

# Online Labor Markets II: Market Power Asymmetries and Possible Mediation

March 22, 2017

Speakers: Kira Goldner and Sid Suri (guest speaker from Microsoft Research NYC)

**Disclaimer:** Many of the notes below include direct quotes from the discussed papers without direct citation.

## 1 Accounting for Market Frictions and Power Asymmetries by Kingsley, Gray, Suri [4]

Mechanical Turk (MTurk or AMT) is an Amazon interface. ‘Requesters’ post jobs, mainly very small and low-payment tasks, e.g, ‘is this a cat or dog’. There are about 10,000 active workers in a month, about half of whom turn over in 6 months.

- Data used in the paper:
  1. MTurk surveys embedded after other tasks. One fundamental problem is getting a representative survey of the platform. Posting tasks that are surveys skews toward the worker population that completes survey tasks. By embedding them at the end of other tasks for extra payment, the authors sidestep this skewing.
  2. Ethnographic data from interviews.
  3. Geographic mapping task: Workers were paid to (1) put a pin on a map wherever they are and (2) report how they find out about the task.
- High-level questions:
  1. Is information perfect – is it equally available to all MTurk participants?
  2. Is there market concentration (a small number of major requesters)?

### 1.1 Information Asymmetry

**Main idea:** One major problem with MTurk is the asymmetry of information in favor of the requesters.

- Visibility of information about the other side of the market:
    - Requesters can view workers’: past work-acceptance-by-requester rates
-

- Workers *cannot* view requesters’: past rejection rates, responsiveness to worker attempts to communicate, and payment history
- Penalties for bad behavior to the other side of the market:
  - Requesters can penalize workers they deem to be bad actors by:
    1. Blocking workers from doing tasks
    2. Withholding payments
    3. Rejecting work without reason (though sometimes keeping the output submitted by workers)
    4. Reporting workers to Amazon.com (which can lead to suspension of worker accounts)
  - Workers can do nothing.
- **Question:** How easy is it for workers to create new profiles?
 

**Answer:** It’s fairly simple to create a new profile, but difficult to create a desirable new profile. (In order to be paid in US dollars rather than Amazon credit, they need to link a bank account.)

**Main idea:** Workers see far less information and have far higher search costs to find good employers, good wages, etc. Workers have developed some solutions to share information about requesters.

- Turkoption (by Irani and Silberman) is a browser plug-in to rate requesters, and it is widely adopted by workers. BUT
  - Requesters can still unilaterally reject work and block workers.
  - Turkoption does not provide any method for workers to have their reputation repaired or to regain lost wages.
- Online forums have been created where workers share information about potential employers
- In the experimental mapping task, just under half of the workers were referred via an online forum. (Figure 1 in [4])
- Forums are an aberration in a well-functioning market, but they do more than offer information; they also:
  - Offer social cohesion, mentorship, help workers build a sense of identity, and offer them entertaining breaks from work routines.
  - Forums represent the trade-offs between efficiency and social interaction found in a marketplace (Lehdonvirta & Castronova, 2014, p. 130, figure 7.2)

**Main idea:** Even with these solutions, there’s still power asymmetry and problems with the market.

- Forums aren’t directly on the platform, so not all workers know about them, and workers face a significant impediment for market participation.
-

- Recruitment costs are shifted to workers: job seekers pay this cost to secure future employment, and it is usually considered that their future wage/salary is the return for this cost. However, because MTurk work is short term, workers are not receiving this return and bearing too much of the cost.
- Search time isn't free! It's a cost or rent that workers pay to find decent work online, and this should be factored into their compensation for jobs.
- Lack of information leads to bad matches! Workers struggle to figure out what type of work they're good at, workers have no work-specific rating, etc. The matches are inefficient, and as a result, there are lower wages, lower quality of work, and more rejected work.
- Prices are well below equilibrium prices, as there is no shortage of work; this may explain why quality of work is not responsive to increased prices (an earlier finding).
- **Remark from Sid:** The evolving narrative for why higher wages don't buy higher quality work is 'workers are dumb!' From ethnographic interviews, it's clear that these workers are actually quite intelligent, and are successfully doing what many people can't – making money on the internet. A better explanation: there is no incentive to provide higher quality work, as different people doing the same work get the same pay.
- **Question:** Do you see these workers unionizing in any form?  
**Answer:** They coalesce around the online forums, which are sort of a pseudo-union. There have been efforts by people to unionize the workers, but the workers seem to see through it, and feel like the union organizer isn't there to help them.
- **Question:** What are the problems we hope to fix?  
**Answer:** Bad matches! Workers can't find work, wages are too low, and quality of work is low. Workers are incentivized to read instructions once and then perform the same tasks because they don't get compensated for "training" or instruction-reading. This leads to workers sticking with jobs that are not necessarily a good match for their skill set, resulting in bad equilibria. Observe: MTurk uses fixed wages and there is no bargaining. There are high search costs and frictions.
- **Question:** Are there third party platforms that offer better recruiting of workers for requesters, easier search for workers, or better quality matching?  
**Answer:** Crowdfunder – they post a company's tasks on MTurk for them, but they didn't do a good job and eventually mutually parted from MTurk.

## 1.2 Market Concentration

**Main idea:** [4] shows that market power is super concentrated!

- .1% of requesters (~10-50) request 30% of activity
    - The top 10% of requesters are defined by the frequency at which they posted tasks to the market.
    - Figure 2 of [4] shows that 10% of all requesters post approximately 98-99% of all tasks to the MTurk platform.
-

- Figure 3 of [4] shows that approximately 98-99% of all tasks on MTurk are posted to only 10% of HIT Groups.
  - A HIT Group is a collection of similar tasks posted by the same requester each for the same pay.
  - **Remark from Sid:** The fact that requesters are concentrated is not surprising, but the implications are that they have market power over wages.
  - **Question:** Are the tasks substitutable?  
**Answer:** Some are (less skilled tasks, e.g. image matching), but skilled ones (translation) are not so much substitutable.
  - Requester concentration implies that:
    - Workers have fewer wage-earning opportunities outside this top percent of requesters.
    - In economic terms, this means that a small fraction of the total number of requesters likely capture the majority of labor supply on MTurk.
  - Ex-ante wages:
    - Economic theory says that market distortions which limit the ability of workers to negotiate wages produce market power that favors employers, since employers rather than the market set wage rates (Ashenfelter, Farber, & Ransom, 2010a, 2010b; Manning, 2003; Staiger, Spetz, & Phibbs, 2008).
    - There’s no evidence to this, but combined with market concentration, it leads to belief this does favor requesters.
    - Workers have few alternatives to accepting: either search for other tasks (and pay added search costs) or walk away from MTurk altogether.
    - Reputations and skills from MTurk are not transferrable to other labor markets, so workers can’t capture surplus profits or extract rent from their experience.
    - High quality workers get no benefit. They’re not paid according to job performance or higher education/skills.
  - **Question:** Is there a reputation system?  
**Answer:** Yes for workers (percentage of accepted work), not for requesters on platform (but yes via the Turkopticon plug-in). Workers are also extremely protective of their approval rating.
  - **Question:** It seems like then that there are good incentives to do a good job.  
**Answer:** Yes, but only to do just enough work for it to be marked ‘accepted’ by the requester. Sid has a follow-up paper [3]:  
*Control:* Find as many typos as you can.  
*Experiment:* Find as many typos as you can, as long as you find one it will not be rejected. Performance-based payment (if you find 75% of what other searchers find) boosted performance.
  - **Question:** Do workers set a gold standard?  
**Answer:** Yes, sometimes.
-

- **Question:** Employers dominate the offline labor market in some markets. How much greater is the skew in some offline labor markets compared to online labor markets, or specifically MTurk?

**Answer:** Unknown.

- **Question:** Has anyone built tools for performance based payment? The fact that there's no ground truth isn't unique to MTurk. There's Duolingo, ESP Game, and a lot of research from peer prediction people. Why is the absence of ground truth specifically problematic here?

**Answer:** Note that Duolingo and ESP game use gamification and intrinsic motivators, as opposed to extrinsic. There has been work on intrinsic motivators on MTurk, e.g. tasks that ask workers to classify a cell as cancerous or not for the greater scientific good. They've found that this motivates higher quantity of this work, but not higher quality.

- Note: There *is* a way to pay for quality, or for speed, by offering bonuses for certain metrics. However, these are still all set ex ante.

- **Question:** Why are incentives to do higher quality work not usually set up?

**Answer:** The literature on how to do so is not clear. The finding that you can incentivize for quality is only 1-2 years old. Also, this only holds for tasks that are 'effort-responsive', and Sid suspects that most tasks on MTurk are not effort-responsive. (E.g. handwriting recognition is not effort-responsive, whereas finding typos is.)

- **Question:** What's the distribution of skill levels, e.g. compared to ODesk? Is quality an important dimension?

**Answer:** Sometimes the problem is that the work cannot be done. Many requesters join MTurk, expect a certain baseline quality that is not met, and then stop using MTurk. This is the primary reason that requests walk away. Sometimes, however, the solution is as simple as restricting the task to those who are qualified, e.g. English speakers if the task requires English.

### 1.3 Implications of Market Power Imbalance

Imbalance issues:

- Asymmetric reputation systems
- Concentration of requesters
- Requesters set wages
- Requesters unilaterally reject work

No existing online labor market alternatives fix these issues, so the workers must either pay search costs by using Turkopticon or browsing forums, or they must take on risky requesters with unknown reputations. However, the workers always end up bearing this cost, not the requesters.

Claim: The following contribute to poor matches:

- That higher wages cannot be set for good workers.
-

- Concentration of requesters  $\implies$  less diversity  $\implies$  harder to find tasks suited to workers.

Evidence this is happening:

- Mason and Suri [5] advocate for recruiting a panel of trusted workers for research studies. The idea is to maintain a ‘trusted worker pool’. Pre-recruit and pre-test a set of workers for tasks based on their performance in previous tasks designed by the same requester. Then, contact them to let them know when the tasks will be posted, asking them to participate. This idea of using a trusted worker pool was used a lot in follow-up work that ran experiments on MTurk.
  - Crowdfunder + SpeechInk (now SpeechPad) used to post on MTurk for other companies but now implement their own reputation systems for workers (because MTurk worker reputations were insufficient).
  - **Question:** Are these third-party recommendation systems, recommending workers to requesters? How did they operate?  
**Answer:** Yes, e.g. Crowdfunder, but they were very opaque in how they evaluated workers. Essentially they would post the tasks on MTurk on behalf of the real requesters and maintain their own evaluations/recommendations of workers.
  - **Question:** Are there theory papers about competing reputation systems, studying the equilibrium when competition is different? Do third party reputation systems do better than in-house reputation systems (different incentives)?  
**Answer:** Sid: Not to his knowledge.  
**Example:** Lyft’s built-in driver reputation system might do things differently and have different outcomes than a third party’s. What are the differences? How would having the competition change the system? Would Amazon buy a good third-party recommendation system?  
**Comment:** There’s literature on settings where revealing all information isn’t always optimal in terms of both revenue and welfare.
  - **Question:** Shouldn’t there be more fine-grained reputation in a labor market?  
**Answer:** Yes, especially because there is evidence that workers specialize (i.e. in translation) and should have specialty reputations.
  - **Question:** Why doesn’t Amazon fix the problems with Mechanical Turk?  
**Answer:** It doesn’t make Amazon a lot of money. The value of MTurk for Amazon is that it solves problems that they can’t get solved any other way. E.g. there’s a rumor that Amazon was trying to disambiguate between two products / two addresses. Amazon also has a history of building internal tools and externalizing the API.
  - **Follow up question:** Why let any external requesters post jobs?  
**Answer:** Volume. That is, to maintain a consistent pool of workers on the platform. Amazon also has a history of building tools for internal use and then externalizing the API.
  - **Question:** Are there other companies that let other players use their ‘platform’ in order to maintain volume?  
**Answer:** Power companies. Smelting companies use a lot of power—regular customers use
-

power during the day, smelting companies during the night. Also Bitcoin miners generate “free electricity” by running their mining rigs all the time, even when they don’t need to. Another example is free programming on network TV.

## 1.4 Proposed Solutions

Solutions for wages:

- Institute mechanisms that allow for negotiations. One solution would be a double auction between workers and requesters, encoding the bargaining over wages and jobs into a well-studied mechanism.
- Second solution: Dynamo. A platform built by researchers to support crowdworkers’ efforts to share information, collaborate, and determine guidelines for academic requesters in setting wages and designing tasks. Workers are given a central role on Dynamo in deciding what constitutes fair pay.
- Speed is an essential factor in OLMs – information and communication delays are costly. One idea that takes this into account is to incorporate chat service between requester/workers, but this doesn’t scale. An alternative is to communicate information quickly to all parties working in a virtual system. Answers to a prompt about what constitutes fair pay for a particular task could rapidly circulate opinions among participants. Some researchers are investigating systems-level visualizations playing a role here (e.g. Dynamo).

Solutions for imperfect information:

- (1) Give the workers equal opportunity to compete and (2) hold requesters accountable for (a) their on-platform behavior and (b) the quality of tasks they design.
- Build a platform for workers to review requesters on MTurk itself, plus a platform for workers to weigh in on the metrics or standards by which these ratings are constructed.
- Develop better communication tools – The Lead Genius has an in-platform real-time chat for workers who are doing projects together.

## 1.5 Policy Concerns

- Platform providers (like Amazon) do *\*NOT\** want to be defined as employers, or they’ll opt out of the market. Laws about the definition of employer and employer responsibilities impact the market-making decisions that the platform can make.
  - Requesters do *\*NOT\** want to be defined as employers, as they’d face prohibitive costs associated with the legal obligations of being an employer. This would harm both the requesters and workers depending on this income.
  - If platforms integrate technical fixes that support workers through training, collaboration, and information sharing, such enhancements might suggest that the platform curates a workforce, which policy governs.
-

- A class action lawsuit [19] brought against Crowdfunder (editorial crowdsourcing site) suggests we have not yet legally categorized the type of employment that crowdwork is.
- The platform usually doesn't set wage rates, but leaves it to requesters. (See Horton's work on minimum wage from our first meeting.) BUT the platform decides the space of allowable actions, so they really are the regulator here, and "at least in the US, it is hard to think of many workplace environments that are not at least minimally regulated to ensure the well-being and safety of both employers and their employees."
- Many platforms eliminate worker interactivity by delegating workflow through an API, but this paper claims that professional and personal relationships associated with work environments might still be necessary to constitute meaningful employment.
- This paper advocates for treating crowdsourcing labor markets according to their needs, and not those of traditional offline markets!
  - Enact new rules defining employment relationships and protect crowdworkers.
  - Institute enforcement mechanisms to hold bad actors on platforms responsible for their actions, especially requesters who commit cybercrimes, violate best practices (e.g. researchers not abiding by ethics standards of universities/IRB).
  - Consider mechanisms to make the role of the platform provider similar to those of a fiduciary, acting in the best interest of all parties on the platform and not a select few.
- **Remark from Sid:** The crowd is not a crowd, but a network. In another paper [2], we built "Facebook lite" for MTurk workers and mapped out their relationships.
- **Question:** What are the fundamental inefficiencies? Is bad matching the underlying problem? Is this independent of the other problems, e.g. wage problems.  
**Answer:** Yes. Workers can't find the jobs they're good at. They can't even figure out what jobs they're good at. Requesters can't find workers who are good at the jobs they're posting. This is the primary problem, but it's compounded by the addition factors: inability to negotiate wages, not being paid for quality of work, etc.

## 2 Can Reputations Discipline the Gig Economy by Benson, Sojourner, Umyarov [1]

### 2.1 Labor Market Basics (based on a conversation with Aaron Sojourner)

- The labor market is important because it's the most important thing that most people have to survive, earn a living, and thrive is their own time and skills. At the end of the day, most people have their time and skills to sell, not a car or house or condo like in the sharing economy. How well you can do in the labor market is the foundation for communities, families, and countries.
  - Information problem: Decentralization.  
 The labor market is not a classical perfect market. A lot of this is due to information problems. There's not a central clearinghouse, information is very widely distributed about
-



supply and demand, and there is no central posting of prices or openings. This has only very recently started to get systematized, organized, and standardized, but the centralized solutions are local and network-based.

- Information problem: Not knowing about what you're getting.  
The differences in jobs, workers, and employers are very hard to observe, and the differences matter to the other side of the market. This is more of an experience good than something you could observe. You might have an impression of the differences, but you can only learn about them over time through engagement. There's unobserved heterogeneity on both sides.
- Suppliers of labor have more bargaining power and better economic circumstances.  
Demanders of labor tend to be quite big relative to suppliers of labor. Suppliers are often in hard circumstances, whereas buyers are in better circumstances economically and have better bargaining power. Companies pool together capital from a lot of investors and hire management to collectively bargain for them with workers, suppliers, etc. Capital organizes itself, hires representation, and acts as a strong actor. On the other side, we have decentralized individual people who are largely disorganized and are squaring up against more powerful bargaining partners. (This demander-of-labor having more market power is what we've been discussing as observed in MTurk as well.)
- Online labor markets have helped with the decentralized aspects, helping more with the search process, driving down search costs. Websites such as indeed.com and glass door let workers go and observe things about companies and observe some of the previously unobserved heterogeneity. This addresses the second problem from the worker side.
- There is a lot of literature on the boss's problem of unobserved heterogeneity of workers, called "personnel economics." However, there's been very little attention to the reverse problem. How do workers build solidarity and share information?
- Turkopticon is essentially creating a public good by individual workers taking their private information from a risky experience with a private employer and contributing that to everyone else. A worker's narrow interest is to not share that. Yet a worker takes the costly action of reporting, and then other people benefit, while the worker doesn't benefit from this specific act.
- Sojourner's background: He started in organizing the carpenter's union. Musicians are the original gig economy, but construction is an old-school gig economy job. There's no expectation of a long-run employment relationship. The employers don't need the same number of employees every day, perhaps just 100 for three days and that's it. Unions improve jobs by bargaining, getting health insurance, creating apprenticeship and training programs for the next generation, act as the intermediary who run the hiring haul and assignments, maintain a list of who is out of work, set standards, etc.

## 2.2 Overview of the Paper

- A large literature discusses how to identify good workers. Economists and legal scholars theorize that reputational concerns constrain opportunistic employers, but there has been little empirical work. OLMs enable this. It's clear that employees care about employer reputations (beyond just information available in a contract).
-

- US Government Accountability Office (2015):
  - 40% of workers are employed in “contingent work arrangements”
  - “Online clearinghouses for obtaining ad hoc jobs are attempting to obscure or eliminate the link between the worker and the business..., which can lead to violations of worker protection laws.”
- For jobseekers lacking experience with an employer, can workers aggregate their private experiences into shared memory in order to discipline opportunistic employers?
- Useful things for an empirical study:
  - (1) ALL contracts have no enforcement or possibility for workers to seek recourse. This doesn’t vary.
  - (2) No native employer-reputation system, unlike other OLMs.
- Turkopticon: Workers compete for jobs, but this requires that workers voluntarily contribute accurate, private information to a common pool, which costs time and directs other workers to scarce, high-paying tasks. Also informativeness varies, and anyone can post.

This study

- Test employer’s blindly from a worker’s perspective and see if reputations align with employer/wage quality. (They do. This is surprising because the ratings are inflated.)
    - “One research assistant (RA) randomly selects tasks from employers who have good reputations, bad reputations, or no reputation and sends them to a second RA who is blind to employers’ reputations. We find that effective wages while working for good-reputation employers is 40 percent greater than effective wages while working for bad-reputation employers”
  - Test employer’s ability to recruit workers for different artificial employers and wage rates and posted reputations on Turkopticon.
    - “We create 36 employers on M-Turk. Using Turkopticon, we endow them with (i) 8-12 good ratings, (ii) 8-12 bad ratings, or (iii) no ratings. We then examine the rate they attract workers to posted jobs.”
    - “We find that employers with good reputations attract work about 50 percent more quickly than our otherwise-identical employers with no ratings and 100 percent more quickly than those with bad reputations.”
    - “Using estimates of M-Turk wage elasticities published elsewhere, we estimate that posted wages would need to be almost 200 percent greater for bad-reputation employers and 100 percent greater for no-reputation employers to attract workers at the same rate as good-reputation employers do.”
  - **Remark from Sid:** For the experiment with artificial employers, workers strongly opposed being experimented on.
-

### 3 Final Discussion

- **Open questions:**

- Designing a “good” reputation system.
- The effect of competing reputation systems.
- Fixing information asymmetry: what is symmetric information, what’s necessary/sufficient on either side for both parties to find a good match?
- Worker-requester matching: There are already many theory papers, but there’s very little empirical work on how well it’s done and how much it matters.
- (Panos Ipeirotis) How can we ensure fair payment schemes in the presence of uncertainty about true performance? How should you pay someone when you are not confident about the quality of the deliverable?
- (Panos Ipeirotis) How can we migrate people from a highly variable task-based / freelancer / gig-based payment approach into a salary-based one? And vice versa? What is the proper mechanism that allows the two choices to be interchangeable given the same amount and quality of work?
- How do you let workers bargain for wages on a task? Do a double auction for every task? What’s better than a posted price?

- **Question:** What is the appeal of MTurk? It is so broken.

**Answer:** First mover advantage.

- **Question:** Is Amazon passive with respect to the third party platforms, or have they shown resistance?

**Answer:** Passive.

### References

- [1] Alan Benson, Aaron J Sojourner, and Akhmed Umyarov. Can reputation discipline the gig economy? experimental evidence from an online labor market. 2015.
  - [2] Mary L Gray, Siddharth Suri, Syed Shoaib Ali, and Deepti Kulkarni. The crowd is a collaborative network. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 134–147. ACM, 2016.
  - [3] Chien-Ju Ho, Aleksandrs Slivkins, Siddharth Suri, and Jennifer Wortman Vaughan. Incentivizing high quality crowdwork. In *Proceedings of the 24th International Conference on World Wide Web*, pages 419–429. ACM, 2015.
  - [4] Sara Constance Kingsley, Mary L Gray, and Siddharth Suri. Accounting for market frictions and power asymmetries in online labor markets. *Policy & Internet*, 7(4):383–400, 2015.
  - [5] Winter Mason and Siddharth Suri. Conducting behavioral research on amazon’s mechanical turk. *Behavior research methods*, 44(1):1–23, 2012.
-