IT-enabled Monitoring and Labor Contracting in Online Platforms: Evidence from a Natural Experiment

Chen Liang, Yili Hong, Bin Gu
W. P. Carey School of Business, Arizona State University

Abstract

Two-sided online platforms are typically plagued with hidden information (adverse selection) and hidden actions (moral hazard), limiting market efficiency. Situated in the context of the increasingly popular online platforms for labor contracting (herein referred to as "online labor markets"), this paper investigates whether the implementation of an IT-enabled monitoring system mitigates moral hazard in online platforms by providing direct information on workers' effort. Our identification hinges on a natural experiment at *Freelancer* when it first introduced an IT-enabled monitoring system for time-based projects but not for fixed-price projects in February 2014. Based on a unique dataset comprising 6,464 fixed-price projects and 3,120 time-based projects matched on observable characteristics, we employ a difference-in-differences (DID) approach to identify the treatment effect of the monitoring system implementation on various outcomes from both the employer (demand) side and the worker (supply) side, including employers' worker choice, workers' entry decisions and reputation premium. We observe that the implementation of the monitoring system lowers the employers' preference for bidders with a high effort-related reputation in time-based projects, and thus reduces reputation premiums and lowers the entry barrier for workers who have not yet established a reputation on the platform. However, there is no significant change in employers' preference for bidders with a high capability-related reputation. Further, using fixed-price projects as the control group, the implementation of the monitoring system increased the number of bids on time-based projects by 23.7% (primarily from bidders with no prior experience on the platform) and reduced the transaction price in time-based projects by 6.9%. Our results indicate a partial substitution relationship between reputation systems and monitoring systems, and suggest that IT-enabled monitoring systems have a significant effect on alleviating moral hazard, reducing agency costs, and intensifying supply-side platform competition.

Keywords: platforms, online labor market, moral hazard, monitoring systems, reputation systems, entry barrier, contract type

1. Introduction

By developing unique digital infrastructures that tap into underused physical and human resources, the sharing economy is transforming business processes and individual activities (Parker et al. 2016). Online labor markets, two-sided platforms that connect employers with freelance workers, are at the forefront of this phenomenon. Over the past decade, online labor markets have experienced tremendous growth. As a prominent example, by July 2016, about 20 million registered users have posted (employers) or bid (workers) for millions of projects at Freelancer,[1] one of the major online labor markets.

Despite the tremendous growth, online labor markets are plagued with agency problems, i.e., adverse selection due to hidden information and moral hazard due to hidden action and, both amplified by spatial and temporal separations between the employers and the workers (Hong and Pavlou 2017; Horton 2017). Given the substantial information asymmetry between employers and workers, online labor markets are subject to adverse selection, wherein employers could not separate high-quality workers from low-quality workers (Lin et al. 2016). At the same time, compared with a traditional employment, monitoring and control mechanisms to ensure work performance (Srivastava and Teo 2012) in online labor markets are much weaker and mostly indirect (Horton and Golden 2015). Therefore, opportunistic workers may exploit this limitation by misrepresenting their qualification or over-reporting their effort.

To mitigate information asymmetry, major online platforms have developed reputation systems that allow buyers (or employers) to share their experiences with sellers (or workers) (e.g., Dellarocas 2006; Moreno and Terwiesch 2014). Thanks to reputation systems, online platforms from auction sites such as eBay and Taobao, to online labor markets such as Freelancer and Upwork, have experienced substantial growth despite the prediction that information asymmetry will inevitably lead to market failures (Pauly 1974; Bockstedt and Goh 2011; Dellarocas 2005, 2006; Ba and Pavlou 2002). Reputation systems allow employers to share their experiences about workers, which help other employers screen for capable and

[1] https://www.freelancer.com/community/articles/20-million-users-things-that-made-this-milestone-remarkable-for-freelancer-com

trustworthy workers who are willing to expend commensurate effort for projects, thus mitigating both adverse selection and moral hazard. Given that adverse selection arises in online labor markets because workers have private information about their capabilities, reputation information regarding workers' capabilities lowers the likelihood that workers could misrepresent their capability to win contracts. In addition, by enabling employers to share information on workers' effort in previous projects, effort-related reputation serves as a sanctioning device that deters workers' shirking behavior when employers could not observe workers' actual effort (Banker and Hwang 2008). Taking reputation ratings as effective signals, employers construct their beliefs about the capability and effort of workers for differentiation. As a result, workers with high capability-related and effort-related ratings enjoy higher winning probabilities and price premiums (Ba and Pavlou 2002; Moreno and Terwiesch 2014). One unintended consequence of reputation systems, however, is that they create an entry barrier for qualified workers who have not yet established a reputation on a particular platform (Pallais 2014).

Another strategy used by employers to mitigate information asymmetry is through contract design. Two contract forms are available in online labor markets: time-based contracts and fixed-price contracts. In time-based contracts, compensation is determined based on the hourly wage set in the contract and the number of hours the worker has spent on the contracted project (Mani et al. 2012). While time-based contracts provide a stronger incentive to achieve better project performance (Dey et al. 2010; Mani et al. 2012), they are more susceptible to moral hazard as workers' compensation is not directly linked to the project outcome and the risk in time-based contracts is mainly allocated to employers (Dey et al. 2010; Mani et al. 2012). In fixed-price contracts, workers' compensation is based on the outcome of a project, such that the worker receives payment only when the project has been successfully completed (Mani et al. 2012). Therefore, fixed-price contracts have the potential to mitigate issues due to information asymmetry. However, fixed-price contracts involve high contracting costs and suffer from a lack of flexibility in accommodating changing requirements, leading to high ex-post costs of maladaptation and renegotiation (Benaroch et al. 2016). Taken together, both time-based contracts and fixed-price contracts are limited in their efficacy in fully resolving information asymmetry in online labor markets.

With the advance in information technology and improvement in network bandwidth, a new mechanism to mitigate information asymmetry – online monitoring – has gained popularity among online platforms (Aron et al. 2007; Agrawal et al. 2014). Specifically, with the use of a suite of IT-enabled monitoring technologies, employers can observe workers' working progress through screenshots and webcams, and even keystroke recordings from automatically archived log files, which serve as the first-hand information about workers' effort and help to alleviate employers' concerns about moral hazard from workers. However, these log files and tracked work hours only happens after the contract is written, therefore, they are not informative in pre-contractual screening on worker capability and could not alleviate hidden information. In summary, monitoring is more effective in mitigating hidden action than alleviating hidden information.

While a significant amount of research effort has been afforded to the design, evaluation, and optimization of reputation systems (Banker and Hwang 2008; Bockstedt and Goh 2011; Dellarocas 2006; Yoganarasimhan 2013), few studies have considered the role of monitoring, particularly its interactions with reputation signals and its implications for competition in online platforms. In this paper, we address the following three research questions:

- *How does IT-enabled monitoring moderate the effect of worker reputation on employer contracting decisions? Does the moderation effect vary for effort-related versus capability-relate reputation?*
- *How does IT-enabled monitoring influence entry decisions by workers? Does it affect less experienced workers differently from more experienced workers?*
- *How IT-enabled monitoring influence market equilibrium price and effective reputation premium in online labor markets?*

To answer the above research questions, our analyses leverage a natural experiment when Freelancer first introduced an IT-enabled monitoring system on February 5th, 2014. This natural experiment offers us an appropriate research design to identify the effects of IT-enabled monitoring systems in online labor markets. Our econometric identification hinges on the fact that monitoring was only implemented for time-based projects but not for fixed-price projects, which allows us to use time-based projects as the treatment group and fixed-price projects as the control group. Using a dataset including 12,467 projects posted at Freelancer, we first performed propensity score matching (and also coarsened exact matching) to match fixed-price projects to time-based projects. The resulting matched sample of 6,464 fixed-price projects and 3,120 time-based projects are comparable in terms of any observable characteristic. We then use a difference-in-differences (DID) approach to identify the treatment effect of the implementation of the monitoring system on employers' worker choice, worker entry decisions and market equilibrium price and implied reputation premium. Our analyses suggest that after the implementation of the IT-enabled monitoring system, employers place less weight on workers' effort-related reputation information, but remain unchanged in their weight on capability-related reputation information. Further, using fixed-price projects as the baseline, the monitoring system implementation increases number of entries in time-based projects by 23.7% on average, primarily from bidders with no prior experience on the platform. Finally, the market equilibrium price in time-based projects drops by 6.9%, likely because of the reduced reputation premium and increased competition on the supply-side for time-based projects.

Our study contributes to the literature on IT-enabled monitoring on three fronts. First, most prior studies focus on the performance effect of monitoring in offline contexts (Duflo et al. 2012; Hubbard 2000; Pierce et al. 2015; Staats et al. 2016; Ranganathan and Benson 2017), whereas this study focuses on the impact of an IT-enabled monitoring artifact on both demand-side (employer) preference and supply-side (worker) competition in online platforms. Second, our study advances prior literature on the relationship between monitoring systems and reputation systems in online platforms by showing that the implementation of monitoring systems only reduces employers' preference for workers with high effort-related reputation but not those with high capability-related reputation (Demiroglu and James 2010; Diamond 1991; Lin et al. 2016). Building on the recent work that found high reputation cannot significantly increase the prospect of winning time-based projects in an online labor market with a monitoring system (Lin et al. 2016), our study's setting allows us to identify the causal effects of the implementation of the monitoring system on both the supply and demand sides of an online labor market, and further to disentangle effort-related reputation from capability-related reputation. Third, numerous studies have attested to the positive role of reputation systems in addressing information asymmetry (Dellarocas 2005, 2006; Kokkodis and Ipeirotis 2015; Moreno and Terwiesch 2014; Pallais 2014; Yoganarasimhan 2013), this study demonstrates the unintended consequence of reputation systems in creating an entry barrier for workers who have not established a reputation on a focal platform, and how the implementation of an IT-enabled monitoring system can reduce such entry barriers for time-based projects by reducing the need for effort-related reputation screening.

2. Literature Review

Platform-based businesses are thriving in today's economy (Anderson et al. 2013; Parker et al. 2016; Eisenmann et al. 2006, 2011; Van Alstyne et al. 2016). Numerous new business models have been developed based on platform paradigm, ranging from platforms that enable transaction of physical products (e.g., eBay, Taobao, Amazon Marketplace), skilled labor (e.g., Upwork, Freelancer) and local tasks (e.g., TaskRabbit, PostMates), to sharing of car rides (e.g., Uber, Lyft) and temporary lodging (e.g, AirBnB, CouchSurfing). Due to asymmetric information between the two sides of the platform, one major theme in platform research is to tackle the agency problems with reputation systems (e.g., Brown and Morgan 2006; Dellarocas 2005, 2006; Duflo et al. 2013; Kokkodis and Ipeirotis 2015; Zacharia et al. 2000). Studies show that reputation systems improve disclosure and serve as sanctioning devices that deter agents from "shirking" (Bockstedt and Goh 2011; Dellarocas 2005, 2006; Moreno and Terwiesch 2014; Pavlou and Dimoka 2006). Our study contributes to this stream of literature by showing that IT-enabled monitoring systems are also effective IT artifacts that alleviate information asymmetry between the two sides of the platform. Further, we elaborate that there is a nuanced substitutive relationship between IT-enabled monitoring and reputation, the most prevalent mechanism to alleviate information asymmetry on online platforms.

Online labor markets are prime examples of platforms subject to moral hazard problems due to spatial and temporal separations of employers and workers. In order to mitigate moral hazard in online labor markets, a stream of research finds that reputation systems provide pre-contractual screening (e.g., Banker and Hwang 2008; Yoganarasimhan 2013), and serve as a sanctioning device to deter the shirking behavior (e.g., Dellarocas 2006; Moreno and Terwiesch 2014). A number of studies consider the effect of reputation on employers' worker choice (e.g. Banker and Hwang 2008; Pallais 2014). They found that reputation scores posted by previous employers can help workers obtain significantly better employment (Pallais 2014). With a better reputation, workers can obtain price premiums (Bajari and Hortacsu 2004), receive more employment opportunities, and are less likely to exit the platform (Moreno and Terwiesch 2014). Reputation in online labor markets is also shown to be transferable, such that previous ratings in related tasks can also indicate workers' category-specific quality in other similar projects (Kokkodis and Ipeirotis 2015). In summary, a great number of studies suggest that reputation systems effectively address adverse selection problems. We seek to advance this stream of literature by disentangling reputation information into capability-related reputation and effort-related reputation, and examining the different substitutive relationships between these two types of reputation and monitoring systems.

Prior studies have also analyzed monitoring systems but mainly in offline employment contexts. (Duflo et al. 2012; Hubbard 2000; Pierce et al. 2015; Staats et al. 2016; Ranganathan and Benson 2017). Monitoring systems provide a means for employers to obtain direct information on the actions of workers, thus mitigating moral hazard by addressing the hidden action issue (Bolton and Dewatripont 2005). The effectiveness of monitoring in increasing workers' effort and subsequently leading to better performance has been shown in multiple offline employment contexts, such as, the trucking industry (Hubbard 2000), schools (Duflo et al. 2012), restaurants (Pierce et al. 2015), and hospitals (Staats et al. 2016). In online labor markets, monitoring enables employers to check the project progress directly with monitoring records, such as random screenshots of workers' screens. Our study contributes to this body of literature by empirically studying the effect of monitoring in the online spot labor market versus the traditional offline employment-based labor market.

Even though ample literature has found evidence that the monitoring system can mitigate moral hazard problems (Drago 1991; Duflo et al. 2012; Hubbard 2000; Pierce et al. 2015), most prior studies analyze monitoring systems and reputation systems separately, without considering how they jointly mitigate information asymmetry regarding hidden action and whether they are substitutes or complements to each other (Demiroglu and James 2010; Diamond 1991). Therefore, a key research gap in the literature is to disentangle the roles of monitoring systems and reputation systems in respectively addressing the ex-post information asymmetry (hidden action) from ex-ante information asymmetry (hidden information). Further, it is also important to theoretically differentiate two types of reputation: effort-related reputation and capability-related reputation, and recognize that monitoring can only substitute for effort-related reputation in mitigating moral hazard but not for capability-related reputation.

Lastly, related to this study, Lin et al. (2016) find evidence that the effectiveness of reputation depends on contract type, such that time-based contracts benefit less from reputation systems. This paper builds on Lin et al. (2016) to leverage a natural experiment for causal identification, and consider both the capability-related and effort-related reputation, and separate their effects on alleviating hidden information and hidden action. Further, our paper theoretically proposed and empirically evaluated an important unintended benefit of monitoring systems: lowering entry barrier for new users. As Pallais (2014) suggests, mainly relying on reputation systems to alleviate information asymmetry tends to lead to inefficient hiring in entry-level jobs, which is also referred as the “cold-start” problem. Our study suggests that the implementation of monitoring systems reduces the importance of effort-related reputation and has the potential to increase hiring efficiency.

3. Research Context and Hypotheses Development

3.1 Research Context

The research context of this study, online labor markets, are web-based two-sided platforms that facilitate the contracting of labor services around the globe (Chan and Wang 2017; Lin et al. 2016). In recent years, online labor markets have grown significantly. It is reported that 25 percentage of jobs in the

US are offshore outsourced (Blinder and Kruger 2013), with a substantial portion delegated through online labor markets.[2] Because of spatial and temporal separations between employers and workers, information asymmetry persists in such platforms (Hong and Pavlou 2017), as workers' quality is difficult to observe and their actual effort is difficult to monitor. Therefore, similar to many other online platforms, information asymmetry issues are prevalent in online labor markets, making the agency problem a major theme in the literature on online labor markets.

Two forms of information asymmetry exist on online platforms, namely, hidden information and hidden action. Hidden information refers to the scenario wherein workers possess ex-ante private information about their capability and skills (Bolton and Dewatripont 2005, Horton 2017). Hidden action relates to ex-post information asymmetry regarding workers' actual actions, such as amount of time and effort spent on projects. Due to spatial and temporal separations between employers and workers, online labor markets are plagued with both hidden information and hidden action, which lead to adverse selection, moral hazard and subsequently market inefficiency.

First, adverse selection occurs due to hidden information. Adverse selection is driven by the asymmetric information and the difficulties in evaluating workers' capabilities and skills (Eisenhardt 1989). In order to address adverse selection problems, most online labor markets provide a reputation system, which aggregates workers' ratings supplied by their prior employers. Second, once a worker is awarded the contract, the moral hazard problem follows due to hidden actions. Moral hazard occurs when workers opportunistically misrepresent their effort level to maximize their own utility, to the detriment of employers (Eisenhardt 1989). Moral hazard is caused by the employer's inability to observe the worker's actual effort level and the misalignment between the employer's and worker's interests.

One way to partially address information asymmetry is through contract design. Most online labor platforms offer two contract types, fixed-price contracts and time-based contracts to employers (Banerjee

[2] http://www.forbes.com/sites/groupthink/2014/10/21/the-next-big-thing-in-e-commerce-online-labor-marketplaces

and Duflo 2000).[3] Fixed-price contracts are outcome-driven, such that the worker receives a fixed payment based on the output (Chen and Bharadwaj 2009; Mani et al. 2012). On the other hand, time-based contracts, also known as cost-plus contracts, require that the payment is calculated based on the amount of time the worker spent in the work process (Clemons and Chen 2011; Mani et al. 2012). Typically, fixed-price contracts have higher ex-ante costs to collect information and to negotiate the provision, plus higher ex-post maladaptation costs and renegotiation costs (Susarla et al. 2009; Susarla and Krishnan 2014). In comparison, time-based contracts entail higher ex-post monitoring and auditing costs (Bajari and Tadelis 2001; Dey et al. 2010; Susarla and Krishnan 2014).

In terms of incentive alignment, for fixed-price contracts, workers are contracted for the final output of the projects, which incentivizes them to efficiently finish the projects. Thus, both adverse selection and moral hazard problems in such contracts are less severe (Fama 1991). For time-based contracts, workers' payments are based on the amount of time they have spent on the projects. In this case, on the one hand, workers might misrepresent their capability and effort as well as engage in other opportunistic behaviors, since the employer bears all the project risks (Bolton and Dewatripont 2005). On the other hand, time-based contracts could provide more flexibility and yield better performance and higher client validation quality than fixed-price projects if the monitoring and auditing process is efficient (Dey et al. 2010). Apart from these two contract types, there are other optional contract types, including performance-based contracts, profit-sharing contracts (Dey et al. 2010), and hybrid contracts (Banerjee and Duflo 2000). Our paper focuses on the comparison between fixed-price and time-based contracts as these are the only two contract type options in major online labor markets.

3.2. Hypothesis Development

We propose three hypotheses. First, we propose a nuanced substitution effect between monitoring and reputation. Second, we propose the role of monitoring systems in lowering entry-barriers for

[3] In Banerjee and Duflo (2000), the counterpart of fixed-price contracts are termed "time and materials contracts." Since the costs are measured by the agent's efforts and time in the projects, time and materials contracts herein refer to "time-based contracts" in online labor platforms.

inexperienced workers and intensifying supply side competition for time-based projects. Finally, we propose the effect of monitoring systems in reducing market equilibrium price on the online labor market.

Nuanced Relationship between Monitoring and Reputation

IT-enabled monitoring and effort-related reputation help employers to alleviate information asymmetry via different mechanisms. IT-enabled monitoring converts workers' private information about their actual effort into procedural and tractable records that employers could observe (Lin et al. 2016). Therefore, it decreases the likelihood of "shirking" going unnoticed, and in turn increases workers' effort (Drago 1991). This mechanism determines that monitoring could be very effective to address post-contractual moral hazard, but is less effective at addressing pre-contractual hidden information.

Reputation systems, on the other hand, provide signals of workers' capability and effort based on their past performance evaluated by previous employers (Banker and Hwang 2008). This allows employers to construct beliefs about both workers' expected capability and their expected effort, and thus the expected value of their work. Reputation thus serves as a sanctioning mechanism by imposing a potential penalty on both hidden information and "shirking" behavior by reducing the probability of landing contracts in the future (Dellarocas 2006; Moreno and Terwiesch 2014).

In summary, based on the previous literature, monitoring systems and reputation systems are two prevalent mechanisms in alleviating information asymmetry. Specifically, monitoring systems are mainly found to effectively mitigate moral hazard and hidden action in multiple offline employment contexts (Duflo et al. 2012; Hubbard 2000; Pierce et al. 2015; Ranganathan and Benson 2017; Staats et al. 2016). Meanwhile, reputation systems not only help mitigate moral hazard and hidden action by deterring the "shirking" behavior but also help alleviate hidden information by providing pre-contractual screening. Given differential effects of monitoring systems and reputation systems, instead of exploring the effect of general reputation, we segment reputation into two types: capability-related reputation which helps to alleviate hidden information and effort-related reputation which is effective in mitigating moral hazard and hidden action. Additionally, we propose the nuanced substitution relationship between reputation and monitoring: monitoring can substitute for effort-related reputation by alleviating hidden action, but not for capability-related reputation.

Table 1. A Comparison between Reputation and Monitoring in Alleviating Information Asymmetry

Mechanisms	Hidden Information	Hidden Action
Reputation systems	Provide pre-contractual screening on online service markets (e.g., Banker and Hwang 2008)	Deter the “shirking” behavior on online trading and service markets (e.g., Dellarocas 2006)
Monitoring systems	Not applicable	Mitigate moral hazard in multiple offline employment contexts (e.g., Duflo et al. 2012; Staats et al. 2016)

Given that monitoring systems enable employers to directly observe workers’ effort, we argue that they reduce employers’ reliance on effort-related reputation to deter “shirking” behavior. Without monitoring systems, employers have to rely on workers’ past effort-related reputation to extrapolate the expected effort level of workers for the focal project (Kokkodis and Ipeirotis 2015). However, such extrapolation could be inaccurate for a number of reasons. First, effort-related reputation is based on accumulated reports written by previous employers. As a worker’s true effort level is not observed, the reporting could be inaccurate (Pallais 2014). Second, even if the reporting of past experience were accurate, past effort level does not guarantee future effort level on a particular project. For example, workers may have limited capacity when they concurrently work on multiple projects (Horton 2017), or they may become more strategic or opportunistic over time as they accrue more experience on the platform. With the IT-enabled monitoring system, employers can verify a worker’s actual effort level based on direct observations of monitoring records and, if needed, condition continued employment based on the observed level of effort. In sum, the monitoring system enables a more direct and precise observation of worker effort than what can be inferred from workers’ prior work histories. Therefore, monitoring systems at least partially substitute for effort-related reputation.

Second, the monitoring system reduces cost uncertainty. Monitoring systems allow employers to observe workers’ current performance and progress, thereby facilitating meaningful interactions between employers and workers. Without the monitoring system, employers prefer workers with the high effort-related reputation not just because of their higher expected effort, but because of lower cost uncertainty as well (Mani et al. 2012). Since workers with high effort-related reputation are expected to be more responsible and trustworthy, they are less likely to cause budget overrun in time-based projects. With monitoring systems, employers are informed in real time about workers' performance (e.g. timely update of project progress, workflow, etc.). Employers can improve workers' performance by providing feedback based on the automatically archived log files. Therefore, irrespective of the workers' effort-related reputation, employers can reduce cost uncertainty by monitoring worker performance, which reduces the disparities between the workers with high effort-related reputation and those with low effort-related reputation. In summary, the monitoring system lowers employers' concerns about workers' "shirking" behavior and cost uncertainty, and thus substitutes for the signaling effect of effort-related reputation.

We also argue that the monitoring records and accumulated work hours generated by monitoring systems could not effectively alleviate hidden information, such that monitoring could not substitute for the capability-related reputation. First, monitoring can after the hiring decision. Therefore, despite a monitoring system in place, employers still need to rely on workers' capability-related reputation to infer their capability in making the hiring decision. Second, the disparities between the workers with high capability-related reputation and those with low capability-related reputation could not be reduced by monitoring systems. After the monitoring system is implemented, with same effort, the work quality of the low capable worker would still be inferior to that of the highly capable workers. Therefore, the implementation of the monitoring system would not significantly reduce employers' preference for workers with high capability-related reputation, given that monitoring could not effectively alleviate hidden information. Therefore, we propose:

H1: *Implementation of an IT-enabled monitoring system leads employers to place less emphasis on workers' effort-related reputation for time-based projects, but not on capability-related reputation.*

Monitoring, Entry Barrier and Worker Competition

Now we consider the effect of IT-enabled monitoring systems on the worker (supply) side. Before the implementation of monitoring systems, employers rely on reputation systems to provide pre-contractual screening and to mitigate moral hazard. This leads to significant advantage for workers who entered the platform early, as they were able to accrue platform-specific experience and establish a reputation. Given the employers' preference for workers with platform reputation (Pallais 2014), inexperienced workers, who have not established their reputation yet, are less likely to bid for either fixed-price or time-based contracts because their likelihood of landing a contract is slim. The same argument applies to workers who have accumulated relatively low effort-related reputation. Consequently, the high entry barrier due to accumulated capability-related reputation and effort-related reputation discourages workers (especially inexperienced workers and workers with low effort-related reputation) to participate in the market. However, after the implementation of the IT-enabled monitoring system, employers can observe workers' effort from the procedural track records rather than rely on workers' effort-related reputation. Therefore, although the entry barrier due to capability-related reputation is not affected by the implementation of the monitoring system, the barrier to entry due to accumulated effort-related reputation decreases (Demiroglu and James 2010) in time-based contracts relative to fixed-price contracts. Therefore, more workers are likely to bid for those time-based contracts. Therefore, we propose:

H2a: *Implementation of an IT-enabled monitoring system leads more bidders to bid for time-based projects.*

Bidders tend to be heterogeneous in online markets (Lu et al. 2016). The reduction in entry barrier mentioned above is especially strong for inexperienced workers. Consistent with the logic of the substitution relationship between monitoring systems and effort-related reputation, we argue that as monitoring systems alleviate employers' concerns on workers' moral hazard, they are more likely to hire inexperienced bidders. As the difference between the workers with little platform experience and those with extensive experience narrows in terms of their probability to shirk and their expected effort levels, employers will be more neutral about workers' platform experience. Prior to the implementation of the monitoring system, effort-related reputation serves as an important signal that facilitates the differentiation of workers with the high-effort tendency (Brynjolfsson et al. 2004; Pallais 2014). With the monitoring system implemented that allows direct observation of workers' effort, the signaling value of effort-related reputation decreases. Therefore, the lower entry barrier tends to disproportionately attract more bids from inexperienced bidders. We thus expect that the monitoring system will intensify supply-side competition by attracting a large number of inexperience workers who are qualified but have not yet established their effort-related reputation. Therefore, we propose:

H2b: *Implementation of an IT-enabled monitoring system leads to a higher percentage of workers with no platform experience bidding for time-based projects.*

Monitoring and Marketing Equilibrium Price

Without monitoring systems, reputation serves as an important signal to facilitate the differentiation of workers in terms of both capability and effort (Brynjolfsson et al. 2004; Pallais 2014). With monitoring systems, the signaling value of effort-related reputation decreases as employers have access to a more effective tool to mitigate moral hazard. Given the partial substitution effect between monitoring and effort-related reputation, employers will be less willing to pay premiums for effort-related reputation. Findings from Allgulin and Ellingsen (2002) support this argument, wherein authors find that when the monitoring system is precise, efficient and low-cost, the agent's utility reaches its minimum level and the agent becomes less capable of earning information rents regarding hidden action. When the agent can be monitored perfectly, any effort can only be paid at his corresponding reservation wage. Similarly, the Efficiency Wage Model, which predicts that more intense monitoring leads to lower wage premiums (Ewing and Payne 1999; Leonard 1987; Shapiro and Stiglitz 1984), also lends support to our argument. Furthermore, as the implementation of monitoring systems increases market entries from inexperienced workers, we expect market competition to increase, further eroding market salary for workers. Therefore, we propose:

H3: *Implementation of an IT-enabled monitoring system leads employers to pay a lower price for time-based projects.*

4. Data

4.1. Data Source

Our data is obtained from www.freelancer.com (Freelancer), one of the largest online labor market platforms. At Freelancer, an employer can post a project with a description, estimated budget and skills required. The employer can choose between two contract types for the project: fixed price contract (Figure 1-a) for which the employer provides the estimated budget for the entire project; or time-based contract (Figure 1-b) for which the employer provides the estimated hourly budget for the project in dollars per hour.

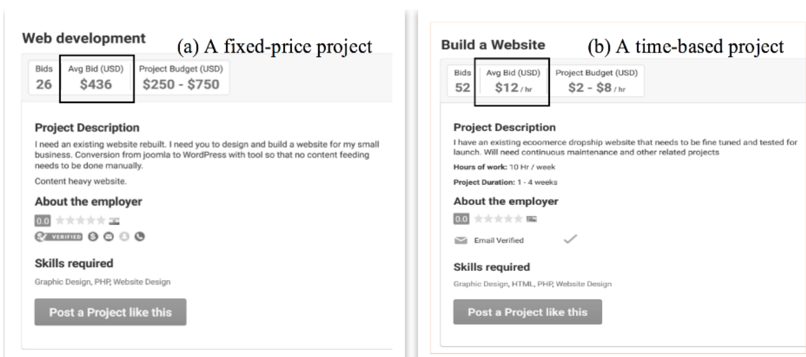

Figure 1. Screenshots of the Web pages of a Fixed-price versus a Time-based Project

Typically, a project is open for bidding for a week and any worker who is interested can bid on the project. For fixed price projects, each bidder submits his bid amount for the entire project, while for time-based projects, each bidder submits his bid in terms of hourly rate. At the end of the bidding period, the employer reviews bidders' information, including their skills, bid amount, past project experience and their former employers' ratings. Additionally, filtering tools are available to enable the employer to sort or filter bidders according to the number of reviews and average rating. Once the employer finds the bidder who satisfies his or her requirements, the employer could award that worker a contract with details.

4.2. Sample and Variables

We obtained a unique archival dataset from Freelancer that includes all project information and worker information from September 1st, 2013 to August 31st, 2014. We follow Lin et al. (2016)'s approach to construct our sample. In particular, we limited our sample to awarded projects that reflect realistic labor demand without the contamination of resubmitted projects. Further, to reduce possible selection bias and the relationship between various pretreatment covariates and contract choices, we matched fixed-price and time-based projects (Abadie 2005; Ho et al. 2007) based on distributions of important covariates suggested by the previous literature (details reported in Table 5) using propensity score matching (PSM). Our final sample includes 6,464 fixed-price projects and 3,120 time-based projects. The descriptive statistics of our dataset are shown in Table 2, Table 3 and Table 4. The dataset includes the following attributes: 1) project-level information (e.g., project description, project budget, contract type, number of bidders, average bid price, etc.); 2) worker-level information (i.e. ratings, the amount of reviews, average hourly wage, etc.); 3) bid-level information (i.e. bid price, etc.).

Table 2. Definitions and Summary Statistics of Project-level Variables

Variable	Variable definition	Mean	SD	1st perc	99th perc
Budget_min	The minimum of project budget set by the employer	50.81	136.62	2.00	750.00
Budget_max	The maximum of project budget set by the employer	8169.17	895608.90	2.00	1181.00
Bid_min	The minimum of bid prices for each project	64.08	129.29	2.00	750.00
Bid_max[4]	The maximum of bid prices for each project	18656.99	1990642.00	5.00	2631.00
Time-based	A dummy variable; =1 if the project is a time-based project; =0 if the project is a fixed-price project	0.25	0.43	0.00	1.00
Bid_count	Total number of bids received by the project	12.64	13.74	1.00	64.00
Bid_mean[5]	Average bid price for each project	6087.56	663420.60	3.14	1081.50

[4] The large variation in Bid_max is driven by outliers. In rare cases, workers asked for unreasonably high prices.

[5] Here, Bid_Mean refers to the mean among all the bids. But it is disproportionately influenced by the maximum of the bid price, we recalculate the average bid price by dropping both the maximum and minimum of bid prices before including it into our DID models.

Paid_amount	Amount of dollars paid by the employers after the project was completed	130.47	242.56	4.00	1464.00
Project_title_length	Number of characters in the project title	5.60	3.16	2.00	16.00
Project_desc_length	Number of characters in the project description	16.33	4.29	3.00	23.00

Table 3. Definitions and Summary Statistics of Worker-level Variables

Variable	Variable definition	Mean	SD	1st perc	99th perc
User_developed	A dummy variable; =1 if the worker comes from a developed country	0.19	0.39	0.00	1.00
Quality	Average quality rating given by all the employers (ranging from 0 to 5)	4.83	0.40	3.00	5.00
Communication	Average communication rating given by all the employers (ranging from 0 to 5)	4.83	0.40	3.00	5.00
Expertise	Average expertise rating given by all the employers (ranging from 0 to 5)	4.83	0.40	3.00	5.00
Professionalism	Average professionalism rating given by all the employers (ranging from 0 to 5)	4.85	0.39	3.00	5.00
Hire-again rating	Average hire-again rating given by all the employers (ranging from 0 to 5)	4.83	0.43	2.84	5.00
Overall	Average overall employer-entered ratings for the worker	4.83	0.38	3.00	5.00
Review_count	Total number of reviews which were written by previous employers	54.52	148.52	1.00	592.00
Completion_rate	Percentage of awarded projects which were successfully completed	0.79	0.20	0.21	1.00

Table 4. Definitions and Summary Statistics of Bid-level Variables

Variable	Variable definition	Mean	SD	1st perc	99th perc
Bid_Price	Bid price submitted by the worker	1895.33	623022.10	2.00	1526.00
Hire_before	A dummy variable; =1 if the worker has been hired by the employer before	0.23	0.42	0.00	1.00
No_rating	A dummy variable; =1 if the worker has not received any ratings when he/she submitted the bid	0.12	0.32	0.00	1.00
Bidder_tenure_month	The worker's tenure at *Freelancer* measured in months	34.72	25.89	3.00	115.00
Bid_rank	The bidder's ranking among all the candidates, *Freelancer* automatically sorts all the bidders according to its own ranking algorithm which is mainly based bidders' employer-entered reviews	14.61	14.64	1.00	68.00
Bid_order_rank	The sequence in which the bidders' bids were submitted	14.93	14.95	1.00	70.00
Preferred_freelancer	A dummy variable; =1 if the worker gets a special Preferred *Freelancer* Badge because their	0.18	0.39	0.00	1.00

	workmanship and customer service abilities				
Local_freelancer	A dummy variable; =1 if the worker works for offline jobs nearby	0.02	0.13	0.00	1.00

5. Research Methodology

5.1. Identification: A Quasi-Natural Experiment

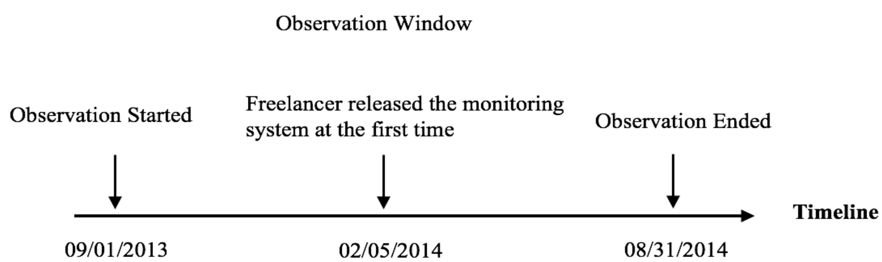

Figure 2. A Timeline of Our Observation Window

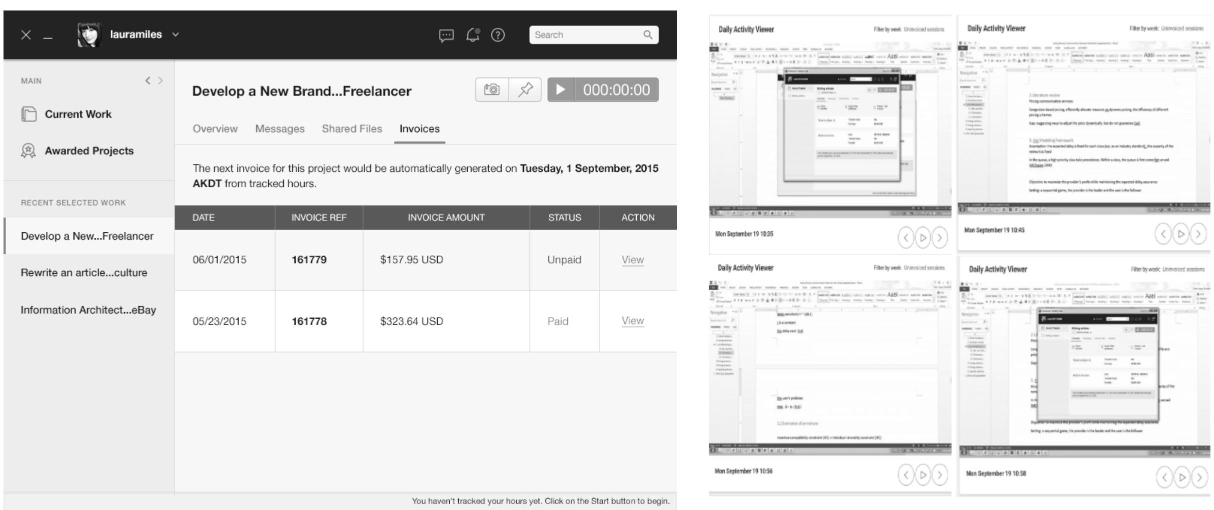

Figure 3. Screenshots of the Freelancer Monitoring System

On February 5th, 2014, Freelancer released its monitoring system for the first time and the feature was only available for employers and workers in time-based contracts. The monitoring system is a software application that allows employers to effortlessly monitor freelancers. Freelancer requires all workers with time-based contracts to download and install the application. Once installed,[6] this monitoring system

[6] If workers who work for time-based projects do not use this monitoring application to track their work hours, they would not get paid for their work. Therefore, time-based workers are required to install this monitoring system after

randomly takes several screenshots about every ten minutes, and tracks the number of minutes the worker has spent on each time-based project.[7] Specifically, it automatically tracks when and for how long the worker has worked, the accumulated compensation the worker has earned, and the corresponding screenshots with timestamps. Therefore, it effectively keeps a detailed record of workers' effort, providing the employer up to date information on the projects' progress. The employer can file a dispute with Freelancer regarding the worker's effort and claimed hours based on records kept by the monitoring system. Figure 3 is a screenshot of the monitoring application provided by Freelancer.

The fact that the use of the monitoring system is mandatory for all time-based projects and the system is not usable for fixed-price projects provides us a unique research opportunity. In this study, we leverage a difference-in-differences (DID) design with fixed-price projects as the control group to examine the effect of the IT-enabled monitoring system on time-based projects relatively to fixed-price projects. In doing so, we compare employers' and workers' choices and market equilibrium prices across the two types of projects before and after the introduction of the monitoring system. The DID model is a widely used causal model, which been extensively used in IS research when exogenous changes are available (e.g., Chan and Ghose 2014, Huang et al. 2017; Wang et al. 2017; Zhang and Li 2017).

5.2. Econometric Analyses

Propensity Score Matching (PSM)

In order to satisfy the common support requirement and reduce the potential heterogeneity across time-based projects and fixed-price projects, we use the propensity score matching (PSM) method to generate a comparable sample. The PSM approach for matching has been widely applied in the information systems literature (Hong et al. 2016; Lin et al. 2016; Xu et al. 2016; Xue et al. 2010). First, according to prior

---

it was rolled out. Moreover, it is noted that no matter whether employers install this monitoring application or not, they could always check the monitoring records from the Freelancer website.

[7] The application does not track time spent on fixed-price projects, because workers could only find time-based contracts but not fixed-price contracts within the application. Therefore, they could not use this application to take any screenshots or track the time for fixed-price projects.

literature (Banerjee and Duflo 2000; Gopal and Sivaramakrishnan 2008; Lin et al. 2016; Roels et al. 2010), we identify project characteristics and employer characteristics that might correlate with the contract type (Table 5). Then we predict the propensity scores and match fixed-price projects with time-based projects. Furthermore, we also compare the distribution of propensity score and carry out balance checks for all observed covariates (Xu et al. 2016). As Table A1 in Appendix A shows, the matching process has significantly reduced the difference between the control and treatment groups and the means of all covariates are statistically the same across the two types of projects after the matching process. Based on the full sample with 12,467 projects posted on Freelancer, we generate our final matched sample which includes 6,464 fixed-price projects and 3,120 time-based projects.

Table 5. Pre-treatment Covariates Used to Adjust for Potential Selection Bias

Covariates	Variable	Variable Description
Task complexity, risk of project (Gopal and Sivaramakrishnan 2008)	Project category dummies	Dummy variables for various project categories, including software, design, marketing, data-entry, etc.
Project Title Length (Lin et al. 2016)	Project_title_length	Number of characters in the project title
Project Description Length (Lin et al. 2016)	Project_desc_length	Number of characters in the project description
Project Size (Lin et al. 2016)	Paid_amount	The amount of dollars paid by the employers after the project was completed
External environment (Banerjee and Duflo 2000)	Project_submit_month	The month dummy, which is used to control all the changes in economic cycles, platform policy, labor supply, and so on
Client Level of Knowledge (Lin et al. 2016)	Employer_tenure_month; Employer_overall_rating	Employer's tenure at Freelancer measured in months, which is also a proxy of employers' experience and relevant knowledge. Employers' overall rating.

Principal Component Analysis for Dimension Reduction

Freelancer has a multi-dimensional reputation system during our observational period, which presents multiple indicators prominently shown when the cursor is on the worker profile. As high correlations were observed among some rating dimensions, we employed the Principal Component Analysis (PCA) for dimension reduction, which generated two principal components by using ~1 as the cutoff for eigenvalues and 80% as the threshold of the cumulative variance explained, as shown in Table 7. The first component (PC1) comprises dimensions of ratings entered by employers after the transactions, which largely indicates a worker's capability. The second component (PC2) is workers' project completion rate, which was computed by the system based on the percentage of projects completed for all contracted projects. This component largely indicates a worker's effort at work because project incompletion is typically due to insufficient efforts. Further, the five items with significant loading on the first component help to mitigate ex-ante information asymmetry (hidden information), while the one item with significant loading on the second component helps to alleviate ex-post information asymmetry (hidden action). Therefore, we label PC1 as "Capability" and PC2 as "Effort". This label assignment was further confirmed by interviewing a number of Freelancer employers on how they perceive the reputation signals of workers. We also conducted a survey of employers with the help of Freelancer, which confirms our PCA results that employers are generally concerned with freelancers' capability and their service effort. We report the item loadings and eigenvalues/variance explained in Table 6 and Table 7, respectively.

Table 6. Item Loadings of Two Principal Components with Rotations

Variable	**Eigenvectors**	
	1	**2**
Quality	0.449	-0.023
Communication	0.432	-0.011
Expertise	0.451	-0.023
Professionalism	0.452	-0.019
Hire-again rating	0.450	-0.022
Completion Rate	0.043	0.999

Table 7. Eigenvalues and Variance Explained by Two Principal Components

Label	**Component**	**Eigenvalue**	**Diff**	**Proportion of variance explained**	**Cumulative variance explained**
Capability	1	4.374	3.381	0.729	0.729
Effort	2	0.994	0.741	0.166	0.895

Estimating Employer Preference

To estimate employer preference for workers (H1) in terms of their observable characteristics, we formulate a model for each worker's winning probability within each project. Specifically, we estimate the probability of bidder k being awarded in project j as P_{jk}. Denote U_{jk} for the employer's utility from hiring bidder *k* for project *j*.

$$U_{jk} = \beta_1 t_j C_{jk} + \beta_2 T_j C_{jk} + \beta_3 t_j T_j C_{jk} + \beta_4 t_j E_{jk} + \beta_5 T_j E_{jk} + \beta_6 t_j T_j E_{jk} + \beta_7 t_j P_{jk} + \beta_8 T_j P_{jk} + \beta_9 t_j T_j P_{jk} + \gamma B_k + \delta Z_{jk} + \varepsilon_{jk} \quad (1)$$

In Equation (1), t_j is the period dummy variable, which is set to 1 if project j is posted after the implementation of the monitoring system. T_j is the contract type dummy variable, which is set to 1 if project j is a time-based project. C_{jk} denotes bidder k's reputation related to his or her capability based on the principal component analysis. E_{jk} denotes bidder k's reputation related to his or her effort based on principal component analysis. P_{jk} denotes the bid price submitted by bidder k. Z_{jk} represents a set of other project-bidder paired characteristics, including the bidder k's ranking based on their reputation and experience among all the competitors, the order of bidder k's bid based on the sequence of all the bids were submitted, whether the bidder k has worked for this employer before. B_k captures bidder k's individual characteristics, including whether bidder k has received any ratings or not, the number of ratings entered by bidder k's previous employers, whether he or she is from a developed country, bidder k's tenure at *Freelancer* measured in months, whether bidder k gets a special *Preferred Freelancer Badge*, and whether bidder k also works for local projects.[8] The employer's utility model could be estimated with a linear probability model (Heckman and Snyder 1997; Lin et al. 2016) or a logit model (Lin et al. 2016; Liu et al. 2015). In our main analyses, we assume that ε_{jk} follows Type-I extreme value distribution (Train 2009) and use a conditional logit model. We also estimated linear probability models and observed highly consistent results.

Difference-in-Differences Models

To assess workers' bidding activity (for H2) and market price equilibrium (for H3), we estimate standard difference-in-differences models (Bertrand et al. 2004; Angrist and Pischke 2008):

[8] Based on our review data, a worker's average rating is almost constant during our observational period. Therefore, we didn't treat the worker rating as a time-variant variable here.

$$Perc\_Norating_{ij} = \alpha + \beta_1 After_j + \beta_2 Time\_based_j + \beta_3\, Time\_based_j \times After_j + \gamma_i + \delta_j + \tau_t + \varepsilon_{ij} \quad (2)$$

$$Log\_Bid\_Count_{ij} = \alpha + \beta_1 After_j + \beta_2 Time\_based_j + \beta_3\, Time\_based_j \times After_j + \gamma_i + \delta_j + \tau_t + \varepsilon_{ij} \quad (3)$$

$$Award\_Price\_Premium_{ij} = \alpha + \beta_1 After_j + \beta_2 Time\_based_j + \beta_3\, Time\_based_j \times After_j + \gamma_i + \delta_j + \tau_t + \varepsilon_{ij} \quad (4)$$

In equation (2), the dependent variable $Perc\_Norating_{ij}$ denotes the percentage of inexperienced bidders (i.e. bidders without ratings) in project j posted by employer i. In equation (3), the dependent variable is the log value of the total number of bids for each project j posted by employer i, $Log\_Bid\_Count_{ij}$. $After_j$ is the dummy variable indicating whether project j is awarded after monitoring system implementation. The contract type is indicated by $Time\_based_j$, which equals to 1 if project j is a time-based project and 0 if it is a fixed-price project. The coefficient of the interaction term $Time\_based_j \times After_j$ (β_3) thus identifies the effect of the implementation of the IT-enabled monitoring system on time-based projects relative to fixed-price projects. To control for project heterogeneity, we also add other project characteristic controls (δ_j), a vector of employer fixed-effects (γ_i), and a vector of time fixed-effects (τ_t) into the DID model and ε_{ij} denotes the robust error term clustered on employers.

Equation (4) assesses the effect of monitoring mechanism on market equilibrium price (H3). For its dependent variable, we use a normalized measure of market price premium to control for unobserved heterogeneity across projects. The price premium is evaluated at the awarded bid price of project j, $Award\_Price\_Premium_j$, relative to the average bid price for the project. This normalization allows us to cancel out unique project characteristics that may influence bid prices.

$$Award\_Price\_Premium_j = \frac{(Award\_BidPrice_j - Bid\_Mean_j)}{Bid\_Mean_j} \qquad (6)^{9}$$

5.3. Empirical Results

Employer Preference Estimation

The results of the conditional logit model are reported in Table 9.[10] We observe that, before and after the IT-enabled monitoring system was implemented, the coefficients for the reputation of workers' capability, *Capability_of_worker*, remain unchanged at 0.188. The finding indicates that employers' preference towards the capability-related reputation does not change because of the monitoring system. Notably, we observe a different pattern regarding the coefficients for the other dimension of reputation, the reputation of workers' effort level, *Effort_at_work.* As Table 8 attested, for fixed-price projects, the employer preference remains at the similar level (β=0.593), while for time-based projects, the employer preference shows a relatively large decrease (from 0.876 to 0.608), indicating employers' preference for workers with high reputation on their effort level decrease after the introduction of the monitoring system. This significant decrease in employers' weight on workers' effort-related reputation and the insignificant change in employers' weight on workers' capability-related reputation suggests that the implementation of the monitoring system helps alleviate moral hazard problems and lower information rent acquired by workers with high effort-related reputation, but has limited effects on solving hidden information problems, as predicted in H1. We also find that employers' price-sensitivity for time-based projects increased (from -0.777 to -1.715) after the implementation of the monitoring system while their price-sensitivity for fixed-price projects remains unchanged, further confirming that the alleviation of moral hazard problems makes employers more sensitivity to bid prices. Overall, the findings suggest that there

[9] Because Bid_Mean_j is disproportionately influenced by the maximum of the bid price, we recalculate the average bid price by dropping both the maximum and minimum of bid prices.

[10] Besides estimating the conditional logit model, we also estimate the coefficients with linear probability models with fixed effects and evaluate the significance levels and effect sizes based on the cluster-robust standard errors as the linear model helps to ensure consistency of the estimation results and provides a meaningful interpretation of coefficients for the interaction terms (Greenwood and Agarwal 2015). The result is highly consistent with that of the conditional logit model.

exists a partial substitution relation between monitoring and reputation such that monitoring substitutes for some of the effort-related reputation but does not substitute for the capability-related reputation.

Table 8. Estimation Results of the Conditional Logit Model

Variable	Bid_Selected
Capability_of_worker	0.188***(0.033)
Capability_of_worker* Time_based	0.069 (0.052)
Capability_of_worker*After	0.002 (0.048)
Capability_of_worker* Time_based *After	-0.094 (0.075)
Effort_at_work	0.593***(0.049)
Effort_at_work* Time_based	0.283***(0.075)
Effort_at_work* After	0.103 (0.077)
Effort_at_work* Time_based *After	-0.268** (0.119)
Log_Bid_Price	-1.815***(0.063)
Log_Bid_Price* Time_based	1.038***(0.096)
Log_Bid_Price*After	0.326***(0.090)
Log_Bid_Price*Time_based *After	-1.264***(0.165)
Hire_before	2.861***(0.070)
No_rating	-0.659***(0.079)
Log_bidder_rank	-0.385***(0.019)
Log_bid_order_rank	0.282***(0.021)
Preferred_freelancer	0.085** (0.034)
Local_freelancer	-0.535***(0.107)
Observations	87,482
Clusters(projects)	6,956
Log likelihood	-12031

Notes: a) We limit our sample to those projects with more than one bid and awarded to only one worker; and the results is based on all the workers who bid for both fixed-price and time-based projects (named as "dual-typed workers") (Lin et al. 2016). b) Since we do not have any capability-related or effort-related reputation information for those workers who have not received any ratings from employers, we add the No_rating dummy and set their capability-related and effort-related reputation component scores as zeros. We also estimate the model with only those workers with reputation and add the Review_count variable instead of the No_rating dummy. The results are highly consistent with our main model. v) Results are highly consistent when we estimate the treatment effect with linear probability models. d) * p<0.1, ** p<0.05, *** p<0.01

Bidding Behavior and Entry Barrier

As employers are less willing to pay high price premiums to bidders with high effort-related reputation when the monitoring system is in place, we expect that the barrier for new or inexperienced workers to enter the time-based project becomes lower, which leads to more bids and higher competition among bidders for a given time-based project. Columns (1) and (2) of Table 9 reports the DID regression results on the effect of the monitoring system on the percentage of new bidders and the total number of bids for time-based projects relative to fixed-price projects at the project level.

Column (1) of Table 9 reports the regression analysis on the effect on number of bidders. We first note that the coefficient is insignificant for *After*, indicating that the implementation of the monitoring system has no significant impact on the number of bidders for fixed-price projects. At the same time, the coefficient (β_3) of the interaction term $After_j \times Time\_based_j$ is significantly positive, which suggests that the implementation of the monitoring system significantly boost the number of bids (Bid_Count_{ij}) for time-based projects. The finding that the interaction term is 0.213 indicates that the number of bids increases 23.7%.[11] Therefore, H2a is supported.

We further assess the conjecture that the monitoring system reduces entry barrier for new bidders in time-based projects. We compare the percentage of inexperienced workers (workers with no reputation score) among all the bidders with time-based contracts before and after the implementation of the IT-enabled monitoring system. We create a binary variable, No_rating, which denotes whether the worker has received any ratings before (Lin et al. 2016). Then we use the percentage of workers ($Perc\_Norating_{ij}$) who haven't accumulated any reputation records from employers (Lin et al. 2016), as a proxy of the entry barrier to inexperienced workers. We include employer-level fixed effects to control for unobserved heterogeneity across employers and project characteristics to control for heterogeneity across projects. The estimation results are reported in Table 9. The marginal effect of the $Time\_based_j$ dummy is insignificant, indicating that the percentage of inexperienced bidders for time-based projects is about the same as the percentage of new bidders for fixed-price projects before the implementation of the monitoring system. However, after the implementation of the monitoring system, the coefficient of $Time\_based_j$ increases significantly (by 0.076). This increase suggests that other things equal, the percentage of workers with no ratings increases more in time-based projects than fixed-price projects. Specifically, marginal effect estimates based on the delta method indicate the percentage increases by

[11] Based on the estimation results in Column (2) of Table 10, before the implementation of monitoring systems, the partial correlation $Time\_based_j$ dummy and Log_bid_count is 0.410. This partial coefficient becomes 0.623 after the implementation. Since the dependent variable takes the log transformation, we transform the change in the coefficient with the exponential function to obtain the actual increase in the number of bids. Exp(0.213) - 1=23.7%

14.6%. The fact that relatively more participation of inexperienced workers validates H2b that the implementation of the monitoring system lowers the entry-barrier for inexperienced workers entering into time-based projects. As a further analysis, we break down number of workers into number of experienced workers and number of inexperience workers as dependent variables and run the same DID models. We do not observe a significant increase in number of experience workers (p= 0.105), yet we observe a significant increase for number of inexperience workers (p<0.001). The analyses are omitted for brevity.

Overall, the results regarding number of bids and the percentage of inexperienced bidders provide support for our hypothesis that the monitoring system will attract more bids by lowering the entry barrier for inexperienced workers.

Table 9. Estimation Results of the DID Models

Model	(1)	(2)	(3)
Dependent Variable	Log_Bid_Count	Perc_Norating	Award_Price_Premium
Time_based	0.410***(0.103)	-0.019 (0.016)	-0.065 (0.049)
After	-0.354 (0.293)	0.070 (0.074)	0.037 (0.142)
Time_based*After	0.213** (0.093)	0.076***(0.015)	-0.106** (0.052)
Log_bid_count		0.042***(0.005)	-0.071***(0.023)
Log_paid_amount	-0.031 (0.026)	0.001 (0.004)	0.120***(0.020)
Log_budget_max	0.201***(0.034)	-0.015***(0.005)	-0.120***(0.020)
Log_title_length	-0.036 (0.056)	0.015* (0.009)	-0.017 (0.040)
Log_preview_desc_length	0.327***(0.070)	-0.007 (0.010)	0.056 (0.046)
Category dummies	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes
Employer dummies	Yes	Yes	Yes
Clusters(employers)	1,290	1,290	880
Observations	3,042	3,042	1,995
R-squared	0.284	0.153	0.074

Notes: a) As a standard approach, we calculate the average bid price after dropping both the maximum and minimum of bid prices. Therefore, projects with fewer than three bids are dropped in the DID model with Award_Price_Premium as the dependent variable. As such, the sample size in Column (3) is smaller than the first two columns. b) Results are highly consistent when we recalculate the price premium with various benchmark prices, such as the minimum of bid price, the minimum of project budget set by the employer. The result is also highly consistent when we use the bid price as the dependent variable. c) Robust standard errors clustered on employers are reported in parentheses. d) * p<0.1, ** p<0.05, *** p<0.01.

We now look at whether the increased number of bids and more entries from inexperienced workers affect the market equilibrium prices on the online labor market. Column (3) of Table 9 reports the results. We find that the coefficient of the interaction term $After_j \times Time\_based_j$ in the Equation (5) is significantly negative, which suggests that after the implementation of the monitoring system, employers paid significantly lower prices for time-based projects. We calculate the marginal effect of the interaction term at the mean values of all the covariates and find that the price premiums paid by employers of time-based projects decline by 6.9%. Hence, Hypothesis 3 is supported. In summary, results of the three DID models lend support to Hypothesis 2a, Hypothesis 2b, and Hypothesis 3. In Appendix B, we report results based on alternative measures of price premiums, which shows high consistency.

5.4. Robustness Checks

Alternative Matching Method

In the above analysis, we conduct the Propensity Score Matching to generate the matched sample to balance the distribution of observed characteristics across the treatment and control group. To further verify the stability of our results, we employ another matching algorithm – coarsened exact matching (CEM) to regenerate a comparable sample (Iacus et al. 2012; Blackwell et al. (2009); Subramanian and Overby 2016). CEM enables us to explicitly match fixed-price projects with time-based projects within the same category and the same submission quarter. As such, CEM increases the homogeneity between two types of projects from a multivariate perspective and lends support to the causality of our findings. We rerun the DID model on the CEM matched samples and report the results in Table 10. Overall, the results based on the CEM matched sample are consistent with our main results. Again, we find that number of bids and the percentage of inexperienced bidders significantly increase after the implementation of the monitoring system, while the equilibrium price decreases. In Appendix C, we report results based on inverse probability of treatment weighting method (Blackwell 2013), which again shows consistent findings.

Table 10. Estimation Results of the DID Models based on the CEM Matched Sample

Model	(1)	(2)	(3)
Dependent Variable	Log_Bid_Count	Perc_Norating	Award_Price_Premium
Time_based	0.311***(0.091)	-0.003 (0.013)	-0.054 (0.048)
After	-0.191 (0.267)	0.100 (0.065)	-0.069 (0.144)
Time_based*After	0.292***(0.082)	0.077***(0.014)	-0.110** (0.050)
Log_bid_count		0.038***(0.004)	-0.071***(0.021)

Log_paid_amount	-0.011 (0.025)	0.002 (0.004)	0.141***(0.018)
Log_budget_max	0.170***(0.029)	-0.009** (0.004)	-0.120***(0.018)
Log_title_length	-0.049 (0.053)	0.013 (0.008)	-0.015 (0.036)
Log_preview_desc_length	0.275***(0.068)	-0.004 (0.010)	0.035 (0.041)
Category dummies	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes
Employer dummies	Yes	Yes	Yes
Clusters(employers)	1,604	1,604	1,088
Observations	3,719	3,719	2,428
R-squared	0.300	0.149	0.084

Notes: Robust standard errors clustered on employers are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Placebo Tests

To assess the parallel trends assumption of the DID models, we conduct a series of placebo tests. First, we reassign the intervention to the middle of our pre-treatment period (November 1st, 2013) and check the existence of a pre-treatment tendency in the observation window before the actual introduction of the monitoring system. As the placebo treatment did not exist, we should not observe a significant effect from that placebo treatment. As Table 11 shows, the interaction between the placebo treatment time ($After\_placebo$) and the contract type ($Time\_based$) is insignificant.

Second, following Abadie et al. (2015), we conduct another placebo test by randomly reassigning the treatment to projects within our sample. Again, since only projects that are actually treated (time-based projects) should be affected by the implementation of the monitoring system, if we randomly assign treatment to projects, we should not see a treatment effect. We simulate this permutation procedure 1000 times and capture the distribution of the placebo effects based on the randomly assigned placebo treatments. After comparing the estimated coefficient of the actual treatment to the distribution of placebo effects, we find that the probability of similarly sized treatment effect happening by chance is near zero (outside the 99% confidence interval), indicating that the significant finding is robust to alternative variance-covariance specifications.

Lastly, as a last robustness check, we conduct a dynamic DID analysis, reported in Appendix D. We observe that all the relative time parameters are insignificant prior to the implementation while most of the relative time parameters in three models are significant after February 2014 wherein Freelancer introduced the IT-enabled monitoring system. As such, the result of the relative time model lends further support to the validity of the parallel trend assumption and also to our main findings.

Table 11. Estimation Results of the DID Models based on Placebo Treatment Time

Model	(1)		(2)		(3)	
Dependent Variable	Log_Bid_Count		Perc_Norating		Award_Price_Premium	
Time_based	0.045	(0.185)	-0.024	(0.028)	-0.199**	(0.099)
After_placebo	0.084	(0.162)	0.020	(0.023)	0.179	(0.196)
Time_based* After_placebo	0.025	(0.146)	-0.020	(0.022)	0.023	(0.080)
Log_bid_count			0.059***	(0.008)	-0.058*	(0.029)
Log_paid_amount	-0.027	(0.038)	0.006	(0.006)	0.144***	(0.037)
Log_budget_max	0.084	(0.055)	-0.016*	(0.009)	-0.129***	(0.029)
Log_title_length	0.100	(0.090)	0.007	(0.014)	-0.041	(0.055)
Log_preview_desc_length	0.350***	(0.117)	-0.015	(0.015)	0.091	(0.111)
Category dummies	Yes		Yes		Yes	
Month dummies	Yes		Yes		Yes	
Employer dummies	Yes		Yes		Yes	
Clusters(employers)	557		557		491	
Observations	1,261		1,261		918	
R-squared	0.248		0.196		0.120	

Notes: Robust standard errors clustered on employers are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 12. Placebo Effects of Random Implementation Model

	(1)	(2)	(3)
Dependent Variable	Log_Bid_Count	Perc_Norating	Award_Price_Premium
Placebo effects (mean)	0.000	0.000	0.001^{ns}
Placebo effects (st.d.)	0.067	0.011	0.045
Actual treatment effects	0.213** (0.093)	0.076***(0.015)	-0.106** (0.052)
Replication	1000	1000	1000
z-score (H$_0$: placebo = actual effect)	3.179	6.912	-2.390
p-value	0.001***	0.000***	0.017**

6. General Discussion

In this research, we report evidence that shows the implementation of the IT-enabled monitoring system can lower entry barriers for new workers to online labor platforms, reduce employers' preference for workers with high effort-related reputation, drive supply-side competition and lower market prices for employers. Our estimation results are based on a unique quasi-natural experiment at Freelancer.com that implemented a monitoring system for time-based projects but not for fixed-price projects. This allows us to use a DID framework to estimate the causal effects of implementing a monitoring system. We report three main findings. First, after the implementation of the IT-enabled monitoring system, while employers' preference for reputation that signals worker capability for both fixed-price and time-based projects do not change, employers care less about reputation that signals worker effort for time-based projects (but not fixed-price projects). Second, the introduction of the IT-enabled monitoring system intensifies supply-side platform competition by lowering the entry barrier for inexperienced workers and it attracts more bids for time-based projects. Finally, our analysis shows that market equilibrium price for time-based projects decreases after the implementation of the monitoring systems. This finding identified a nuanced substitution relationship between monitoring and reputation such that monitoring partially substitutes effort-related reputation but does not substitute capability-related reputation.

Our study contributes to several streams of IS research. First, this is the first large-scale empirical research to examine the effect of deploying an IT system on information asymmetry (Bardhan et al. 2014). Specifically, we examine the role of IT-enabled monitoring system on both the demand and supply side of an online labor platform. Unlike the previous literature mainly examining the effect of monitoring systems in a firm setting (Gopal and Koka 2010; Pierce et al. 2015; Ranganathan and Benson 2017), we analyze the impact of a monitoring system on a two-sided online labor platform, which enables us to identify unique aspects of online platforms and systematically study the effect of the IT-enabled monitoring system on the demand, supply and equilibrium outcomes in online labor platforms. Second, our study extends the previous research on the effect of reputation systems in digital platforms (Ba and Pavlou 2002; Bockstedt and Goh 2011; Dellarocas 2005, 2006; Lin et al. 2016; Moreno and Terwiesch 2014). The previous literature on reputation systems commonly view reputation acts as a signal of workers' future performance (Banker and Hwang 2008), and motivates workers to spend more effort (Horton and Golden 2015). This paper adds to our understanding of reputation by underscoring the heterogeneous effect of capability-related reputation and effort-related reputation. Our result suggests that IT-enabled monitoring has no significant impact on the importance of capability-related reputation while it can substitute for the signaling effect of effort-related reputation, which alleviates moral hazard problems by providing more precise and timely information about workers' effort (Agrawal et al. 2014; Pierce et al. 2015). This suggests that future research on reputation systems should also take the availability of monitoring systems as a critical contingency factor. Third, this research suggests that the IT-enabled monitoring system is not simply a partial substitution for reputation systems. By substituting for effort-related reputation, IT-enabled monitoring systems reduce agency costs by lowering the entry barrier for workers who have no prior experience on a focal platform. Therefore, our finding suggests the role of IT-enabled monitoring in overcoming a significant limitation of reputation systems that has hitherto been ignored in the IS literature: they create entry barriers for qualified workers who have not established a reputation on a platform.

Our research also provides important managerial implications on the design of online labor markets (Hong et al. 2016), and online platforms in general (Ghasemkhani 2017; Kumar and Tan 2015). There is a large body of research suggesting reputation helps to mitigate moral hazard by acting as both a stimulus for high effort (Horton and Golden 2015) and a sanctioning mechanism (Dellarocas 2006). Meanwhile, monitoring systems are found to be highly effective in improving agents' performance (Duflo et al. 2012; Hubbard 2000; Pierce et al. 2015). Our study suggests that there exists a nuanced substitutional relationship between monitoring and reputation. Specifically, monitoring partially substitutes effort-related reputation but does not substitute capability-related reputation. Hence, our study deepens our understanding of the optimal design of online labor platforms (Hong et al. 2016) by emphasizing the potential interaction effect between effort-related reputation and monitoring.

We acknowledge a number of limitations of this study, which opens up avenues for future research. First, we note that due to data limitation, employers' actual usage of records from monitoring systems is not available. However, given that this is mandatory for workers of time-based projects, the potential selective attrition does not appear to be a serious concern. Second, we only focused on testing the effect of the IT-enabled monitoring system on employer preference and workers' bidding behaviors. Future research shall consider exploring the long-term effect of the IT-enabled monitoring system on workers' skill investment. Finally, our study was conducted in the context of online labor markets and our finding could be limited in its generalizability to other online platforms. Although moral hazard is a universal issue in online platforms, the IT artifact examined in this study – a monitoring system – may not be applicable to platforms that focus on transactions of physical products, such as eBay. Further research should explore the effects of other monitoring systems that are suitable for other online platforms.

7. Concluding Remark

Using a large-scale data set from one of the major platforms that facilitate labor contracting, we utilize matching methods (PSM and CEM) in tandem with a quasi-natural experimental difference-in-differences analysis to identify and quantify the effects of implementing an IT-enabled monitoring system. Our results demonstrate a nuanced substitution relationship between monitoring and reputation such that monitoring partially substitutes for effort-related reputation but not for capability-related reputation. Our findings further suggest implementing a monitoring system lowers the entry barrier for inexperienced workers with no prior reputation and thus driving supply-side competition. Overall, our results provide support for the effectiveness of IT-enabled monitoring in addressing moral hazard issues in online labor markets, and carry important implications for the design of two-sided platforms.

References

Abadie, A. (2005) Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*. 72(1): 1-19

Abadie, A., Diamond, A., & Hainmueller, J. (2015) Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*. *59*(2): 495-510.

Agrawal, A., Lacetera, N., and Lyons, E. (2014) Does Standardized Information in Online Markets Disproportionately Benefit Job Applicants from Less Developed Countries?. Mimeo.

Allgulin, M., and Ellingsen, T. (2002) Monitoring and Pay. *Journal of Labor Economics.* 20(2): 201-216.

Anderson Jr, E.G., Parker, G.G. and Tan, B. (2013) Platform Performance Investment in the Presence of Network Externalities. *Information Systems Research.* 25(1): 152-172.

Angrist, J. D., & Pischke, J. S. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Aron, R., Jayanty, S. and Pathak, P. (2007) Impact of Internet-based Distributed Monitoring Systems on Offshore Sourcing of Services. *ACM Transactions on Internet Technology (TOIT)*. 7(3): 16.

Autor, D. H. (2003) Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. *Journal of Labor Economics*. *21*(1): 1-42.

Ba, S. and Pavlou, P.A. (2002) Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Quarterly.* 26(3): 243-268.

Bajari, P. and Hortacsu, A. (2004) Economic Insights from Internet Auctions. *Journal of Economic Literature.* 42(2): 457-486.

Bajari, P., and Tadelis, S. (2001) Incentives versus Transaction Costs: A Theory of Procurement Contracts. *RAND Journal of Economics.* 32(3): 387-407.

Banerjee, A. V., and Duflo, E. (2000) Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry. *The Quarterly Journal of Economics.* 115(3): 989-1017.

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

Bardhan, I., Oh, J.H., Zheng, Z. and Kirksey, K. (2014) Predictive Analytics for Readmission of Patients with Congestive Heart Failure. *Information Systems Research*. 26(1): 19-39.

Benaroch, M., Lichtenstein, Y., & Fink, L. (2016) Contract Design Choices and The Balance of ex ante and ex post Transaction Costs in Software Development Outsourcing. *MIS Quarterly*. 40(1): 57-82.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004) How Much Should We Trust Differences-in-Differences Estimates?. *Quarterly Journal of Economics.* 119(1): 249-275.

Blackwell, M., Iacus, S. M., King, G., & Porro, G. (2009) Cem: Coarsened Exact Matching in Stata. *The Stata Journal*. *9*(4): 524-546.

Blackwell, M. (2013). A Framework for Dynamic Causal Inference in Political Science. *American Journal of Political Science*. *57*(2): 504-520.

Blinder, A. S., & Krueger, A. B. (2013) Alternative Measures of Offshorability: A Survey Approach. *Journal of Labor Economics*. 31(2): 97-128.

Bockstedt, J. and Goh, K.H. (2011) Seller Strategies for Differentiation in Highly Competitive Online Auction Markets. *Journal of Management Information Systems.* 28(3): 235-268.

Bolton, P., & Dewatripont, M. (2005). *Contract theory*. MIT press.

Brown, J. and Morgan, J. (2006) Reputation in Online Auctions: The Market for Trust. *California Management Review.* 49(1): 61-81.

Brynjolfsson, E., Dick, A.A. and Smith, M.D. (2004) Search and Product Differentiation at an Internet Shopbot (No. 4441-03). Working paper, MIT Sloan School of Management.

Chan, J. and Wang, J. (2017) Hiring Preferences in Online Labor Markets: Evidence of a Female Hiring Bias. *Management Science* (forthcoming).

Chan, J. and Ghose, A. (2014) Internet's Dirty Secret: Assessing the Impact of Online Intermediaries on HIV Transmission. *MIS Quarterly*. 38(4): 955-976.

Chen, Y., and Bharadwaj, A. (2009) An empirical analysis of contract structures in IT outsourcing. *Information Systems Research.* 20(4): 484-506.

Clemons, E. K., and Chen, Y. (2011) Making the Decision to Contract for Cloud Services: Managing the Risk of an Extreme Form of IT Outsourcing. System Sciences (HICSS), the 44th Hawaii International Conference: 1-10.

Dellarocas, C. (2005) Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard. *Information Systems Research.* 16(2): 209-230.

Dellarocas, C. (2006) Reputation Mechanisms. *Handbook on Economics and Information Systems:* 629-660.

Demiroglu, C., and James, C. M. (2010) The Role of Private Equity Group Reputation in LBO Financing. *Journal of Financial Economics.* 96(2): 306-330.

Dey, D., Fan, M., and Zhang, C. (2010) Design and Analysis of Contracts for Software Outsourcing. *Information Systems Research.* 21(1): 93-114.

Diamond, D. W. (1991) Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt. *Journal of Political Economy.* 99(4): 689-721.

Drago, R. (1991) Incentives, Pay, and Performance: a Study of Australian Employees. *Applied Economics.* 23(9): 1433-1446.

Duflo, E., Greenstone, M., Pande, R. and Ryan, N. (2013) What Does Reputation Buy? Differentiation in a Market for Third-Party Auditors. *The American Economic Review.* 103(3): 314-319.

Duflo, E., Hanna, R., and Ryan, S. P. (2012) Incentives Work: Getting Teachers to Come to School. *The American Economic Review.* 102(4): 1241-1278.

Eisenhardt, K. M. (1989) Agency Theory: an Assessment and Review. *Academy of Management Review.* 14(1): 57-74.

Eisenmann, T., Parker, G. and Van Alstyne, M.W. (2006) Strategies for Two-Sided Markets. *Harvard Business Review*. 84(10): 92.

Eisenmann, T., Parker, G., and Van Alstyne, M. (2011) Platform Envelopment. *Strategic Management Journal.* 32(12): 1270-1285.

Ewing, B. T., and Payne, J. E. (1999) The Trade-off between Supervision and Wages: Evidence of Efficiency Wages from the NLSY. *Southern Economic Journal.* 66(2): 424-432.

Fama, E. F. (1991) Time, Salary, and Incentive Payoffs in Labor Contracts. *Journal of Labor Economics.* 9(1): 25-44.

Ghasemkhani, H., Li, Y.M., Moinzadeh, K. and Tan, Y. (2017) Contracting Models for P2P Content Distribution. *Production and Operations Management* (forthcoming).

Gopal, A., and Koka, B. R. (2010) The Role of Contracts on Quality and Returns to Quality in Offshore Software Development Outsourcing. *Decision Sciences.* 41(3): 491-516.

Gopal, A., and Sivaramakrishnan, K. (2008) Research Note-On Vendor Preferences for Contract Types in Offshore Software Projects: The Case of Fixed Price vs. Time and Materials Contracts. *Information Systems Research.* 19(2): 202-220.

Greenwood, B.N. and Agarwal, R. (2015) Matching Platforms and HIV Incidence: An Empirical Investigation of Race, Gender, and Socioeconomic Status. *Management Science.* 62(8): 2281 - 2303.

Heckman, J. J., & Snyder Jr, J. M. (1997). Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators. *RAND Journal of Economics*. 28: 142-189.

Ho, D.E., Imai, K., King, G. and Stuart, E.A. (2007) Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis.* 15(3):199-236.

Hong, Y., Wang, C., & Pavlou, P. A. (2016) Comparing Open and Sealed Bid Auctions: Evidence from Online Labor Markets. *Information Systems Research.* 27(1): 49-69.

Hong, Y. and Pavlou, P.A. (2017) On Buyer Selection of Service Providers in Online Outsourcing Platforms for IT Services. *Information Systems Research*. 28(3): 547-562.

Horton, J. J. (2017) Buyer Uncertainty about Seller Capacity: Causes, Consequences, and A Partial Solution. *Management Science* (forthcoming).

Horton, J. J., and Golden, J. M. (2015) Reputation Inflation: Evidence from an Online Labor Market. Working paper, New York University, New York.

Huang, N., Hong, Y., Burtch, G. (2017) Social Network Integration and User Content Generation: Evidence from Natural Experiments. *MIS Quarterly*. 41(4):1035–1058.

Hubbard, T. N. (2000) The Demand for Monitoring Technologies: the Case of Trucking. *Quarterly Journal of Economics.* 115(2): 533-560.

Iacus, S. M., King, G., Porro, G., & Katz, J. N. (2012) Causal Inference without Balance Checking: Coarsened Exact Matching. *Political analysis*. 20: 1-24.

Kokkodis, M., and Ipeirotis, P. G. (2015) Reputation Transferability in Online Labor Markets. *Management Science.* 62(6): 1687 - 1706.

Kumar, N., Qiu, L. and Kumar, S. (2017) Exit, Voice, and Response in Digital Platforms: An Empirical Investigation of Online Management Response Strategies. *Information Systems Research* (forthcoming).

Leonard, J. S. (1987) Carrots and Sticks: Pay, Supervision and Turnover. *Journal of Labor Economics*. S136-S152.

Lin, M., Liu, Y., and Viswanathan, S. (2016) Effectiveness of Reputation in Contracting for Customized Production: Evidence from Online Labor Markets. *Management Science* (forthcoming).

Liu, D., Brass, D., Lu, Y., & Chen, D. (2015) Friendships in Online Peer-to-Peer Lending: Pipes, Prisms, and Relational Herding. *MIS Quarterly*. 39(3): 729-742.

Lu, Y., Gupta, A., Ketter, W. & Heck, E. van (2016) Exploring Bidder Heterogeneity in Multi-channel Sequential B2B Auctions: Evidence from the Dutch Flower Auctions. *MIS Quarterly*. 40 (3): 645-662.

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): 618-634.

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

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

Parker, G., Van Alstyne, M., and Choudary, S. (2016) *Platform Revolution*. W.W. Norton & Company, New York, NY.

Pauly, M. V. (1974). Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection. *The Quarterly Journal of Economics*. 88(1): 44-62.

Pavlou, P.A. and Dimoka, A. (2006) The nature and role of feedback text comments in online marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Information Systems Research.* 17(4): 392-414.

Pierce, L., Snow, D. C., and McAfee, A. (2015) Cleaning House: The Impact of Information Technology Monitoring on Employee Theft and Productivity. *Management Science.* 61(10): 2299-2319.

Ranganathan, A. and Benson, A. (2017) Hemming and Hawing over Hawthorne: Work Complexity and The Divergent Effects of Monitoring on Productivity. *Management Science* (forthcoming).

Roels, G., Karmarkar, U. S., and Carr, S. (2010) Contracting for Collaborative Services. *Management Science.* 56(5): 849-863.

Shapiro, C., and Stiglitz, J. E. (1984) Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review.* 74(3): 433-444.

Srivastava, S. C., and Teo, T. S. (2012) Contract Performance in Offshore Systems Development: Role of Control Mechanisms. *Journal of Management Information Systems.* 29(1):115-158.

Staats, B.R., Dai, H., Hofmann, D. and Milkman, K.L. (2016) Motivating Process Compliance through Individual Electronic Monitoring: an Empirical Examination of Hand Hygiene in Healthcare. *Management Science* (forthcoming).

Subramanian, H. and Overby, E., 2016. Electronic Commerce, Spatial Arbitrage, and Market Efficiency. *Information Systems Research*. 28(1): 97-116.

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): 205-240.

Susarla, A., and Krishnan, R. (2014) Contracting in the Shadow of The Future. *Available at SSRN 1975215*.

Train, K. E. (2009) *Discrete Choice Methods with Simulation*. Cambridge University Press.

Van Alstyne, M., Parker, G., Choudary, S. (2016) Pipelines, Platforms, and the New Rules of Strategy. *Harvard Business Review.* 94(4): 54.

Wang, A., Zhang, M. and Hann, I.H .(2015) Socially nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings. *Information Systems Research* (forthcoming).

Werner, R. M., Konetzka, R. T., Stuart, E. A., Norton, E. C., Polsky, D., & Park, J. (2009) Impact of Public Reporting on Quality of Postacute Care. *Health Services Research*. *44*(4): 1169-1187.

Xu, K., Chan, J., Ghose, A. and Han, S.P. (2016) Battle of the Channels: The Impact of Tablets on Digital Commerce. *Management Science* (forthcoming).

Xue, M., Hitt, L. M., & Chen, P. Y. (2010) Determinants and Outcomes of Internet Banking Adoption. *Management Science.* 57(2): 291-307.

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

Zacharia, G., Moukas, A. and Maes, P. (2000) Collaborative Reputation Mechanisms for Electronic Marketplaces. *Decision Support Systems.* 29(4): 371-388.

Zhang, Z. and Li, B., (2017) A Quasi-experimental Estimate of the Impact of P2P Transportation Platforms on Urban Consumer Patterns. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 1683-1692.

Online Supplementary Appendices

A. Balance Check for Propensity Score Matching

Table A1. Balance Check for Propensity Score Matching[12]

Variable	Sample	Mean		%bias	% reduced |bias|	t-test	
		Treated	Control			t	p>|t|
Project_desc_length	Unmatched	16.50	16.27	5.50		2.62	0.01
	Matched	16.50	16.47	0.90	84.00	0.36	0.72
Paid_amount	Unmatched	133.22	129.56	1.40		0.73	0.47
	Matched	133.22	134.64	-0.50	61.30	-0.20	0.84
Title_length	Unmatched	5.68	5.57	3.50		1.68	0.09
	Matched	5.68	5.59	2.90	15.80	1.16	0.25
Software	Unmatched	0.32	0.33	-1.00		-0.48	0.63
	Matched	0.32	0.32	1.80	-87.20	0.73	0.46
Design	Unmatched	0.09	0.09	-1.50		-0.73	0.46
	Matched	0.09	0.09	-0.40	70.90	-0.18	0.86
Writing	Unmatched	0.15	0.12	7.80		3.88	0.00
	Matched	0.15	0.14	1.80	76.50	0.70	0.48
Marketing	Unmatched	0.05	0.04	3.10		1.54	0.12
	Matched	0.05	0.05	-0.80	73.50	-0.31	0.75
Data-entry	Unmatched	0.06	0.04	11.40		5.87	0.00
	Matched	0.06	0.07	-1.70	85.30	-0.59	0.56
Translation	Unmatched	0.02	0.03	-1.10		-0.50	0.61
	Matched	0.02	0.02	0.50	51.10	0.21	0.84
Engineering	Unmatched	0.02	0.02	-1.30		-0.62	0.54
	Matched	0.02	0.02	-0.40	70.40	-0.15	0.88
Manufacturing	Unmatched	0.00	0.00	-0.40		-0.19	0.85
	Matched	0.00	0.00	-0.90	-122.80	-0.34	0.73
Mobile	Unmatched	0.01	0.01	-0.30		-0.14	0.89
	Matched	0.01	0.01	-1.10	-277.90	-0.43	0.66
Other	Unmatched	0.28	0.32	-9.90		-4.74	0.00
	Matched	0.28	0.28	-1.50	84.40	-0.62	0.54
Employer_tenure_month	Unmatched	30.13	30.17	-0.10		-0.06	0.95
	Matched	30.13	30.46	-1.20	-901.10	-0.49	0.63
Employer_overall_ratin	Unmatched	4.92	4.92	-0.60		-0.25	0.80
	Matched	4.92	4.92	-0.50	6.30	-0.21	0.84

Notes: a) Results of Nearest Neighbor (4) Matching Method are presented. We also conducted robustness checks with other matching algorithms in the additional analysis section. The result is qualitatively consistent. b) Within the matched sample, the group means of all the month dummies are not significantly different between time-based projects and fixed-price projects. Balance checks of all the month dummies are omitted for brevity.

[12] We match fixed-price projects with time-based projects by using the Nearest Neighbor (4) matching method. In order to construct a more homogenous sample, we limit our sample to projects with the common public auction format. Therefore, those projects which require NDA contracts, are featured or sealed, are fulltime jobs, use a non-dollar currency, are not written in English are dropped.

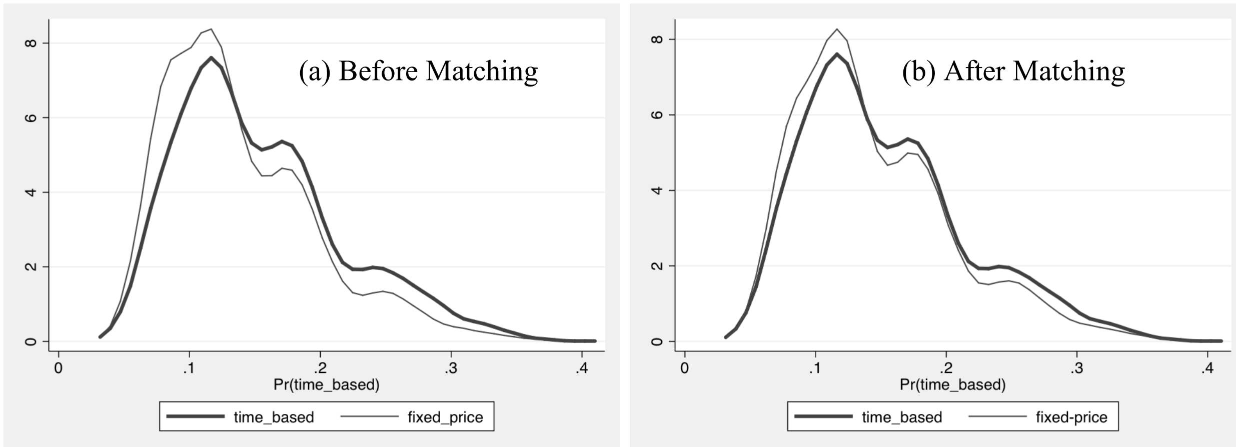

Figure A1. Distribution of Propensity Scores for Time-based Projects and Fixed-price Projects

B. Alternative Measures of Price Premium

Because of the two different kinds of reputation signals (Capability and Effort), the reputable workers are expected to produce a higher value of output. As such, employers are willing to pay workers a markup which is regarded as information rent. In Table 9, we take the recalculated mean of the bid price by dropping the maximum and minimum of bid prices as the benchmark and construct the price premium accordingly. In order to check the robustness of our conclusion, we rerun the DID models with alternative measures of price premium and the log-transformed bid price. As Table A2 shows, the results are highly consistent with our conclusion that the implementation of the monitoring system intensifies market competition and leads to a lower price premium.

$$Price\_Premium\_Min\_Bid_{jk} = \frac{(Bid\_Price_{jk} - Bid\_Min_j)}{Bid\_Min_j} \tag{A1}$$

$$Price\_Premium\_Min\_Budget_{jk} = \frac{(Bid\_Price_{jk} - Min\_Budget_j)}{Min\_Budget_j} \tag{A2}$$

Table A2. Estimation Results of the DID Models with Alternative Measures

Model	(1)	(2)	(3)
Dependent Variable	Price_Premium_Min_ Bid	Price_Premium_Min_ Budget	Log_Bid_Price
Time_based	-0.137 (0.105)	-0.566***(0.208)	-0.582***(0.063)
After	-0.485* (0.256)	-0.577 (0.387)	-0.195* (0.101)
Time_based* After	-0.435***(0.113)	-0.512***(0.178)	-0.163***(0.052)
Log_bid_count	0.268***(0.031)	0.126* (0.074)	-0.042***(0.015)
Log_paid_amount	0.184***(0.043)	0.446***(0.077)	0.248***(0.021)
Log_budget_max	-0.158***(0.042)	-0.197** (0.091)	0.563***(0.027)

Log_title_length	0.066	(0.064)	0.127	(0.120)	-0.026	(0.029)
Log_preview_desc_length	0.084	(0.058)	0.099	(0.163)	-0.002	(0.036)
Category dummies	Yes		Yes		Yes	
Month dummies	Yes		Yes		Yes	
Employer dummies	Yes		Yes		Yes	
Clusters(employers)	1,290		1,290		1,290	
Observations	3,042		3,042		3,042	
R-squared	0.076		0.073		0.768	

Notes: a) The results of these category dummies and month dummies are suppressed for brevity; b) Robust standard errors clustered on employers are reported in parentheses; c) * $p<0.1$, ** $p<0.05$, *** $p<0.01$

C. Matching in Causal Inference

In order to lower the potential model dependence and extrapolation, we apply two different additional matching methods to make the pre-treatment and post-treatment sample more comparable (Ho et al. 2007) and control for the potential change in the group composition (Stuart et al. 2014).

First, by applying the inverse probability of treatment weighting (IPTW) method (Blackwell 2013), we could extend it to the difference-in-differences setting (Stuart et al. 2014). Specifically, we consider four different groups (the treatment group in the pre-period, the treatment group in the post-period, the control group in the pre-period, the control group in the post-period) as multiple treatments (McCaffrey et al. 2013) and reweight all the groups to reflect the covariate distribution in the treatment group before the monitoring system was implemented (McCaffrey et al. 2013; Stuart et al. 2014). Based on the standard propensity score theorems (Rosenbaum and Rubin 1983), when the treatment assignment and the covariate matrix are conditionally independent given the matched propensity score, we could estimate the treatment effect of interest consistently (Stuart et al. 2014).

Second, instead of balancing the sample both across time periods and treatment group assignment simultaneously, we also try to avoid extrapolation by matching the pre-period observations with the post-period observations (Werner et al. 2009). Specifically, we first use a propensity score matching approach to generate a matched sample by pairing the treatment group with the control group. Further, we prune post-treatment pairs that are outside of the “convex hull” of the pre-treatment pairs (Ho et al. 2007; Keele et al. 2016). As suggested by Keele et al. (2016), we use the pair mean of covariates to match pairs across two periods and drop those pairs if there are not similar pairs in the other corresponding period.

By applying the above two different methods, we rerun the DID models and estimate the treatment effect on the three dependent variables of interest ($Perc\_Norating$, Bid_Count, $Award\_Price\_Premium$). As both Table A3 and Table A4 suggest, the implementation of the monitoring system leads to a significant boost in number of bids, a significant increase in the percentage of bidders without ratings, and a significant drop in price premium. Overall, all our three hypotheses are supported.

Table A3. Estimation Results of the DID Models based on IPTW Across Four Groups

Model	(1)	(2)	(3)
Dependent Variable	Log_Bid_Count	Perc_Norating	Award_Price_Premium
Time_based	0.458***(0.063)	0.008 (0.009)	0.046 (0.029)
After	0.017 (0.238)	0.032 (0.046)	-0.037 (0.087)
Time_based* After	0.234***(0.069)	0.063***(0.011)	-0.123***(0.036)
Log_bid_count		0.029***(0.003)	-0.080***(0.016)
Log_paid_amount	0.002 (0.015)	-0.006** (0.002)	0.088***(0.010)
Log_budget_max	0.172***(0.023)	-0.010***(0.003)	-0.076***(0.011)
Log_title_length	-0.088**(0.039)	0.010 (0.006)	-0.015 (0.019)
Log_preview_desc_length	0.482***(0.056)	0.001 (0.009)	-0.012 (0.041)
Log_employer_overall_rating	0.007 (0.049)	-0.003 (0.009)	-0.025 (0.039)
Log_employer_reviews_count	-0.015 (0.014)	0.007***(0.002)	0.003 (0.007)
Log_employer_tenure_month	0.005 (0.021)	-0.000 (0.003)	0.008 (0.012)
Employer_developed	0.063* (0.034)	-0.036***(0.006)	-0.040* (0.023)
Category dummies	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes
Observations	3,647	3,647	2,754
R-squared	0.195	0.235	0.074

Notes: a) Because Bid_Mean_j is disproportionately influenced by the maximum of the bid price, we recalculate the average bid price by dropping both the maximum and minimum of bid prices. Therefore, projects with less than three bids are dropped in the DID model with Award_Price_Premium as the dependent variable. As such, the sample size in Column (3) is smaller than the first two columns. b) Results are highly consistent when we recalculate the price premium with various benchmark prices, such as the minimum of bid price, the minimum of project budget set by the employer. The result is also highly consistent when we use the bid price as the dependent variable. c) The results of these category dummies and month dummies are suppressed for brevity. d) Robust standard errors clustered on employers are reported in parentheses. e) * p<0.1, ** p<0.05, *** p<0.01.

Table A4. Estimation Results of the DID Models based on Matched Pairs Across Two Time Periods

Model	(1)	(2)	(3)
Dependent Variable	Log_Bid_Count	Perc_Norating	Award_Price_Premium
Time_based	0.118** (0.050)	-0.001 (0.009)	0.046 (0.034)
After	0.021 (0.189)	-0.041 (0.079)	-0.028 (0.111)
Time_based* After	0.126** (0.051)	0.054***(0.010)	-0.177***(0.042)

Log_bid_count		0.006* (0.003)	-0.077***(0.012)
Log_paid_amount	0.030** (0.013)	-0.004* (0.002)	0.120***(0.009)
Log_budget_max	0.095***(0.016)	-0.010***(0.003)	-0.098***(0.009)
Log_title_length	-0.087***(0.030)	0.003 (0.005)	-0.004 (0.016)
Log_preview_desc_length	0.167***(0.052)	-0.003 (0.009)	-0.050* (0.030)
Log_employer_overall_rating	-0.000 (0.036)	-0.004 (0.005)	0.012 (0.009)
Log_employer_reviews_count	-0.029***(0.010)	0.002 (0.002)	0.003 (0.005)
Log_employer_tenure_month	0.040***(0.014)	-0.004* (0.003)	0.010 (0.009)
Employer_developed	0.090***(0.028)	-0.033***(0.005)	-0.049** (0.022)
Category dummies	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes
Observations	4,369	4,369	4,369
R-squared	0.190	0.066	0.094

Notes: a) Because Bid_Mean_j is disproportionately influenced by the maximum of the bid price, we recalculate the average bid price by dropping both the maximum and minimum of bid prices. b) Results are highly consistent when we recalculate the price premium with various benchmark prices, such as the minimum of bid price, the minimum of project budget set by the employer. The result is also highly consistent when we use the bid price as the dependent variable. c) The results of these category dummies and month dummies are suppressed for brevity. d) Robust standard errors clustered on employers are reported in parentheses. e) * p<0.1, ** p<0.05, *** p<0.01.

D. Relative Time Model

In order to further test the parallel trend assumption of the DID model (Angrist and Pischke 2008), we employ the relative time model test to assess whether time-based projects and fixed-price projects have a common trend during the pre-treatment period. This analysis also allows us to check at what time the effects start to emerge. We specify the relative time model as follows:

$$y_{ij} = \alpha + \rho\tau_t + \mu Time\_based_j + \beta(\tau_t \times Time\_based_j) + \gamma_i + \delta_j + \varepsilon_{ij} \quad (6)$$

where y_{ij} represents the dependent variables of our interest, including $Perc\_Norating_{ij}$, Bid_Count_{ij}, $Award\_Price\_Premium_{ij}$. τ_t represents a vector of time dummies and $\{\beta\}$ denotes the matrix of relative time parameters to be estimated at time t for project j posted by employer i. If there exists a pre-treatment trend, we should observe significant relative time parameters before the implementation of the monitoring system. Following Autor (2003)'s approach, we use the month before the change (January 2014) as the baseline since the monitoring system implementation happened on February 5th, 2014. We visualize the results in Figure A2. The analysis shows that all the relative time parameters are insignificant prior to the implementation while most of the relative time parameters in three models are significant after February 2014 wherein *Freelancer* introduced the IT-enabled monitoring system. As such, the result of the relative time model lends further support to the validity of the parallel trend assumption and also to our main findings.

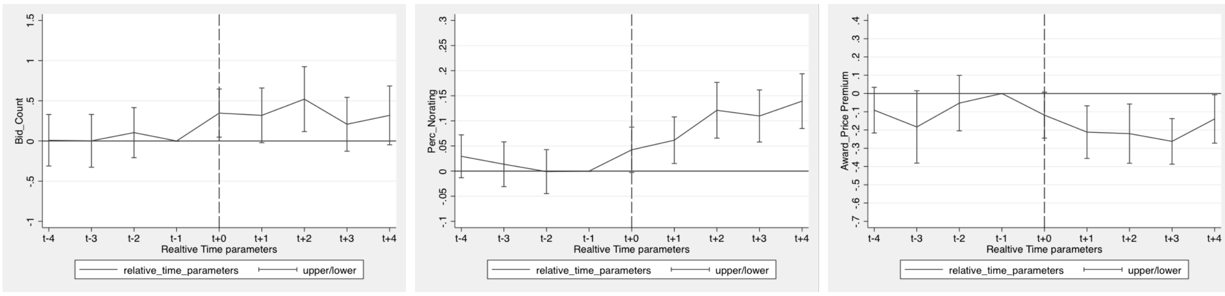

Figure A2. Coefficients of the Monthly Dynamic Difference-in-Differences Estimates

Note: The dash vertical line denotes the month in which *Freelancer* first introduced the monitoring system (February 2014). Error bars represent the 90% confidence intervals of clustered standard errors.

References

Angrist, J. D., & Pischke, J. S. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton university press.

Blackwell, M. (2013) A Framework for Dynamic Causal Inference in Political Science. *American Journal of Political Science*. 57(2): 504-520.

Ho, D.E., Imai, K., King, G. and Stuart, E.A. (2007) Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis.* 15(3):199-236.

Keele, L., Small, D. S., & Hsu, J. Y. (2016) Patterns of Effects and Sensitivity Analysis for Differences-in-Differences. Working paper.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013) A Tutorial on Propensity Score Estimation for Multiple Treatments Using Generalized Boosted Models. *Statistics in Medicine*. 32(19): 3388-3414.

Rosenbaum, P. R., & Rubin, D. B. (1983) The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. 70(1): 41-55.

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M. E., & Barry, C. L. (2014) Using Propensity Scores in Difference-in-Differences Models to Estimate the Effects of a Policy Change. *Health Services and Outcomes Research Methodology*. 14(4): 166-182.

Werner, R. M., Konetzka, R. T., Stuart, E. A., Norton, E. C., Polsky, D., & Park, J. (2009) Impact of Public Reporting on Quality of Postacute Care. *Health Services Research*. 44(4): 1169-1187.