Determinants of Matching in Online Labor Markets: A Structural Two-Sided Matching Model

Jing Gong

Fox School of Business, Temple University

gong@temple.edu

Abstract

In the past decade, IT has facilitated the shift from permanent employment to need-based outsourcing and from local labor market to global online labor markets. While prior studies have examined how global frictions affect employers' hiring decisions on online labor markets, we have limited understanding of the inter-dependence between workers and employers and the economic impact of IT-enabled globalization on matching outcomes such as the number of matched projects, freelancer wages, and project values generated from matching. This study is an attempt to fill in the gap by examining the dual roles of IT-enabled globalization, i.e., (1) in determining the formation of matches between employers and freelancers, and (2) in affecting market outcomes. From a market perspective, we take into account two-sided decision making, competition on each side, complementarities between employer and freelancer attributes, and endogenous money transfers between employers and freelancers.

In our empirical analysis, we estimate a structural two-sided matching model of the online labor market from a revealed preference perspective. The estimation is based on a dataset from a major freelancing website that connects freelancers and employers from

more than 200 countries. We then conduct counterfactual analysis to quantify the economic impact of IT-enabled globalization in online labor market by comparing the current scenario with a counterfactual scenario where employers can only match with freelancers from the same country. The results from our estimation suggest that employers tend to match with freelancers from the same country, and that employers from developed countries tend to match with freelancers from developing countries. The results from the counterfactual analysis suggest that IT-enabled globalization leads to more employers and freelancers with successful matches, lower average wage among matched freelancers, and higher total project values generated on the market.

Keywords: *Online labor markets, globalization, online freelancing, two-sided matching*

1 Introduction

Online labor markets are intermediaries that match workers with employers in exchange for services or labor (Hong et al. 2016). In the past decade, IT has facilitated the shift from permanent employment to need-based outsourcing and from local labor market to global online labor markets. While traditional labor markets have been segmented by geography and national borders, online labor markets match workers and employers across borders, and create a flat world (Gefen and Carmel 2008). Many of these online marketplaces allow employers to post projects, and freelancers to propose bids and compete for projects. Employers choose among freelancers that vary in both prices and non-price attributes such as reputation and geographic location.

In a typical online labor marketplace, both employers and workers make decisions based on not only rival choices, but also the other side. The interaction between workers and employers determines whether an employer and a freelancer match with each other. However, prior studies have mainly focused on analyzing employers' hiring decisions while taking the workers' decisions as given (e.g., Banker and Iny 2008; Gefen and Carmel 2008; Mill 2011; Agrawal et al. 2014; Hong and Pavlou 2014; Moreno and Terwiesch 2014; Pallais 2014) . Therefore, we have limited understanding of how workers participate on online labor markets, and how workers and employers make inter-dependent decisions. A better understanding of the inter-dependence between the two sides may provides implications for policy makers to evaluate the impact of IT-enabled labor markets, and for platform owners to design mechanisms to improve match efficiency and quality.

This study is an attempt to fill in the gap by examining the determinants of matching between employers and freelancers. We are particularly interested in understanding how geographic and economic differences affect the matching between freelancers and employers. While several studies have examined how geographic attributes affect employers' hiring decisions and freelancers' bidding decisions and found that global frictions exist in online labor markets (e.g., Yoganarasimhan 2013; Agrawal et al. 2014; Hong and Pavlou 2014), most of the studies use single-agent choice models and only consider the decisions of one side, taking the other side as given. To the best of our knowledge, our study is the first to examine the impacts of both geographic and eco-

nomic differences in online labor markets using a two-sided matching framework. Furthermore, we examine how IT-enabled globalization, specifically the availability of foreign workers and job opportunities, affects aggregate market outcomes such as the number of matched projects, freelancer wages, and project values generated from online labor market. Because our empirical analysis is from a market perspective and takes into account two-sided decision making we are able to utilize our empirical strategy to answer questions related to aggregate matching outcomes that would not be answered with single agent choice models.

In this study, we build on an empirical two-sided matching framework (Fox 2010) by jointly modeling the decisions of both freelancers and employers from a revealed preference perspective. The online labor market is a type of two-sided matching market, where each side (freelancers or employers) takes into account the decisions of the other side, competition on each side, complementarities between employer and freelancer attributes, and endogenous money transfers between employers and freelancers. In our model, we assume that the matches between freelancers and employers represent an equilibrium outcome called *pairwise stability*. Following Fox (2010), we then use inequalities derived from necessary conditions for pairwise stability to form maximum score estimators (Manski 1975, 1985), which are semiparametric and computationally simple.

We utilize a dataset from *Freelancer*, one of the largest freelancing platforms that connects freelancers and employers from over 200 countries. The dataset includes detailed history of projects posted in March 2014. The data demonstrate substantial evidence of globalization. the majority of the matched projects consist of employers from high income regions and freelancers from low income regions. If employers can only match with freelancers within their own country, a large percentage of employers may not be able to find a freelancer to match.

By estimating a two-sided matching model, we recover structural parameters to understand how partner attributes interact with each other to create increasing joint value (Mindruta et al. 2016). Specifically, we find that employers tend to match with freelancers from the same country, suggesting that same country matches tend to generate higher match values. Our results also suggest that employers from developed countries tend to match with freelancers from developing

countries, suggesting that the economic growths of employer country and freelancer country are substitutes. Furthermore, we find positive assortative matching pattern between freelancer reputation and employer reputation. The estimates we obtain from the two-sided matching model allow us to conduct counterfactual analysis to quantify the economic impact of IT-enabled globalization on matching outcomes such as match efficiency and contract price. Specifically, we find that IT-enabled globalization results in more successful matches, lower average wage among matched freelancers, and higher total project values generated on the market.

This paper makes several key contributions. First, it contributes to the literature on online labor markets by understanding how geographic and economic differences affect the matching between employers and workers, and more importantly, by quantifying the impact of IT-enabled globalization on aggregate market outcomes. Second, the findings offer both managerial and policy implications. The insights from this study can help employers and freelancers understand the pros and cons of online freelancing platforms, and offer managerial implications for both platform owners and employers to develop strategies to improve matching efficiency and quality and for policy makers to better respond to changes as a result of IT-enabled labor globalization. Third, this study is one of the first studies that utilize an empirical two-sided matching model to to the context of online labor markets. To the extent that researchers in the field of information systems focus on a variety of technology-mediated two-sided markets, matching models can be applied to various contexts where two sets of participants engage in bilateral exchange in these markets.

2 Literature Review

This paper is mainly related to three streams of literature. The first stream of literature analyzes online labor markets (Snir and Hitt 2003; Carr 2003; Agrawal et al. 2013). Early research predicts that internet reduces search cost and transaction cost and increases both labor demand and supply, thus may increase match quality and efficiency (Autor 2001). However, online markets

may suffer from the information asymmetry problem (Akerlof 1970), and many studies focus on analyzing how various types of information or quality signals affect employer and worker behavior and reduce informational frictions. For example, prior studies have examined the impact of reputation systems (Banker and Iny 2008; Gefen and Carmel 2008; Mill 2011; Yoganarasimhan 2013; Moreno and Terwiesch 2014; Pallais 2014), different auction designs (Hong et al. 2016), different contract regimes (Lin et al. 2010), third party certifications (Goes and Lin 2012), and agency affiliations (Stanton and Thomas 2012).

Despite early prediction that the internet creates a level play field where employers can connect freelancers from anywhere in the world (Autor 2001), a few recent studies find that global frictions still exist in online labor markets (Yoganarasimhan 2013; Agrawal et al. 2014; Hong and Pavlou 2014). For example, Gefen and Carmel (2008) find that on an online programming market employers prefer domestic workers, though ceteris paribus, they prefer workers from developing countries. In contrast, Yoganarasimhan (2013) finds that employers prefer workers from developed countries, and employers from developing countries prefer domestic workers than foreign workers. Agrawal et al. (2014) find that freelancers from developing countries have lower winning probability compared to freelancers from developed countries. Hong and Pavlou (2014) examine how differences in language, culture and time zone affect freelancers' pricing strategy and employers' hiring choice. In sum, while prior studies have consistently find that employers tend to prefer domestic workers to foreign workers, there is no consensus on whether employers prefer workers from developing regions or developed regions.

However, the prior empirical literature on online labor markets is mainly descriptive. Most of the prior empirical studies analyze employer decisions while taking freelancers' decisions as given, and hypotheses are tested using single-agent choice models (such as logit models), thus making an implicit assumption that an agent in the market makes choices independent of other agents. Therefore, the research gap relates to how to model an online labor market as a two-sided matching market. Our study attempts to fill in the gap by using a two-sided matching model that take into accounts the decisions of both freelancers and employers jointly, and quantify the

globalization effect of online labor markets on key players.

The second stream of literature relates to international trade, immigration and labor market. There has been significant debate on the impact of international trade and immigration on local labor markets. On one hand, offshoring or hiring immigrants cause a direct substitution effect and may reduce the demand for domestic workers (Harrison and McMillan 2011; Autor et al. 2013; Ebenstein et al. 2014). On the other hand, the cost savings associated with offshoring and hiring immigrants may lead to increased efficiency of production, and thus increasing the demand for domestic workers (Peri 2012; Ottaviano et al. 2013; Kovak et al. 2015).

The third stream of literature relates to two-side matching models (Gale and Shapley 1962; Roth and Sotomayor 1990). Our work is closely related to recent developments in empirical two-sided matching models (e.g., Choo and Siow 2006; Sorensen 2007; Fox 2010; Agarwal 2015; Yang and Goldfarb 2015) that have applied to the marriage market (Choo and Siow 2006), medical residency match (Agarwal 2015), merger-acquirer match (Akkus et al. 2014), executive-firm match (Pan 2015), automobile supplier-assembler match (Fox 2010), and inter-firm alliance (Yang et al. 2009; Mindruta 2013; Mindruta et al. 2014). However, two-sided matching models have received limited attention from researchers in information systems. As researchers in information systems focus on a variety of technology-mediated two-sided markets, matching models can potentially be applied to various contexts such as online dating, buyer-supplier relationships, advertiser-publisher match, and inter-firm alliance.

3 Data and Setting

3.1 Research Context

In this study, we use a dataset obtained from *Freelancer*. *Freelancer* is one of the world's largest outsourcing and freelancing platforms established in 2009. By March 2015, *Freelancer* has over 14 million registered users from over 240 countries, with over 7 million projects posted. *Freelancer* uses a reverse auction mechanism for service procurement that allows employers to post projects

with additional information about required skills and budget, typically for a week. Freelancers then decide whether and how much to bid on the projects, and the original employers then choose a winner among the bidders to complete the work. Upon completion, the employer pays the freelancer the pre-determined price.

3.2 Data

Our dataset includes all projects in the *Data Processing* category posted on *Freelancer* between March 3, 2014 and March 30, 2014. In this paper, we restrict our analysis to projects with fixed rates and remove projects with hourly rates (13% of the projects), and focus our analysis on projects that received at least one bid and awarded a winner. Also, we also remove projects with maximum budget higher than $1,000 and projects that awarded an outlier bid as the winner.[1] In our final sample, we are left with 270 projects matched with 248 freelancers. The descriptive statistics of the sample dataset are summarized in Table 1.

[Table 1 about here.]

For the 270 projects with matched freelancers, we observe attributes at employer level, freelancer level, and employer-freelancer pair level. The description of key variables is presented in Table 2.

[Table 2 about here.]

[Table 3 about here.]

Table 3 presents the summary statistics of key variables. Among the 270 projects, 79% are posted by employers from developed countries. In comparison, only 17% of the 248 freelancers are from developed countries. In addition, only 7% of the employers are matched with freelancers from the same country. If employers are restricted to search within their own country, a large

---

[1]We define outlier bids based on the interquartile range of bids for each project. For each project, we compute the lower quartile (Q1) and upper quartile (Q3) of the bidding prices, we define outliers as bids with price outside the range [Q1-3(Q3-Q1), Q3+3(Q3-Q1)]. We then remove projects that awarded an outlier bid as the winner.

percentage of them may not be able to find a match. Therefore, the data demonstrate substantial evidence of globalization on *Freelancer.*

4   A Two-Sided Matching Model

In this section, we model the online labor market as a two-sided market (Gale and Shapley 1962; Roth and Sotomayor 1990) where the market outcome is the joint decision of both freelancers and employers, and the interaction determines whether an employer and a freelancer match with each other and the money transfer (i.e. contract price) between them. We follow the empirical framework of Fox (2010) and estimate a structural two-sided model of the online labor market. In our model, we assume that the matches between freelancers and employers represent an equilibrium outcome called *pairwise stability*, where matched pairs have no incentive to deviate from current partners.

4.1   Model Setup

To formally model the joint decision of both sides, we denote a freelancer by i, an employer by j, and a market by t. [2] Both freelancers and employers are utility maximizing agents. For simplicity, we use employer and project interchangeably.

Let the pre-transfer utility of the ith freelancer from working for the jth employer be u_{ij}, the pre-transfer utility of jth employer from hiring the ith freelancer be v_{ij}, and the money transfer between freelancer i and employer j be w_{ij}. Then the net payoff freelancer i receives by working for employer j is $U_{ij} = u_{ij} + w_{ij}$, and the net payoff employer j receives is $V_{ij} = v_{ij} - w_{ij}$. We define the match value function of freelancer i and employer j as the total payoff received by the matched pair, as follows:

$$f(i,j) = u_{ij} + v_{ij}. \tag{1}$$

---

[2]In our context, we treat each week as a different market, because projects are typically posted for a week.

An outcome of the market t consists of a matching μ_t, which is a set of matches, and w_t, which is a vector of money transfers between matched pairs. $(i, j) \in \mu_t$ if and only if freelancer i and employer j are matched.The equilibrium concept used here is called *pairwise stability*. The intuition behind pairwise stability is that any pair of employer and freelancer who are currently not matched each other should have no incentive to deviate from their current partner to match with each other. Formally, in a pairwise stable equilibrium, the following holds for any two observed matches (i, j), $(i', j') \in \mu_t$,

$$u_{ij} + w_{ij} \geq u_{ij'} + \widetilde{w}_{ij'}, \tag{2}$$

where $\widetilde{w}_{ij'} = v_{ij'} - V_{i'j'}$ for all $i' \neq i$. That is, suppose freelancer i and employer j are matched, and freelancer i' and employer j' are matched. Inequality (2) considers a deviation of freelancer i and employer j' from their current partners to form a new match with each other. $\widetilde{w}_{ij'}$ can be interpreted as the maximum transfer that employer j' is willing to pay to freelancer i by switching from freelancer i' to freelancer i. Therefore, the left hand side of (2) is simply the net payoff freelancer i gets from the current matched employer j, and the right hand side is net payoff freelancer i gets from switching to employer j'. Inequality (2) means that even if employer j' is willing to pay $\widetilde{w}_{ij'}$ to freelancer i, freelancer i would not want to deviate from employer j to employer j'. In other words, for the proposed deviation to take place, both freelancer i and employer j' need to be better off by matching with each other than their current partner. Substituting $\widetilde{w}_{ij'} = v_{ij'} - V_{i'j'}$ and $V_{i'j'} = v_{i'j'} - w_{i'j'}$ into the inequality in (2), we get

$$u_{ij} + w_{ij} \geq u_{ij'} + v_{ij'} - (v_{i'j'} - w_{i'j'}). \tag{3}$$

Following the same logic, we obtain the following inequality by considering a deviation of freelanceri' and employer j from their current partners to form a new match with each other:

$$u_{i'j'} + w_{i'j'} \geq u_{i'j} + v_{i'j} - (v_{ij} - w_{ij}). \tag{4}$$

Fox (2010) derives the following inequality by combining (3) and (4), thus canceling the transfers w_{ij} and $w_{i'j'}$.

$$(u_{ij} + v_{ij}) + (u_{i'j'} + v_{i'j'}) \geq (u_{ij'} + v_{ij'}) + (u_{i'j} + v_{i'j}). \tag{5}$$

By replacing $u_{ij} + v_{ij}$ with $f(i,j)$, we get:

$$f(i,j) + f(i',j') - f(i,j') - f(i',j) \geq 0. \tag{6}$$

The inequality in (6), referred to as the local production maximization inequality (Fox 2010), implies that the total match value of the two observed matched pairs is greater than the total match value if the two matched pairs switch partners.

4.2   Smoothed Maximum Score Estimator

Fox (2010) develops a maximum score estimator (Manski 1985, 1975) based on the inequality in (6). The estimator does not require transfer data, and is proved to be consistent for two-sided matching games under certain conditions. Let ε_{ij} be an unobserved pair-specific error term that affects $f(i,j)$, the match value of freelancer i and employer j. The match value function can be written as $f(i,j) = f_0(i,j) + \varepsilon_{ij}$. Given a functional form for $f_0(i,j|\beta)$, where β is a vector of parameters to be estimated, we estimate β by maximizing the following objective function:

$$S(\beta;\sigma_N) = \frac{1}{N} \sum_{t=1}^{T} \sum_{(i,j)\in\mu_t} \sum_{(i',j')\in\mu_t} 1[i \neq i'] \cdot 1[j \neq j'] \cdot \Phi\left[\frac{f_0(i,j|\beta) + f_0(i',j'|\beta) - f_0(i,j'|\beta) - f_0(i',j|\beta)}{\sigma_N}\right],$$
$$\tag{7}$$

where N is the total number of inequalities included in $S(\beta;\sigma_N)$, and T is the total number of markets The smoothed maximum score estimator chooses parameters that maximize $S(\beta;\sigma_N)$ (Horowitz 1992), which uses $\Phi(\boldsymbol{.})$, the cumulative density function of standard Normal distribution, as the smoothing function. σ_N is the bandwidth of the smoothing function.

The maximum score estimator that maximizes (7) has several advantages. First, the estimator is semiparametric and does not make assumptions about the distribution of error terms. Second, the estimator does not suffer from the "curse of dimensionality", which is often a challenge when decisions are inter-dependent. When estimating from a market perspective, one has to consider the interplay between different players in the game, and the estimation often requires computing all possible matches. The maximum score estimator addresses the computational challenge of modeling two-sided matching games by relying only on inequalities based on (6). Furthermore, because we do not explicitly solve the equilibrium outcome, the maximum score estimator allows for the existence of multiple equilibria.

We use the plug-in method described in Horowitz (1992) to find the optimal bandwidth σ_N. Horowitz (1992) discusses the consistency and asymptotic properties of the smoothed estimator.

4.3 Determinants of Match Value Function

Following existing work on empirical matching models (e.g., Fox 2010; Fox and Bajari 2013; Yang et al. 2009; Yang and Goldfarb 2015), we include pair-specific attributes and interactions between freelancer and employer attributes in the match value function $f(i, j)$. We assume that

$$f(i, j) = f_0(i, j) + \varepsilon_{ij} = \beta_1' X_i Y_j + \beta_2' Z_{ij} + \varepsilon_{ij}, \tag{8}$$

where $X_i Y_j$ is a vector of interaction terms between freelancer and employer attributes, Z_{ij} is a vector of characteristics specific to an employer-freelancer pair (such as non-transfer bidding attributes), and ε_{ij} is an unobserved pair-specific error term assumed to be independent across matches. In our main model, $X_i Y_j$ includes $Budget \times FreelancerTenure$, $EmployerGDP\_PPP \times FreelancerGDP\_PPP$, $EmployerAvgRating \times FreelancerAvgRating$. Z_{ij} includes $SameCountry$ and $PriorInteraction$. β_1 and β_2 are vectors of parameters to be estimated. One may notice that we do not include freelancer-specific or employer-specific attributes in the match value function $f(i, j)$ (Equation 8). The reason is that, if $f(i, j)$ includes any freelancer-specific or employer-

specific attributes, such attributes will appear in both sides of Inequality (6), and thus will be canceled out. Therefore, any coefficient that is freelancer-specific or employer-specific cannot be identified with the smoothed maximum score estimator based on (7). Similarly, any freelancer-specific, employer-specific, or market-specific fixed effects cannot be identified as well, because they will also be canceled out in Inequality (6). Because our main interest is to examine how interactions between freelancer and employer such as geographic differences impact the matching outcomes, the model we use here is appropriate for our purpose.

For identification, we do not include an intercept term in $f(i,j)$ because adding any constant term to $f(i,j)$ would result in the same value for the objective function $S(\beta)$. In addition following Fox(2010), we normalize the coefficient of a continuous interaction variable, $Budget \times FreelancerTenure$ to be ± 1. We estimate other parameters with the coefficient of $Budget \times FreelancerTenure$ normalized to 1 and -1 separately. We then pick the set of parameter estimates with the higher number of inequalities satisfied.

5   Estimation Results

In the estimation, we group all projects and freelancers who find a match in the same week into one market, as most of the projects were posted on the website for one week.[3] There are 9,016 inequalities used for estimation. We obtain the estimates by first applying the generalized simulated annealing algorithm to maximize the objective function (7), followed by a Newton-Raphson algorithm to obtain convergence. The confidence intervals are obtained using the bootstrap procedure described in Horowitz (2002).

Table 4 presents the estimates of the match value function $f(i,j)$ based on the smoothed maximum score estimator discussed in Subsection 4.2. The 95% bootstrap confidence intervals are reported in parentheses. Model 1 is a baseline model without variables that measure the economic growth of a country. Model 2 uses $EmployerDeveloped$ to indicate whether the employer comes

---

[3]The definition of a market ensures that only observed matches and counterfactual matches within the same market are compared.

from a developed country, and *FreelancerDeveloped* to indicate whether the freelancer comes from a developed country. Model 3 uses GDP PPP (per capita) to measure the economic growth of a country.

The maximum score values reported in Table 4 represents the percentage of inequalities satisfied based on our estimates, which measure the fitness of the proposed models. Sixty-nine percent of the inequalities are satisfied, indicating reasonable predictive performance.

As can be shown in Table 4, the coefficient of the interaction between employer rating and freelancer rating is positive and statistically significant across three different models, suggesting that there is a positive assortative matching pattern between employers and freelancers. That is, employers with higher ratings tend to match with freelancers with higher ratings.

The positive coefficient for *SameCountry* indicates that employers tend to match with freelancers from the same country. This result suggests that being in the same country leads to higher match value than coming from different countries, and that there are geographic complementarities between employer and freelancer. In addition, we find that match value is higher when the employer hires a freelancer with prior interactions.

As can be shown in Column 2 of Table 4, the coefficient for the interaction between *EmployerDeveloped* and *FreelancerDeveloped* is negative, which means that employers from developed countries tend to match with freelancers from developing countries, suggesting substitution between the economic statuses of freelancer and employer's origin countries.The results in Column 3 are qualitatively similar. In particular, the coefficient for the interaction between employer and freelancer's GDP PPP (per capita) is negative and statistically significant, suggesting that the economic growths of employer country and freelancer country are substitutes. The maximum score of Model 3 (0.6881) is slightly higher than Model 1 (0.6812) and Model 2 (0.6863), which shows that estimates based on Model 3 provides better predictive performance.

[Table 4 about here.]

14

6 Counterfactual Analysis

What would the market look like if freelancers can only match with employers from the same country? To quantify the impact of IT-enabled globalization on the matching outcomes of the market, we consider a counterfactual scenario where matches are restricted to be within country. Under the counterfactual scenario, employers can only match with domestic freelancers. Based on our estimates as listed in Column 3 of Table 4, we conduct the counterfactual analysis by following a similar procedure to Yang and Goldfarb (2015).

Step 1. Compute match values for each possible pair of employer and freelancer with estimates from Column 3 of Table 4. Step 2. Assign optimal matches under the current scenario. An assignment problem is formulated to maximize the total match value from all matches given freelancer capacity c_{it}.[4] The optimal assignment is obtained through linear programming (Shapley and Shubik 1971) with constraints specified above.

Step 3. Assign optimal matches under the counterfactual scenario. An assignment problem similar to Step 3 is formulated, with an additional constraint that employers are only allowed to match with freelancers from the same country.

Table 5 compares the observed matches with optimal matches computed from Step 3 (i.e. under the current scenario), as well optimal matches from Step 4 (i.e. under the counterfactual scenario where only within-country matches are allowed). As can be seen from Column 1 and Column 2 of Table 5, the observed matches and optimal matches under the current scenario are similar, indicating that our model has reasonable model fit. However, the differences between the optimal matches under the current scenario and counterfactual scenario are substantial. In particular, if we restrict employers to only match with freelancers from the same country, we see a 75% decrease in the number of matches and a 57% decrease in total project value. Therefore, without IT-enabled globalization, the market results in fewer matches, which is both socially suboptimal and a loss for the platform owner. We also find that by restricting matches

---

[4] c_{it} denotes the capacity constraint of freelancer i in market t. Within each market, we compute c_{it} for each freelancer by counting the number of projects that freelancer i is currently matched with in market t, based on observed data.

to be within country, the mean contract price among matched freelancers increases by 56%, which may be explained by the decreased competition among freelancers. Therefore, IT-enabled globalization has greatly benefited employers with increased availability of foreign workers and lower wages.

[Table 5 about here.]

7 Conclusion

In this study, we apply a structural two-sided matching model to the context of online labor market and examine determinants of employer-freelancer matching. First, we find employers tend to match with freelancers from the same country. Second, ly, we find that employers from developed countries tend to match with freelancers from developing countries, suggesting that the economic growths of employer country and freelancer country are substitutes. Furthermore, we find positive assortative matching pattern between freelancer reputation and employer reputation (as measured by average rating received by freelancer and employer).

By studying the online labor market from a two-sided matching perspective, we are able to examine the inter-dependence between employers and freelancers by incorporating two-sided decision making, competition on each side, complementarities and substitutions between employer and freelancer attributes, and endogenous money transfers. These key features of our model differentiate our paper from prior literature on online labor market, which mainly focuses the decision of one side with the other side as given.

The estimates we obtain from the two-sided matching model allows us to conduct counterfactual analysis to quantify the economic impact of IT-enabled globalization on matching outcomes such as match efficiency and contract price. Our counterfactual analysis demonstrates the importance of IT-enabled globalization in the context of labor market and offers implication for employers, freelancers, platform owners as well policy makers. The insights from this study can help employers and freelancers understand the pros and cons of online freelancing platform and

offer managerial implications for both platform owners and policy makers to develop strategies to improve matching efficiency and quality.

Our study also highlights the usefulness of two-sided matching models in understanding other contexts that involve the joint decisions of two or more groups of participants. With the emergence of new business models, we have witnessed a variety of technology-mediated two-sided markets. This paper opens up several potential new directions of empirical research on two-sided markets in the information systems area.

Our study suffers from some limitations. First, we model the online market as a static game, and do not consider entry or exit decisions. Our counterfactual analysis may underestimate the impact of globalization if availability of potential partners affects the overall attractiveness of such markets and thus leads to fewer participants on the market. Second, we only use a subset of projects with successful matches, and do not use data on unmatched projects and freelancers. Future research may investigate the reverse auction mechanism and study inter-temporal decisions made by freelancers and employers. Third, our counterfactual analysis assumes that, when only same-country matches are allowed, there are no additional participants entering the market.

References

N. Agarwal. An Empirical Model of the Medical Match. 2015.

A. Agrawal, J. Horton, N. Lacetera, and E. Lyons. Digitization and the Contract Labor Market: A Research Agenda. 2013.

A. K. Agrawal, N. Lacetera, and E. Lyons. Does Information Help or Hinder Job Applicants from Less Developed Countries in Online Markets? 2014.

G. A. Akerlof. The market for" lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

O. Akkus, J. A. Cookson, and A. Hortac. The Determinants of Bank Mergers: A Revealed Preference Analysis. 2014.

D. H. Autor. Wiring the Labor Market. *Journal of Economic Perspectives*, 15(1):25–40, 2001.

D. H. Autor, D. Dorn, and G. H. Hansan. The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review*, 103(6):2121–2168, 2013.

R. D. Banker and H. Iny. Importance of Measures of Past Performance: Empirical Evidence on Quality of e-Service Providers. *Contemporary Accounting Research*, 25(2):307–337, 2008.

S. M. Carr. Note on Online Auction with Costly Bid Evaluations. *Management Science*, 49(11): 1521–1528, 2003.

E. Choo and A. Siow. Who Marries Whom and Why. *Journal of Political Economy*, 114(1): 175–201, 2006.

A. Ebenstein, A. Harrison, M. McMillan, and S. Phillips. Estimating the impact of trade and offshoring on american workers using the current population surveys. *Review of Economics and Statistics*, 96(4):581–595, 2014.

J. T. Fox. Estimating Matching Games with Transfers. 2010.

J. T. Fox and P. Bajari. Measuring the efficiency of an FCC spectrum auction. *American Economic Journal: Microeconomics*, 5(1):100–146, 2013.

D. Gale and L. S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69(1):9–15, 1962.

D. Gefen and E. Carmel. The World Really Flat? A Look at Offshoring at an Online Programming Marketplace. *MIS Quarterly*, 32(2):367–384, 2008.

P. Goes and M. Lin. Does Information Really "Unravel"? Understanding Factors That Motivate Sellers to Seek Third-Party Certifications in an Online Labor Market. 2012.

A. Harrison and M. McMillan. Offshoring jobs? multinationals and us manufacturing employment. *Review of Economics and Statistics*, 93(3):857–875, 2011.

K. Y. Hong, A. C. Wang, and P. A. Pavlou. Comparing Open and Sealed Bid Auctions: Evidence from Online Labor Markets. *Information Systems Research*, 2016.

Y. Hong and P. A. Pavlou. Is the World Truly "Flat"? Empirical Evidence from Online Labor Markets. 2014.

J. L. Horowitz. A smoothed maximum score estimator for the binary response model. *Econometrica*, 60(3):505–31, 1992.

J. L. Horowitz. Bootstrap critical values for tests based on the smoothed maximum score estimator. *Journal of Econometrics*, 111(2):141–167, 2002.

B. Kovak, L. Oldenski, and N. Sly. The labor market effects of offshoring by us multinational firms: Evidence from changes in global tax policies. Technical report, Working paper, 2015.

M. Lin, S. Viswanathan, and R. Agarwal. An Empirical Study of Online Software Outsourcing: Signals under Different Contract Regimes. 2010.

C. F. Manski. Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3):205–228, 1975.

C. F. Manski. Semiparametric analysis of discrete response:: Asymptotic properties of the maximum score estimator. *Journal of Econometrics*, 27(3):313–333, 1985.

R. Mill. Hiring and Learning in Online Global Labor Markets. 2011.

D. Mindruta. Value creation in university-firm research collaborations: A matching approach. *Strategic Management Journal*, 34(6):644–665, 2013.

D. Mindruta, M. Moeen, and R. Agarwal. A Two-sided Matching Approach for Partner Selection and Assessing Complementarities in Inter-firm Alliances. 2014.

D. Mindruta, M. Moeen, and R. Agarwal. A two-sided matching approach for partner selection and assessing complementarities in partners' attributes in inter-firm alliances. *Strategic Management Journal*, 37(1):206–231, 2016.

A. Moreno and C. Terwiesch. Doing Business with Strangers: Reputation in Online Service Marketplaces. *Information Systems Research*, 25(4):865–886, 2014.

G. I. Ottaviano, G. Peri, and G. C. Wright. Immigration, offshoring, and american jobs. *American Economic Review*, 103(5):1925–59, 2013.

A. Pallais. Inefficient Hiring in Entry-Level Labor Markets. *American Economic Review*, 104 (11):3565–3599, 2014.

Y. Pan. The determinants and impact of executive-firm matches. 2015.

G. Peri. The effect of immigration on productivity: Evidence from us states. *Review of Economics and Statistics*, 94(1):348–358, 2012.

A. E. Roth and M. A. O. Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press, 1990.

L. S. Shapley and M. Shubik. The assignment game i: The core. *International Journal of game theory*, 1(1):111–130, 1971.

E. M. Snir and L. M. Hitt. Costly Bidding in Online Markets for IT Services Costly Bidding in Online Markets for IT Services. *Management Science*, 49(11):1504–1520, 2003.

M. Sorensen. How Smart Is Smart Money? a Two Sided Matching Model of Venture Capital. *The Journal of Finance*, 62(6):2725–2762, 2007.

C. Stanton and C. Thomas. Landing the First Job: The Value of Intermediaries in Online Hiring. 2012.

Y. Yang and A. Goldfarb. Banning Controversial Sponsors: Understanding Equilibrium Outcomes When Sports Sponsorships Are Viewed as Two-Sided Matches. 2015.

Y. Yang, M. Shi, and A. Goldfarb. Estimating the Value of Brand Alliances in Professional Team Sports. *Marketing Science*, 28(6):1095–1111, 2009.

H. Yoganarasimhan. The Value of Reputation in an Online Freelance Marketplace. *Marketing Science*, 32(6):860–891, 2013.

List of Tables

Table 1: Sample Description

# projects	663
# bids	11,914
# freelancers	4,665
# matched projects	270
# matched freelancers	248
# freelancer source countries	130
# employer source countries	72
Top employer source countries	U.S., India, U.K., Australia, Canada,
Top freelancer source countries	India, Pakistan, Philippines, Bangladesh, U.S.

Table 2: Variable Description

Employer (Project) Attibutes

- Budget (in 100 USD): the maximum budge normalized by the median project length (by day) proposed by bidders

- AvgRating: average rating received by the employer

- Country

- Developed: whether the employer comes from a developed country, based on World Bank classification following Agrawal et al. (2014)

- GDP_PPP (per capita, in 1000 USD): gross domestic product of employer's country converted to U.S. dollars using purchasing power parity rates, per capita

Freelancer Attributes

- Tenure: logarithm of number of days since registration

- AvgRating: average rating received by the freelancer

- Country

- Developed: whether the freelancer comes from a developed country, based on World Bank classification following Agrawal et al. (2014)

- GDP_PPP (per capita, in 1000 USD): gross domestic product of freelancer's country converted to U.S. dollars using purchasing power parity rates, per capita

Employer-Freelancer Attributes

- SameCountry: whether employer and freelancer are from the same country

- PriorInteraction: whether the freelancer has worked for the employer in the past

- ContractPrice: the winning contract price normalized by the median project length (by day) proposed by bidders

Table 3: Summary Statistics

Variable	Obs	Mean	SD	Median	Min	Max
Employer specific attributes						
Budget	270	0.56	0.49	0.31	0.04	5
Developed	270	0.79	0.41	1	0	1
GDP_PPP	270	38.49	18.08	44.35	2.22	67.04
AvgRating	270	4.33	1.63	5	0	5
Freelancer specific attributes						
Tenure	248	5.81	1.74	6.31	0	8.08
Developed	248	0.17	0.38	0	0	1
GDP_PPP	248	11.97	15.41	4.31	1.37	96.99
AvgRating	248	4.54	1.21	4.89	0	5
Employer-Freelancer specific attributes						
SameCountry	270	0.07	0.26	0	0	1
PriorInteraction	270	0.09	0.28	0	0	1
ContractPrice	270	0.32	0.27	0.25	0.04	2.5

Table 4: Estimation Results

Variable	Model 1	Model 2	Model 3
Budget*Ln(FreelancerTenure)	1	1	1
EmployerDeveloped*FreelancerDeveloped		**-1.724**	
		(-38.058, -0.820)	
EmployerGDP_PPP*FreelancerGDP_PPP			**-0.001**
			(-0.024, -0.001)
EmployerAvgRating*FreelancerAvgRating	**0.299**	**0.701**	**0.613**
	(0.089, 6.795)	**(0.240, 7.165)**	**(0.205, 10.036)**
SameCountry	**1.278**	**3.396**	**3.134**
	(1.130, 4.321)	**(1.564, 49.789)**	**(2.583, 74.395)**
PriorInteraction	**49.302**	**15.102**	**99.573**
	(12.642, 195.271)	**(10.331, 96.435)**	**(32.748, 289.480)**
Maximum Score	0.6812	0.6863	0.6881
Smoothed Maximum Score $S(\beta)$	0.6739	0.6804	0.6864

Note: The results are obtained from smoothed maximum score estimation.

The 95% bootstrap confidence intervals are reported in parentheses.

Table 5: Comparing Observed Matches, Optimal Matches, and Counterfactual Matches

Variables	Observed Matches	Optimal Matches			
		Under current scenario		Under counterfactual scenario	
		Matched Freelancer	All Freelancer	Matched Freelancer	All Freelancer
Number of Matched Employers	270	267		67	
Number of Matched Freelancers	248	245		62	
Total Project Value	24,312.2	23,671.2		10,189.3	
Mean Contract Price	32.2	33.9	33.5	52.9	13.2
Median Contract Price	29.2	31.1	30.9	46.6	0
Gini Coefficient	0.31	0.40	0.41	0.25	0.81

Note: the contract prices for the optimal matches are predicted based on linear regression that includes employer specific, freelancer specific, and pair specific attributes. Project value is computed as the product of contract price and project length.