Works in Progress of John J. Horton

This statement describes my in-progress research projects. These are projects for which I have identified research questions and methods but not yet drafted abstracts or working papers. The projects are extensions of my current research, which focuses on using online labor markets to understand core problems in labor markets. I have divided the list of in-progress projects into core projects and non-core projects. Core projects are those that I conceived and for which I will take the lead on the analysis. Non-core projects are those where my role may be more limited.

Core Projects

Understanding and Fixing Marketplace Congestion

Labor matching markets of all kinds often have congestion problems: would-be trading partners have too many choices to make in too short a period of time. This can create a host of bad outcomes and inefficiencies such as unraveling and exploding offers. Unraveling is often a justification for imposing a strong "center" that implements a static matching mechanism. However, for many—probably most—markets, a centralized approach is infeasible.

My research suggests that congestion problems in decentralized matching markets typically originate from the fact that it is too cheap to send applications. A stylized version of the model is this:

1. Because it is very easy for job-seekers to send applications, job-seekers send many applications
2. Because each employer receives so many applications, the per-application probability of success is very low for everyone.
3. Because the per-application probability is low, job-seekers *might* send even more applications than they otherwise would, or they invest less time and effort into each application. Each application also carries less signal value
4. If job-seekers invest less in each application, it is harder for employers to assess whether a candidate is a match.
5. Because job-seekers apply without fully understanding the job, they commonly cannot actually perform the job once they learn the full details.

Electronic commerce is particularly beset with these problems because the cost of sending another application is almost zero. The causal chain outlined above suggests a marketplace intervention: change the costs of applying to jobs. We are currently launching an experiment that will do this by requiring applicants to

complete a proof-of-work-type exercise that will illustrate that they understand the job. At a high level, the experiment will tell us whether introducing some friction into the application process can improve outcomes.

In addition to this empirical work (the experiment and observational data), we will also develop a model of the congestion/under-investment phenomena. Some basic research questions the model should answer include:

1. What is the right number of applicants per worker?
2. How do applicants decide how much effort to put into each application?
3. How do the evaluation costs and application costs affect these investment decisions?
4. How does application quality affect employer screening choices?
5. When, if ever, can efficiency be improved by increasing the cost of applying for vacancies?

Although the context for this experiment is an online labor market, the issue appears to be much more general. Workers in traditional markets also wonder what is the "right" number of applications to send. See Figure 1, which is a screenshot of the auto-suggested queries (which are determined by frequency of that query) from Google. The problem of too-cheap applications arises in law review journal submissions, college admissions, online dating sites and academic job markets (the signaling mechanism was a response to this).
Several recent news articles have documented efforts by firms to try to screen candidates algorithmically, suggesting that they receive more applicants than they can reasonably handle. A recent "experiment" where a person posting a fake job on Craigslist received over 600 applications in less than 24 hours.

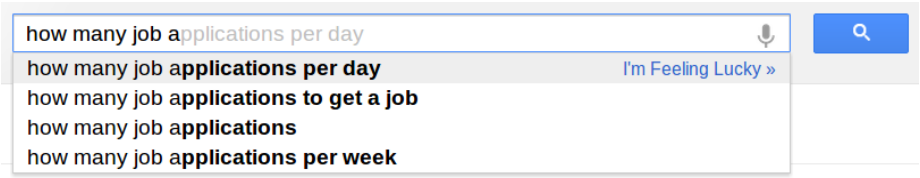

Figure 1: How many applications should I send?

This project is joint with Dana Chandler and Ramesh Johari.

Algorithmic Enhancement of Markets

A key motivation of the job market paper is the notion that matching efficiency in labor markets can be improved with algorithmic interventions. The paper

finds that these recommendations were very successful, despite being relatively algorithmically unsophisticated. While at oDesk, I have been involved with several projects to advance these ideas and reduce frictions more generally. Some of these software features are supply-focused, such as new tools that allow workers to more easily monitor the flow of jobs and learn about a worker's preferences. There are also several employer-oriented features, including one that shows employers to similar clusters of contractors, with clustered determined by a novel application-to-vacancy graph measure. Each of these features is being introduced experimentally and thus will allow us to understand how changes in match-making technology affect choices. The key research questions are: How do reduced search frictions affect search intensity and specificity, reservation wages and bargaining? To answer these questions, we will develop a model or augment an existing model from the search literature. These on-going experiments should provide many new opportunities to understand how technological change affects match-making.

Human Capital Acquisition and Recommendations

The basic idea of the Becker model of human capital is that workers acquire education until the marginal cost of more education is equal to the marginal increase in NPV of lifetime earnings [@becker1993human]. The endogenous choice is "years of education." This high level of abstraction hides the reality that most workers must continually make choices about what skills to learn over the course of their careers.

This choice is complex and multi-faceted. A worker considering learning some new skill X must consider:

1. What are the market returns to knowing X, given my current collection of skills?
2. How long are these returns likely to last, i.e., will this skill go out of date?
3. How much effort will it take to learn X, given my current skills?
4. What are the new skills I will be to more easily acquire if I have skill X, and what are the market returns and (and stability of those returns) for follow-on skills?

Instead of the one-dimensional human capital decision, workers are essentially trying to expand along a dynamic weighted graph. In traditional markets, these decisions are difficult to observe. In online labor markets, these choices are far more visible and measurable. At oDesk, one of my initiatives was the creation of a controlled skills vocabulary for use within the marketplace. This cannonical list of skills is used on all contractor profiles and jobs. The skills are also cross-referenced to the numerous skills tests that contractors can take (and that

employer can require of applicants). A side effect of this normalization is that we now have an enormous dataset illustrating which skills tend to go together, the wages associated with those skills and even common career trajectories. Within oDesk, we also have the power to make recommendations or perform other skill-related interventions. Some empirical research questions include:

1. Can exposing workers to information about the demand for various skills change their human capital accumulation decisions?
2. Can we make context-aware recommendations about skills that workers might find attractive to learn?
3. How does the market evolve—on both the demand and supply sides—as new-technologies emerge?

We will also develop a theory of skills that explains:

1. How workers decide which skills to acquire
2. How different skills are "related" to each other, in terms of market returns and ease of acquiring tjese skills

Aside from the importance that skills play in explaining labor market outcomes, there are other reasons to pursue this research. Technological changes should dramatically increase the demand for guidance on which skills to learn. We are witnessing a tremendous growth in the tools for the scalable delivery of education. If this technology unleashes a new era of algorithmically-assisted auto-didactism, people will still need to learn *what* to lean. This seems like an area ripe for marketplaces to use their holistic, bird's-eye-view of the marketplace to help inform these decisions.

This work is joint with Peter Coles.

Building the Online Laboratory

During graduate school, I developed a number of tools for conducting experiments using Amazon's Mechanical Turk, an online labor market focusing on micro-work. These tools were used in several of my papers [@horton2010onlinelab]. The tools include software for conducting randomizations, making payments, messaging workers and having them play simple economic style games. This part of my research agenda has been on hiatus during my time at oDesk; however as I return to research full time, I would like to continue this project and create more and better tools. One of the perspectives of my *Experimental Economics* paper is that the Internet should make it increasingly possible for the social sciences to become more of a "bench" science. This transation will require more tool

building, which although presently underemphasized compared to the natural sciences, should become more commonplace and valued as the new research possibilities of online markets and communities are realized more broadly.

In particular, I would like to:

1. Build a pool of paid subjects available for experiments and create tools for the building an managing of online experiments using oDesk workers
2. Create the ability to conduct targeted surveys on oDesk when users make certain marketplace decisions to more fully understand their decision-making
3. Create tools for managing large-scale and simultaneous games

The end goal of this research is to augment my other research projects, by making it easier to add an experimental or survey component to existing projects.

Allocation of Visibility

A key responsibility (and source of revenue) for two-sided platforms is the allocation of visibility. In traditional markets, buyers and sellers are responsible for finding each other in physical space and time. In computer-mediated markets, the platform—by choosing the interface and design—wholly determines how buyers and sellers find each other. This power has not gone unnoticed—the multi-billion-dollar paid positional search industry is direct consequence of this power [@edelman2005internet]—but it has received little attention in the particular context of online labor markets. Labor markets are different in that sellers are inherently supply-constrained. Because of this constraint, even a worker that is the "best" match for a particular job might be a very poor match in practice because their availability is limited at the particular time. Ideally, workers would self-report availability, but because job offers are useful even if not accepted (they can be used for bargaining power) and workers have free disposal on offers, getting them to honestly report availability is challenging. One can try to infer availability, but reliance on such inference can create bad incentives for workers to imitate the signals.

This paper will answer several research questions:

1. Assuming some ability to determine availability exists, how should the platform allocate visibility?
2. How should the platform allocate visibility subject to the constraint of assortativity (i.e., more able workers are always given more visibility than less able workers)?
3. Should some workers get "buried," never receiving any visibility?

4. Conditional upon wanting to give workers some expected amount of visibility, how should this be done in practice? Are there polynomial-time algorithms for doing so?

5. When would the platform like to auction off visibility and when does it want to reserve the right to decide for itself?

This work would likely be joint with Panos Ipeirotis and other members of the oDesk Research team.

Non-Core Projects

The Effects of Temperature and Environmental Pollution on White Collar Productivity and Labor Supply

In order to guarantee payment, workers on oDesk install special time- and activity-tracking software on their computers. When they are working, this software periodically (but randomly) reports a worker's intensity, as measured by keyboard presses and mouse-movements. These measurements are made down to the millisecond and are highly precise. We propose to use this data to measure how temperature and environmental pollution affect white-collar productivity. In particular, we will examine whether:

1. Do extreme temperatures affect productivity?

2. To what extent can workers foresee and adjust to these environmental changes, such as by working at other times of the day or purchasing an air conditioner?

3. Do workers increase their labor supply when the intensity of their work goes down, in order to keep output constant, or do they decrease labor supply in response to their lower productivity (and hence wages in a competetive market)?

This would is joint with Tal Gross, Matt Neidell and Josh Graff Zivin.

Online Work as Virtual Migration

There is substantial evidence that liberalization in migration policies could have enormous welfare consequences [@clemens2011economics]. However, there is no great well-spring of political support for easing migration restrictions—where enthusiasm exists at all, it is usually only with regards to high-skilled workers (c.f., Silicon Valley agitating for more H1-B visas). I believe that online work could serve as a "hack" around these restrictions, allowing the world to feasibly

obtain much of the benefit of liberalization without the required sea change in public policy or opinion. Online work is more like virtual migration and less like traditional off-shoring. Workers work directly for employers in rich countries, taking advantage of the better opportunities and more mature institutions offered in these richer countries. Online work requires very little overhead—essentially an Internet connection and a way to receive payments. There is a literature that suggests the share of services that can be "sent down a wire" is enormous—on the order of 25% of the total wage bill [@blinder2006offshoring].

I have partially explored this argument in [@horton2011condition], but a research project quantifying these effects could be very informative. The actual research project would seek to understand the welfare impacts of online work.

1. How does it affect income and consumption?
2. How does it affect human capital investment decisions?
3. What are the second-order consumption effects of online work? In particular, do workers spend less on:

- transportation,
- child care
- prepared foods
- work clothes?

1. Do they invest more in IT?

The ideal experiment would to be survey workers joining the marketplace and then randomize some to get jobs and others to not get jobs and then conduct a follow-up survey. Unfortunately, creating a large number of online jobs might be both unrealistic to the actual trajectory of workers and be prohibitively expensive.
Pallais shows that that even a small initial job can help many workers get started [@pallais2012ineffiient], so perhaps the cost would not be too great. There are natural experiment approaches—namely idiosyncratic factors that influenced whether or not a worker gets hired—though concerns about statistical power might make these difficult to use.

This project would be joint with Dina Pomeranz, who has extensive experience in doing development-oriented RCT.

References