomes (Burks et. al., 2015; Kassenboehmer et. al. 2018) and risk tolerance (Rustichini et al. 2016). However, relevant evidence from online labour markets is rare. To the best of our knowledge, our study is the first attempt to use an online experiment to directly test the relationship between the personality traits and individual performance in online labour markets. An exception is the preliminary evidence provided by Kazai et al. (2012), who show that personality characteristics vary with micro-workers performance and classify as effective workers middle-aged females with high scores in openness and conscientiousness. Our work is mostly linked to Müller & Schwieren (2012), Cubel et al. (2016), and Larkin & Leider (2012) who empirically analyse task specific outcomes in offline lab experiments.

We measure performance with two variables. The first is the raw number of total correct answers and the second is the logarithm of the ratio of correct answers per minute in the task. The first performance indicator is an unadjusted measure which captures the cross-sectional variation in performance between micro-workers while the second is a measure of effort-adjusted performance since is trace out the cross-sectional variation in the effort allocated by each micro-worker. For both measurements, our estimation strategy relies on linear regression models which provide estimates for the effects of demographic characteristics, human capital endowments, effort and personality traits on indicators of performance. In addition, for interpretation purposes we adopt a fractional probit model where the ratio of total correct answers over the total number of words in the music sample with lyrics is utilized as the dependent variable.

As a first step in the analysis we aim to establish a relationship between performance and demographic and human capital characteristics. We find that performance is higher for females, non-Asian, those with higher education, computer competence and crowdsourcing experience (tenure). Then, we proceed by estimating the effects of personality traits on performance without taking into consideration the effects of human capital variables. We find that performance is positively related to extraversion and agreeableness and these results are independent of micro-workers demographic characteristics. We continue by estimating models of performance in which both cognitive and personality traits are included in the empirical specifications. Our results regarding the impact of personality traits on performance are confirmed in all cases. More specifically, an increase by a standard deviation in extraversion is followed by around 4.0% higher performance in the task and analogously in the case of the agreeableness the performance is higher by around 2.0%. These results are slightly reduced when human capital variables are added during estimation while a more noticeable drop in these estimates appears with the inclusion of micro-workers effort. Lastly, taken into consideration the cross-sectional variation in the effort distribution we found that our adjusted performance indicator is positively related to the above two domains of personality with the effect of extraversion to be robust to alternative model specifications. Furthermore, higher performance is identified for females, micro-workers with higher educational levels and with increased experience in online labour markets. According to our findings, platform designers could incorporate the idea that soft skills are important contributors to workers' performance and thus, mechanisms that enhance workers' soft skills could be embedded into online platforms to improve performance outcomes (Difallah et. al., 2015). This should be especially a priority for platforms in which micro-workers have a highly heterogenous profile regarding the observed screening mechanism (relying on education, experience etc.).

The structure of the paper is the following. Section 2 presents the online experiment and it provides information on the task's attributes, design and the measurement of our performance indicators. In Section 3 we present the utilized empirical model and the estimated results are presented in Section 4. Section 5 concludes.

2. The experiment

2.1 Task design and implementation

The experiment is implemented on the microWorkers crowdsourcing platform (see Hirth et al. 2011 for a survey on the demographics of microWorkers platform). The task was programmed using the music platform Soundcloud. The task belongs to a family of performance-based tasks and measures the ability of workers to provide results that meet specific requirements to get paid, replicating the context of a real-world workplace. Usually online tasks are classified into four main categories (e.g. Microtasks, Macrotasks, Contests, and Crowdfunding). Our task belongs to the category of Microtasks (paid online micro-jobs), where the crowd is asked to perform short and temporary tasks (Human Intelligence Tasks) that do not require creativity, high knowledge specialization and much time for their completion. In these tasks, micro-workers are paid a small monetary reward per unit. Popular examples include spell checking of short paragraphs, sentiment analysis of tweets, rewriting product reviews, image tagging, translation or transcription. Relevant tasks are also implemented in experimental economics (Chandler & Kapelner, 2013; Balasubramanian et. al., 2017).

The microWorkers platform initially provides to the requester, before he publishes the job, a sort list of twenty discrete categories that the job may belong (e.g. surveys, testing, translation, transcription, bookmark a page, Google search, Youtube listen, Facebook like, Twitter repost, promotion, Yahoo Answers, forums, download, install, comment on a Blog, write a review, write an article, mobile applications, Website owners, leads). Requesters choose a specific category of tasks and follow tasks with a unique ID campaign. In this way, workers can identify the characteristics of the published task when they search for online tasks[3]. In our experiment we choose a combination of the "surveys" and "transcription" categories. After we proposed to the microWorkers our task categories, we received after one day, a validation email by the platform that our online job matches the chosen categories characteristics.

Then, the microWorkers platform, based on the average job completion time proposed us the offered wage amount. The offered wage equals to the average wage in tasks with similar characteristics that take place in the platform. This payment is also offered to the requester by the platform to avoid self-selection biases regarding participation in the task. Potential workers can scroll down all the active online jobs in the platform and behold the ID, type, wage and how many recruiters the task requires for its final completion. Hence, they have basic information regarding the task characteristics before they proceed. Upon selection, additional information and instructions are provided to workers.

In our experiment, respondents were asked to take part in a demographic and personality questionnaire and afterwards to listen a music sample and write down as many words of the lyrics as they were able to catch in the designated area. The task required a minimum equipment, speakers or headphones. Thus, a microWorkers "open call" campaign was launched for 15 hours and interested workers could participate. Each participant (N=250) is paid by $0.82 cents and the overall cost of the project counts to $205. At a first stage, workers were obliged to fulfil an online survey regarding demographic characteristics, cognitive skills and personality traits. At the second stage, workers were driven through a link to the online job, called the task. In this stage, workers have access to a player of a music sample with lyrics (56 words in English), taking place on the well-known music platform https://soundcloud.com/

[3] Unfortunately, we are not able to know the population of micro-workers who see the announcement of our task and thus to observe two-groups of micro-workers (assigned to the task and not-assigned to the task). From this point of view potential self-selection issues may arise which however cannot be addressed effectively in the present study.

and they should provide in a text box as many correct words of the music sample, as they could. The workers had no time limit and they had the opportunity to listen to the music sample as many times as they desire. Once a worker pressed the button of final submission under the answer box, it could not be changed and thus we are able to directly measure individual performance. The wording of the music sample is the following:

"You enjoy the vibe, that runs through your body, It sets the pace and you dance, The music feels like we move off the ground, It's all the sense about sound, Our eyes meet despite the crowd, And I drift to your side, The air makes us feel so high, Makes us feel so alive"

The data on demographic characteristics refer to respondent's gender, age at the time of survey and country of birth. Information on cognitive skills is captured through indicators for whether the worker has finished tertiary education and has high computer competence. To collect information on personality traits we constructed a "short length & quick to fulfil" questionnaire with the aim to avoid participants to get bored and to spend a lot of time for providing answers. Thus, we administrated the well-known 44-item Big Five Inventory survey (John & Srivastava, 1999)[4] which is a mid-sized questionnaire ensuring an accurate measure of each personality trait using responses on a five-point Likert scale, i.e., from 1: Disagree to 5: Agree. The Big Five Inventory is designed through factor analysis, so each trait is orthogonal to the rest (McCrae & Costa, 1999). Furthermore, it is quite brief for a multidimensional personality inventory (44 items total) and consists of short phrases with relatively accessible vocabulary (John et al. 1991). This five-factor structure questionnaire of personality can capture much of the variance in the personality trait ratings and represents personality at the broadest level of abstraction, and each dimension includes a large number of the individual's characteristics (Goldberg, 1992).

[4] A complete reference of the utilized questionnaire can be found in John & Srivastava (1999, p. 132).

Furthermore, we collect data for each worker for two work effort variables by using the provided metrics of the https://soundcloud.com/ platform. The first, refers to the time required to complete the task by measuring the minutes between the time of signing in the task until the time of submitting the final answer. The second, measures the number of times each worker played the music sample during the task. In addition, we collect data on workers' experience in microWorkers platform (registration year, number of completed tasks and total number of tasks) by using a web-crawler. Using this information, we measure worker's experience in crowdsourcing micro-task activities, and it is defined as the distance in years between the year of registration in microMorkers and the year in which the experiment is conducted (2015).

2.2 Summary statistics

Table 1 presents summary statistics. On average, the amount of correct answers is 33.86 (with a median of 34.5) denoting that the success rate in our experiment is 60% (out of 56 words). Dividing the amount of correct answers by the minutes allocated in the task we found that, on average, micro-workers provide 6 correct answers per minute. A vivid representation of the distribution of the two performance indicators (raw number of correct answers and the logarithm of the amount of correct answers per minute) is provided by Figure 1 and give us the ability to further investigate their observed variation across several individual-specific characteristics. For example, according to the observed variation in personality traits we see that the mean score for Openness is 3.54 suggesting a high tendency toward creativity and active imagination, for Conscientiousness is 3.60 signifying thoroughness, for Extraversion is 3.22 exhibiting energetic behaviour, for Agreeableness is 3.56 indicating that our workers seem to be more empathetic and altruistic and for Neuroticism is 2.78 suggesting that workers don't tend to experience negative emotions. Regarding cognitive skills, we observe that the half workers have completed tertiary education (50%) and have advanced computer skills (52%).

We also observe that workers have on average 2.87 years of experience in crowdsourcing jobs within the platform and that the average completion time of the task is 5.82 minutes. Lastly, most of our workers are males (67%) with an average age of 30 years and 62% of our workers come from Europe, 12% from South Asia, 10% from North America, 7% from Middle East & North Africa, 5% from Latin America & Caribbean and 5% from East Asia & Pacific.

-Insert Table 1 here-

-Insert Figure 1 here-

In order to understand the relationship between performance and personality traits we present in Figure 2, the relationship between performance (A: correct answers, B: (ln) correct answers per minute) and the distribution of each personality trait using non-parametric local polynomial smoothing techniques. We observe that performance (amount of correct answers) is positively related with openness, extraversion and agreeableness and negatively with conscientiousness and neuroticism. When the amount of correct answers per minute is utilized as the performance indicator, the above positive relation is confirmed in the cases of openness, extraversion and agreeableness. However, for conscientiousness and neuroticism we are not able to present a clear pattern with effort-adjusted performance. In all cases there is no evidence of extreme values across the distributions of personality traits. Thus, personality traits seem to constitute a source of variation regarding micro-workers' performance in online labour markets.

-Insert Figure 2 here-

The profile of workers' performance (correct answers and correct answers per minute) across several sources of workers' heterogeneity is presented in Table 2 and Figure 3 for several categorical and continuous variables, respectively. Focusing in Table 2, we observe that on average, women perform better than men (3.6 more correct answers in total or 0.6 more correct answers per minute), workers with tertiary education -compared to those with lower education-exert higher performance levels (10.8 in total and 0.6 per minute) and micro-workers with high computer competence provide 9.3 more correct answers compared to those with lower level of computer competence. However, there is no evidence, in terms of statistical significance, that computer competent workers provide more correct answers per minute. Regarding the differentiation of performance across country grouping we observe that Asian micro-workers provide a smaller amount of correct answers compared to their non-Asian counterparts (-12.9 in total and -0.5 per minute). Indeed, micro-workers from South Asia and East Asia & Pacific seem to have the lowest amount of correct answers compared to workers from other regions (Europe, Latin America & Caribbean, Middle East & North Africa and North America). The low performance of micro-workers from South Asia is also confirmed when we use the amount of correct answers per minute as a performance indicator.

Turning now to the results shown in Figure 2 (A: correct answers) we observe that workers' performance seems to exert a non-linear relationship with age and crowdsourcing experience (tenure) and to be positively related with completion time. The same profiles (age and tenure) are identified when we focus on the amount of correct answers per minute (Panel B). Overall, these results imply that performance in online labour markets is an outcome that varies across several sources of individual heterogeneity such as demographics, cognitive skills and effort variables. This conclusion confirms the evidence provided in Kokkodis & Ipeirotis (2018).

-Insert Table 2 here-

-Insert Figure 3 here-

3. Empirical analysis

3.1 Empirical model

To examine the individual performance in online micro-tasks we assume that the micro-task designer (requester) is willing to pay a specific amount of monetary values to a specific micro-worker in order to produce a specific outcome (Q) which in our study is the amount of words in our music sample. As far as the expected amount of correct answers is achieved by the micro-worker, the aim regarding the efficient allocation of the budget constraint (compensation paid) is totally met. The expected outcome (Q) is assumed to be a function depending on observed productive characteristics (e.g., cognitive and non-cognitive skills, crowdsourcing experience), effort variables (e.g., time allocated in the task) and other unobserved factors (e.g., innate ability, worker's motivation in task participation). We note that in our empirical model, worker compensation is rather fixed, and thus performance is not affected by any kind of performance-related-pay scheme (e.g., Dohmen & Falk, 2011). Therefore, the expected outcome (Q) is simply the summation of the amount of correct answers (Q^C) and the amount of wrong answers. High quality workers are those with the higher ratio Q^C/Q and consequently higher values in this ratio ensures efficiency in the allocation of resources regarding the total amount of money spent for the specific micro-task. Thus, low performance workers are those with lower values in the ratio Q^C/Q and consequently the allocation of resources for the task becomes inefficient.

For estimation purposes we rely on typical regression models where specific performance indicators are utilized as the dependent variable. To isolate the effects of personality traits and other "environmental" factors we adopt the following specification:

$$P_i = \alpha + \beta^k NC_i^k + \gamma C_i + \delta E_i + \zeta D_i + \theta \xi_i + e_i \qquad (1)$$

where, P_i is a specific performance indicator for micro-worker i, NC_i^k is a k-vector of non-cognitive skills of worker i (where k=1,...,5 corresponds to Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism, respectively), C_i is a vector that includes indicators for micro-worker i regarding tertiary education and high computer competence, E_i is a vector of task effort-specific variables of worker i, D_i is a vector of the demographic characteristics of micro-worker i, ξ_i is a vector of individual specific location fixed effects and e_i an i.i.d. disturbance term.

This specification described by equation (1), although general, is expected to provide evidence on the role of Big Five variables (standardized to have a zero mean and a standard deviation of one) on workers' performance through the vector of the estimated coefficients β^k (using linear regression techniques). The dependent variable P refers to (a) the amount of total correct answers and (b) to the amount of total correct answers per minute (in order to trace out differences in performance due to differences in the effort allocated in the task by each worker). Moreover, since the amount of correct answers that a certain worker provides is bounded between 0 and 56, P also refers to the proportion of correct answers (Q^C/Q) for each micro-worker. The values of this variable range between 0 and 1 and thus a fractional response model (e.g., fractional probit) is adopted, instead of linear regression techniques. We note that our empirical strategy is expected not only to provide robust evidence regarding the role of non-cognitive skills in worker's performance but also to offer insights regarding the distributional implications for the allocation of resources in online labour markets when micro-tasks are developed.

4. Estimation results

In this section we present estimates regarding the determinants of worker's performance for different measures of performance (linear and proportional outcomes) and alternative model specifications. Firstly, we show estimates on the effect of cognitive skills on performance. Secondly, we present results regarding the impact of personality traits on performance net of human capital variables. Thirdly, we provide estimated coefficients from an augmented model of performance in which, demographics, personality traits and human capital variables are taken into consideration. We continue the interpretation of our results by showing whether our model estimates are robust to the inclusion of the effort variable. Lastly, we test the ability of our model specification to explain the variation in micro-workers' effort-adjusted performance.

4.1 Cognitive skills

Table 3 presents the estimation results of 4 different model specifications. Panel A includes linear coefficient estimates and Panel B average marginal effects. In the first specification (Column 1), the set of explanatory variables includes only demographics characteristics (female, age and region of origin). We observe that performance depends on gender and region while the effect of age in not statistically different from zero. More specifically, females provide 2.7 more correct answers (compared to males) which corresponds -according to the average marginal effect- to 4.8% higher performance. Furthermore, Asian micro-workers provide on average a lower amount of correct answers (12.6) compared to non-Asian ones which corresponds to 22.4% lower performance in our microtask. When an indicator for tertiary education is included (Column 2) a positive and statistically significant correlation is identified between performance and education. Micro-workers with tertiary education provide 8.4 correct answers more than those with less than tertiary education, which is equivalent to a 15.1% higher performance. The effects of demographic variables remain that same regarding their statistical significance and direction of the effects but of a reduced magnitude. This implies that the observed gender and region differences in performance are, to some extent, due to differences in human capital endowments. These findings are further confirmed when an indicator for computer competence is included as an additional argument (Column 3). The corresponding estimate is positive and statistically different than zero. Specifically, micro-workers with high computer competence provide 6.5 more correct answers (or 11.7% higher performance in Panel B) compared to the less computer competent ones. Lastly, the inclusion of tenure does not affect the previous estimates in terms of statistical significance, direction and magnitude. According to the estimated coefficients, workers with more than 5 years of crowdsourcing experience provide 5.2 more correct answers (9.2% higher performance in

Panel B) than those with lower tenure. Overall, our results imply that microtask performance in online labour markets varies with several human capital endowments (i.e., education, computer competence and crowdsourcing experience). This conclusion appears to be robust to the inclusion of demographic variables.

-Insert Table 3 here-

4.2 Personality traits

In this sub-section we aim to investigate whether the cross-sectional variation in personality traits explain the observed variation in performance. Column 1 of Table 4 present estimates from a model specification that includes in the set of explanatory variables the five personality dimensions and demographic indicators. In this exercise we are interested in identifying possible correlations between non-cognitive skills and performance -net of human capital variables- since personality traits are inherently present to the individual and thus, they may affect choices such as educational achievement. We observe that across specifications, the estimates of gender and region of origin are statistically significant and are very close to the estimates of Table 3 (Panels A and B). Regarding the estimated coefficients of personality traits, we show that extraversion and agreeableness are positively related to performance while openness, conscientiousness and neuroticism appear to be statistically insignificant. More specifically, rising extraversion (Panel A) by a one standard deviation results in 3.4 more correct answers (6.1% higher performance in Panel B). Furthermore, an increase of a standard deviation in agreeableness (Panel A) is associated with 1.2 more correct answers (2.0% higher performance in Panel B). Thus, the distribution of personality traits constitutes a hidden source of congenital heterogeneity between micro-workers which seems to affect their performance in online micro-tasks. In the following sub-section, we test whether the estimated impact of personality traits on performance is due to other sources of observed heterogeneity that relate to cognitive skills and task-specific effort.

4.3 Cognitive skills and personality traits

Table 4 (Columns 2-4) includes the estimated coefficients of three different model specifications regarding the impact of personality traits, demographics and human capital variables on microtask performance. We find that extraversion and agreeableness exert a positive and statistically significant impact on performance. These findings corroborate our previous results presented in Column 1 of Table 4. This implies that more extroverted micro-workers and those with higher levels of agreeableness perform better in online crowdsourcing microtasks. More specifically, the inclusion of the tertiary education indicator (Column 2) does not seem to affect the impact of agreeableness but it contributes to the reduction of the estimated coefficient of extraversion by 25% (compared to Column 1 of Table 4). This implies that the effect of extraversion on performance enhances to some extent the impact of education on performance. Indeed, extraversion and education are positively correlated with a coefficient of 0.35. When computer competence is added in the performance equation (Column 3), the estimated coefficients are slightly reduced but all of them continue to exert the same impact in terms of direction and significance. The robustness of our results is further verified by the inclusion of tenure indicators (Column 4). Again, higher levels of extraversion and agreeableness are associated with increased performance, although the effect of agreeableness is reduced from 1.05 (Column 3 of Table 4) to 0.92.

The above findings are further confirmed when a fractional probit models is adopted (Panel B). Based on the average marginal effects (Column 4), we find that performance is higher by 4.0% (1.5%) due to a one standard deviation increase in extraversion (agreeableness), by 3.0% for females (compared to males), by 10.7% for workers with tertiary education (compared to those with less than tertiary education), by 10.0% for those with high computer competence (compared to those micro-workers with lower levels of computer competence), and by 8.7% in the case where micro-workers have more than 5 years of crowdsourcing experience (compared to those with lower experience). In contrast, the performance of Asian micro-workers is lower by 11.7% compared to non-Asians.

-Insert Table 4 here-

4.4 Task completion time and effort-adjusted performance

In this sub-section we investigate the robustness of our results when effort is taken into consideration during estimation of equation (1). If micro-workers with different personality traits allocate a different amount of time in the task and performance is related to non-cognitive skills and effort, then the estimated coefficients of personality traits on performance will be different from those estimates obtained from models which do not incorporate the observed distribution of effort. Table 5 presents two model specifications for the two dependent variables. The first specification (Columns 1 and 3) is the same as in Column 4 of Table 3 (Panels A and B) with the effort variable to be the only additional explanatory variable while the second (Columns 2 and 4) is an augmented specification by including the set of personality traits. According to the results presented in Columns 1 and 3 we observe that effort (minutes to complete the task) is statistically significant and is positively related to performance. More specifically, one additional minute is associated with 1.68 more correct answers (Column 1) or 3.0% higher performance (Column 3). Comparing the results with those presented in Column 4 of Table 3 we observe that in all cases the estimated coefficients and standard errors are now reduced, except for the estimated coefficient of gender which is marginally higher. Regarding the effects of personality traits (Column 2) we find that extraversion and agreeableness continue to be statistically significant and positive and very close to the estimated coefficients shown in Column 4 of Table 4. Thus, we conclude that the effects of personality traits are robust to alternative model specifications and independent of the micro-workers' effort. A one standard deviation increase in extraversion is associated with 2.0 more correct answers (3.6% higher performance) and the corresponding impact on agreeableness is 0.8 more correct answers (1.2% higher performance).

-Insert Table 5 here-

Table 6 presents estimates for the effects of personality traits, demographics and human capital variables on effort-adjusted performance. We show results from 6 different model specifications. Columns [1]-[4] include estimates from specifications that ignore the impact of personality traits while in Columns [5]-[6] these traits are taken into consideration. We find that (Column 6) the effort-adjusted performance of females is 9% higher than that of males, of micro-workers with tertiary education is 8.0% higher than that of micro-workers with lower education and that micro-workers with more than 5 years of experience exert 8.3% higher level of effort-adjusted performance compared to those workers with lower experience. We note that although Asian micro-workers seem to perform worse than the non-Asian ones (Column 1), this finding is not robust when cognitive skills and experience is taken into consideration (Columns 2-6). This implies that differences in online microtask effort-adjusted performance between Asian and non-Asian workers are due to differences in human capital endowments. In addition, the effect of agreeableness does not seem to remain significant when we use the total amount of answers per minute (in logs). The above findings suggest that the impact of personality traits and other individual specific characteristics is more accurately estimated when the performance indicator is adjusted for the time allocated by each worker in the execution of the task.

-Insert Table 6 here-

5. Conclusions

Nowadays soft skills are emerging as a critical factor in achieving high quality of work not only in traditional labour markets but in crowdsourcing tasks, as well. In this paper we empirical investigate this argument by focusing on the role of non-cognitive skills on micro-workers' performance regarding online tasks. Our results provide implications for personality-oriented job analysis, work motivation and management policies regarding the processes of recruiting workers on crowdsourcing platforms. For analytical purposes, we conducted an online experiment using the microWorkers platform where the performance of micro-workers is based on their correct answers regarding the listening of a music sample with lyrics. We were also able to also collect information on cognitive skills, personality traits and several socio-demographic characteristics in order to have a clear insight of the individuals' characteristics.

According to our results extraversion and agreeableness exert a statistically significant and robust effect regarding the micro-workers' performance in online micro-tasks. However, when we use measures of effort-adjusted performance only extraversion is positively related with performance. Therefore, our study provides evidence that personality traits matter as far as the performance of micro-workers is concerned. We also found that performance (either unadjusted or adjusted) depends on gender, education and tenure. In addition, micro-workers' effort in online micro-tasks is associated with higher performance. Furthermore, we found that the observed differences in effort-adjusted performance between Asian and non-Asian micro-workers is due to differences in human capital endowments. However, this is not the case when performance is not adjusted for the task completion time. While our results are robust to alternative specifications and tests, we should acknowledge that the external validity of our findings is difficult to be tested with the utilized dataset due to difficulties in the identification of potential confounders. In addition, since our tasks do not require high level of creativity, the generalization of our findings for other microtasks might come up with misleading messages. To this direction, future research is further warranted. Furthermore, we acknowledge that our estimates have not be corrected for possible self-selection biases regarding the mechanism that micro-workers use regarding their choices to be assigned to the task. Nevertheless, future research could include additional experiments in online labour markets that replicate well-known experiments from offline labour market environments and relevant surveys.

Our findings provide implications regarding the integration of selection mechanisms in online labour matching platforms aiming in uncovering micro-workers soft skills to improve performance and consequently the allocation of resources in online microtasks. Such attempts are expected to enrich our understanding on the mechanisms behind the relationship between personality traits and task performance and to contribute towards the standardization of the basic "building block" tasks that would make crowdsourcing more scalable and easier to set prices, spread best practices, build meaningful reputation systems and track quality. Finally, the results illustrate the benefits of using the Big-Five model of personality as a mechanism for the selection for micro-workers in crowdsourcing activities.

References

Acemoglu, D., & Restrepo, P. (2018). Artificial Intelligence, Automation and Work (No. w24196). National Bureau of Economic Research.

Agrawal A., Horton J., Lacetera, N., and Lyons, E., (2015). "Digitization and the Contract Labour Market: A Research Agenda," NBER Chapters, in: *Economic Analysis of the Digital Economy*, pp. 219-250 National Bureau of Economic Research, Inc.

Autor, H.D., (2001). Wiring the labour market. *Journal of Economic Perspectives*, Vol. 15, No.1, pp. 25-40.

Autor, H.D., & M. J. Handel (2013) Putting Tasks to the Test: Human Capital, Job Tasks, and Wages, *Journal of Labour Economics,* Vol. 31, No. 2, pp. 59-96.

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, 21(1), 99-131.

Baert S. & L. Decuypere (2014) Better sexy than flexy? A lab experiment assessing the impact of perceived attractiveness and personality traits on hiring decisions, *Applied Economics Letters*, 21:9, 597-601,

Balasubramanian, P., Bennett, V. M., & Pierce, L. (2017). The wages of dishonesty: The supply of cheating under high-powered incentives. *Journal of Economic Behavior & Organization,* 137, 428-444.

Barrick, M.R., & Mount, M.K. (1991). "The Big Five personality dimensions and job performance: a meta-analysis", *Personnel Psychology,* vol. 44(1), pp. 1–26.

Brañas-Garza, P., Capraro, V., & Ramírez, E. R. (2018). Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour. *Economics Letters*. Vol. 170, pp. 19-23.

Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. (2014). Does working from home work? Evidence from a Chinese experiment. *The Quarterly Journal of Economics*, 130(1), 165-218

Borghans, L., Duckworth, A. L., Heckman, J. J. & Ter Weel, B. (2008) The economics and psychology of personality traits. *Journal of Human Resources*, Vol. 43, pp. 972-1059.

Burks, S. V., Lewis, C., Kivi, P. A., Wiener, A., Anderson, J. E., Götte, L., & Rustichini, A. (2015). Cognitive skills, personality, and economic preferences in collegiate success. *Journal of Economic Behavior & Organization*, 115, 30-44.

Chandler, D., & Horton, J. J. (2011). Labour allocation in paid crowdsourcing: Experimental evidence on positioning, nudges and prices. *Human Computation: Papers from the 2011 AAAI Workshop*, Vol. 11, No. 11, pp. 14-19.

Chandler, D., & Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization,* 90, 123–133.

Chen, D., & Horton, J. (2016). Research Note-Are Online Labour Markets Spot Markets for Tasks? A Field Experiment on the Behavioral Response to Wage Cuts. *Information Systems Research*, Vol. 27: 403-423.

Chen, Y., & Konstan, J. (2015) Online field experiments: A selective survey of methods. *Journal of the Economic Science Association*, 1(1):29–42.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11-15.

Cubel, M., Nuevo-Chiquero, A., Sanchez-Pages, S. & Vidal-Fernandez, M. (2016), Do Personality Traits Affect Performance? Evidence from the Labouratory. *Economic Journal*, 126: 654–681.

Dube, A., Jacobs, J., Naidu, S., & Suri, S. (2018). Monopsony in Online Labour Markets (No. w24416). *National Bureau of Economic Research*.

Difallah, D. E., Catasta, M., Demartini, G., Ipeirotis, P. G., & Cudré-Mauroux, P. (2015, May). *The dynamics of micro-task crowdsourcing: The case of amazon mturk.* In Proceedings of the 24th international conference on world wide web (pp. 238-247). International World Wide Web Conferences Steering Committee.

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, 101(2), 556-90.

Farrell, M.A., J.H., Grenier, and J., Leiby (2017) Scoundrels or Stars? Theory and Evidence on the Quality of Workers in Online Labour Markets. *The Accounting Review*, Vol. 92, No. 1, pp. 93-114.

Feitosa, J., Joseph, D. L., & Newman, D. A. (2015). Crowdsourcing and personality measurement equivalence: A warning about countries whose primary language is not English. *Personality and Individual Differences*, 75, 47-52.

Filiz-Ozbay, Emel., John C. Ham, John H. Kagel, and Erkut Y. Ozbay (2018) "The Role of Cognitive Ability, Personality Traits and Gender in Gift Exchange Outcomes" *Experimental Economics*, Vol. 21(3), pp 650-672.

Fletcher, J. M. (2013). The effects of personality traits on adult labour market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, 89, 122-135.

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, 51, 170-183.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.

Gow, I. D., Kaplan, S. N., Larcker, D. F., & Zakolyukina, A. A. (2016). CEO personality and firm policies (No. w22435). National Bureau of Economic Research.

Heckman, J.J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, 91(2), 145-149.

Heckman, J.J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labour market outcomes and social behavior. *Journal of Labour Economics*, 24(3), 411-482.

Heineck, G., & Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, 17(3), 535-546.

Hirth, M., Hoßfeld, T., & Tran-Gia, P. (2011, June). Anatomy of a crowdsourcing platform-using the example of microworkers.com. *In Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, 2011 Fifth International Conference on (pp. 322-329). IEEE.

Horton, J.J., Rand, D. G., & Zeckhauser, R. J. (2011). The online labouratory: Conducting experiments in a real labour market. *Experimental Economics*, 14(3), 399-425.

Ipeirotis, P. (2010). Demographics of Mechanical Turk. New York University. Unpublished Manuscript

John, O. P., & Srivastava, S. (1999), "The Big-Five trait taxonomy: History, measurement, and theoretical perspectives." In L.A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and Research*, vol. 2, pp. 102–138, New York: Guilford Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The "Big Five" Inventory--Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

Kazai, G., Kamps, J., & Milic-Frayling, N. (2012, October). The face of quality in crowdsourcing relevance labels: Demographics, personality and labeling accuracy. In Proceedings of the 21st ACM international conference on Information and knowledge management (pp. 2583-2586). ACM.

Kokkodis M. and Ipeirotis G. P. (2018), Career Development Paths in Online Labour Markets. Unpublished manuscript.

Kassenboehmer, S. C., Leung, F., & Schurer, S. (2018). University education and non-cognitive skill development. *Oxford Economic Papers*, 70(2), 538-562.

Larkin, I., & Leider, S. (2012). Incentive schemes, sorting, and behavioral biases of employees: Experimental evidence. *American Economic Journal: Microeconomics*, 4(2), 184-214.

Litman, L., Robinson, J., & Rosenzweig, C. (2015). The relationship between motivation, monetary compensation, and data quality among US-and India-based workers on Mechanical Turk. *Behavior research methods*, 47(2), 519-528.

Mason, W. A., & Watts, D. J. (2009). Financial incentives and the "performance of crowds". *Proceedings of the Human Computation Workshop. Paris: ACM*, pp.77-85.

McCrae, R. R., & Costa Jr, P. T. (1999). A five-factor theory of personality. *Handbook of Personality: Theory and Research*, 2, 139-153.

Milland Kristy (2017), The Unsupported Crowd: Exclusion of Indian Workers in Amazon Mechanical Turk Communities, Unpublished Manuscript.

Mueller, G., & Plug, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labour Relations Review*, 60(1), 3-22.

Müller, J. and Schwieren, C. (2012). 'Can personality explain what is underlying womens' unwillingness to compete?', *Journal of Economic Psychology*, vol. 33(3), pp. 448–60.

Müller, J. and Schwieren, C. (2017): Using Personality Questionnaires in Experiments-Limits and Potentials. MPRA paper: 78132. Unpublished Manuscript

Ng, T. W., & Feldman, D. C. (2010). Human capital and objective indicators of career success: The mediating effects of cognitive ability and conscientiousness. *Journal of Occupational and Organizational Psychology*, 83(1), 207-235.

Nyhus, E. K., & Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, 26(3), 363-384.

Pallais A, Sands E.G. (2016) Why the Referential Treatment? Evidence from Field Experiments on Referrals. *Journal of Political Economy*. 124 (6):1793-1828.

Pallais, A. (2014). Inefficient hiring in entry-level labour markets. *American Economic Review*, 104(11), 3565-3599.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322.

Rothmann, S., & Coetzer, E. P. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1), 68-74.

Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, 64, 122-137.

Salgado, J.F. (1997). 'The Five Factor Model of personality and job performance in the European Community', *Journal of Applied Psychology*, vol. 82(1), pp. 30.

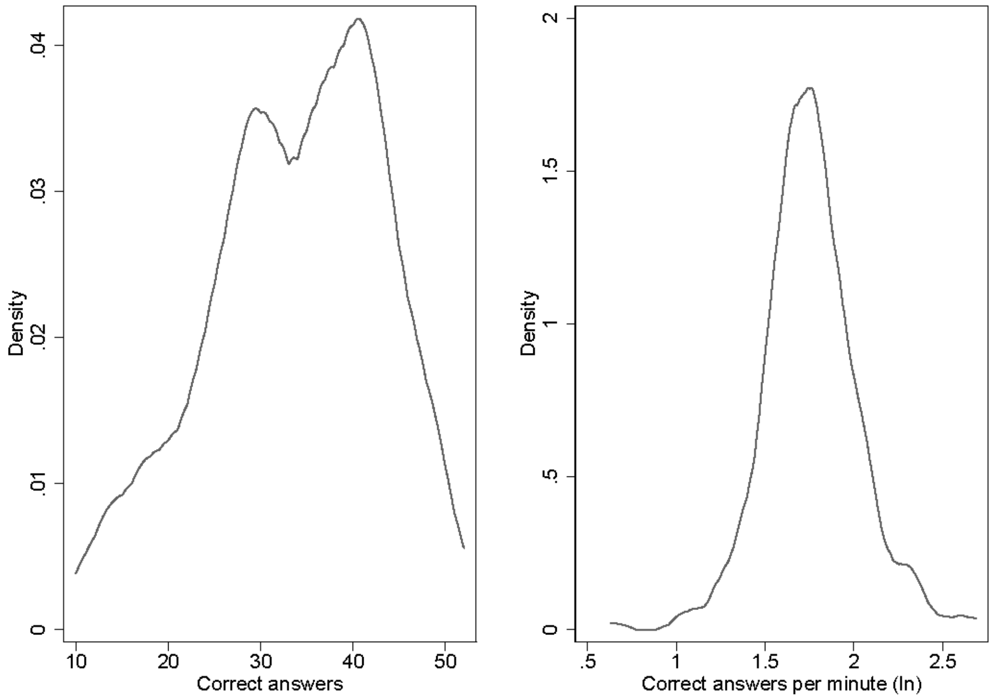
Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94(1), 168.

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, 43(1), 84-90.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: a meta-analytic review. *Personnel psychology*, 44(4), 703-742.

Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology*, 87(1), 164.

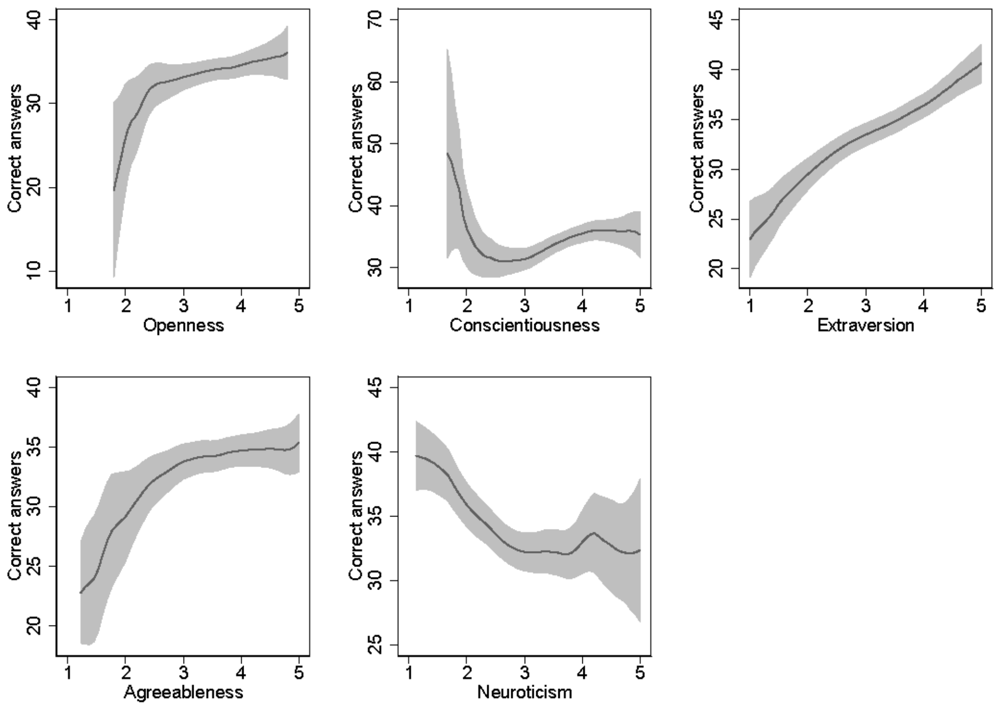
Figure 1. Distribution of performance indicators in online microtask



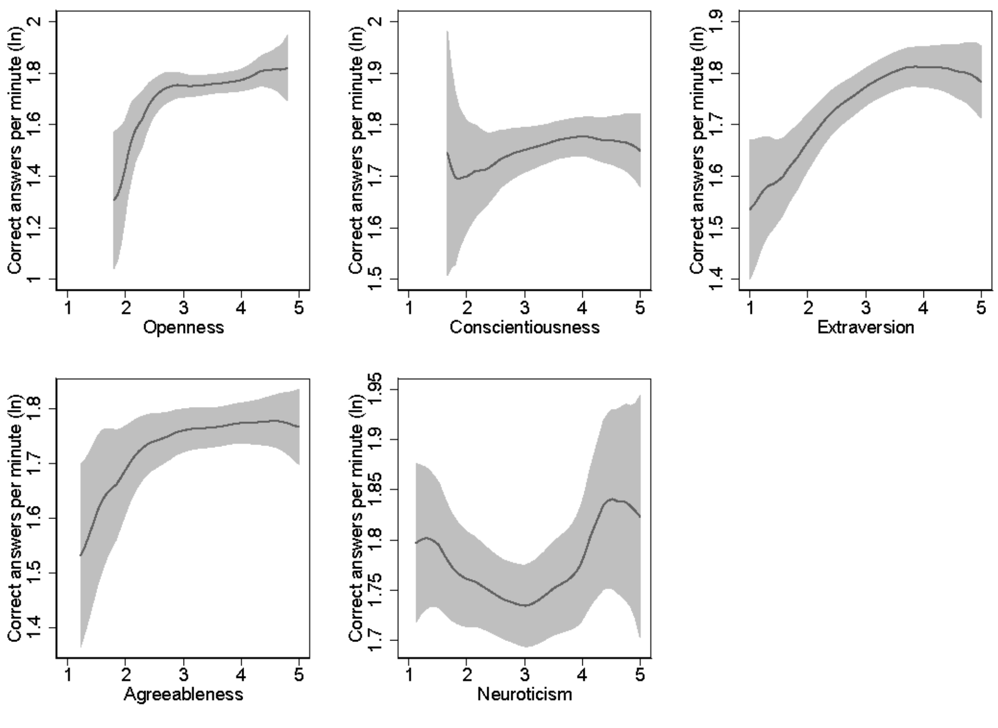
Source: Dataset with results drawn from microworkers.com (N=250). Author's calculations.

Figure 2. Performance across the distribution of personality traits

(A: Correct answers)



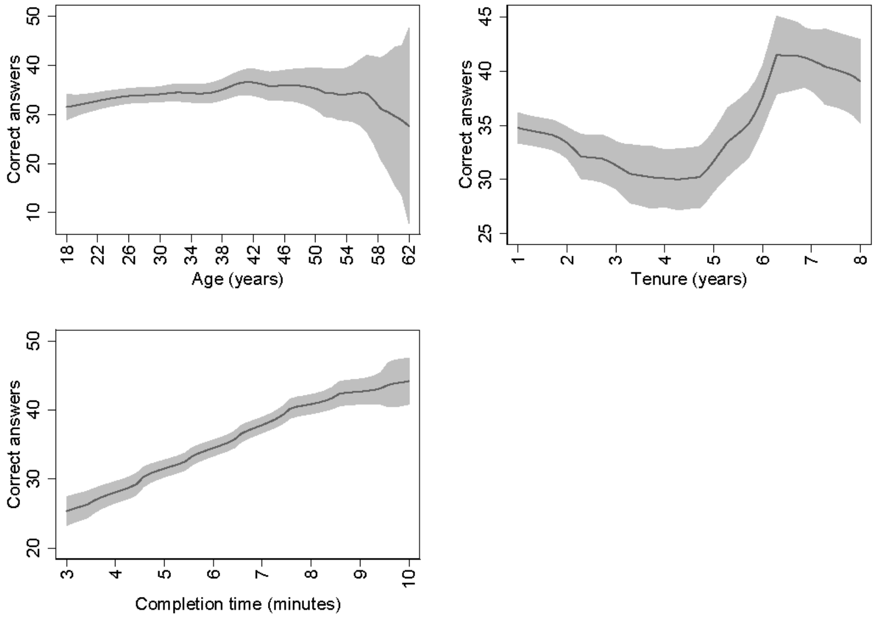
(B: Correct answers per minute)



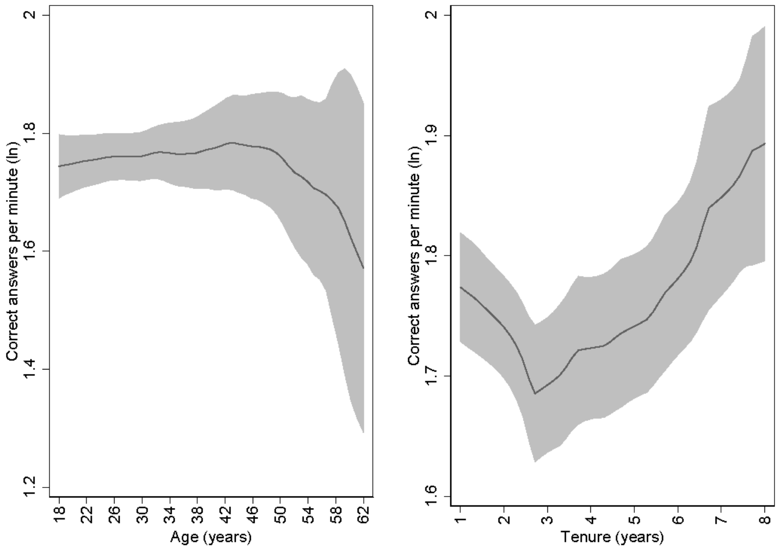
Source: Dataset with results drawn from microworkers.com (N=250). Author's calculations.
Notes: Local polynomial smoothing with confidence bands (shaded area).

Figure 3. Performance across the distribution of selected indicators

(A: Correct answers)



(B: Correct answers per minute)



Source: Dataset with results drawn from microworkers.com (N=250). Author's calculations.
Notes: Local polynomial smoothing with confidence bands (shaded area).

Table 1. Summary Statistics

	Mean	S.D.	Min.	Max.
Total correct answers	33.86	9.34	10.00	52.00
Total correct answers per minute	6.02	1.69	1.87	14.67
Ratio of correct answers (out of 56 words)	0.60	0.17	0.18	0.93
Openness	3.54	0.47	1.80	4.80
Conscientiousness	3.60	0.61	1.66	5.00
Extraversion	3.22	0.82	1.00	5.00
Agreeableness	3.56	0.67	1.22	5.00
Neuroticism	2.78	0.73	1.12	5.00
Tertiary education (0/1)	0.50	0.50	0	1
High computer competence (0/1)	0.52	0.50	0	1
Tenure (years)	2.87	1.99	1.00	8.00
Completion time (minutes)	5.82	1.67	3.00	10.00
Female (0/1)	0.33	0.47	0	1
Age (years)	30.10	8.28	18.00	62.00
East Asia & Pacific (0/1)	0.04	0.20	0	1
Europe (0/1)	0.62	0.49	0	1
Latin America & Caribbean (0/1)	0.05	0.22	0	1
Middle East & North Africa (0/1)	0.07	0.25	0	1
North America (0/1)	0.10	0.30	0	1
South Asia (0/1)	0.12	0.33	0	1

Source: Dataset with results drawn from Micro-workers.com (N=250). Author's calculation.

Table 2. Performance across groups of micro-workers

	Total correct answers			Total correct answers per minute		
	Mean	S.D.	Mean Difference [2]-[1]	Mean	S.D.	Mean Difference [2]-[1]
Male [1]	32.66	9.16	3.60	5.81	1.54	0.62
Female [2]	36.26	9.29	(2.91)	6.43	1.90	(2.77)
Less than tertiary education [1]	28.42	9.06	10.87	5.70	1.59	0.63
Tertiary education [2]	39.29	5.80	(11.29)	6.33	1.74	(2.95)
Low computer competence [1]	29.01	8.32	9.38	5.88	1.79	0.26
High computer competence [2]	38.40	7.87	(9.16)	6.15	1.59	(1.24)
Non-Asian [1]	35.98	7.90	-12.93	6.10	1.65	-0.55
Asian [2]	23.05	8.65	(9.42)	5.55	1.81	(1.92)
East Asia & Pacific	28.63	5.18		7.03	1.47	
Europe	35.64	7.68		6.04	1.63	
Latin America & Caribbean	34.38	5.85		6.17	1.56	
Middle East & North Africa	37.35	5.33		6.74	2.21	
North America	37.96	11.06		6.04	1.43	
South Asia	21.00	8.82		5.02	1.63	

Source: Dataset with results drawn from Microworkers.com (N=250). Author’s calculations.
Note: |t|-statistic in parentheses.

Table 3. Microtask performance, demographics and human capital

	[1]	[2]	[3]	[4]
Panel A. Dependent variable: Total correct answers (OLS)				
Female	2.684** (1.120)	1.991** (1.004)	1.619* (.907)	1.705* (.902)
Age	.511 (.403)	-.003 (.349)	-.140 (.314)	-.213 (.315)
Age-squared	-.007 (.005)	-.001 (.004)	.002 (.004)	.003 (.004)
Asian	-12.546*** (1.445)	-8.962*** (1.342)	-8.225*** (1.337)	-7.540*** (1.315)
Tertiary education	-	8.438*** (1.036)	6.958*** (1.014)	7.056*** (.959)
Computer competence	-	-	6.543*** (.872)	6.154*** (.812)
Tenure 3-5 years	-	-	-	-.491 (.883)
Tenure 6+ years	-	-	-	5.198*** (1.300)
Constant	26.802*** (6.719)	30.816*** (5.703)	29.967*** (5.173)	30.731*** (5.194)
F-test (Overall)	21.76***	48.98***	71.57***	56.60***
R-squared	.288	.463	.575	.610
Panel B. Dependent variable: Ratio of correct answers (Fractional Probit, AME)				
Female	.048** (.019)	.036** (.017)	.030* (.016)	.030* (.016)
Age	.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Asian origin	-.224*** (.025)	-.157*** (.024)	-.144*** (.023)	-.130*** (.023)
Tertiary education	-	.151*** (.018)	.124*** (.018)	.126*** (.017)
Computer competence	-	-	.117*** (.015)	.110*** (.014)
Tenure 3-5 years	-	-	-	-.009 (.015)
Tenure 6+ years	-	-	-	.092*** (.022)
Wald-test (Overall)	83.80***	233.30***	406.49***	428.63***
Pseudo R-squared	.024	.040	.050	.054

Source: Dataset with results drawn from Microworkers.com (N=250). Author's calculations.
Notes: Robust standard errors in parentheses (Panel A). Standard errors in parentheses estimated by Delta method (Panel B). AME: Average marginal effects. Tenure of less than 3 year is the reference category.
*** p<0.01, ** p<0.05, *p<0.10

Table 4. Microtask performance, personality traits, demographics and human capital

	[1]	[2]	[3]	[4]
Panel A. Dependent variable: Total correct answers (OLS)				
Openness	.203 (.502)	.060 (.479)	.001 (.438)	.001 (.429)
Conscientiousness	-.467 (.612)	-.188 (.544)	-.176 (.479)	-.361 (.461)
Extraversion	3.420*** (.493)	2.582*** (.423)	2.301*** (.393)	2.274*** (.394)
Agreeableness	1.197** (.522)	1.277*** (.450)	1.051*** (.386)	.923** (.379)
Neuroticism	-.137 (.583)	.359 (.475)	.564 (.456)	.616 (.453)
Female	2.325** (1.001)	1.774* (.933)	1.457* (.855)	1.604* (.850)
Age	.461 (.374)	.009 (.327)	-.128 (.294)	-.205 (.295)
Age-squared	-.006 (.005)	-.001 (.004)	.002 (.004)	.003 (.004)
Asian	-10.071*** (1.355)	-7.918*** (1.251)	-7.463*** (1.286)	-6.918*** (1.268)
Tertiary education	-	6.923*** (1.040)	5.880*** (1.022)	6.055*** (.921)
Computer competence	-	-	5.856*** (.815)	5.87*** (.767)
Tenure 3-5 years	-	-	-	-.110 (.862)
Tenure 6+ years	-	-	-	4.918*** (1.318)
Constant	27.307*** (6.224)	31.161*** (5.362)	30.491*** (4.888)	31.065*** (4.889)
F-test (Overall)	25.27***	40.47***	51.67***	45.03***
R-squared	.445	.551	.638	.667
Panel B. Dependent variable: Ratio of correct answers (Fractional Probit, AME)				
Openness	.003 (.008)	.001 (.008)	.001 (.007)	.001 (.007)
Conscientiousness	-.008 (.010)	-.003 (.009)	-.003 (.008)	-.006 (.008)
Extraversion	.061*** (.008)	.046*** (.007)	.042*** (.007)	.040*** (.007)
Agreeableness	.020** (.009)	.022*** (.007)	.017*** (.006)	.015*** (.006)
Neuroticism	-.002 (.010)	.005 (.008)	.009 (.008)	.009 (.008)
Female	.043** (.018)	.033** (.016)	.028* (.014)	.030** (.014)
Age	.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Asian origin	-.178*** (.023)	-.137*** (.022)	-.128*** (.022)	-.117*** (.022)
Tertiary education	-	.123*** (.017)	.104*** (.017)	.107*** (.016)
Computer competence	-	-	.104*** (.014)	.099*** (.013)
Tenure 3-5 years	-	-	-	-.001 (.014)
Tenure 6+ years	-	-	-	.087*** (.022)
Wald-test (Overall)	215.00***	384.28***	537.73***	557.84***
Pseudo R-squared	.039	.048	.056	.059

Source: Dataset with results drawn from Microworkers.com (N=250). Author's calculations.
Notes: Robust standard errors in parentheses (Panel A). Standard errors in parentheses estimated by Delta method (Panel B). AME: Average marginal effects. Tenure of less than 3 year is the reference category.
*** p<0.01, ** p<0.05, *p<0.10

Table 5. Microtask performance, personality traits, demographics, human capital and effort

	Total correct answers (OLS)		Ratio of correct answers (Fractional Probit, AME)	
	[1]	[2]	[3]	[4]
Openness	-	.172 (.413)	-	.003 (.007)
Conscientiousness	-	-.400 (.457)	-	-.006 (.008)
Extraversion	-	2.020*** (.396)	-	.036*** (.007)
Agreeableness	-	.788** (.351)	-	.012** (.006)
Neuroticism	-	.673 (.426)	-	.010 (.007)
Female	2.025** (.851)	1.880** (.788)	.037** (.015)	.035*** (.014)
Asian	-5.500*** (1.239)	-5.091*** (1.226)	-.091*** (.021)	-.082*** (.021)
Tertiary education	5.539*** (.930)	4.830*** (.856)	.097*** (.016)	.082*** (.015)
Computer competence	4.540*** (.840)	4.220*** (.788)	.080*** (.015)	.074*** (.013)
Tenure 3-5 years	-.158 (.819)	.154 (.822)	-.003 (.014)	.003 (.014)
Tenure 6+ years	4.717*** (1.125)	4.574*** (1.139)	.083*** (.020)	.080*** (.019)
Completion time (minutes)	1.684*** (.291)	1.530*** (.281)	.030*** (.005)	.027*** (.005)
Constant	20.947*** (5.042)	21.951*** (1.835)	-	-
F-test (Overall)	69.71***	56.30***	-	-
R-squared	.668	.712	-	-
Wald-test (Overall)	-	-	575.15***	692.25***
Pseudo R-squared	-	-	.059	.063

Source: Dataset with results drawn from Microworkers.com (N=250). Author's calculations.
Notes: Robust standard errors in parentheses (OLS). Standard errors in parentheses estimated by Delta method (Fractional Probit). AME: Average marginal effects. Tenure of less than 3 year is the reference category. All models include a second order polynomial of age.
*** p<0.01, ** p<0.05, *p<0.10

Table 6. Effort-adjusted performance, personality traits, demographics and human capital

	[1]	[2]	[3]	[4]	[5]	[6]
Openness	-	-	-	-	.023 (.018)	.022 (.018)
Conscientiousness	-	-	-	-	-.014 (.023)	-.014 (.023)
Extraversion	-	-	-	-	.055*** (.020)	.048*** (.019)
Agreeableness	-	-	-	-	.030 (.018)	.026 (.018)
Neuroticism	-	-	-	-	.028 (.021)	.033 (.021)
Female	.098** (.039)	.092** (.039)	.091** (.039)	.091** (.039)	.088** (.038)	.086** (.038)
Asian	-.104** (.051)	-.069 (.054)	-.067 (.054)	-.056 (.054)	-.065 (.054)	-.038 (.056)
Tertiary education	-	.081** (.037)	.076** (.036)	.078** (.036)	-	.063* (.035)
Computer competence	-	-	.024 (.034)	.018 (.034)	-	.008 (.033)
Tenure 3-5 years	-	-	-	-.004 (.037)	-	.004 (.039)
Tenure 6+ years	-	-	-	.083** (.041)	-	.080* (.043)
Constant	1.662*** (.193)	1.701*** (.195)	1.698*** (.196)	1.709*** (.198)	1.663*** (.198)	1.702*** (.199)
F-test (Overall)	3.41***	4.16***	3.62***	3.47***	3.58***	3.15***
R-squared	.055	.075	.077	.088	.118	.138

Source: Dataset with results drawn from Microworkers.com (N=250). Author's calculations.
Notes: OLS estimates. The dependent variable is the natural logarithm of the ratio of total correct answers to the number of minutes in the task. Robust standard errors in parentheses. Tenure of less than 3 year is the reference category. All models include a second order polynomial of age.
*** p<0.01, ** p<0.05, *p<0.10