Research Statement of John J. Horton

Advances in information and communications technology are profoundly changing labor markets. One remarkable result is the emergence of new, fully global, online labor markets. In these markets, buyers hire workers to perform work amenable to remote completion, such as computer programming, data entry and clerical work, graphic design, writing, translating and so on. Unlike previous examples of outsourcing, online labor markets facilitate true employment relationships between market participants. In addition, for the first time in history we have labor markets that are not meaningfully segmented by geography.

Online labor markets are worthy of serious study in their own right, as they provide a proving ground for ideas in market design. They also provide powerful new tools for gaining basic insights in the social sciences. In particular these markets offer two advantages for research: they are ideal for running experiments, and they generate unprecedented quantities of high-quality, contextualized data about economic relationships and transactions.

My research is rooted in both of these aspects: some of my research is in the "economist-as-engineer" vein, focused on identifying and designing solutions to market problems, while the rest of my research is basic social science. This statement describes my completed research projects, my works in progress and the themes that connect them.

Prior Research

Some of my earliest research on online labor markets describes the markets and their key economic properties (J. J. Horton 2010a). In joint work with Richard Zeckhauser and Dave Rand, we explain how researchers can conduct internally valid experiments in an online labor market. Although there are several "how-to" papers in this vein, this paper is important because it discusses precisely how to do valid causal inference—in the Rubin Causal Model sense—in online labor markets. The paper also replicates several well-known experimental results.

Beyond simply describing online labor markets and how to use them for experiments, I have also used them for basic research. In joint work with Lydia Chilton, we provide a model of workers supplying labor to paid crowdsourcing projects (Horton and Chilton 2010). We also introduce a novel method for estimating a worker's reservation wage. Experiments confirmed some of the key predictions of the model, though at least some subjects appear to be "target earners," contrary to the assumptions of the rational model. Another paper that examines why people provide labor to crowdsourcing projects is my joint work with Aaron Shaw and Daniel Chen, where we essentially conduct a "horse race" among various theories of motivation (Shaw, Horton, and Chen 2011).

In addition to motivation and incentives for participation, I am also interested in how workers decide how much effort to put into tasks. One notion is that workers can be strongly influenced by their peers, though showing this causally is often difficult. In online settings, however, introducing "narrow channel" peer effects is relatively easy. In a series of five field experiments, I explore the effects that peers can have in team production settings. Workers labeled photographs and evaluated their peers' performances at the same task (J. J. Horton 2010b). Evaluating high-output work made workers more productive, with stronger effects observed for higher-productivity workers. Even very explicit employer instructions were unable to stamp out these productivity peer effects. In their evaluations, workers punished workers who demonstrated low effort, but low output alone was insufficient to trigger punishment. Willingness to punish was strongly correlated with a worker's own productivity, yet this relationship was experimentally mutable, with productivity-reducing treatments also reducing punishment. In another paper (Chen and Horton 2009) we attempt to answer basic questions in economics using experiments that would be difficult to conduct in other contexts. We performed an experiment in which we offered workers lower wages and looked at how workers responded to the framing of wage cuts.

I consider several studies conducted on Amazon's Mechanical Turk to be pre-cursors to my work at oDesk on topics in applied market design. In joint work with Lydia Chilton (Chilton et al. 2010) we investigate how workers on Amazon's Mechanical Turk were actually matched to tasks. We conducted a high-frequency scrape of the website which allowed us to estimate the rate at which different tasks were completed in the marketplace. We found that the page position of a job listing had enormous effects on its attractiveness to workers. In particular, jobs that were at the top of the page when sorted from newest to oldest were far more likely to be actively worked on. We also found evidence that employers are aware of a worker's preference for newly posted jobs and try to "game" the system by continually re-posting their jobs to be at the top of search results. This paper motivated my later work with Dana Chandler on how interface choices can affect labor supply and the effectiveness of incentives (Chandler and Horton 2011).

I have long been interested in how different contractual forms affect production. In a paper I wrote early on in graduate school, I examine how variation in contractual form (between fixed-price and cost-plus contracts) affects the probability of litigation in government procurement contracts (Horton 2008). The paper has a revise and resubmit at the *Journal of Law and Economics*; I have not actively worked on it for some time, as I do not think that I can satisfactorily address the referees' concerns. However, I am planning to re-write the paper using data from oDesk, which has several sources of plausibly exogenous variation in the contractual structure, as well as a well-defined measure of ex post haggling/quasi-litigation.

Aside from their research potential, online labor markets are independently interesting as a powerful tool for economic development. To provide some background, there is substantial evidence that liberalization in migration policies could have enormous welfare consequences (Clemens 2011). Michael Clemens memorably subtitled his *Journal of Economics Perspectives* piece on the economics of migration, "Trillion-Dollar Bills on the Sidewalk?" But, there is no great wellspring of political support for easing migration restrictions—where enthusiasm exists at all, it is usually only with regard to high-skilled workers (e.g., Silicon Valley agitating for more H1-B visas). I believe that online work could serve as a "shortcut" around these restrictions, allowing the world to 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 offshoring or outsourcing. Workers work directly for employers in rich countries, taking advantage of the better opportunities and more mature institutions offered in richer countries. Online work requires very little overhead—essentially an Internet connection and a way to receive payments. A literature suggests that the share of services that can be "sent down a wire" is enormous—on the order of 25% of the total wage bill (Blinder 2006).

Athough I have not yet seriously explored the welfare implications of online work, I wrote a short piece in *Economic Letters* that discusses some these ideas. The article examines some of the basic welfare properties of online work and why they could be powerful tools for economic development (Horton 2011). Currently these markets are small, but that may change. Many strong trends—the proliferation of low-cost IT and high-bandwidth, increasing demand for services, an increasingly educated global population and so on—suggest that online work will grow in size and in importance.

Job Market Paper

My job market paper is a hybrid vehicle, encompassing both basic research into how firms fill vacancies and also applied market design examining the consequences of making recommendations. The paper adds to the labor literature a new model of employer search and screening, as well as causal evidence about how firms fill vacancies. The existing market design literature focuses on a strong "center" that fixes market congestion and market thinness by creating matches directly (Roth 2008). Platforms in decentralized markets like oDesk do not have—nor do they necessarily want—this match-setting power. This paper shows that even without match-setting power, markets can be tilted toward desirable ends

through purely informational interventions. Coles et al. (2010) is a recent example of work in this decentralized market design approach.

Works in Progress

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 little time. This can create a host of bad outcomes and inefficiencies, such as unraveling and exploding offers and poorly considered matches. Unraveling is often a justification for imposing a strong "center" that implements a static matching mechanism. However, for many—possibly most—markets, a centralized approach is infeasible.

One hypothesis is 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 as follows:

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 might 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.

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 (determined by frequency of query) from Google. The problem of cheap applications arises in law review journal submissions, college admissions, online dating sites and academic job markets (the economics job market 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. In a recent "experiment", a fake job posting on Craigslist received over 600 applications in less than 24 hours.
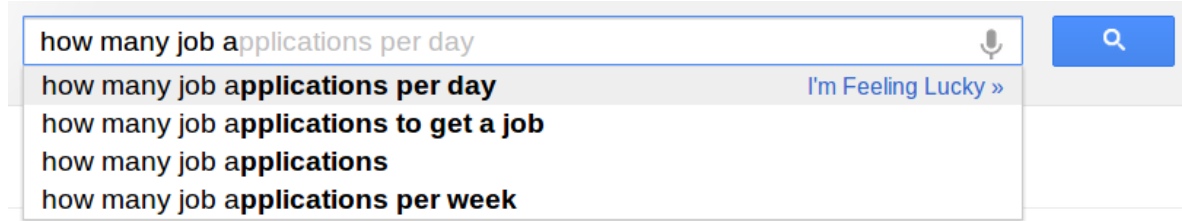

Figure 1: How many applications should I send?

This project is joint with Dana Chandler and Ramesh Johari.

Algorithmic Enhancement of Markets

A key theme in my job market paper is the idea that matching efficiency in labor markets can be improved with algorithmic interventions. The paper finds that these recommendations are very successful, despite being relatively algorithmically unsophisticated. At oDesk, I have been involved with several projects to advance these ideas and reduce frictions more generally. Some 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 clusters determined by a novel application-to-vacancy graph measure. Each of these features is being introduced experimentally and will allow us to understand how changes in match-making technology affect choices. In particular, we examine how reduced search frictions affect search intensity and specificity, reservation wages and bargaining. To do so, 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 (Becker 1993). The endogenous choice is "years of education." This high level of abstraction hides the reality that most workers must continually make choices about which skills to learn over the course of their careers.

This choice is complex and multi-faceted. A worker considering acquiring 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 and ability?

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 solve a complex, multi-dimensional optimization problem replete with poorly-understood pay-offs and costs. 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 which contractors can take (and which employers may 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-based recommendations about skills that workers are interested in learning?

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 and

2. How different skills are "related" to each other, in terms of market returns and ease of acquiring skills.

Aside from the importance of skills 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 tools for the scalable delivery of education, such as Khan Academy, Coursera and MITx. If this technology unleashes a new era of algorithmically-assisted auto-didactism, people will still need to learn *what* to learn. 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. 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 transition 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 building and managing online experiments using oDesk workers.

2. Create a tool for conducting 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 a direct consequence of this power (Edelman, Ostrovsky, and Schwarz 2005)—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 a particular time. Ideally, workers would self-report availability, but because job offers are useful even if they are not accepted (e.g., for bargaining power) and workers have free disposal on offers, getting them to honestly report availability is challenging. We 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.

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 work 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:

1. Do extreme temperatures affect productivity?

2. To what extent can workers foresee and adjust to environmental changes, such as by working at other times of day or using air conditioning?

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 competitive market)?

This work is joint with Tal Gross, Matt Neidell and Joshua Graff Zivin.

Online Work as Virtual Migration

I have partially explored the idea of online work as virtual migration in (Horton 2011), but a research project quantifying the welfare effects of this kind of intervention could be very informative. In particular, we would examine:

1. How does virtual migration affect income and consumption?

2. How does it affect human capital investment decisions?

3. Do workers invest more in IT?

4. What are the second-order consumption effects of online work? In particular, do workers spend less on transportation, child care, prepared foods and work clothes?

The ideal experiment would to be survey workers joining the marketplace, randomize some to get jobs and others to not get jobs and conduct a follow-up survey. It is possible that creating a large number of online jobs would be unrealistic to the actual trajectory of workers and prohibitively expensive. However, Pallais shows that that even a small initial job can help many workers get started (Pallais 2012), suggesting that the cost may not be manageable. There are also natural experiment approaches—namely idiosyncratic factors that influence 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 randomized controlled trials.

References

Becker, G. S. 1993. "Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education: University Of Chicago Press." *Chicago, IL*.

Blinder, A. S. 2006. "Offshoring: The next industrial revolution." *Foreign Affairs* 85: 113.

Chandler, D., and J. Horton. 2011. "Labor Allocation in Paid Crowdsourcing: Experimental Evidence on Positioning, Nudges and Prices." In *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*.

Chen, D., and J. Horton. 2009. "The wages of pay cuts: Evidence from a field experiment." *Available at SSRN 1443526*.

Chilton, L. B., J. J. Horton, R. C. Miller, and S. Azenkot. 2010. "Task search in a human computation market." In *Proceedings of the ACM SIGKDD workshop on human computation*, 1–9.

Clemens, M. A. 2011. "Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?." *The Journal of Economic Perspectives* 25: 83–106.

Coles, Peter, J. Cawley, P. B. Levine, M. Niederle, Alvin E. Roth, and J. J. Siegfried. 2010. "The job market for new economists: A market design perspective." *The Journal of Economic Perspectives* 24: 187–206.

Edelman, B., M. Ostrovsky, and M. Schwarz. 2005. "Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords."

Horton, J. 2008. "Procurement, Incentives and Bargaining Friction: Evidence from Government Contracts." *Available at SSRN 1094622.*

Horton, J. J. 2010a. "Online labor markets." *Internet and Network Economics*: 515–522.

–——. 2010b. "Employer expectations, peer effects and productivity: Evidence from a series of field experiments." *SSRN eLibrary.*

–——. 2011. "The condition of the Turking class: Are online employers fair and honest?." *Economics Letters* 111: 10–12.

Horton, J. J., and L. B. Chilton. 2010. "The labor economics of paid crowdsourcing." In *Proceedings of the 11th ACM conference on Electronic commerce*, 209–218.

Pallais, A. 2012. "Ineffiient Hiring in Entry-Level Labor Markets." *Available at SSRN 2012131.*

Roth, Alvin E. 2008. "What Have We Learned from Market Design?." *The Economic Journal* 118: 285–310. doi:10.1111/j.1468-0297.2007.02121.x. http://dx.doi.org/10.1111/j.1468-0297.2007.02121.x.

Shaw, A. D., J. J. Horton, and D. L. Chen. 2011. "Designing incentives for inexpert human raters." In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 275–284.