Call for Bids to Improve Matching Efficiency: Evidence from Online Labor Markets

Short Paper

Xue Guo
Temple University
1801 N Broad St,
Philadelphia, PA 19122
xueguo@temple.edu

Jing Gong
Temple University
1801 N Broad St,
Philadelphia, PA 19122
gong@temple.edu

Paul Pavlou
Temple University
1801 N Broad St,
Philadelphia, PA 19122
pavlou@temple.edu

Abstract

In online labor markets, a Call for Bids (CFB) serves as a form of contracts with the description of services required from service providers. It helps service providers understand and perform the project by reducing the uncertainty about the required services. In this study, we (1) theorize the nature of description uncertainty in CFBs from three dimensions—codifiability, requirements, and flexibility, and (2) examine their respective role in matching efficiency between employers and service providers. We use content analysis and deep learning algorithms to analyze unstructured textual data and test our model using archival data from a major online labor platform. The preliminary results show that different dimensions of description uncertainty have different empirical effects on a project's matching outcome. Our findings provide rich implications for employers, service provider, and platform owners. Also, our text mining approach can be applied in other fields that involve analyzing large-scale textual data.

Keywords: Online Labor Markets, Call for Bids, Codifiability, Requirements, Flexibility, Matching Efficiency

Introduction

Online labor markets have profoundly changed the way people work and employers hire workers. These markets serve as intermediaries to connect potential employers and service providers. In 2016, more than 55 million Americans participated in the freelance economy, which accounts for 35% of the total workforce, thus creating more than $1 trillion in revenues; these numbers have been steadily increasing every year (Freelancers Union 2016). However, one of the biggest challenges faced by labor markets is how to increase the matching efficiency between employers and service providers. Studies find that more than 60% of the projects failed to reach a contract (e.g., Snir and Hitt 2003; Carr 2003; Zheng et al 2015), which is a great waste of time and effort. Hence, it is important to explore the reasons of the low matching efficiency on labor markets.

While prior studies found that the matching efficiency of online labor markets depends on project size (Snir and Hitt 2003), bid evaluation cost (Carr 2003), reputation (Yoganarasimhan 2013), and auction format (Hong et al. 2015), this study focuses on the role of the *Call For Bids* (CFB), or the project descriptions, in affecting matching efficiency. In online labor markets, CFBs are created by employers to elicit bids from service providers who offer to perform the project. Generally, a CFB contains the required procedures, deliverables, acceptable standards, and flexibility provisions as specified by the employer who posted the project. The main goal of our study is to analyze the content of CFBs and examine its role in the matching efficiency in online labor markets.

A CFB plays an important role in determining employers' and service providers' behaviors during the matching process. It serves as a contract to regulate the behaviors of both sides. In the focal context of online labor markets where a reverse auction mechanism is used for the procurement of services, the detailed process is as follows: 1) an employer posts a CFB to describe the required services, deliverables, and duration; 2) potential service providers browse the website to find suitable CFBs that match their skills and propose bids by entering a reverse auction; 3) the employer (the buyer of the project) decides whether to award a winner from the set of bidding service providers; 4) if a winner is chosen, the winning service provider finishes the project, and the employer and provider then give feedback to each other. As shown in Figure 1, the entire process is completed online without face-to-face communication. Employers thereby mainly rely on CFBs to describe their required services and attract bids from potential service providers who can perform the required services. Service providers also rely on CFBs to learn about the project requirements and evaluate whether their skills match the project requirements.

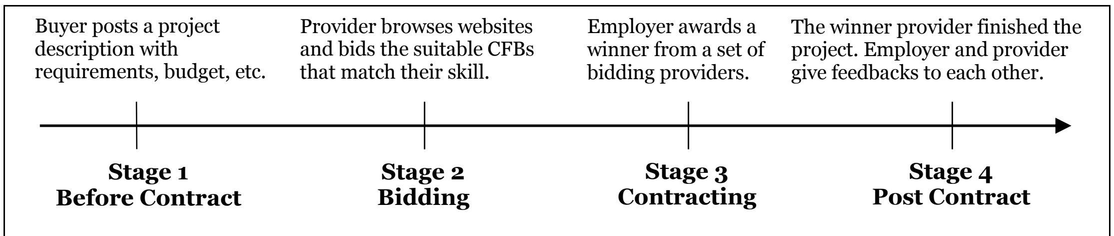

Figure 1. The Matching Process in Online Labor Markets

Accordingly, the extent to which service providers cannot assess the nature of the CFBs, which we broadly refer to as *description uncertainty* in the CFBs, may affect matching efficiency. Description uncertainty comes from individuals' inability to foresee all possible future states and all costs associated with stating all possible outcomes (e.g., Hart and Moore, 1988; Segal 1999; Simon 1981). Specifically, we define description uncertainty as the extent to which providers cannot accurately predict the detailed procedures and final outcomes based on the CFBs. Building on the literature on IT contracts (e.g., Chen and Bharadwaj 2009; Fitoussi and Gurbaxani 2012), we propose and analyze three dimensions of description uncertainty—**codifiability** (Mithas and Whitaker 2007; Liu and Aron, 2015), **requirements** (Jensen and Meckling 1992; Goo et al, 2009), and **flexibility** (Susarla 2012) (Table 1). Accordingly, we ask the following research questions:

- What is the nature of description uncertainty (i.e., codifiability, requirements, and flexibility) in Call for Bids (CFBs) in online labor markets?
- How do the proposed dimensions of description uncertainty affect matching efficiency?

To answer these two research questions, we theorize the nature of the description uncertainty in CFBs, and how the proposed dimensions of description uncertainty affect the contract success rate (matching efficiency). Specifically, the first dimension of description uncertainty, *codifiability*, is the extent to which a CFB contains detailed procedures. CFBs with high codifiability helps service providers better understand the necessary procedures required for the project before they propose bids (Mani et al. 2012). The second dimension of description uncertainty, *requirements*, is the extent to which a CFB contains clear deliverables and specific quality and quantity standards. Clear deliverables and standards could help employers and service providers establish common goals for the project (Choudhury and Sabherwal 2003; Viswesvaran and Ones 2000). The third dimension of description uncertainty, *flexibility*, is the extent to which a CFB contains provisions that permit changes in project requirements (Masten and Crocker 1985;

Joskow 1988). Flexibility provisions provide service providers renegotiation opportunities during project execution (Susarla 2012).

Then, we theorize separately how each of these three dimensions of description uncertainty in the CFBs may affect the final contract success rate. In general, description uncertainty may affect matching efficiency from two perspectives. First, description uncertainty due to low codifiability and a lack of clear requirements may execrate the information asymmetry between employers and service providers, thus reducing the matching efficiency (Golan 2005; Liu and Aron 2015). Specifically, low codifiability and a lack of specific requirements in CFBs may increase the unforeseen risks associated with the project, which may deter service providers from proposing bids. In addition, these two types of description uncertainty may increase the difficulty of conducting the project, thus preventing bids. Thus, service providers may choose to bid on projects that they have a high probability of accomplishing, which lead to low matching efficiency.

Second, description uncertainty related to flexibility may affect matching efficiency in both directions. The nature of IT service is complex and constantly changing. Flexibility provisions in CFBs offer both employers and providers the opportunity to update contract terms during the execution process. Thus, flexibility provisions reduce employer's cost of specifying each possible outcome during the execution process (Bajari and Tadelis 2001). It is often an economic choice by the employer to choose whether allows the flexibility or not. On the one hand, flexibility reduces the ex-post bargaining cost of service providers (Susarla 2012) and the possibility of being locked in the relationship (Tadelis 2002). Hence, flexibility provisions in CFBs may attract better bids and thus lead to higher matching efficiency. On the other hand, flexibility also increases the ex-post risks for service providers by the requirements instability (Nidumolu 1995; Banerjee and Duflo 2000). In this case, employers can modify the project requirements under changed circumstances, which cause unexpected work or investment for providers. Therefore, the flexibility provisions may lead to lower contract success rate. The detailed research framework in shown in Figure 2.

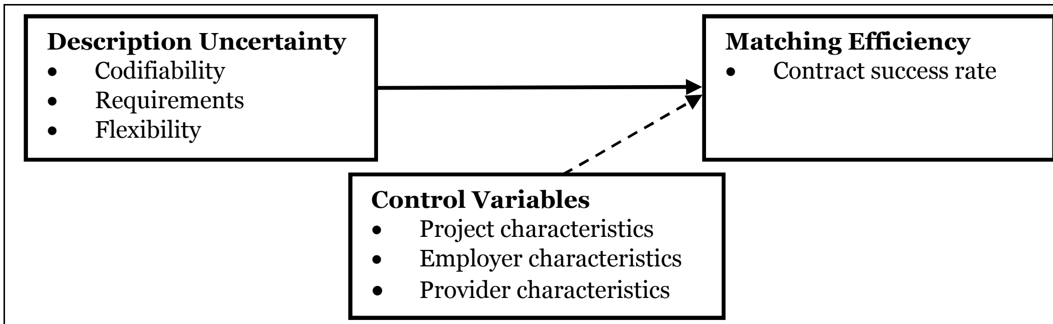

Figure 2. Research Framework

To capture the three dimensions of description uncertainty in CFBs, i.e., codifiability, requirements, and flexibility, we integrate content analysis, state-of-the-art deep learning algorithms, and supervised learning to analyze the textual content of CFBs. First, we use content analysis (Berelson, 1952; Holsti, 1969) to classify each word in our training set into different levels of codifiability, requirements, and flexibility. Second, we use a deep learning algorithm called Word2Vec (Mikolov et al. 2013a), which is used to learn vector representation of words and documents, to convert each word in the CFBs into a vector that best predicts each word's occurrence in certain context. Third, we use the vectors to train a binary classifier and predict the categories for all words in our dataset.

Our research context is Freelancer.com, one of the largest online labor/freelancing platforms in the world, which has more than 8 million registered users in total, earning 229 million gross payment volume in 2016. Our unique and rich data set contains all projects posted on the platform from March 2, 2014 to March 31, 2014. The dataset includes detailed characteristics at the project, employer, and service provider levels. These data features allow us to empirically test our hypotheses.

By treating each CFB in our dataset as a unique document for text analysis, our deep learning algorithm, Word2Vec (Mikolov 2013a), achieves high accuracy of classifying each document into the right level of codifiability, requirements, and flexibility. Our preliminary results confirm our hypotheses on the dimensions of description uncertainty. First, high codifiability and clear requirements of a CFB are positively associated with the contract success rate. Second, the presence of flexibility provisions in CFBs is positively associated with the contract success rate. Our study shows that different dimensions of description uncertainty have different empirical effects on a project's matching outcome.

This study makes several contributions. First, our research theorizes the nature of description uncertainty into three dimensions, i.e., codifiability, requirements, and flexibility, and analyzes their effects on matching efficiency. Although considerable research has been devoted to studying the factors that affect matching efficiency in online labor markets (e.g., Snir and Hitt 2003; Pallia, 2014; Hong et al, 2015), few studies have examined the nature of CFBs in terms of description uncertainty. Extending the IT contracting literature (e.g., Chen and Bharadwaj 2009; Susarla et al 2010), we developed three dimensions of CFBs (codifiability, requirements, and flexibility) and empirically tested their corresponding effects on matching efficiency. Second, we measured the three dimensions of description uncertainty using textual data. Our methodology that uses deep learning algorithms to extract features from large-scale textual data is relatively new in the IS literature. The deep learning algorithm that we use in this study, Word2Vec, provides us more accurate measurements than traditional N-gram models by considering the word context and multiple degrees of similarity (Mikolov et al. 2011; Mikolov et al. 2013a; Mikolov et al. 2013b). Our method can be easily combined with other supervised methods such as classification or unsupervised methods such as clustering models and applied to many other fields that require analyzing large scaled textual information.

Related Literature

The first stream of literature relates to uncertainty in online markets. There is a rich body of literature that examines the antecedents and consequences of seller and product uncertainty in the context of online markets. Pavlou et al. (2007) defines (buyer's) perceived uncertainty as "*the degree to which the outcomes of a transaction cannot be accurately predicted by the buyer due to the seller- and product-related factors*." In terms of antecedents of buyer's perceived uncertainty, studies find that feedback systems (Dellarocas 2003), third party structures (Pavlou and Gefen 2004) and textual reviews (Pavlou and Dimoka 2006) affect buyer's perceived uncertainty about sellers, and online product descriptions (Dimoka 2012) affect buyer's perceived uncertainty about sellers and products. These factors provide information about the sellers and products, which reduces the uncertainty that buyers perceive. In terms of the consequences of buyer's perceived uncertainty, studies find that it is related to trust building (Ba and Pavlou 2002), price premium (Dimoka et al. 2012), and product returns (Hong and Pavlou 2014). For instance, Dimoka et al. (2012) found that reducing uncertainty enables seller differentiation, which helps sellers attract more bids. In our study, extending the literature, we focus on *description uncertainty* in the CFBs that may affect service providers' ability to assess the requirements of the projects, which may affect their bidding behavior and contract success rate.

The second stream of literature relates to online labor markets. On the one hand, transaction cost theory suggests that online labor markets facilitate the exchange of labor and service by reducing search costs (Freund and Weinhold 2002), which increases matching efficiency (Snir and Hitt 2003). On the other hand, according to information asymmetry theory, online markets and the complexity of IT services increase the degree of information asymmetry. The nature of transacting with strangers on online markets may increase the risk of opportunism behaviors, which will lower matching efficiency. With the aim of improving the matching efficiency of online labor markets, previous literature has analyzed various factors that affect employers' hiring decisions, including service providers' uncertainty (Banker et al. 2002; Pallais 2014) and employers' evaluation cost (Snir and Hitt 2003). For example, the reputation scores of service providers (Moreno and Terwiesch 2014), prior working interactions (Gefen and Carmel 2008), and provider location (Mithas and Whitaker 2007; Hong at el. 2017; Gong 2017) may affect the service providers' probability of being hired. However, we have limited understanding of how CFBs may play a role in affecting the service providers' bidding behaviors. CFBs serve as a form of contracts that specify the services and deliverables required by employers, thus provide a guideline for potential service providers.

Hypotheses

We theorize the role of description uncertainty in matching efficiency by examining three dimensions of CFBs: codifiability, requirements, and flexibility. The detailed definitions of the three dimensions are presented in Table 1.

Table 1. Dimensions of CFBs and Definition

Dimension	Definition	Proposed Effects
Codifiability	The extent to which a CFB contains detailed procedures.	1. Increase easiness and accuracy of process execution (Mani et al 2012). 2. Facilitate cost evaluation process and reduce search cost (Snir and Hitt 2003).
Requirements	The extent to which a CFB contains clear deliverables and specific quality and quantity standards.	1. Establish common goals (Goo et al 2009). 2. Facilitate cost evaluation process and reduce search cost (Snir and Hitt 2003, Carr 2003).
Flexibility	The extent to which a CFB contains provisions that permit changed circumstances in the project requirements.	1. Reduce ex-post bargaining cost (Susarla 2012). 2. Reduce risk of being locked in a specific relationship (e.g., Tadelis 2002). 3. Increase risk of requirements instability (e.g., Gopal and Koka 2012)

In the literature on IT contracts, ***codifiability*** has been used to describe the ability to create a set of task execution rules for how a process is to be executed (Liu and Aron 2015). In our context, we define codifiability as the extent to which a CFB contains detailed procedures. Codifiability may affect matching efficiency from two perspectives. First, describing tasks in steps increases the ease and accuracy of contract execution (Mani et al. 2012). Thus, CFBs with high codifiability are more likely to attract suitable providers and have higher probability to contract with a service provider than those with low codifiability. Specifying detailed procedures also reduces the unforeseen risk during the execution process. Without specific procedures, service providers may misunderstand the project requirements and perform the task differently from what is expected by the employer. Second, it helps service providers better evaluate the cost of task execution and whether their skills match the project. Therefore, a project with high codifiability can attract more suitable bids, resulting in higher matching efficiency. Thus, we hypothesize:

Hypothesis 1: Codifiability is positively associated with the contract success rate.

Another important dimension of description uncertainty is the presence or lack of detailed requirements in a CFB. We define ***requirements*** from three perspectives—deliverables, quality, and quantity. Specifically, we define requirements as the extent to which a project description contains clear deliverables and specific quality and quantity standards. Requirements provide information about a project final goals and detailed standards. First, project with clear requirements are more likely to attract bids since it reduces the service providers' search costs (e.g., Snir and Hitt 2003; Carr 2003) and help them evaluate their cost and match their skills faster. Second, requirements regulate providers' bidding behaviors by providing outcome control (Goo et al 2009). If a provider cannot achieve the specific standard, the provider is unlikely to bid on the project. Therefore, the CFBs are more likely to attract bids from qualified service providers, which reduces the employers' evaluation cost. For these two reasons, CFBs that contain detailed requirements about deliverables, quality, and quantity will deter unqualified bids and reduce employers' bid evaluation costs, which leads to higher matching efficiency. In sum, we hypothesize:

Hypothesis 2: Requirements are positively associated with the contract success rate.

The third dimension of CFBs on which we focus is ***flexibility***. Flexibility is an essential part of formal contract structures (Chen and Bharadwaj 2009). Its main role is to provide renegotiation opportunities about the contract terms, which include provisions about price adjustments (Joskow 1988b) and quantity changes (Tsay 1999). Building on previous literature (e.g., Susarla 2012), we define flexibility as the extent to which a CFB contains provisions that permit changed circumstances in the project requirements. The existence of flexibility provisions efficiently reduces the contingencies needed to be written in the CFBs, which reduce the employers' cost of writing the CFBs. Thus, it may be an economic choice for employers to determine whether include this provision or not.

Flexibility affects the matching efficiency from both directions. On the one hand, flexibility reduces the ex-post bargaining cost (Susarla 2012) and ex-post risk of being locked in a specific relationship (Tadelis 2002). It provides necessary solutions to the unexpected changes during the service execution process. Service providers could update or terminate the contract based on environment change. Thus, projects with flexibility provisions in the CFB are more likely to result in a contract. On the other hand, flexibility increases the risks of requirement instability (e.g., Nidumolu 1995; Wallace et al 2004). It increases unforeseen ex-post risks by allowing potential rent-seeking behaviors of service employers (e.g., Bahli and Rivard 2005; Susarla 2012). For instance, employers can update the project requirements under changed circumstances, which entail costs for providers such as require rework (Gopal and Koka 2012) or specific investment (Nidomolu 1995; Pressman 2005). Hence, potential providers are less likely to bid the CFB with flexibility provisions to avoid potential risks, which lead to lower contract success rate. Therefore, we hypothesize that:

> *Hypothesis 3a: Flexibility is positively associated with the contract success rate.*
> *Hypothesis 3b: Flexibility is negatively associated with the contract success rate.*

Measurement Development

To measure the three dimensions of description uncertainty in CFBs, i.e., codifiability, requirements, and flexibility, we combine content analysis and newly-developed deep learning algorithms to develop our measurements. Our approach contains three steps: 1) manual coding, 2) using deep learning for feature extraction, and 3) binary classification, as elaborated below.

First, we employed manual content analysis methods to classify words in our dataset. We adapted the measurements from literature for codifiability (Liu and Aron 2015), requirements (e.g., Krisch et al 2002; Goo et al. 2009; Chen and Bharadwaj 2009; Gopal and Koka 2012), and flexibility (Susarla 2012). Our dataset contains 59,054 CFBs, with 2,541,354 words[1], and 7,560 unique words with frequency above 20. We randomly selected 20% of the unique words from our dataset and manually coded each word into one of the two levels (high or low) for each dimension (i.e., codifiability, requirements, and flexibility). We used the labeled data as the training set for the binary classifier, which we describe later. The first column of Table 3 presents some results of the coded word for high codifiability, high requirements, or high flexibility.

The second step was to use a deep learning algorithm, i.e., Word2Vec (Mikolov et al. 2013a) to extract textual features of each CFB, which are later used as inputs to build classifiers that learn the relationship between these textual features of a CFB and the three dimensions of description uncertainty. Word2Vec is a machine learning technology that converts each word into a dense vector that best predicts context words in the document. This method is developed from neural networks to mimic the process of the human brain to relationships between words automatically. It outperforms other text mining algorithms (e.g., bag of words and skip gram) by considering the order and meaning of the context words and becomes the new state-of-the-art (Mikolov et al. 2013a; Mikolov et al. 2013b).

The intuition of Word2Vec is that words with similar syntactic and semantic meanings tend to occur in the same context. The algorithm mainly takes two steps. First, it maps each word and document into a unique vector using unsupervised training. Second, it assigns random probabilities to each vector and updates the probabilities to maximize the ability of context and documents to predict the word in the context. The algorithm iterates for every word and document in the data until the maximum conditional probability is achieved. Thus, the word and document with similar features move towards each other in a multi-dimensional feature space. We utilized this algorithm to capture the textual features with high accuracy and efficiency. Following Socher et al. (2013), we used 300-feature dimensionality and 10 context words to train the model, and we iterated the algorithm 20 times to achieve the best accuracy. In this step, each word is represented by a vector that captures its syntactic and semantic meanings.

Third, we used the labeled data to train a binary classifier. From the deep learning algorithm, we turned each word into a 300-dimensional vector. We used the coding sample (20% of the unique words) as a

[1] Before content analysis, we pre-processed the documents and removed stop words.

training set for a binary classifier and then used the classifier to predict the level of codifiability, requirements, and flexibility of the remaining words in the dataset. We utilized three commonly used classifiers—Naïve Bayes, logistic regression, and support vector machine (SVM). The three classifiers are trained and tested on the coded results. The corresponding cross-validation accuracies are shown in Table 2. Based on the accuracy results, we chose the logistic classifier in our final analysis. Thus, we predicted the level (high or low) of codifiability, requirements, and flexibility for all unique words in the dataset. The second column of Table 3 presents some examples of the predicted words that are associated with high codifiability, high requirements, or high flexibility.

Table 2. Cross-Validation Accuracy

Dimension	**Naïve Bayes**	**Logistic**	**Support Vector Machine**
Codifiability	0.82	0.83	0.81
Requirements	0.93	0.95	0.93
Flexibility	0.94	0.96	0.95

Table 3. Examples of Words

Dimension	**Input Examples**	**Output Examples**
Codifiability	procedures, guidelines, rules, steps	instructions, criteria, followed, outlines, explanations
Requirements	deliverables, goals, milestones, guarantee, quality, quantity	output, aim, achieve, assurance, amount, standards
Flexibility	flexible, negotiable, adjustable, updated, change, options	negotiation, modifiable, changeable, remuneration, charges, urgency

Data

Our data come from Freelancer.com, one of the largest freelancing platforms that bring together buyers and service providers of IT services and other professional services. By August 2016, there are more than 20 million registered users and 9 million jobs posted on Freelancer.com. Our dataset includes all the projects posted between March 2, 2014 and March 31, 2014 on the platform. The detailed definitions and statistics of the key variables are shown in Table 4.

Table 4. Definition and Summary Statistics of Key Variables

Variable	**Variable Definition**	**Mean**	**SD**	**Min**	**Max**
Codifiability	Frequency of words belong to high codifiability	1.32	3.15	0	101
Requirements	Frequency of words belong to high requirements	2.77	4.50	0	118
Flexibility	Frequency of words belong to high flexibility	2.31	3.59	0	57
Number of Bids	Total number of bids received in an auction	17.01	20.01	0	303
Contract success rate	Whether a contract is reached	0.37	0.48	0	1
Word count	Word count of the CFB	85.74	96.60	2	1289
Control Variables					
Number of skills	The number of skills a project requires	3.28	1.43	1	8
Project min budget	The lower bound of the buyer's budget	214.71	662.87	0	30000
Project max budget	The upper bound of the buyer's budget	508.96	1913.15	10	200000
Auction duration	Number of days an auction was active	17.59	17.76	0	101
Buyer experience	Number of reviews the buyer has received	16.09	50.83	0	1682
Avg buyer rating	The average rating of a buyer	2.92	2.42	0	5
Is quickly hired	Whether the project is quickly hired	0.12	0.33	0	1
Budget fixed	Whether the project has fixed budget	0.86	0.35	0	1
Urgent	Whether the project is urgent	0.02	0.12	0	1
Feature	Whether the project is featured	0.03	0.17	0	1
Project categories					
PC1	Websites, IT& Software				
PC2	Mobile Phones & Computing				
PC3	Data Entry & Admin				

Empirical Model

Empirically, we used a logistic regression to analyze the role of description uncertainty in matching efficiency (contract success rate). Our unit of analysis is at the CFB level. The main dependent variable is $Contracted$, a binary variable indicating whether a project resulted in a contract. The main independent variables are the frequency of the words in the CFB that are related to high codifiability, requirements, and flexibility. In addition, we also controlled for other textual features, such as word count. Lastly, we controlled for buyer-, provider-, and project-specific characteristics, including category fixed effects ($CategoryFE$) in the model. Equation (1) outlines our detailed empirical model. To address the within-employer variations, we also included employer fixed effects ($EmployerFE$) in the model.

$$\Pr(Contracted = 1) = Logistic(\beta_0 + \beta_1 Codifiability + \beta_2\, Requirements + \beta_3 Flexibility + \beta_4 WordCount + \beta_6 Control\ variables + CategoryFE + EmployerFE + \varepsilon) \qquad (1)$$

Table 5 shows the results from the logistic regression. Model 1 is the base model with only controls, Model 2 shows the model include the three dimensions of description uncertainty, and Model 3 includes the employer fixed effects. From three models, we find relative consistent and statistically significant results. To interpret the results, a one-unit increase in the frequency of words related to codifiability, requirements, and flexibility, on average, is associated with 0.33%, 1.59%, and 0.80% higher probability of contract success rate, respectively. To compare the effects among three dimensions, requirements have the largest magnitude on the contract success rate. The results are consistent with our hypotheses that description uncertainty due to low codifiability and requirements is negatively associated with the contract success rate[2], while description uncertainty related to flexibility is positively associated with the contract success rate.

Table 5. Regression Results

VARIABLES	(1) Contracted	(2) Contracted	(3) Contracted
Codifiability		0.0243***	0.0191***
		(0.00447)	(0.00643)
Requirements		0.0254**	0.0988***
		(0.0116)	(0.0274)
Flexibility		0.0366**	0.04163**
		(0.00287)	(0.0153)
Word count	0.00277***	0.00403***	0.00495***
	(0.000353)	(0.000489)	(0.000571)
Word count square	-3.81e-06***	-3.89e-06***	-4.30e-06***
	(7.54e-07)	(8.40e-07)	(9.38e-07)
Number of skills	-0.0287***	-0.0264***	-0.0202
	(0.00999)	(0.0100)	(0.0138)
ln (Project max budget)	0.117***	0.119***	0.0809***
	(0.0267)	(0.0267)	(0.0297)
ln (Avg buyer rating)	0.986***	0.985***	1.427***
	(0.0171)	(0.0171)	(0.0306)
Constant	0.545***	0.545***	
	(0.0704)	(0.0706)	
Observations	59,054	59,054	59,054
Category FE	YES	YES	YES
Buyer FE	No	No	YES
Number of buyer_id	37,481	37,481	37,481
Pseudo R^2	0.5442	0.5452	0.5492

Notes: *p<0.1, **p<0.05, ***p<0.01
Other control variables include *is quickly hired*, *auction duration*, *urgent*, *featured*. These variables are included in the regression models but not reported in the table because of page limitation.

[2] Higher frequency of words related to codifiability and requirements represent lower description uncertainty.

Manuscript Status and Future Work

Currently, we have analyzed the textual data at the word level. Our text mining methods show high accuracy and the results are promising. In the next steps, we plan to conduct document-level analysis to capture richer textual information. Specifically, we have designed the coding instructions and training samples for the manually coding, and have conducted several pilot studies to revise the coding instructions. We are in the process of recruiting workers from Amazon Mechanical Turk (AMT) to code a subset of the CFBs into two levels (high or low) for the three dimensions of description uncertainty. The AMT workers will be asked to code 5,000 descriptions, which we will later use as the training set for training a binary classifier. Next, we plan to train the deep learning algorithm (Doc2Vec) at the document level, which will represent each CFB using a vector. Third, we plan to use the vectors to train a binary classifier and predict the levels of description uncertainty for the remaining CFBs. Fourth, we plan to analyze the role of the dimensions of description uncertainty in CFBs, as measured using document level analysis, in contract success.

Contributions

This study aims to make the following contributions to the IS literature. First, our research theorizes the nature of description uncertainty in CFBs into three dimensions, i.e., codifiability, requirements, and flexibility, and analyzes their respective role in matching efficiency between employers and service providers in the context of online labor markets. While there is limited knowledge about the underlying role of description uncertainty embedded in CFBs and how the nature of the CFBs may affect matching efficiency, this study contributes to this nascent body of research on online labor markets by focusing in the inherent nature of CFBs and the key underlying dimensions embedded in CFBs.

Second, the methodology that we propose in this research, i.e., the use of deep learning algorithms to extract features from large-scale textual data, is relatively new in the IS literature. Different from other classic text mining algorithms, deep learning algorithms aim to mimic the process of the human brain to capture the sequence and semantic meanings of each word in the text. Thus, it provides a more accurate and meaningful measurement than other classic algorithms such as traditional N-gram models (e.g., Bengio et al. 2003; Schwenk 2007; Mikolov et al. 2011). Our method can be easily combined with other classification or clustering models and applied in other fields that involve analyzing large scaled textual information.

Third, our research provides practical implications. For employers, the findings shed light on how to write CFBs to attract more qualified bids and increase the probability of finding qualified service providers. For service providers, the findings help them better understand the structure of CFBs which facilitate their cost evaluation process and reduce research cost. For the platform owners, our results help them design guidelines and regulations for CFBs to assist employers to write CFBs to reduce uncertainty, which may increase matching efficiency. In addition, the platform takes 13% (3% from employer and 10% from freelancer) of payments for matched project as an introduction fee. Thus, the increased matching efficiency directly lead to higher platform revenue.[3]

[3] https://www.freelancer.com/feesandcharges/

References

~~Anderson, S. W., and Dekker, H. C. 2005. "Management Control for Market Transactions," *Management Science* (51:12), pp. 1734–1752.~~

Ba2002 Ba, S., and Pavlou, P. A. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets," *MIS Quarterly*, pp. 243–268.

~~Baccarini, D. 1996. "The Concept of Project Complexity—a Review," *International Journal of Project Management* (14:4), pp. 201–204.~~

~~Bahli, B., and Rivard, S. 2005. "Validating Measures of Information Technology Outsourcing Risk Factors," *Omega* (33:2), pp. 175–187.~~

~~Bajari, P., and Tadelis, S. 2001. "Incentives versus Transaction Costs," *Rand Journal of Economics*, pp. 387–407.~~

~~Banerjee, A. V., and Duflo, E. 2000. "Reputation Effects and the Limits of Contracting," *The Quarterly Journal of Economics* (115:3), pp. 989–1017.~~

~~Banker, R. D., Hwang, I., and Mishra, B. K. 2002. "Product Costing and Pricing under Long-Term~~ Capacity Commitment," *Journal of Management Accounting Research* (14:1), pp. 79–97.

Bapna2010 Bapna, R., Barua, A., Mani, D., and Mehra, A. 2010. "Research Commentary—Cooperation, Coordination, and Governance in Multisourcing," *Information Systems Research* (21:4), pp. 785–795.

Barki1993 Barki, H., Rivard, S., and Talbot, J. 1993. "Toward an Assessment of Software Development Risk," *Journal of Management Information Systems* (10:2), pp. 203–225.

~~Bengio, Y., Ducharme, R., Vincent, P., and Jauvin, C. 2003. "A Neural Probabilistic Language Model," *Journal of Machine Learning Research* (3:Feb), pp. 1137–1155.~~

~~Berelson, B. 1952. *Content Analysis in Communications Research*. (http://pubman.mpdl.mpg.de/pubman/item/escidoc:2281462/component/escidoc:2283003/Berelson_1952_content_analysis.pdf).~~

~~Boehm, B. 1989. "Software Risk Management," *ESEC'89*, pp. 1–19.~~

~~Carr, S. M. 2003. "Note on Online Auctions with Costly Bid Evaluation," *Management Science* (49:11), pp. 1521–1528.~~

Chen2009 Chen, Y., and Bharadwaj, A. 2009. "An Empirical Analysis of Contract Structures in IT Outsourcing," *Information Systems Research* (20:4), pp. 484–506.

Choudhury2003 Choudhury, V., and Sabherwal, R. 2003. "Portfolios of Control in Outsourced Software Development Projects," *Information Systems Research* (14:3), pp. 291–314.

~~Dellarocas, C. 2003. "The Digitization of Word of Mouth," *Management Science* (49:10), pp. 1407–1424.~~

~~Dewan, S., and Hsu, V. 2004. "Adverse Selection in Electronic Markets," *The Journal of Industrial Economics* (52:4), pp. 497–516.~~

Fitoussi2012 Fitoussi, D., and Gurbaxani, V. 2012. "IT Outsourcing Contracts and Performance Measurement," *Information Systems Research* (23:1), pp. 129–143.

~~Freund, C., and Weinhold, D. 2002. "The Internet and International Trade in Services," *The American Economic Review* (92:2), pp. 236–240.~~

~~Freelancers Union, and Elance-o-Desk. 2016. *Freelancing in America. A National Survey of the New Workforce*. New York.~~

Gefen2008 Gefen, D., and Carmel, E. 2008. "Is the World Really Flat?," *MIS Quarterly*, pp. 367–384.

Ghose2009 Ghose, A. 2009. "Internet Exchanges for Used Goods," *Mis Quarterly*, pp. 263–291.

~~Ghosh, M., and John, G. 1999. "Governance Value Analysis and Marketing Strategy," *The Journal of Marketing*, pp. 131–145.~~

~~Golan, L. 2005. "Counteroffers and Efficiency in Labor Markets with Asymmetric Information," Journal of Labor Economics (23:2), pp. 373–393.~~

~~Gong, J. 2016. *A Structural Two-Sided Matching Model of Online Labor Markets*. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2745450).~~

Goo2007 Goo, J., Kishore, R., Nam, K., Rao, H. R., and Song, Y. 2007. "An Investigation of Factors That Influence the Duration of IT Outsourcing Relationships," *Decision Support Systems* (42:4), pp. 2107–2125.

Goo2009 Goo, J., Kishore, R., Rao, H. R., and Nam, K. 2009. "The Role of Service Level Agreements in Relational Management of Information Technology Outsourcing," *MIS Quarterly*, pp. 119–145.

Gopal2012 Gopal, A., and Koka, B. R. 2012. "The Asymmetric Benefits of Relational Flexibility," *MIS Quarterly* (36:2). (http://aisel.aisnet.org/cgi/viewcontent.cgi?article=3035&context=misq).

~~Gurbaxani, V., and Whang, S. 1991. "The Impact of Information Systems on Organizations and Markets," *Communications of the ACM* (34:1), pp. 59–73.~~

~~Hancock, J. T., Curry, L. E., Goorha, S., and Woodworth, M. 2007. "On Lying and Being Lied To," *Discourse Processes* (45:1), pp. 1–23.~~

~~Hart, O., and Moore, J. 1998. "Foundations of Incomplete Contracts," National bureau of economic research. (http://www.nber.org/papers/w6726).~~

~~Holsti, O. R. 1969. *Content Analysis for the Social Sciences and Humanities*. (http://admn502awiki.pbworks.com/Content-and-Discourse-Analysis).~~

~~Hong, Y., and Pavlou, P. A. 2013 "Online Labor Markets: An Informal Freelancer Economy," *IBIT Report*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2132869.~~

Hong2014a Hong, Y., and Pavlou, P. A. 2014. "Product Fit Uncertainty in Online Markets," *Information Systems Research* (25:2), pp. 328–344.

Hong2017a Hong, Y., and Pavlou, P. A. 2017. "On Buyer Selection of Service Providers in Online Outsourcing Platforms for IT Services," *Information Systems Research*. (http://pubsonline.informs.org/doi/abs/10.1287/isre.2017.0709).

Hong2016a Hong, Y., Wang, C., and Pavlou, P. A. 2015. "Comparing Open and Sealed Bid Auctions," *Information Systems Research* (27:1), pp. 49–69.

~~Horton, J. J. 2010. "Online Labor Markets," in *International Workshop on Internet and Network Economics*, Springer, pp. 515–522. (http://link.springer.com/chapter/10.1007/978-3-642-17572-5_45).~~

~~Hosios, A. J. 1990. "On the Efficiency of Matching and Related Models of Search and Unemployment," *The Review of Economic Studies* (57:2), pp. 279–298.~~

~~Ipeirotis, P. G., Provost, F., and Wang, J. 2010. "Quality Management on Amazon Mechanical Turk," in *Proceedings of the ACM SIGKDD Workshop on Human Computation*, ACM, pp. 64–67. (http://dl.acm.org/citation.cfm?id=1837906).~~

~~Jensen, M. C., and Meckling, W. H. 1992. *Specific and General Knowledge and Organizational Structure*. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6658).~~

~~Joskow, P. L. 1988a. "Asset Specificity and the Structure of Vertical Relationships," *Journal of Law, Economics, & Organization* (4:1), pp. 95–117.~~

~~Joskow, P. L. 1988b. "Price Adjustment in Long-Term Contracts," *The Journal of Law and Economics* (31:1), pp. 47–83.~~

~~Kirsch, L. J., Sambamurthy, V., Ko, D.-G., and Purvis, R. L. 2002. "Controlling Information Systems Development Projects," *Management Science* (48:4), pp. 484–498.~~

Liu2015b Liu, Y., and Aron, R. 2014. "Organizational Control, Incentive Contracts, and Knowledge Transfer in Offshore Business Process Outsourcing," *Information Systems Research* (26:1), pp. 81–99.

~~Loughran, T., and McDonald, B. 2011. "When Is a Liability Not a Liability?," *The Journal of Finance* (66:1), pp. 35–65.~~

Mani2012 Mani, D., Barua, A., and Whinston, A. B. 2012. "An Empirical Analysis of the Contractual and Information Structures of Business Process Outsourcing Relationships," *Information Systems Research* (23:3-part-1), pp. 618–634.

~~Mason, W., and Suri, S. 2012. "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavior Research Methods* (44:1), pp. 1–23.~~

~~Masten, S. E., and Crocker, K. J. 1985. "Efficient Adaptation in Long-Term Contracts," *The American Economic Review* (75:5), pp. 1083–1093.~~

~~Melnik, M. I., and Alm, J. 2005. "Seller Reputation, Information Signals, and Prices for Heterogeneous Coins on EBay," *Southern Economic Journal*, pp. 305–328.~~

~~Mikolov, T., Deoras, A., Kombrink, S., Burget, L., and Černockỳ, J. 2011. "Empirical Evaluation and Combination of Advanced Language Modeling Techniques," in *Twelfth Annual Conference of the International Speech Communication Association*.~~

~~Mikolov, T., Chen, K., Corrado, G., and Dean, J. 2013a. "Efficient Estimation of Word Representations in Vector Space," *ArXiv Preprint ArXiv:1301.3781*. (https://arxiv.org/abs/1301.3781).~~

~~Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. 2013b. "Distributed Representations of Words and Phrases and Their Compositionality," in *Advances in Neural Information Processing Systems*, pp. 3111–3119. (http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality).~~

Mithas2007a Mithas, S., and Whitaker, J. 2007. "Is the World Flat or Spiky?," *Information Systems Research* (18:3), pp. 237–259.

Moreno2014 Moreno, A., and Terwiesch, C. 2014. "Doing Business with Strangers," *Information Systems Research* (25:4), pp. 865–886.

Nidumolu1995 Nidumolu, S. 1995. "The Effect of Coordination and Uncertainty on Software Project Performance," *Information Systems Research* (6:3), pp. 191–219.

~~Palangi, H., Deng, L., Shen, Y., Gao, J., He, X., Chen, J., Song, X., and Ward, R. 2016. "Deep Sentence Embedding Using Long Short-Term Memory Networks," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)* (24:4), pp. 694–707.~~

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

~~Paolacci, G., Chandler, J., and Ipeirotis, P. G. 2010. *Running Experiments on Amazon Mechanical Turk*. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1626226).~~

Pavlou2006 Pavlou, P. A., and Dimoka, A. 2006. "The Nature and Role of Feedback Text Comments in Online Marketplaces," *Information Systems Research* (17:4), pp. 392–414.

Pavlou2004 Pavlou, P. A., and Gefen, D. 2004. "Building Effective Online Marketplaces with Institution-Based Trust," *Information Systems Research* (15:1), pp. 37–59.

~~Pavlou, P. A., Liang, H., and Xue, Y. 2006. *Understanding and Mitigating Uncertainty in Online Environments*. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2380164).~~

~~Poppo, L., and Zenger, T. 1998. "Testing Alternative Theories of the Firm," *Strategic Management Journal*, pp. 853–877.~~

~~Salancik, G. R., and Pfeffer, J. 1980. "Effects of Ownership and Performance on Executive Tenure in US Corporations," *Academy of Management Journal* (23:4), pp. 653–664.~~

~~Schwenk, H. 2007. "Continuous Space Language Models," *Computer Speech & Language* (21:3), pp. 492–518.~~

~~Segal, I. 1999. "Complexity and Renegotiation," *The Review of Economic Studies* (66:1), pp. 57–82.~~

~~Simon, H. A. 1991. "Bounded Rationality and Organizational Learning," *Organization Science* (2:1), pp. 125–134.~~

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

~~Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., Potts, C., and others. 2013. "Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Vol. 1631), p. 1642. (http://www.anthology.aclweb.org/D/D13/D13-1170.pdf).~~

~~Susarla, A. 2012. "Contractual Flexibility, Rent Seeking, and Renegotiation Design," *Management Science* (58:7), pp. 1388–1407.~~

Susarla2009 Susarla, A., Barua, A., and Whinston, A. B. 2009. "A Transaction Cost Perspective of the" Software as a Service" Business Model," *Journal of Management Information Systems* (26:2), pp. 205–240.

Susarla2010a Susarla, A., Subramanyam, R., and Karhade, P. 2010. "Contractual Provisions to Mitigate Holdup," *Information Systems Research* (21:1), pp. 37–55.

~~Tadelis, S. 2002. "Complexity, Flexibility, and the Make-or-Buy Decision," *The American Economic Review* (92:2), pp. 433–437.~~

~~Tsay, A. A. 1999. "The Quantity Flexibility Contract and Supplier-Customer Incentives," *Management Science* (45:10), pp. 1339–1358.~~

~~Viswesvaran, C., and Ones, D. S. 2000. "Perspectives on Models of Job Performance," *International Journal of Selection and Assessment* (8:4), pp. 216–226.~~

~~Williamson, O. E. 1975. "Markets and Hierarchies," *New York*, pp. 26–30.~~

~~Williamson, O. E. 1983. "Credible Commitments," *The American Economic Review* (73:4), pp. 519–540.~~

Xia2005 Xia, W., and Lee, G. 2005. "Complexity of Information Systems Development Projects," *Journal of Management Information Systems* (22:1), pp. 45–83.

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

~~Zheng, A., Hong, Y., and Pavlou, P. 2015. "Value Uncertainty and Buyer Contracting." *Working Paper*.~~