Contracting Decisions on Digital Markets for Knowledge Work Services: A Qualitative Systematic Review

Completed Research Paper

Julian Prester
UNSW Business School
Sydney, Australia
j.prester@unsw.edu.au

Gerit Wagner
HEC Montréal
Montréal, Canada
gerit.wagner@hec.ca

Abstract

A growing amount of research is dedicated to digital markets for knowledge work services and contracting decisions as one of their most critical bottlenecks. Large-scale studies with average samples exceeding 144,000 observations have considered over 70 variables explaining contracting decisions. Our review aims to facilitate the convergence towards commonly agreed categories of variables and overarching models of contracting decisions in this context. To mitigate the risk of propagating deflated p-values associated with large sample studies, we propose a robust version of vote-counting techniques. The principal findings of our research suggest that only a few variables related to the client-worker relationship, the bid and individual worker characteristics have a practically significant effect on worker selection, and that auction success is primarily affected by project value and client experience. These findings lay the groundwork for future research that will lead to better explanations of contracting decisions.

Keywords: Digital markets, knowledge work services, IT services, professional services, digital labor platforms, online labor markets, outsourcing, microsourcing, gigwork, digital platforms, future of work, digital transformation.

Introduction

Digital markets for knowledge work services, such as Upwork, Freelancer, and Fiverr, allow clients to outsource knowledge work services to skilled freelancers who complete projects involving programming, design, translation, and administrative services. Such digital service markets and platform businesses have already attracted a significant workforce. One in every three U.S. workers already participates in online labor markets (Soergel 2016), which experience significant spikes of worker supply after mass layoff events (Huang et al. 2020). According to industry reports, online talent platforms could benefit up to 540 million workers globally over the next five years (Manyika et al. 2015). Global participation in digital service markets is fueled by potential labor arbitrage (Roach 2003), with *completely-digital* labor markets effectively circumventing policies that restrict work-related migration and thus promising trillions dollar gains (Clemens 2011). Client organizations benefit from these markets by gaining access to new sources of skilled digital talent, which is rare and notoriously difficult to acquire in local labor markets (Khan and Sikes 2014; Trost 2014), as well as by benefiting from the greater flexibility of short-term contracts and increasing workforce scalability (Gol et al. 2019).

In practice, contracting on such markets is a critical bottleneck. Contracting rates remain low despite both market sides' interest in entering a contractual relationship and a range of governance mechanisms put in place by the platform provider to facilitate matchmaking. In a recent study of worker survival, Kanat et al. (2018) note that 87% of workers "exit the market because of their inability to survive" (p.6). Similarly, Snir and Hitt (2003) report that less than 60% of projects posted by clients lead to a contract, with substantially lower rates (10%) reported for large projects. Initial analyses point to worker characteristics, such as skills, which are difficult to evaluate in the presence of rating inflation (Kokkodis 2020), market characteristics, such as excessive numbers of worker bids, which lead to *choice overload* (Hong and Zheng 2015; Zheng et al. 2015), or elusive client-worker relationship aspects, such as different forms of signaling (Holthaus and Stock 2018). Overall, practitioners are lacking a coherent understanding of the factors that contribute to low contracting rates.

A growing amount of IS research has been dedicated to understanding contracting decisions on digital markets for knowledge work services (Gefen and Carmel 2008; Wagner and Prester 2019). Synergies between a wealth of publicly accessible data and research approaches suitable to identify causal relationships (e.g., instrumental variables, difference-in-differences approaches) have triggered research explaining worker selection decisions (e.g., Gefen and Carmel 2008), or auction success (e.g., Zheng et al. 2015). However, an overarching and concise model that integrates current findings and guides ongoing research has yet to be proposed. Such a model is needed as a basis for a cumulative body of research and to move toward a more theory-driven research agenda. At the same time, there are critical conceptual and methodological challenges. From a conceptual perspective, the prevalence of many simple variables lacking an overarching theoretical framework makes it challenging to assess how prior findings fit together. Furthermore, the heterogeneity of variable names impedes efforts to disentangle corresponding categories of constructs and relationships. From a methodological perspective, the large-scale samples analyzed in prior research may deflate *p*-values so that very small and practically insignificant effects are reported as statistically significant. Because conventional vote-counting techniques are not designed to distinguish between practically significant and practically insignificant effects, it is unclear how vote-counting techniques can be applied to such evidence without propagating the problem of deflated *p*-values and risking misleading conclusions at the aggregated level. More broadly, the plethora of antecedent variables and the insufficiency of coefficient significance levels to guide the selection of variables (rooted in the *p*-value problem) create unique challenges for the development of theoretical models in the area of digital service platforms and in similar contexts.

In light of these challenges, our work is dedicated to two research questions:

- What has extant research found about the variables predicting contracting decisions, i.e., worker selection and auction success, on digital markets for knowledge work services?
- What are the gaps and shortcomings to be considered in future research?

To address these questions, we conducted a qualitative systematic review of digital markets for knowledge work services, extracted variables that predict contracting decisions, and adapted a vote-counting approach to develop an integrated model. This work makes three contributions. First, we review extant research and provide concise models of contracting decisions to facilitate a transition towards more conceptually integrated research. Second, we recognize particular methodological challenges associated with the large-scale studies in our context and propose *robust vote-counting* as a viable approach to mitigating the propagation of deflated *p*-values in conventional vote-counting techniques. This methodological contribution enables vote-counting studies in contexts where large sample sizes are prevalent, such as ours. It equips researchers for vote-counting studies in similar contexts involving large-scale data sets of micro-level observations, such as digital marketplaces, online social networks, or internet-of-things ecosystems. Finally, we build on the methodological and conceptual insights gained from our analyses and outline critical avenues for future research.

The next section provides the background on digital markets for knowledge work services. The following sections outline our methodological approach, including the new robust vote-counting approach, and findings obtained in our study. The paper concludes with a discussion of future research directions.

Background

Knowledge-intensive services represent an emergent outsourcing arrangement, which often involves contracts with individual workers (Gefen and Carmel 2008; Lacity et al. 2010). Such digital markets, enabled by platform intermediaries, are expected to facilitate efficient transactions by lowering the cost of searching for workers, bargaining, and enforcing contracts (Williamson 1991). Such contingent work arrangements represent a form of relational contracting that require trust, provide viable alternatives to organizational employment relationships, and point to an important area in which hierarchical control is replaced by market coordination (Adler 2001). Because knowledge-intensive services can often be completed digitally, their execution does not face conventional restrictions that labor markets and migration policies impose upon other services requiring physical mobility of workers (Gong et al. 2018). Therefore, knowledge-intensive service markets are considered to have an unprecedented potential for scaling and becoming a significant driver of digital disruption in labor markets worldwide (Clemens 2011).

In outsourcing arrangements, allocation and contractual governance mechanisms depend on the nature of services being sourced. Automated services (e.g., cloud computing) are primarily allocated through on-demand access models offered by large providers and governed through standardized contracts, which involve service-level agreements and demand-based pricing (Yang and Tate 2012). Sourcing of (human) micro-task work primarily relies on fixed-price assignments; micro-task work is characterized by repeatability, anonymity between clients and workers, and complete specification in the task description (Chen and Horton 2016; Wang et al. 2017). Innovative work may be sourced through contests in which multiple workers complete a project, and the client selects the winning submission and awards a fixed price (Majchrzak and Malhotra 2013). In the following, we focus on sourcing of knowledge-intensive services, accomplished through reverse auctions in which clients solicit bids from individual workers (reflecting quality and price) to contract for future services (Hong et al. 2016).

The predominant mechanism for sourcing knowledge-intensive services is the client-determined (non-binding) reverse auction. Instead of requiring clients to explore a large market of workers (service providers), reverse auctions allow clients to post a service description (request for proposal) and evaluate bids from workers who are willing to accept an offer and provide different skill sets under different conditions (e.g., price, availability). The combination of efficient market discovery and provider competition explains why reverse auctions have created significant process efficiencies and cost savings in several areas, such as the industrial procurement of goods and services (Jap 2002). In our context, *client-determined (non-binding)* means that clients can decide whether, to whom, and when to offer a contract; such auctions differ from auctions relying on event-determined award rules (Jap 2002), in which a prespecified rule (e.g., lowest bid price when the auction's time expires) determines the winner of the auction. Clients can decide not to offer a contract, they can close a contract when the auction is still in progress or after it has ended, and they can choose the worker who best meets their multi-criteria preferences (e.g., regarding price, availability, certifications, and communication skills). Navigating this bidding and selection process of client-determined reverse auctions holds substantial challenges for both parties. For instance, clients may face heterogeneous worker signals, including potentially inflated ratings (Holthaus and Stock 2018; Kokkodis 2020), and workers have imperfect information for understanding how competitive their bids are (Jap 2002), because clients may not fully disclose their preferences or choose not to select the bid with the lowest price.

Conceiving the sourcing of knowledge-intensive services as a purely bilateral reverse auction would overlook the vital role of the platform intermediary, who governs the matching, contracting, and executing activities (Wagner and Prester 2019). When sourcing services at the individual level, developing trust in a worker is challenging due to spatial separation and agency problems, such as adverse selection and moral hazard (Adler 2001; Chen and Horton 2016; Du and Mao 2018). The situation is further complicated by the nature of knowledge work, which is often emergent, prone to unanticipated contingencies, and rarely standardized. This means that clients have to rely on incomplete contracts (Hong et al. 2020) and place their trust in the worker's integrity and benevolent behavior (Adler 2001). In this context, the role of the platform intermediary is to facilitate the development of trust by enforcing a range of governance mechanisms across individual contractual episodes, such as rating, monitoring, and arbitration systems (Adler 2001; Du and Mao 2018; Gol et al. 2019).

The reverse auction of posting a request for proposal, bidding, and selecting a worker represents an interactive process, during which a range of information is exchanged and processed (see Figure 1 for an overview). This process is preceded by the participation phase, during which workers and clients register on the platform and create profiles. Most notably, workers can design their profiles, describe their competencies, and present a portfolio of prior projects. To source services, the client can initiate a reverse auction by posting a request for proposal, which involves specifying the modalities of the auction (e.g., auction duration, and bid visibility) as well as the service requirements (e.g., skills, deliverables, and deadlines). Workers can then bid on the project by suggesting a price and submitting a description to summarize their skills and approach to completing the project. Because submitting a bid incurs costs, such as the time spent to develop the bid description and sometimes fees payable for submitting an individual bid, rational workers would offset the expected profits of the transaction against the cost of bidding (Snir and Hitt 2003). The client can then review the bids and select the worker who submitted the best bid, as determined by a multi-criteria utility function (e.g., covering bid price and worker quality). Evaluation processes are informed by a range of information, covering individual and aggregated bid and worker characteristics (Hong and Zheng 2015). Because auctions may receive a large number of bids, and knowledge-intensive services require customized bids that match the project description, this process may create significant costs for bid preparation and evaluation. When the client has made his/her selection and offered the contract, the worker can decline, accept, or negotiate the terms. Auction success, therefore, requires a worker selection decision (contract offer) and the acceptance by the worker. When these conditions are met, both parties can proceed with the service exchange. Naturally, the complexity of the process and the non-binding nature of the auction often result in failed auctions (Kanat et al. 2018; Snir and Hitt 2003). The multi-faceted information cues available at different levels, and the action spaces of the involved parties, motivate our rigorous analysis of empirical research to explain the critical element of contracting decisions on digital markets for knowledge work services.

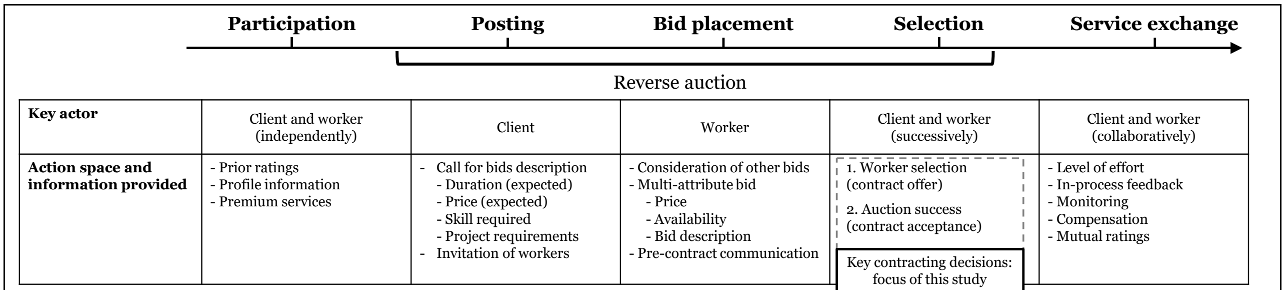

Figure 1. Process of Sourcing Knowledge Work Services on Digital Markets

Our review of the literature on digital markets for knowledge work services differs from the two previous reviews in several ways. In a first review, Gol et al. (2019) synthesized the literature on crowd work platforms to develop a conceptual model of platform governance. The authors offer insights into the governing mechanisms of the contracting process, such as contract management and reputation systems. However, their focus is on the centralized platform provider, with little consideration of worker and client characteristics or market dynamics. In a second review, Wagner and Prester (2019) provide a scoping review on digital platforms for knowledge work. The review develops a definition of the phenomenon and conceptualizes the process of sourcing knowledge work on digital platforms. Although the authors conceptualize a phase that precedes and thus may affect the contracting process, they do not discuss specific factors or variables that influence the contracting decision. Previous reviews on digital markets for knowledge work services have thus focused on a particular mechanism of the digital platform artifact or taken an exploratory perspective to conceptualize the phenomenon. We complement prior work by focusing on empirical findings on worker selection and auction success (contracting decisions) and their independent variables.

Methodology

The purpose of this study is to aggregate the empirical findings and to integrate them into concise models of contracting decisions on digital markets for knowledge work services. Review methods that are suitable for the aggregation and integration of findings are qualitative systematic reviews, meta-analyses, and umbrella reviews (Paré et al. 2015; Schryen et al. 2020). Other review methods that focus on summarizing prior knowledge in the form of general themes, critical assessment, or explanation building through interpretation (Boell and Cecez-Kecmanovic 2014) are less suitable for this purpose. In selecting the most appropriate review method for data aggregation and integration, the maturity of the research stream and the nature of primary studies must be taken into account. In contrast to mature research streams that span decades, the average study in our sample is five years old, which underlines the emergent state of research in this area. This excludes an umbrella review, which would require a more mature research stream that has been examined by several systematic reviews. The evidence in our sample covers both qualitative and quantitative findings. Furthermore, only 25% of studies report standard effect measures such as mean differences or correlation coefficients. While meta-analyses are a powerful research method for data aggregation and integration, their specific data extraction techniques rely on these coefficients. Qualitative systematic reviews offer a suitable approach for contexts like ours.

The qualitative systematic review involves the identification, coding, and analysis of data from primary studies to answer questions related to the categories of independent variables, as well as the directions and consistency of effects reported in prior work (Higgins and Green 2008; Paré et al. 2015). One of its strengths is the ability to include evidence from both quantitative and qualitative studies without relying on correlation coefficients. On the downside, the qualitative systematic review follows a less statistical approach and does not consider construct validity, biases and heterogeneity in the reviewed studies. Taken together, a qualitative systematic review is most appropriate at this stage to aggregate and integrate prior research on digital markets for knowledge work services. Implementing procedures similar to prior reviews of IT outsourcing (Lacity et al. 2010), we identified relevant studies and coded the dependent variables (contracting decisions), the independent variables, as well as their empirical relationships. To build an integrated model that explains contracting decisions, we applied an adapted version of vote-counting.

Vote counting is a technique that uses measures such as variable coefficients and *p*-values to determine the direction and consistency of effects or relationships reported in primary studies. An important shortcoming of conventional vote-counting procedures is that differences in effect sizes are not considered (Higgins and Green 2008; King and He 2005). This may not be a major concern in many traditional research contexts[1], but it can be a particularly critical issue when primary studies apply statistical inference tests designed for small samples to large samples, resulting in high statistical power and deflated *p*-values (Lin et al. 2013; Mertens and Recker 2020; Sawyer and Peter 1983). This is the case in our review, with included studies reporting sample sizes with an average of over 144,000 observations; evidence for deflated *p*-values can be observed with an array of small effect sizes reported as highly significant. Most strikingly, in one of the earliest studies with a sample size exceeding 250,000 observations, the *amount bid* variable was reported to have a highly significant positive effect ($p<.001$) with a coefficient of 0.00 (Gefen and Carmel 2008). Prior research has defined the corresponding *p-value deflation problem* (also referred to as the *p-value problem*) as the problem of claiming support for results that have no practical significance when solely relying on *p*-values in very large samples where *p*-values rapidly approach zero (Lin et al. 2013, p. 906). More formally, when sample sizes approach the size of large populations (or infinity), it is to be expected that "unless the population parameter β is exactly equal to the null value [i.e., $\beta = 0$] with an infinite number of decimals […], the *p*-value [of its estimator $\hat{\beta}$] will approach 0" (Lin et al. 2013, p. 907). In an informative methodological commentary, Lin et al. (2013) suggest that in such cases, primary studies should go beyond statistical significance and report the *practical* significance of effects, which requires the judgment of domain experts. Potential *p*-value deflation at the level of individual studies raises the question of how such problems may affect secondary studies and techniques aimed at aggregating evidence. Meta-analytic techniques extract effect sizes (in addition to other characteristics like sample sizes and construct reliabilities) and explicitly consider them when calculating aggregated effect sizes (Higgins and Green 2008), effectively distinguishing effect sizes at the aggregated level, including small ones that may result from deflated *p*-values. Meta-analyses, thereby, enable readers to dissociate effect sizes and draw informed conclusions (see Sawyer and Peter 1983). In contrast, conventional vote-counting techniques do not distinguish effect sizes once reported as significant (Higgins and Green 2008; King and He 2005). Thus, when conventional vote-counting techniques are applied to large-scale studies, counting very small but significant effects in the same category as the main effects (with medium to large effect sizes) would result in a *propagation of the p-value problem*. This means that small effects that are potentially subject to the *p*-value problem are aggregated and become indistinguishable from medium to large effects that are less likely to result from deflated *p*-values (see Sawyer and Peter 1983).

[1] We note that prior qualitative systematic reviews applying vote-counting procedures, such as the one of Lacity et al. (2010), are unlikely to be affected by the *p*-value problem because they relied on survey data whose average sample sizes presumably range in the low triple-digits, resulting in acceptable levels of statistical power and lower likelihood of detecting statistically significant evidence for small effect sizes.

As a response to these methodological challenges of aggregating findings from large-scale studies, we propose a *robust vote-counting approach* to mitigate the *propagation of p-value problems* in conventional vote-counting approaches. To the best of our knowledge, we are the first to adapt vote-counting techniques to account for the *deflated p-value problem*. Our robust approach builds on the recommendations of Lin et al. (2013) by focusing on the practically significant effect sizes in combination with statistical significance. Specifically, the adapted procedure restricts the vote-counting of quantitative evidence to significant medium to large effects. To determine the practical significance of an effect, researchers must consider the scale of a variable, its standard deviation, the (relative) size of the coefficient, and the mean value of the dependent variable (Ellis 2010; Lin et al. 2013). The aggregated results of the robust vote-counting approach indicate which variables are likely to have a practically significant impact on the dependent variable.

Identification of Studies

To identify relevant studies in Information Systems (IS) research, we cover the scope of 40 IS journals (as identified by Lowry et al. 2013), and five IS conferences (i.e., AMCIS, ECIS, HICSS, ICIS, PACIS). Database searches covered the *AIS Electronic Library (AISeL)*, *ABI/INFORM*, *EBSCO*, the *ACM digital library*, and *Google Scholar*. Search terms included “digital labor platform,” “digital labor market,” “online labor market,” “knowledge work platform,” and “microsourcing,” because digital markets for knowledge work services are a particular type of online labor market. Corresponding queries are consistent with prior reviews (e.g., Wagner and Prester 2019). We conducted additional table-of-content-scans in journals included in the AIS Senior Scholars’ Basket of Journals, covering the last two decades. In addition, we conducted backward searches on all papers in the final sample.

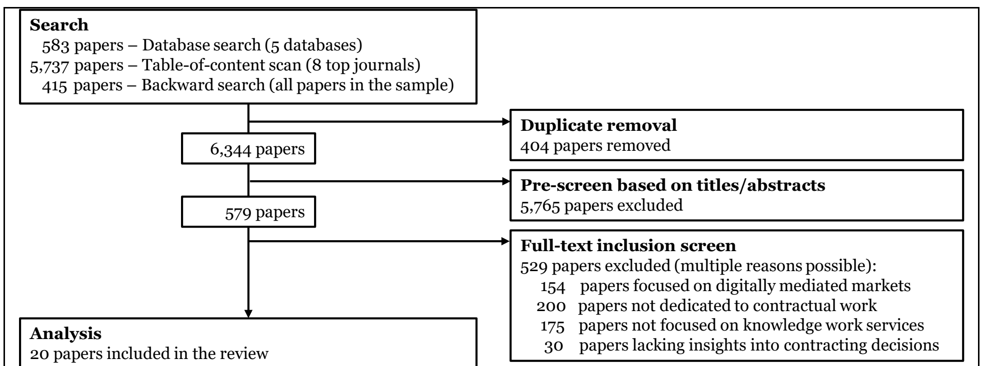

Figure 2. Paper Search and Selection

Overall, 579 papers were retrieved and checked for possible inclusion. We excluded papers that did not analyze digital markets mediated by a platform provider (excluding n=154). We further excluded papers that were not dedicated to contractual work (excluding n=200) and papers that were not focused on knowledge work services (excluding n=175). These criteria mainly excluded research on crowdsourcing contests and clickwork platforms like Amazon Mechanical Turk. The last criterion restricted our sample to papers that provided insights into contracting decisions (excluding n=30). The flow of the papers through the search and screen is illustrated in Figure 2 and the final sample comprises 20 papers (see Table 1).

Coding of Variables and Effects

To code the effects of independent variables on contracting decisions, we proceeded from studies that provided quantified evidence of effect sizes to studies that explored associations qualitatively. The coding was validated in a parallel independent procedure involving both authors, with disagreements reconciled over several meetings (cf. Templier and Paré 2018). All coding results were recorded in a spreadsheet, which was then used to generate aggregated statistics for the analyses.

The primary evidence was compiled by coding the effect sizes of independent variables on dependent variables of contracting decisions. Dependent variables covered contracting decisions on the level of specific workers (explaining why specific workers receive contracts) or on the level of projects (explaining why certain projects lead to a contract). We also included client selection decisions since workers' acceptance of a contract for which they applied (i.e., submitted bids) may be considered a formality in most cases. Next, we coded the independent variables, extracting the variable terms and measurements reported in the paper. In a highly iterative procedure, we compared variables and measures across papers and synthesized them into unified terms for the independent variables. For instance, this involved applying homogeneous terminology of workers (also referred to as contractors, vendors, seller, or freelancers) and clients (also referred to as requester, buyer, job provider, or customer). We also disentangled variables such as *bid rank* and *bid sequence*, which were sometimes used interchangeably, although one referred to the price ranking while the other referred to the time of submission ranking. We grouped over 150 individual variables associated with 67 unified independent variables into overarching categories to provide a concise overview of their relationships. The categories of independent variables include worker characteristics, project characteristics, bid characteristics, market dynamics, client-worker relationship, and client characteristics.

Paper	Sample size	Worker selection	Auction success
Chan and Wang (2014)	194,596	x	
Gefen and Carmel (2008)	263,572	x	
Gong (2017)	NA		x
Guo et al. (2017)	59,054		x
Holthaus and Stock (2018)	1,065	x	
Hong and Zheng (2015)	67,334		x
Hong et al. (2016)	51,887		x
Hong and Pavlou (2017)	117,105	x	
Hong et al. (2018)	442,606	x	
Idowu and Elbanna (2020)	NA	x	
Kabra and Wang (2020)	287,915	x	
Kim (2009)	60,737	x	
Lavilles and Sison (2017)	NA	x	
Liang et al. (2016)	161,994	x	
Liang et al. (2017)	23,438	x	
Liang et al. (2018)	371,968	x	
Schlagwein et al. (2019)	NA	x	
Scholz and Haas (2011)	2,662	x	
Sison and Lavilles (2018)	NA	x	
Zheng et al. (2015)	66,605		x

Table 1. List of Primary Studies

In the next step, we extracted the effect of each independent variable on its dependent variable. We limited the coding exercise to variables with direct relationships and excluded mediator variables. We also checked and reversed scales to ensure comparability (e.g., *cultural similarity* vs *cultural distance*). In some cases, pooling variables with different scales (continuous and categorical) required us to adapt the recorded effects accordingly. Complementary qualitative associations, as well as interaction effects, were recorded separately.

We considered the direction of the effect as well as its effect size during the coding. Conventional vote-counting techniques consider *p*-value thresholds (such as .05) for categorizing effects as *positive*/*negative* (*matters* for categorical variables), and *non-significant* when the study did not report significant evidence (see King and He 2005; Lacity et al. 2010). To gain an overview of the available evidence that is more robust to the *p*-value problem and small but statistically significant effects, we restricted the vote-counting procedures to medium or large effects. As a basis for the coding decisions, we used Lin et al.'s (2013) guidelines as well as further methodological sources, which suggest that the practical significance of an effect should be judged based on the scale of the variable (its standard deviation), the (relative) size of the coefficient, and the magnitude (mean value) of the dependent variable (e.g., Ellis 2010). Within-study and cross-study comparisons of effect sizes were considered to evaluate whether a variable should be classified as having a *small* or a *medium* to *large* (negative, matters, positive) effect. For example, when the effect size of log-transformed *worker ratings* (as the independent variable) is reported as 0.008 in a logit model, this translates into a $100 * \left(e^{0.008 * ln(1.01)} - 1\right) = 0.008\%$ change in the dependent variable for each 1% change in the worker rating. When taking into account a standard deviation of 4 and base rates of worker selection probabilities (the dependent variable) of around 10% in this example, we consider this effect size to be small. With varying base rates (i.e., magnitudes of the dependent variables) within studies predicting worker selection and within studies predicting auction success[2], comparability between studies is limited, and coding decisions on effect sizes cannot rely on the same threshold across studies. Therefore, coding decisions were discussed on a case-by-case basis for the effects of independent variables that could potentially allow for a 1%-10% change in *worker selection* or *auction success* (as the DV). Effects that could not be associated with a 1% change were coded as a separate category of *small* effects, while effects potentially surpassing a 10% change were retained for the vote-counting of effects (*positive*, *matters*, or *negative*). In some cases, insufficient reporting did not allow us to assess the effect size, resulting in a *lack of data* coding.

We further considered qualitative papers. We included studies reporting qualitative evidence by coding whether the paper provides findings for a qualitative association of an independent variable with one of the contracting variables. For example, for the study of Lavilles and Sison (2017), who suggest that ratings and reviews as well as good client relationships enable workers to enter contracts (repeatedly), we coded positive effects for *worker ratings* and *prior working relationships*. We report insights from qualitative associations in a supplementary discussion at the end of the findings section, but exclude them from the principal quantitative findings displayed in the main models.

Data Analysis

Overall, we analyzed 20 papers, with 14 papers reporting quantitative evidence on independent variables affecting contracting decisions, and six papers examining associations qualitatively. We counted the number of times each relationship was studied and calculated the frequency of papers reporting evidence classified as positive, negative, matters, small effect, lack of data, or non-significant. The results were tabulated and integrated into models displaying effects of all independent variables examined (see Figures 3 and 4). Throughout the paper, we report observations on a paper-level, cautioning readers that papers may report results from similar data sets.

Table 2 compares the observations considered in conventional vote-counting to those considered in our robust vote-counting technique. Especially for the dependent variable of worker selection, a considerable number of observations were excluded because effect sizes were small (30 observations) or because missing data did not permit conclusions on effect sizes to be drawn (15 observations). The reduction in the number of observations from 82 to 37 suggests that conventional vote-counting techniques may overstate the conclusiveness of the available evidence in large samples. Differences are less pronounced for independent variables (corresponding observations) of auction success, which may be explained by lower statistical power resulting from smaller sample sizes (61,220 for papers on auction success compared to 175,242 for papers on worker selection).

[2] Several studies did not even report base rates (the average value) of the dependent variable. This further complicates the judgment of effect sizes. In such cases, we relied on average base rates from comparable studies (dependent variable, sampling, empirical context).

Model (DV)	Observations	Conventional vote-counting	Robust vote-counting
Worker selection	**Variable with effect**	82 (86.32%)	37 (38.95%)
	Variable with small effect	*	30 (31.58%)
	Variable lacking data	*	15 (15.79%)
	Variable with no effect	13 (13.68%)	13 (13.68%)
	Total	95 (100.00%)	95 (100.00%)
Auction success	**Variable with effect**	30 (85.71%)	19 (54.29%)
	Variable with small effect	*	4 (11.43%)
	Variable lacking data	*	7 (20.00%)
	Variable with no effect	5 (14.29%)	5 (14.29%)
	Total	35 (100.00%)	35 (100.00%)

Notes. **Bold**: Included in the model. * Not distinguished by conventional vote-counting techniques, i.e., included in the category variables with effects.

Table 2. Observations Considered: Conventional Vote-counting vs Robust Vote-counting

Findings

We organized the principal findings according to the two main dependent variables, worker selection and auction success. In the first part, we integrate the independent variables of **worker selection** (i.e., a contracting decision at the individual worker level), and in the second part, we integrate the independent variables of **auction success** (i.e., a contracting decision at the project level). To provide a concise overview, we group variables in broader categories, as shown in Figures 3 and 4. We summarize the strength and direction of the effects for each variable, distinguishing how many observations were reported as significant with positive, matters, or negative effects, excluded due to a lack of data, reported as not significant, and which effects were coded as small (possibly resulting from deflated *p*-values). Our summaries indicate that the aggregated evidence allows for tentative conclusions when three or more papers examined the variable and more than 60% of the observations were consistently reported as negative, matters, or positive. These thresholds are less restrictive (i.e., lower) compared to prior studies (e.g., on firm-level outsourcing) thereby accounting for the emergent nature of research in our area. Consistent with previous work (e.g., Lacity et al. 2010), the summaries can be considered more conservative conclusions than those suggested in primary studies. At the aggregated level, findings from primary studies are considered as individual and inconclusive data points, which is warranted because primary studies, in the presence of errors such as measurement errors and sampling errors, are always imperfect and may lead to artifactual conclusions (Hunter and Schmidt 2014, chap. 1). At the aggregated level, conclusions should therefore only be proposed when different studies repeatedly report effects in the same direction (King and He 2005; Rosenthal and DiMatteo 2001). We complement the principal findings concerning worker selection and auction success with a summary of the complementary effects concerning interactions and qualitative associations.

Principal Findings on Worker Selection

Worker selection is a binary variable that describes whether an individual worker is offered a contract by the client, i.e., a contracting decision at the individual worker level. It is commonly modeled using logistic regression and estimated in samples restricted to workers who submitted a bid (i.e., not the whole population of workers on the platform). Measures of worker selection include dummy variables indicating whether an applicant is hired for a job (Chan and Wang 2014), a particular bid is a winning bid (Gefen and Carmel 2008), the service provider is selected (Hong and Pavlou 2017), the worker's job application resulted in a job offer (Hong et al. 2018), the freelancer wins the job (Kabra and Wang 2020), the client awards the contract to a bid (Kim 2009), the contractor wins the project (Liang et al. 2016), a bidder is awarded the contract (Liang et al. 2017, 2018), and the acceptance of a bid in online reverse auctions (Scholz and Haas 2011). We pooled studies on client selection (corresponds to a contract offer) and studies on the actual conclusion of a contract with an individual worker because by submitting a bid, workers state their willingness to accept a corresponding offer (Huang et al. 2018).

Overall, 95 individual effects (observations) were extracted from 10 papers. Out of those 95 individual effects, 43 were corrected (i.e., coded as a *small* effect) as part of our robust vote-counting technique. We categorized independent variables into six categories: individual worker characteristics, client-worker relationship, bid characteristics, market dynamics, client characteristics, and project characteristics. A summary model is displayed in Figure 3.

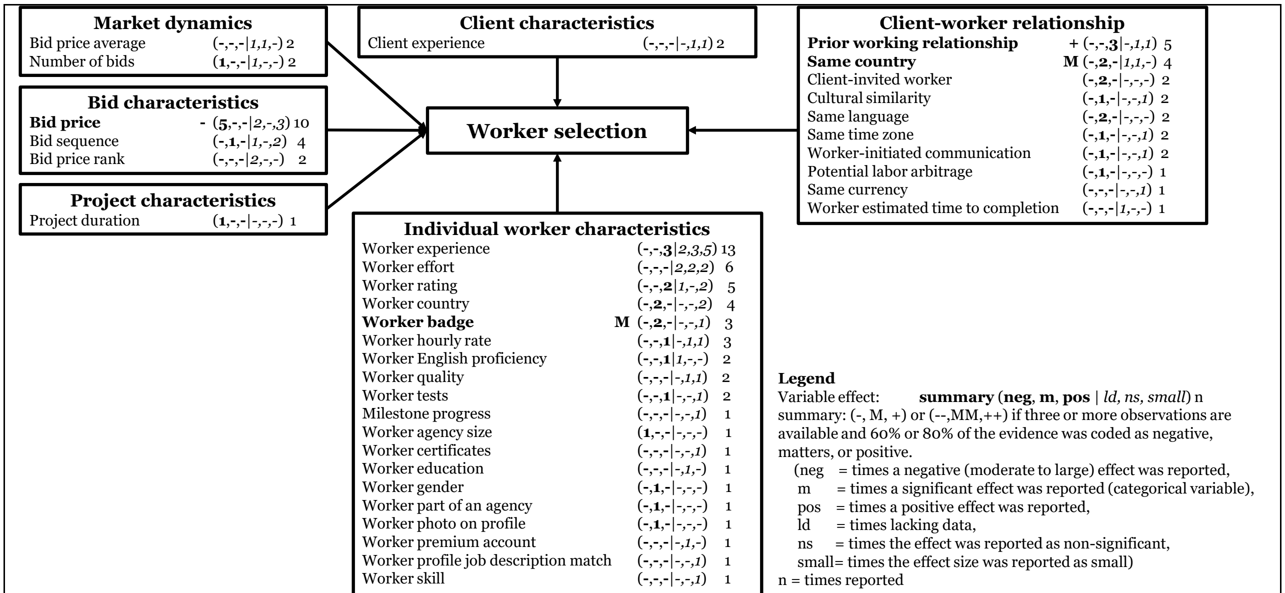

Figure 3. Principal Findings on Worker Selection

Prior research has extensively studied *individual worker characteristics*. This category includes 19 variables that have been examined by nine papers, making it the most studied category among all variables. Frequent classification of observations as *small effects* indicates that *individual worker characteristics* is a category in which findings are particularly prone to deflated *p*-values. Compared to conventional vote-counting techniques, which would identify seven variables as having relatively conclusive effects (five of them with more than 80% of the evidence consistently having positive effects), our more robust analyses indicate that extant research only allows for tentative speculation about variables that affect worker selection (with *worker badge* being the only exception but barely exceeds the two-observation threshold). The most frequently studied individual worker characteristic is concerned with *worker experience* (as measured by the number of projects completed), which has been examined 13 times. With three papers reporting positive effects (two studies lacking data to interpret the effect size and three studies reporting non-significant effects) and five papers reporting small effects, the available evidence does not allow for conclusions on whether *worker experience* substantially improves a worker's chances of being selected. While *worker effort* (as measured by completion rates) has received some attention in prior research (studied six times), no study has found evidence for practically significant effects on worker selection (evidence of the *p*-value problem). The evidence for *worker rating*, *worker country*, and *worker hourly rate* straddles the categories of small vs medium to large effects. It does not exceed thresholds above which the evidence could be considered tentatively conclusive at the aggregated level. For *worker badge*, a feature of some platforms to distinguish high-quality workers, there is initial evidence at the aggregated level for a significant association with worker selection. Individual studies have also included variables such as *worker quality*, *worker tests*, *milestone progress*, *worker agency size*, *worker certificates*, *worker education*, *worker gender*, *worker part of an agency*, *worker photo on profile*, *worker premium account*, *worker profile job description match*, and *worker skill.*

The *client-worker relationship* category includes ten independent variables, which have been studied in eight papers. It is the second most studied category of variables explaining the dependent variable worker selection. The most common variable studied by researchers (examined five times) was a *prior working relationship* between the worker and client, and the preliminary evidence allows tentative conclusion of a practically significant positive effect. This was followed by the variable *same country*, which measures whether or not worker and client are located in the same country. The available evidence indicates that being located in the *same country* substantially increases a worker's probability of being selected. Individual results on *client-invited worker*, *cultural similarity*, *same language*, *same time zone*, *worker-initiated communication*, *potential labor arbitrage*, *same currency*, and *worker estimated time to completion* do not yet allow conclusions to be drawn.

While the *bid characteristic* category has received considerable attention, only three variables that have been considered in ten studies. The *bid price*, which is the price proposed by the worker for the specific project, is one of the most frequently investigated variables overall. It has been studied ten times and negatively associated with worker selection in five papers, allowing a tentative conclusion of a negative effect on worker selection (excluding studies that lack data). *Bid sequence* refers to the position of a worker's bid in the overall bidding sequence, and the available evidence does not yet permit conclusions on its effect. A third variable in this category, *bid price rank*, has only been included in two papers.

The *market dynamics* category is one of the two categories that have been studied for both main dependent variables, worker selection and auction success. Conclusions are not yet possible because none of the variables, *bid price average* and *number of bids* have exceeded the three-observation threshold.

Little research has been conducted on the effects of *client characteristics* and *project characteristics*, which are rarely made public. *Client experience* has been studied in two papers (Hong and Pavlou 2017; Kim 2009), and the *project duration* has been considered in one paper (Kim 2009).

Principal Findings on Auction Success

Auction success is a binary variable that describes whether the client successfully tendered a contract to any worker, i.e., a contracting decision at the project level. It is important to note that the analyses of auction success included in our sample do not analyze contract acceptance decisions conditional upon a contract offer (an auction may fail despite repeated worker selection decisions, i.e., contract offers). Instead, the variable indicates whether a given project successfully leads to a contract. Auction success is commonly modeled using logistic regression. Measures of auction success include contract success or matching (Guo et al. 2017), buyer's contracting decision with one freelancer (Hong and Zheng 2015; Zheng et al. 2015), a variable indicating whether the buyer selected any provider (Hong et al. 2016), and a variable indicating whether a contract is reached (Hong et al. 2016).

Overall, 35 individual effects were extracted from three papers. Out of those 35 effects, 4 were corrected as part of our robust vote-counting technique (seven were excluded due to insufficient data). In the studies, all independent variables categorized as *individual worker characteristics*, *bid characteristics*, and *client-worker relationship* are aggregated across all workers who submitted bids and captured at the level of project contracting decisions. We assembled three categories of independent variables: project characteristics, client characteristics, and market dynamics. A summary model is displayed in Figure 4. It demonstrates that there is limited empirical evidence on the variables explaining auction success.

Project characteristics variables have been examined by four papers (Guo et al. 2017; Hong et al. 2016; Hong and Zheng 2015; Zheng et al. 2015). The evidence supports the conclusion that *project value* is negatively associated with auction success. The evidence on the effect of *auction duration* is not conclusive. Individual studies have also included *auction design (open bids)*, the *number of skills required*, *project codifiability*, *project description length*, *project flexibility*, *project requirements specificity*, and *project type* as variables.

Variables in the *client characteristics* category have been examined by four papers (Guo et al. 2017; Hong et al. 2016; Hong and Zheng 2015; Zheng et al. 2015). The initial evidence indicates that *client experience* may be negatively associated with auction success. This is presumably because inexperienced clients may be more optimistic, and because learning effects may allow clients to better evaluate bids and let auctions fail. Individual studies also covered the variables of *client premium account* and *client ratings*.

Market dynamics variables have been examined in two papers (Hong and Zheng 2015; Zheng et al. 2015). Individual results on *bid arrival dispersion*, *bid arrival rate*, *bid price dispersion*, *number of bids*, *worker experience*, *worker rating average*, and *worker rating dispersion* do not permit conclusions.

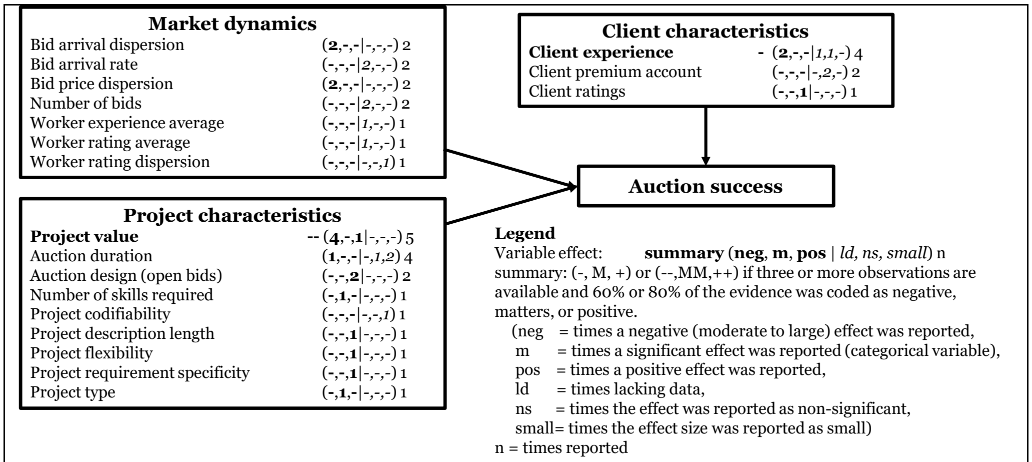

Figure 4. Principal Findings on Auction Success

Complementary Findings on Interaction Effects

Although research has primarily examined the direct effects of independent variables on worker selection and auction success, some studies have also considered interaction effects among independent variables (Hong et al. 2018; Hong and Pavlou 2017; Hong and Zheng 2015; Liang et al. 2016, 2017; Zheng et al. 2015). Because the interacting relationships in these studies were often complex, we report findings at the level of variable categories. Specifically, we summarize two archetype categories of interaction effects.

The first category focuses on the interaction between individual worker characteristics. Researchers have explored interactions between worker characteristics such as *fixed price vs hourly rate contracts*[3] and work effort or worker quality (Liang et al. 2017). The authors tested these interactions in the context of IT-enabled monitoring systems (Liang et al. 2016, 2017), which, after their introduction, caused clients to be less willing to pay higher rates for higher quality workers, suggesting a partial substitution relationship between the monitoring system and standard reputation systems. Further, interaction effects with variables describing the client-worker relationship have been reported. Initial evidence suggests that higher worker ratings can balance the adverse effects of language differences between client and worker (Hong and Pavlou 2017). In addition, worker-initiated communication, or a worker's attempt to establish a relationship with the client before the selection decision, may overcome some of the negative effects of workers' low platform experience or reputation (Hong et al. 2018). Ultimately, as a key worker characteristic, worker ratings may even create feedback effects on recurring and parallel contracts with clients (Du and Mao 2018; Hong and Pavlou 2017).

The second category is concerned with the interaction between different market dynamics. It is the only category of interaction effects that has been examined for the dependent variable auction success. The specific interaction that is significant in this category is between the number of bids and the bids' price dispersion. Initial evidence suggests that the negative effect of price dispersion on auction success decreases when a project receives more bids, presumably because the client is better able to infer the common value of a project (Hong and Zheng 2015; Zheng et al. 2015).

---

[3] The variable *fixed price vs hourly rate contracts* does not occur in the model summaries because Liang et al. (2017) include them as interacting variables without reporting their direct effects.

Complementary Qualitative Findings

We further considered 23 qualitative associations from five studies on worker selection. Most research has focused on *individual worker characteristics*, which have been examined 19 times. The focus of qualitative studies on worker selection settings is not surprising considering the ease of access to individual workers that digital platforms offer to recruit interviewees. In line with the quantitative evidence, the most frequently studied individual worker characteristic is *worker experience*, with qualitative evidence suggesting a positive association with worker selection (examined four times). This was followed by the variable *worker rating*, which has been studied four times and found to be positively associated with worker selection, except for the Holthaus and Stock (2018) study. Individual results on *worker skill*, *worker availability*, *worker country*, *worker education*, *worker English proficiency*, *worker gender*, *worker part of an agency*, *worker tests*, *worker trait extraversion*, and *worker trait openness* do not yet permit conclusions. Variables from other categories that have been explored include *bid price* in the *bid characteristics* category as well as *prior working relationship* and *promotion by worker* in the *client-worker relationship* category. Finally, in the *market dynamics* category, one study (Gong 2017) considers four associations related to the variables of *expectation discrepancy rate*, *size of overall workforce*, *worker experience average* and *worker skill average*.

Avenues for Future Research

The models displayed in Figures 3 and 4 integrate prior research on worker selection and auction success in digital markets for knowledge work services, and thereby offer a foundation for future research. They facilitate a shared understanding of the independent variables and their overarching categories, which should be useful to justify the variables considered in prospective studies. Since the variable names were unified by comparing and integrating terminology across multiple studies, they contribute to establishing a consistent vocabulary. Based on this synthesis, we suggest future studies to avoid terminological ambiguity, such as the *bid order* (which may refer to the order of bids according to time of submission or to bid prices) and instead refer to *bid price rank* and *bid sequence*. Similarly, we recommend that confusion of simple variables with theoretical constructs should be avoided, for example, by distinguishing the variable of *worker ratings* from latent constructs of *worker quality*. To account for the challenges of deflated p-values in primary studies, it is imperative to discuss the practical significance of individual findings, as recommended by Lin et al. (2013). Beyond more consistent terminology and consideration of practical significance, more complete reporting of correlation matrices is important to facilitate future efforts of aggregating empirical findings and achieving a more cumulative progression of knowledge development in this area.

More in-depth conceptual and theoretical grounding is another avenue to further strengthen extant research, which has primarily adopted data-driven approaches and provided many valuable contributions based on empirical identification designs suitable to assess causality (including difference-in-differences, instrumental variables, and fixed-effects regression). In this regard, extant research offers a strong empirical foundation for working towards theory. Considering the early stage of exploratory work, we contend that going beyond simple variables and developing more abstract constructs (Barki 2008) would benefit the maturing research and the shift towards a theory-grounded investigation. With almost no surveys in our sample and potentially correlated variables (e.g., *worker country* and *English proficiency*), establishing constructs and assessing their convergent and discriminant validity would allow this research stream to overcome significant empirical limitations. Future research can build on promising examples, such as the use of text analysis and natural language processing for measurement (e.g., Hong et al. 2018; Hong and Pavlou 2017; Kabra and Wang 2020) or principal component analysis for dimension reduction (e.g., Liang et al. 2016, 2017).

Another important avenue is the platform intermediary's impact on contracting decisions. Our findings and the models displayed in Figures 3 and 4 may suggest the misleading conclusion that sourcing and contracting decisions on digital markets for knowledge work services exclusively result from an exchange between the individual parties. In this line of reasoning, the third-party platform intermediary, who implements trust mechanisms (Du and Mao 2018) and coordinates the exchange in a value network (Stabell and Fjeldstad 1998), may be considered as a boundary condition. For many empirical studies, excluding platform-related variables from consideration may be necessary due to the relatively invariant nature of platform design and governance. However, the vital role of platform providers may provide a starting point to explore how variation in control and coordination systems as well as platform centralization, as proposed by Gol et al. (2019), affect contracting decisions. We encourage future studies to explore platform-related variables of contracting decisions by analyzing differences in how platform providers guide interactions and regulate the flow of information and incentives. Empirically, corresponding research may rely on cross-platform comparisons (see Gol et al. 2019) or analyses of how changes initiated by the platform operator affect the exchange (see Liang et al. 2017, 2018). Ultimately, such work promises to produce insights into the design and governance of digital service markets.

Figures 3 and 4 facilitate the identification and exploration of gaps implied by comparing the independent variables of both models. Given that the independent variables of worker selection and auction success operate at different levels of aggregation, this suggests several directions for future research. For instance, relationship characteristics have rarely been considered an independent variable of auction success despite their significant role in explaining worker selection. Therefore, prospective studies may consider aggregated versions of *prior working relationship* or *same country* (e.g., the number of bids from workers who have successfully completed contracts for the client or who live in the same country) as an independent variable of auction success. Furthermore, practically relevant questions, such as "Does inviting workers help to ameliorate the low contracting rates reported by Snir and Hitt (2003)?" remain to be addressed. Reversely, independent variables of auction success, most notably project characteristics described in the call-for-bids, have rarely been studied when explaining worker selection. There is a general lack of understanding about whether and how worker selection decisions differ depending on the type of clients. Figure 3 covers a plethora of worker-related variables but almost no evidence on client characteristics. In this regard, the insightful analysis of different types of clients offered by Radkevitch et al. (2009) may provide a valuable starting point. Finally, digital markets for knowledge work services provide an interesting context to analyze feedback-loops. Project lengths on these markets occupy a middle ground between micro-tasks, for which repeatability is possible at negligible cost (Wang et al. 2017), and inter-firm outsourcing arrangements in which the length of the outsourcing relationships can outlast the tenure of agents at the firm (Ring and Van de Ven 1994).

Conclusion

Our work makes three contributions to the emergent literature on digital markets for knowledge work services. First, our work provides the first integrated models of contracting decisions in digital markets for knowledge work services. By systematically collecting prior studies and integrating them into two concise models, we offer a consolidated view of the evolving research stream and the aggregated findings of individual studies. Second, we recognize that conventional vote-counting techniques would propagate the deflated *p*-value problem (i.e., the statistical significance of practically irrelevant effects in studies with large samples and excessive statistical power), when integrating findings from large-scale studies like those included in our sample. Instead of adding small effects with little or no practical significance to the same tally as the main effects, we propose a robust vote-counting approach. Based on an individual assessment of each reported effect size, robust vote-counting is restricted to medium to large effects and provides an aggregated view of the evidence that reflects practically significant effects and attenuates the *p*-value problem. We believe that this approach is applicable in further related areas in which large-scale studies lead to a propagation of *p*-value problems when applying conventional vote-counting approaches. It seems reasonable to update methodological recommendations accordingly and caution researchers to consider average sample sizes when selecting conventional or robust vote-counting approaches. Third, we derive specific recommendations for future research and organize them into a research agenda. Considering the challenges and gaps identified in this agenda, our work suggests that there is much to be learned beyond the strides made by prior research.

Our results must be interpreted in light of limitations that are primarily related to the initial effort of integrating emerging research. In pooling findings from heterogeneous studies, our models may be accused of comparing apples and oranges, especially when considering the homogeneous nature of empirical research conducted on mature models. Our study further remains at the level of (robust) vote-counting (King and He 2005), with coding decisions of reported effect sizes as practically insignificant (i.e., *small*), relying on careful judgment of the authors. While conventional vote-counting studies have explored the possibility of combining quantitative and qualitative evidence in the same analysis (e.g., Lacity et al. 2010), we decided to present them separately because detailed corrections according to effect sizes further contribute to uneven treatment of evidence, which is already hard to compare. Concerning the quantitative evidence, more advanced meta-analytic techniques would be capable of considering further characteristics of the evidence (including sample sizes, effect sizes, and construct validity) and addressing additional challenges (including publication bias) while requiring substantial restrictions on the size of the analyzed sample.

Practical implications for workers and clients can be derived from Figures 3 and 4. To improve their contracting rate to a practically significant degree, workers have a few options. As expected from an economic perspective of auction theory, lower bid prices are associated with higher contracting rates. When establishing their profile on the platform, workers can increase their selection rates by pursuing repeated contracts with regular clients (in line with the effect of *prior working relationship*). Finally, another option that does not necessarily jeopardize revenues is to earn a *worker badge* from the platform, signaling superior skills or quality to prospective clients. For clients, extant research offers fewer insights. The evidence confirms that contracts are closed less frequently when the *project value* is high. The finding that a lack of *client experience* is associated with higher rates of contracts may indicate that clients tend to be less cautious in closing contracts when they are new to the platform with unsatisfactory experiences, possibly driving the negative effect of *client experience* through attrition. More comprehensive recommendations on how to achieve practically significant improvements in worker selection rates and auction success rates will require further primary research.

Acknowledgments

We thank Dr. Emrah Yasasin and Dr. Malshika Dias for their feedback on an earlier version of the paper, the anonymous reviewers and the associate editor for their constructive comments.

References

Adler, P. S. 2001. "Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism," *Organization Science* (12:2), pp. 215–234.

Barki, H. 2008. "Thar's Gold in Them Thar Constructs," *ACM SIGMIS Database* (39:3), pp. 9–20.

Boell, S. K., and Cecez-Kecmanovic, D. 2014. "A Hermeneutic Approach for Conducting Literature Reviews and Literature Searches," *Communications of the Association for Information Systems* (34:12), pp. 257–286.

Chan, J., and Wang, J. 2014. "Hiring Biases in Online Labor Markets: The Case of Gender Stereotyping," in *Proceedings of the International Conference on Information Systems*, Auckland, New Zealand.

Chen, D. L., and Horton, J. J. 2016. "Research Note - Are Online Labor Markets Spot Markets for Tasks? A Field Experiment on the Behavioral Response to Wage Cuts," *Information Systems Research* (27:2), pp. 403–423.

Clemens, M. A. 2011. "Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?" *Journal of Economic Perspectives* (25:3), pp. 83–106.

Du, W. D., and Mao, J.-Y. 2018. "Developing and Maintaining Clients' Trust Through Institutional Mechanisms in Online Service Markets for Digital Entrepreneurs: A Process Model," *The Journal of Strategic Information Systems* (27:4), pp. 296–310.

Ellis, P. D. 2010. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*, (1st ed.), Cambridge University Press.

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

Gol, E. S., Stein, M.-K., and Avital, M. 2019. "Crowdwork Platform Governance Toward Organizational Value Creation," *The Journal of Strategic Information Systems* (28:2), pp. 175–195.

Gong, J., Hong, Y., and Zentner, A. 2018. "Role of Monetary Incentives in the Digital and Physical Inter-Border Labor Flows," *Journal of Management Information Systems* (35:3), pp. 866–899.

Gong, Y. 2017. "Estimating Participants for Knowledge-Intensive Tasks in a Network of Crowdsourcing Marketplaces," *Information Systems Frontiers* (19:2), pp. 301–319.

Guo, X., Gong, J., and Pavlou, P. 2017. "Call for Bids to Improve Matching Efficiency: Evidence from Online Labor Markets," in *Proceedings of the International Conference on Information Systems*, Seoul, South Korea.

Higgins, J. P. T., and Green, S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions*, (Vol. 5), John Wiley & Sons.

Holthaus, C., and Stock, R. M. 2018. "Facts Vs. Stories-Assessment and Conventional Signals as Predictors of Freelancers' Performance in Online Labor Markets," in *Proceedings of the Hawaii International Conference on System Sciences*, Hawaii, USA.

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), pp. 547–562.

Hong, Y., Peng, J., Burtch, G., and Huang, N. 2018. "Role of Communication in Online Platforms," in *Proceedings of the Americas Conference on Information Systems*, New Orleans, LA.

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

Hong, Y., and Zheng, A. Z. 2015. "On Service Employers' Hiring Decisions in Online Labor Markets: A Perspective of Price and Quality Discovery," in *Proceedings of the Americas Conference on Information Systems*, Puerto Rico.

Hong, Z., Zhu, H., and Dong, K. 2020. "Buyer-Side Institution-Based Trust-Building Mechanisms: A 3S Framework with Evidence from Online Labor Markets," *International Journal of Electronic Commerce* (24:1), pp. 14–52.

Huang, N., Burtch, G., Hong, Y., and Pavlou, P. A. 2020. "Unemployment and Worker Participation in the Gig Economy: Evidence from an Online Labor Market," *Information Systems Research* (31:2), pp. 431–448.

Huang, N., Burtch, G., and Pavlou, P. 2018. "Local Economic Conditions and Worker Participation in the Online Gig Economy," in *Proceedings of the International Conference on Information Systems*, San Francisco, CA.

Hunter, J. E., and Schmidt, F. L. 2014. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*, (2nd ed.), Thousand Oaks, CA.

Idowu, A., and Elbanna, A. 2020. "Digital Platforms of Work and the Crafting of Career Path: The Crowdworkers' Perspective," *Information Systems Frontiers* (forthcoming).

Jap, S. D. 2002. "Online Reverse Auctions: Issues, Themes, and Prospects for the Future," *Journal of the Academy of Marketing Science* (30:4), pp. 506–525.

Kabra, A., and Wang, Q. 2020. "Hiring in Online Labor Markets: The Role of Job-Specific Experience," in *Proceedings of the Pacific Asia Conference on Information Systems*, Dubai, UAE.

Kanat, I., Hong, Y., and Raghu, T. S. 2018. "Surviving in Global Online Labor Markets for IT Services: A Geo-Economic Analysis," *Information Systems Research* (29:4), pp. 893–909.

Khan, N., and Sikes, J. 2014. "IT Under Pressure: McKinsey Global Survey Results," *McKinsey & Company* (March).

Kim, J. Y. 2009. "Online Reverse Auctions for Outsourcing Small Software Projects: Determinants of Vendor Selection," *e-Service Journal* (6:3), pp. 40–55.

King, W. R., and He, J. 2005. "Understanding the Role and Methods of Meta-Analysis in IS Research," *Communications of the Association for Information Systems* (16:32), pp. 665–686.

Kokkodis, M. 2020. "Dynamic, Multi-Dimensional, and Skillset-Specific Reputation Systems for Online Work," *Information Systems Research* (forthcoming).

Lacity, M. C., Khan, S. A., Yan, A., and Willcocks, L. P. 2010. "A Review of the IT Outsourcing Empirical Literature and Future Research Directions," *Journal of Information Technology* (25:4), pp. 395–433.

Lavilles, R. Q., and Sison, R. C. 2017. "A Thematic Analysis of Software Developers' Experience in Online Sourcing Marketplaces," in *Proceedings of the Pacific Asia Conference on Information Systems*, Langkawi, Malaysia.

Liang, C., Hong, Y., and Gu, B. 2016. "Effects of IT-Enabled Monitoring Systems in Online Labor Markets," in *Proceedings of the International Conference on Information Systems*, Dublin, Ireland.

Liang, C., Hong, Y., and Gu, B. 2017. "Moral Hazards and Effects of IT-Enabled Monitoring Systems in Online Labor Markets," in *Proceedings of the Hawaii International Conference on System Sciences*, Hawaii, USA.

Liang, C., Hong, Y., and Gu, B. 2018. "Home Bias in Hiring: Evidence from an Online Labor Market," in *Proceedings of the Pacific Asia Conference on Information Systems*, Yokohama, Japan.

Lin, M., Lucas Jr, H. C., and Shmueli, G. 2013. "Research Commentary-Too Big to Fail: Large Samples and the p-Value Problem," *Information Systems Research* (24:4), pp. 906–917.

Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D. F., Humpherys, S. L., Barlow, J. B., and Wilson, D. W. 2013. "Evaluating Journal Quality and the Association for Information Systems Senior Scholars'

Journal Basket via Bibliometric Measures: Do Expert Journal Assessments Add Value?" *MIS Quarterly* (37:4), pp. 993–1012.
Majchrzak, A., and Malhotra, A. 2013. "Towards an Information Systems Perspective and Research Agenda on Crowdsourcing for Innovation," *The Journal of Strategic Information Systems* (22:4), pp. 257–268.
Manyika, J., Lund, S., Robinson, K., Valentino, J., and Dobbs, R. 2015. "A Labor Market That Works: Connecting Talent with Opportunity in the Digital Age," *McKinsey Global Institute* (20).
Mertens, W., and Recker, J. 2020. "New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS Research," *Journal of the Association for Information Systems* (21:4), p. 1.
Paré, G., Trudel, M.-C., Jaana, M., and Kitsiou, S. 2015. "Synthesizing Information Systems Knowledge: A Typology of Literature Reviews," *Information & Management* (52:2), pp. 183–199.
Radkevitch, U., Van Heck, E., and Koppius, O. 2009. "Portfolios of Buyer–Supplier Exchange Relationships in an Online Marketplace for IT Services," *Decision Support Systems* (47:4), pp. 297–306.
Ring, P. S., and Van de Ven, A. H. 1994. "Developmental Processes of Cooperative Interorganizational Relationships," *Academy of Management Review* (19:1), pp. 90–118.
Roach, S. 2003. "Coping with the Global Labor Arbitrage," in *Morgan Stanley Global Economic Forum*.
Rosenthal, R., and DiMatteo, M. R. 2001. "Meta-Analysis: Recent Developments in Quantitative Methods for Literature Reviews," *Annual Review of Psychology* (52:1), pp. 59–82.
Sawyer, A. G., and Peter, J. P. 1983. "The Significance of Statistical Significance Tests in Marketing Research," *Journal of Marketing Research* (20:2), pp. 122–133.
Schlagwein, D., Cecez-Kecmanovic, D., and Hanckel, B. 2019. "Ethical Norms and Issues in Crowdsourcing Practices: A Habermasian Analysis," *Information Systems Journal* (29:4), pp. 811–837.
Scholz, M., and Haas, N. 2011. "Determinants of Reverse Auction Results: An Empirical Examination of Freelancer.com," in *Proceedings of the European Conference on Information Systems*, Helsinki, Finland.
Schryen, G., Wagner, G., Benlian, A., and Paré, G. 2020. "A Knowledge Development Perspective on Literature Reviews: Validation of a New Typology in the IS Field," *Communications of the Association for Information Systems* (46), pp. 134–168.
Sison, R., and Lavilles, R. 2018. "Software Gigging: A Grounded Theory of Online Software Development Freelancing," in *Proceedings of the International Conference on Information Systems*, San Francisco, CA.
Snir, E. M., and Hitt, L. M. 2003. "Costly Bidding in Online Markets for IT Services," *Management Science* (49:11), pp. 1504–1520.
Soergel, A. 2016. "1 in 3 Workers Employed in Gig Economy, but Not All by Choice," *U.S. News and World Report*.
Stabell, C. B., and Fjeldstad, Ø. D. 1998. "Configuring Value for Competitive Advantage: On Chains, Shops, and Networks," *Strategic Management Journal* (19:5), pp. 413–437.
Templier, M., and Paré, G. 2018. "Transparency in Literature Reviews: An Assessment of Reporting Practices Across Review Types and Genres in Top IS Journals," *European Journal of Information Systems* (27:5), pp. 503–550.
Trost, A. 2014. *Talent Relationship Management: Competitive Recruiting Strategies in Times of Talent Shortage*, Springer Science & Business.
Wagner, G., and Prester, J. 2019. "Information Systems Research on Platforms for Digital Knowledge Work: A Scoping Review," in *Proceedings of the International Conference on Information Systems*, Munich, Germany.
Wang, J., Ipeirotis, P. G., and Provost, F. 2017. "Cost-Effective Quality Assurance in Crowd Labeling," *Information Systems Research* (28:1), pp. 137–158.
Williamson, O. E. 1991. "Strategizing, Economizing, and Economic Organization," *Strategic Management Journal* (12:S2), pp. 75–94.
Yang, H., and Tate, M. 2012. "A Descriptive Literature Review and Classification of Cloud Computing Research," *Communications of the Association for Information Systems* (31:2), pp. 35–60.
Zheng, A., Hong, Y., and Pavlou, P. 2015. "Value Uncertainty and Buyer Contracting: Evidence from Online Labor Markets," in *Proceedings of the International Conference on Information Systems*, Fort Worth, Texas.