



# The condition of the Turking class: Are online employers fair and honest?

John J. Horton \*

Harvard University, 383 Pforzheimer Mail Center, 56 Linnaean Street, Cambridge, MA 02138, United States

## ARTICLE INFO

### Article history:

Received 19 January 2010  
Received in revised form 11 November 2010  
Accepted 1 December 2010  
Available online 10 December 2010

### JEL classification:

J4  
J8  
D6

### Keywords:

Crowdsourcing  
Experimentation  
Amazon mechanical turk  
Online labor markets

## ABSTRACT

Critics of online labor markets claim that employer abuses are endemic in these markets. Surveying a sample of workers, I find that, on average, workers perceive online employers to be slightly fairer and more honest than offline employers.

© 2010 Elsevier B.V. All rights reserved.

## 1. Introduction

Work conducted over the Internet by workers participating in online labor markets has begun to attract mainstream attention, much of which has been negative. Most of the criticism targets low-skilled, low-paying piece-work sites like Amazon's Mechanical Turk (MTurk), though other remote work sites like oDesk and Elance are also drawing scrutiny [2]. Some critics worry that much of the work is of dubious social value, with many employers using workers to generate spam and write bogus product reviews. Other critics argue that buyers are actively exploiting workers and circumventing labor laws. In a recent *Newsweek* article entitled "Work the New Digital Sweatshops," Harvard Law School Professor and Berkman Center for Internet & Society co-founder Jonathan Zittrain characterized online labor markets as digital sweatshops. At a recent conference on digital labor at the New School, expropriation and exploitation were common themes. Aside from low pay, critics argue that workers do not know the (potentially unethical) purpose of their work and have no ability to organize or appeal the decisions of capricious employers [3,5].

Despite these real and perceived downsides, online labor markets have a tremendous and potentially transformative upside, which is that the markets give people in poor countries access to buyers in rich countries. If this form of increased virtual labor mobility has effects

similar to those of increased real labor mobility, then the emergence of online labor markets should be lauded and supported; the welfare gains from liberalizing restrictions on labor mobility would be truly enormous. Clemens et al. [1] consider the effect relocation to the US would have on the *real* wages of workers from different countries: for the median country (Bolivia), wages would increase by a factor of 2.7, and for the highest country (Nigeria), wages would increase by a factor of 8.4. Even with current strict limits on migration, the World Bank estimates that in 2008 remittances to developing countries totaled over \$305 billion, exceeding both private capital flows and official development aid.

The comparative advantages of the world's poor are that they (either individually or collectively through political institutions) are willing to accept environmental degradation, dangerous working conditions and very low pay. In light of this unpleasant truth, the relative virtues of digital work are obvious: it poses no physical danger to workers, has virtually no environmental impact and does not require robust host country institutions or local entrepreneurial talent. Workers can set their own hours and are not exposed to the elements, dangerous working conditions, the vagaries of agriculture or tyrannical bosses. Unlike labor market access gained through physical migration, workers do not have to live apart from their families or dissipate their earnings by paying developed-country prices for shelter, food and clothing.

The perceived costs and the potential benefits of online labor markets are fairly self-evident; good public policy will require some attempt to quantify the trade-offs under different policy scenarios. Consider a proposal to require Amazon to verify that each new task is not being used

\* Tel.: +1 617 595 2437.

E-mail address: [john.joseph.horton@gmail.com](mailto:john.joseph.horton@gmail.com).

to generate spam. While this may have the intended effect of reducing spam, compliance costs might exceed the per-transaction profits, thus pricing all work out of the market – or it might not, leading to an overall gain in welfare. Fortunately, we do not need to limit ourselves to speculation and conjecture: it is easy to test hypotheses about online phenomena by running experiments.

**2. The experiment**

In MTurk, the decision whether to pay a worker is left to the employer’s discretion; nothing prevents a buyer from expropriating a worker’s product. One might therefore presume that employers in MTurk would regularly cheat their workers. If this is true, then MTurk workers should have dim views of the honesty and fairness of online employers compared to offline employers who operate under a formal framework of sanctions for unethical behavior. To test this proposition, I conducted a simple experiment in which subjects recruited from MTurk were asked to answer one of the following questions:

- What percentage (between 0 and 100) of employers in your home country would you estimate treat workers honestly and fairly?
- What percentage (between 0 and 100) of Mechanical Turk Requesters would you estimate treat workers honestly and fairly?

Subjects were randomly assigned to a question and only saw that particular question. In other words, subjects asked about home country employers were not asked about online employers, and vice versa. I launched the experiment on December 23, 2009, and left it open for 7 days. Exactly 200 subjects completed the HIT, but only 192 responses were usable.<sup>1</sup> I paid 12 cents per response, which gave an hourly average wage of \$5.68, and my total expenditure was \$26.40. Consistent with Amazon’s guidance on what constitutes a “good” feedback score, the HIT was limited to MTurk workers with a 95% approval rating. When asked to name their home countries, 111 subjects reported the US, 58 reported India and 23 reported some other country.

**2.1. Results**

A histogram of worker responses by question type, with a bin width of 5, is shown in Fig. 1. In each panel, the mean response is indicated with a vertical line, and 90% and 95% confidence intervals are shown with red shaded bands around the mean. We can see that (1) means are quite similar and (2) responses in the online case (bottom panel) are slightly more polarized, with a greater number of subjects reporting very positive views of online employers.

We can compare mean responses in the two groups using a linear regression of the reported percentage of honest and fair employers,  $perc_i$ , on an indicator for whether the subject was asked about online employers,  $MTurk_i = 1$  or offline, host country employers,  $MTurk_i = 0$ . The fitted regression line, with robust standard errors, is:

$$perc_i = \underbrace{5.21}_{[3.58]} \cdot MTurk_i + \underbrace{64.38}_{[2.36]}$$

with  $R^2 = 0.011$  and  $N = 192$ . We can see that the mean percentage for offline employers was a little more than 64% and for online employers slightly more than 69%. The difference is not statistically significant.

Confirming the graphical evidence that more workers have very positive views of online employers compared to offline employers, a regression of an indicator for whether the subject perceived that more than 80% of the employers were honest and fair yields a positive and

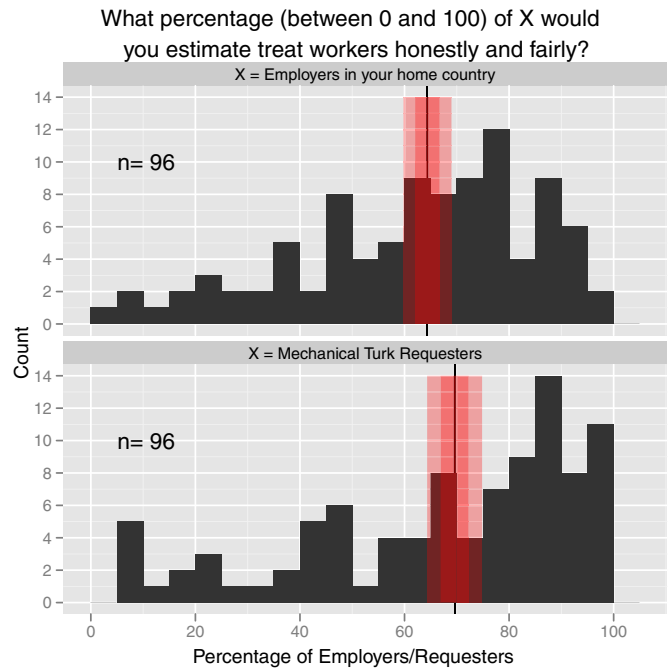


Fig. 1. Worker attitudes about online and offline employers.

highly significant coefficient on the group assignment indicator,  $MTurk_i$ :

$$1\{perc_i > 80\} = \underbrace{0.22}_{[0.07]} \cdot MTurk_i + \underbrace{0.22}_{[0.04]}$$

with  $R^2 = 0.054$  and  $N = 192$ .

Perhaps surprising, given the cross-national differences in institutions, Indian and non-Indians (composed mostly of US-based workers) had very similar responses:

$$perc_i = -\underbrace{3.00}_{[5.57]} \cdot MTurk_i + \underbrace{4.71}_{[4.32]} \cdot India_i + \underbrace{1.81}_{[7.95]} \cdot India_i \times MTurk_i + \underbrace{65.25}_{[2.71]}$$

with  $R^2 = 0.013$  and  $N = 192$ .

**3. Discussion**

These findings come with several caveats. Experimenter effects could matter – subjects might be encouraged to exaggerate how honest and fair they find AMT employers, because the question was asked by an MTurk employer. An unavoidable limitation is that subjects are not a random sample of MTurk workers – the experiment cannot reach workers who may have stopped using MTurk because of bad experiences. Additionally, the 95% cutoff might preclude the participation of a large number of disgruntled workers, though this seems unlikely; in past experiments, I have found very little difference in uptake under different cutoffs, suggesting that most workers have high scores.

The critique of online labor markets goes beyond the perceived fairness of employers. Furthermore, worker perceptions of fairness are not measures of actual fairness. Nevertheless, this experiment offers evidence that MTurk workers view their chances of being treated fairly online as being as good or better than offline. Contrary to our prior expectations, rampant exploitation is a mis-characterization.

<sup>1</sup> The author’s website contains the raw data from MTurk, the code that cleans the data and all code used for the analysis and plots. The main cause was workers offering a range rather than a point estimate.

Future research should investigate other claims related to online markets: How prevalent are dubious tasks? Do workers get repetitive stress injuries, as is often suggested? Do workers feel they are gaining skills? Answers to these questions could help clarify the trade-offs inherent in different policy proposals. Online work is currently a small phenomenon compared to the global trade in services, but it will become far larger and will eventually attract more policy-oriented attention. Given the welfare consequences of online work, it would be a tragedy if supposition and conjecture about easily and cheaply answerable empirical questions determined our policy towards digital markets.

### Acknowledgments

John Horton thanks the Berkman Center for Internet and Society and the NSF-IGERT Multidisciplinary Program in Inequality & Social Policy for generous financial support. Thanks to Aaron Shaw, Roberta

Yerkes Horton, Justin Keenan, Richard Zeckhauser, Carolyn Yerkes and Fara Dina Aquino for helpful comments and suggestions. All plots were made using the open source R package `ggplot2`, developed by Hadley Wickham [4].

### References

- [1] Clemens, M., Montenegro, C., Pritchett, L., Paraguay, D., Floor, T., 2008]. The place premium: wage differences for identical workers across the US border. World Bank Policy Research Working Paper No. 4671.
- [2] Felstiner, A., 2010. Working the crowd: Employment and labor law in the crowdsourcing industry. Working Paper, Boalt Hall Law School, UC Berkeley.
- [3] Irani, L., 2009]. Agency and exploitation in mechanical turk. Internet as Playground and Factory Conference. New School, New York, NY. Nov 13.
- [4] Wickham, H., 2008]. `ggplot2`: an implementation of the grammar of graphics. R package version 0.7, URL: <http://www.CRAN.R-project.org/package=ggplot2>
- [5] Zittrain, J., 2008]. Ubiquitous human computing. *Philosophical Transactions of the Royal Society* 3813–3821.