Trust and Disintermediation: Evidence from an Online Freelance Marketplace

Grace Gu
Harvard Business School
Boston, MA 02163
ygu@hbs.edu

Feng Zhu
Harvard Business School
Boston, MA 02163
fzhu@hbs.edu

Abstract

As an intermediary improves trust between the two sides of its market to facilitate matching and transactions, it faces an increased risk of disintermediation: with sufficient trust, the two sides may circumvent the intermediary to avoid the intermediary's fees. In this paper, we investigate the relationship between increased trust and disintermediation by leveraging a randomized control trial in an online freelance marketplace. We find that enhanced trust increases the likelihood of high-quality freelancers being hired. However, when the trust level is sufficiently high, it also increases disintermediation, which offsets the revenue gains from the increase in hiring high-quality freelancers. We also identify heterogeneity across clients and freelancers in their tendencies to disintermediate. We discuss strategies that intermediaries can use to mitigate the tension between trust building and disintermediation.

Key words: disintermediation, intermediary, trust, online marketplace

1. Introduction

Intermediaries are everywhere in our economy: brokers in the finance and insurance industries, headhunters in the labor market, distributors in retail, housing agents in real estate, and online platforms in the information technology industry, just to name a few. In 2010, intermediaries contributed an estimated 34% of the US gross domestic product (Spulber 2011). Economists have long recognized the importance of intermediaries for providing matching and facilitating transactions (e.g., Parker and Van Alstyne 2005, Armstrong 2006, Rochet and Tirole 2006, Edelman and Wright 2015, Hagiu and Wright 2015). However, all intermediaries face the risk of disintermediation, in which two sides circumvent the intermediary to transact directly and avoid the intermediary's fees.

Disintermediation is prevalent. For example, the traditional role of book publishers as intermediaries was weakened when Amazon enabled authors to sell directly to readers through its self-publishing services. Li & Fung, a supply-chain management company that connects global retail brands with Chinese manufacturers, suffered ongoing decline in revenue as retailers disintermediated to work with manufacturers directly. A survey by ZBJ.com, the largest online freelance marketplace in China, indicates that approximately 90% of transactions are conducted outside the platform after clients and freelancers have been matched on its platform (Zhu et al. 2018). Hotels and airlines offer incentives to lure customers to book directly with them, thereby shrinking the revenue for online travel agencies. A few intermediaries have been unable to sustain their businesses as a result of disintermediation. For example, online platforms such as Homejoy went out of business because their consumers transacted with service providers outside the platforms.

Despite the importance of disintermediation with regard to firms' strategies and survival, there is scant literature on the issue, perhaps because of the difficulty of observing and measuring disintermediated transactions. In this paper, we study the relationship between trust and disintermediation, leveraging a randomized control trial (RCT) in an online platform. We utilize conversations recorded online to provide direct evidence of users' intentions to disintermediate. A number of studies have shown that building trust between two sides is crucial for platforms to facilitate effective matching among users (e.g., Resnick and Zeckhauser 2002). However, significant trust can reduce the perceived importance of platform services such as monitoring transactions, escrow payments, dispute settlements, and refunds for failed transactions (e.g., Edelman and Hu 2016). Users who trust one another feel less of a need for such services and are incentivized to take their transactions off the platform to avoid intermediary fees.

Our research setting is a large outsourcing platform that enables clients to find freelancers who satisfy the clients' job requirements. The platform then provides features through which the two sides contract, collaborate, create invoices, and pay, and it charges a per-transaction service fee that is approximately 10% of each transaction's value. In an RCT, the platform provider shows freelancers' satisfaction scores (SSs) to a random sample of clients. SSs are a newly developed measure of a freelancer's business reputation based on his or her complete work history on the platform. We find that the enhanced trust derived from seeing high SSs increases the likelihood for a high-quality freelancer to be hired. However, it also increases disintermediation between clients and freelancers with high SSs, as evidenced by significantly lower charges, fewer hours reported, and the stronger intention to disintermediate expressed in chat messages between them.

We conduct robustness checks to confirm that the reduction in hours and total charges is driven by disintermediation and not by other factors such as clients' selections of more efficient freelancers. Ultimately, increased disintermediation offsets the revenue gains from the increased hiring of high-quality freelancers. We also find that the tendency to disintermediate increases when users are geographically collocated, jobs are easily divisible, and clients themselves have high ratings.

Our study is related to the literature on trust within online platforms. Existing literature emphasizes the importance of building trust for successful business transactions among strangers online (e.g., Strader and Ramaswami 2002, Dellarocas, 2003, Pavlou and Dimoka 2006, Jin and Kato 2006, Cabral and Hortaçsu 2010, Cai et al. 2013, Moreno and Terwiesch 2014). Although many factors may influence trust—such as an individual's past experiences transacting with the same person, escrow services by the platform, and certifications from trusted parties—in most studies, trust is reduced to the use of reputation systems (e.g., ter Huurne et al. 2017). Many experimental and observational studies provide evidence that reputation systems can effectively enhance trust (e.g., Ba and Pavlou 2002, Resnick and Zeckhauser 2002, Bohnet and Huck 2004, Pavlou and Dimoka 2006, Utz et al. 2009, Charness et al. 2011, Bolton et al. 2013). Such systems operate by mitigating the uncertainty involved in transactions, which stems from information asymmetry and potential opportunism between two transacting parties (e.g., Pavlou et al. 2007).

Further, a number of studies have identified the shortcomings of existing reputation systems in building trust. Research on eBay's feedback mechanism has shown that disappointed buyers often do not leave feedback (Nosko and Tadelis 2015), while buyers who have had a good experience are more likely to leave feedback (Masterov et al. 2015). Bolton et al. (2013) argue that reciprocity in providing feedback distorts reputation information, as fear of retaliation may deter users from truthful reporting. Even in a simultaneous-reveal system, in which reviews are not revealed until both parties have submitted their ratings, users may still be reluctant to provide negative feedback if they suspect it would discourage other parties from transacting with them (Luca 2017). These factors collectively result in reputation inflation. Ert et al. (2015) find that 97% of Airbnb ratings are between 4.5 and 5 stars; consequently, online reviews have no effect on the prices of Airbnb listings. Horton and Golden (2015) find similar patterns using data from Upwork.

Scholars have proposed various ways to develop better reputation measures in order to improve trust and transaction quality (e.g., Hui et al. 2014, Kapoor and Tucker 2017, Dai et al. 2018). However, our study suggests that under the disintermediation threat, attempts at enhancing the trust by providing a more accurate reputation system may actually harm the platform's ability to capture value.

Our study also improves our understanding of how different types of trust affects users' behavior. Bapna et al. (2017), building on Sobel (2005) and Cabral et al. (2014), categorize trust into intrinsic trust and instrumental trust. Intrinsic trust is motivated by the psychological benefits that an individual derives from being kind to others, while instrumental trust is backed by the option of rewarding and punishing a trustee in the future. Consistent with instrumental trust, we find evidence that even after clients and freelancers choose to disintermediate, they still prefer to initiate jobs on the platform so that they have the option to provide each other feedback or seek help from the platform.

Our study also adds to the small literature on disintermediation. Waldfogel (2012), Waldfogel and Reimers (2015), and Peukert and Reimers (2018) are three related studies that study the impact of digital disintermediation on product variety and quality in the music and book publishing industries. A few papers in the supply-chain management literature examine supplier encroachment as a means of circumventing intermediaries (e.g., Arya et al. 2007). However, none of these studies discuss how the degree of disintermediation changes with increased trust.

The remainder of the paper is organized as follows. Section 2 describes our empirical setting and design. Section 3 describes the data and variables. Sections 4 and 5 present empirical results and various robustness checks. Section 6 concludes by discussing limitations and managerial implications of this study.

2. Background and Empirical Design

The empirical context of our study is a major online outsourcing platform. A number of studies have examined the value of such platforms in online hiring (e.g., Agrawal et al. 2016, Stanton and Thomas 2016) and for conducting experiments (e.g., Horton et al. 2011). Jobs posted on the platform encompass a wide range of categories, such as Web, Mobile & Software Development, Design & Creative, Translation, Administrative Support, Accounting & Consulting, Writing, and Customer Service.

As soon as a client posts a job, a *job opening* is created, which typically includes a name, work description, requirements, and deadline. Any freelancer can submit a proposal to the client to bid for the job. Once the client selects a freelancer, the job is *filled* and a service contract (referred to as a *job assignment* hereafter) is created. A job assignment remains active until both parties agree to close it. Figure 1 illustrates the process flow of a typical job on the platform.

Clients can post either fixed-price or hourly jobs. The price for a fixed-price job is negotiated and determined between a client and a freelancer at the time of contracting; they can agree that the client will pay the total amount upon project completion or pay in stages according to agreed-upon milestones. For an hourly job, an hourly rate is decided at the time of contracting. Thereafter, the freelancer can begin working on the job and record working hours. After the freelancer makes a request for payment, the platform charges the client and holds the payment in escrow. The client has four days to review and dispute the amount. Once the dispute period ends, the escrow fund is released to the freelancer.

The platform charges freelancers a service fee of approximately 10% of the amount billed to the client. Disintermediation can take place in two ways. First, clients and freelancers can "chat" with each other on the platform at any time. Thus, they could agree to take jobs off the platform before initiating any projects to avoid the service fee. Second, durative transactions enable them to begin part of the job on the platform and then disintermediate for the remainder of the job to reduce the service fee. The latter approach to disintermediation enables clients and freelancers to leave each other reviews after they mark the job as complete. Our study focuses on the latter scenario.

Since its founding, the platform has used a basic five-star rating system to reflect user satisfaction. Star ratings are shown for both clients and freelancers on their user profiles, which reflect the average of all ratings received from completed jobs. This system experiences the shortcomings documented in the literature. First, the average rating from clients' past reviews does not take into account jobs for which no rating was given, nor does it allow for weight differentiation between older ratings and more recent ratings (e.g., Dai et al. 2018). Second, such a system may encourage reciprocity of positive reviews. Consequently, the average rating of freelancers on the platform is very high (above 4.5 out of 5). Thus, ratings do not accurately identify high- and low-quality freelancers. Finally, the five-star rating system rewards freelancers who have completed a large number of small and short-term projects and disfavors those who have worked on extensive and long-term projects.

For these reasons, the platform designed SSs as a new measure of freelancers' reputation, which represents a more complete picture of a freelancer's business. To avoid strategic manipulation by freelancers, the company does not disclose how the score is determined. However, the company does make it explicit that in addition to the ratings of past jobs, the SSs capture the following information that five-star ratings do not capture: 1) private feedback that clients have provided to the platform,[1] 2) disputes that freelancers have had with past clients, and 3) the number of times clients chose not to provide ratings. While clients can observe point 3) by browsing a freelancer's past work history, clients have no means of obtaining the first two information sets before the introduction of SSs. Research has shown that private ratings can help overcome concerns of retaliation and reciprocity and, thus, encourage clients to reveal truthful information (e.g., Luca 2017). Disputes frequently lead to project cancellation, in which case, the cancelled projects are not included in freelancers' work history, and neither clients nor freelancers leave reviews for each other. However, this information can be very valuable to clients for evaluating the risks associated with working with a given freelancer. Because of these advantages over five-

[1] Private feedback is collected at the same time as public feedback when the job is closed, and both parties are informed that the private ratings are visible only to the platform.

star systems, a high SS can significantly boost a client's confidence that the freelancer can successfully complete the job.[2]

We leverage an RCT that the platform conducted from February 13 to March 10, 2015. It included a random sample (approximately 3%) of registered clients on the platform. The randomization is at the client level, and 50% of the sample clients were selected as the treatment group. When clients in the treatment group logged on to the website and browsed for freelancers, they were shown an SS in addition to the star rating on each freelancer's profile, while clients in the control group saw only the star ratings (Figure 2). The platform did not disclose the use of SSs to clients in the control group or to any freelancers. Moreover, freelancers were unable to observe their own SSs. Given the short duration of the experiment, we expect information leakage to be negligible. In addition, because freelancers were not notified of the introduction of SSs nor could they observe their own scores, we could disregard the possibility of signaling in their job applications. Figure 3 illustrates the distribution of SSs in the assignment sample. We collect data for all jobs initiated during the trial, even if a job is completed after the trial ends.

3. Data and Variables

We collect all job openings and assignments created by 24,732 clients from the treatment group and 24,458 clients from the control group during the trial. The analysis sample is at the job-assignment level, consisting of each assignment's outcome and the characteristics of the corresponding client and the hired freelancer. Jobs that were not filled or observations in which the freelancer's SS was not available are dropped, leaving a final sample of 33,561 job assignments.[3]

We employ two approaches to measure disintermediation. As an indirect approach, we identify disintermediation using job outcomes that might imply jobs that ended prematurely or with partial payment. We collect the number of working hours for each assignment and denote it as *Hours*. For this variable, we have observations only for hourly jobs. Similarly, jobs with small payments could signal disintermediation if clients paid only a small amount on the platform and

[2] SSs are only implemented for freelancers, because providing better services to clients is a priority for this platform.
[3] By design, an SS is only available after a freelancer has sufficient historical job data on the platform, which is usually after five projects or having worked with at least three clients. Freelancers without SSs appear the same to clients in the treatment and control groups and thus are not considered part of the study.

conducted most of the transaction off-platform. We use *Total_Charge* to represent the total amount paid on the platform once a job is closed.

As a direct approach to measuring disintermediation, we quantify clients' and freelancers' intentions to disintermediate by leveraging a text-analysis tool the platform developed to detect sensitive words that imply disintermediation. The list of sensitive words along with their relative weights was developed by the company based on extensive data from past transactions. Over the years, the platform refined its algorithm and dictionary based on data collected from actual disintermediation and from its other trials that were aimed at deterring disintermediation.[4] Table 1 presents examples of sensitive words and phrases. For each message associated with a given job assignment, we sum up the numeric values of sensitive words and use the maximum value among all messages as the disintermediation score for that assignment (*Disintermediation_Score*). Compared to approaches that add up sensitive keywords in all messages or that take an average of all messages, our approach has two advantages. First, users who communicate more are likely to use more sensitive keywords; our measure is independent of the frequency of communication. Second, because users typically express their desire to disintermediate in only a few sentences and hence not all messages are useful for detecting disintermediation, our approach allows us to focus on messages that are most likely related to disintermediation. Out of the 33,561 assignments, 29,690 have historical messages, for which the platform attempts to detect sensitive words that imply users' intent to disintermediate. For job assignments whose messages have no sensitive words, *Disintermediation_Score* is 0; otherwise, it is a positive integer.

For each assignment, the dummy variable *Treated* equals 1 if the client is in the treatment group and 0 otherwise. To account for different levels of trust among high-quality vs. low-quality freelancers, we create a dummy variable, *SS_High,* which is 1 if SS $\geq$ 90%, based on the fact that the platform explicitly informed the clients in the treatment group that an SS above 90% is considered "excellent."[5]

Table 2 provides summary statistics for all variables. The unit of analysis is a job assignment. The average SS in our sample is 0.739 and the mean for *SS_High* is 0.364, meaning

[4] For example, the platform learns about disintermediation when users seek its help with disputes or payment enforcement for transactions off the platform. In other trials, the platform issued warning messages based on disintermediation scores of chat messages and has improved the accuracy of disintermediation scores based on user feedback.

[5] As a robustness check, we replace the threshold for *SS_High* with 75%. All our findings still hold. The use of this dummy variable allows us to take the non-linear effect of SSs into account.

that 36.4% of the freelancers hired by clients have an SS of 90% or higher. In contrast, the five-star ratings of freelancers have a mean of 4.77 and a median of 5. Because the five-star ratings do not contain sufficient variation to discern freelancer quality, unsurprisingly, the SS has a low correlation with five-star ratings (0.284). The two indirect measures for disintermediation, *Hours* and *Total_Charge,* have substantial variation across job assignments and are highly skewed; thus, we use their logarithms in our regression analysis. The mean *Disintermediation_Score* is 7.528, with a maximum value of 34.1 and the distribution is skewed, thereby suggesting that most users in our sample have a relatively low tendency to disintermediate. Among all assignments in our sample, 8% of the client-freelancer pairs are located in the same country as each other.

To confirm the randomness of assignment into either group, we compare the transaction data of clients in the treatment and control groups from the six-month period immediately preceding the study. The balance check, shown in Table 3, confirms that the assignment is indeed random.

Table 4 compares the three key dependent variables for the treatment and control groups after the treatment. We divide our sample by *SS_High*. The values of the three dependent variables for high-SS assignments differ significantly between the treatment and control groups. The treated job assignments have fewer total hours, lower total charges, and higher disintermediation scores, all of which suggest a greater likelihood of disintermediation. For low-SS assignments, we observe no significant differences in the three dependent variables between the two groups.

4. Empirical Results

4.1 Job Fill Rates and Platform Revenue

We first analyze the impact of disclosing SSs on the job fill rate. We begin by comparing the job fill rate in the treatment group versus the control group for all job openings posted by clients. We compute the job fill rate as the ratio of the total number of filled jobs for a group to the total number of job openings for that group during the study. We also calculate the number of days before a job opening is filled. We find that the number of days taken to fill a job is 0.48% shorter for the treatment group than the control group, and the fill-rate difference is 0.51%; neither difference is statistically significant.[6] In addition, a similar number of jobs are posted for each group, with the clients in the treatment group posting only 1.3% more jobs than the clients in the control group clients; a t-test mean comparison reveals that the average number of jobs offered by each client does not differ significantly between the treatment and control groups ($p = 0.88$).

These results indicate that revealing SSs does not have a significant effect on job postings or job fill rate, which is consistent with the intuition that, when there is a sufficient number of freelancers, the job fill rate will not increase with better reputation measures.

Further, we also compare the characteristics of the jobs posted between the two groups. We find that there are no significant differences in the distributions of jobs posted across categories ($p = 0.204$ from a Chi-squared test), in the length of job descriptions ($p = 0.303$), and in clients' posted job prices ($p = 0.553$) across the two groups. These results—along with the findings that the days to fill, fill rate, and the average number of jobs by each client are not significantly different—suggest that clients' needs are exogenously determined and are not influenced by the availability of SSs. Therefore, our findings are unlikely to be driven by job-level differences between the two groups.

Next, we investigate whether revealing SSs affects the probability of hiring high- versus low-quality freelancers. For this, we obtain data on all freelancers who submitted proposals for jobs posted by clients in the study. Simple summary statistics show that the percentage of high-SS freelancers hired in the treatment group is 4.1% higher than in the control group. Then, we use a linear probability model regressing a *Hired* dummy on *SS_High*, *Treated*, and the interaction between them. The regression results shown in Table 5 indicate that freelancers with higher SSs are significantly more likely to be hired than freelancers with low SSs and that revealing SSs increases the likelihood of a high-quality freelancer being hired by 0.48%, from a base of 2.8%.

Finally, we examine the impact of the treatment on the platform's revenue. The results of the two-sample t-test show that the average revenue from each job does not significantly change for the treatment group as compared to the control group (the ratio between the treatment and control groups is 0.996). This result holds regardless of job type. We also check the percentage of successful jobs for the two groups—jobs that were completed with no abnormal status, such as

[6] As requested by the company, the actual values of the measures are not reported in order to protect the company's data confidentiality.

"Inactivity," "No response," or "Cancelled"—and find that there is only a 0.41% difference in the percentages of successful jobs, which is not statistically significant.

These findings raise an interesting question: while revealing SSs leads to increased hiring of high-quality freelancers who often command higher prices (in the control group, the average charge for jobs involving high-SS freelancers is 2.23 times that of jobs involving low-SS freelancers), and given that job fill rates are the same, why does the platform not earn more revenue from the treatment group than from the control group? Next, we provide evidence that disintermediation, not the selection of freelancers, is the key factor that offsets the potential gain.

4.2 Evidence of Disintermediation

We investigate whether clients in the treatment group are more likely to disintermediate when freelancers' SSs are high, using the following regression specification at the job-assignment level:

$$(1) \qquad Y = \beta_0 + \beta_1 Treated + \beta_2 SS\_High + \beta_3 Treated \times SS\_High + \varepsilon.$$

Table 6 reports the regression results. *Log(Hours)* is the dependent variable in Models (1) and (2). Model (1) shows that, on average, fewer working hours are reported for jobs in the treatment group than in the control group, and more working hours are reported for jobs with high-SS freelancers than with low-SS freelancers. Model (2) shows that displaying a freelancer's SS reduces the hours reported by high-quality freelancers by 13.3% for the treatment group relative to the control group.[7] Models (3) and (4) show similar patterns using *Log(Total_Charge)* as the dependent variable. Revealing a high-quality freelancer's SS to the client decreases the total charge by 8.2% relative to the control group.

One might be concerned that the higher probability of clients in the treatment group hiring high-quality freelancers would confound our results. However, we find no correlation between a freelancer with a high SS and the freelancer charging less or working faster (see Section 5.2).

[7] We also add category controls to each model in Table 6, and the results remain virtually unchanged. We also use the difference between each job's posted price and actual price as an alternative dependent variable. Because posted prices are similar between the two groups, unsurprisingly, we find that the treatment group's actual charge is significantly less than the posted charge for jobs with high-SS freelancers, relative to the control group.

Further, if clients in the treatment group are matched with freelancers who are more productive and charge less, they should have lower incentives to disintermediate. In Models (5) and (6), we use *Log(Disintermediation_Score)* as the dependent variable. We find that clients in the treatment group are significantly more likely to disintermediate when the freelancer has a high SS.[8]

Overall, the evidence suggests that increased trust leads to more disintermediation. As a result, although providing SSs makes clients more likely to work with high-quality freelancers, the expected revenue increase is offset by disintermediation.[9]

We perform a back-of-the-envelope analysis to estimate revenue loss due to disintermediation. If the treatment were rolled out to our control group as well, the control group's percentages of high- and low-SS freelancers hired—31.4% and 68.6%, respectively—would switch to the treatment group's 35.5% and 64.5%. In other words, the proportion of high-quality freelancers hired by the control group would increase by 4.1% (35.5% - 31.4%). The average total charge for a job by a high-SS freelancer in the control group is 2.23 times that of a job by a low-SS freelancer in the control group; thus, replacing one job by a low-SS freelancer with one job by a high-SS freelancer in the control group would increase the job's total charge by 223% - 1 = 123%. Putting all of this together, the hypothetical rollout of SSs to all clients would create 4.1% more jobs by high-SS freelancers, each generating 123% more revenue; thus, if it were not offset by disintermediation, the total revenue should increase by approximately 123% × 4.1% = 5.0% when SSs are introduced.

Note that there is a certain level of disintermediation even in the control group. Our estimated revenue loss represents only the additional revenue loss due to disintermediation beyond the baseline.

[8]As the company-assigned weights for the sensitive words could affect the accuracy of the *Disintermediation_Score*, we also run a robustness check by repeating the main analysis using a dummy dependent variable to indicate whether the score is above the median score. We find that our results continue to hold.

[9]To demonstrate that SSs are indeed superior to five-star ratings, we repeat the analysis separately for freelancers with high SSs and high five-star ratings and freelancers with high SSs but low five-star ratings. We find similar results in each case, thereby suggesting that once SSs are offered, clients depend much less on five-star ratings (see Appendix Table A2).

4.3 Heterogeneous Tendencies to Disintermediate

We examine various factors that moderate the impact of trust on disintermediation. Appendix Table A1 provides the summary statistics of all moderators.

Geographical proximity. When the client and freelancer are in the same country, they tend to have similar cultural backgrounds that allow them to build trust more easily than those from different countries. Proximity also reduces the cost of collaborating outside the platform, since they may have many other convenient channels for payment and interaction. Thus, the impact of SSs on disintermediation should be higher for jobs involving clients and freelancers from the same country. We create the dummy variable *Same_Country*, which equals 1 when clients' and freelancers' user profiles show them to be in the same country, and 0 otherwise.

Model (1) of Table 7 reports the regression results using our direct measure, the logarithm of *Disintermediation_Score*, as the dependent variable and including *Same_Country* as the moderator.[10] The coefficient for the three-way interaction suggests that clients and high-quality freelancers from the same country are indeed more affected by the treatment.

Job divisibility. The tendency to disintermediate may also vary for different job categories. Certain job categories are modular by nature and can be divided into independent parts. If a job is modular—that is, if it is more likely to be divided into parts without affecting the overall quality of the outcome (e.g., Baldwin and Clark 2000)—it is easier to perform a portion of the job on the platform, have the client check the output, and then complete the remainder off the platform. Thus, we expect the impact of treatment to be greater for more divisible jobs.

Two types of jobs are considered to be more divisible than others: (a) hourly jobs and (b) fixed-price jobs with more than one hired freelancer. We compute the percentage of such jobs in each of the platform's 13 job categories and rank the categories by that percentage from high to low; the percentages range from 33% to 88%. We create dummy variables to place all job categories into three groups based on the extent to which the jobs are divisible: the *Divisible_High* = 1 group, the *Divisible_Med* = 1 group, and the baseline group. *Divisible_High* equals 1 for the three categories with a percentage of divisible jobs in the top 10% of the divisibility distribution: Customer Service, Sales & Marketing, and Accounting & Consulting. The benchmark group

[10] The results obtained using the two indirect measures, *Log(Hours)* and *Log(Total_Charge)*, as the dependent variables were qualitatively similar.

includes the job categories with the least divisible jobs, with a percentage of divisible jobs in the bottom 10% of the distribution: Translation and Design & Creative. For the remaining categories, whose divisibility is between the top and the bottom 10% of the distribution, *Divisible_Med* equals 1.[11]

Model (2) of Table 7 reports the results. As expected, the increase in disintermediation scores for clients in the treatment group who hired high-quality freelancers is greater for divisible jobs.

Expected duration. Jobs that last a long time tend to have higher costs, thereby generating the most value for the platform. They also face the highest risk of disintermediation because clients and freelancers have the greatest incentive to take the transaction off the platform. Therefore, we expect to see a greater impact of the treatment on long-term versus short-term jobs.

We use the expected job duration, which is selected by the client when posting the job and ranges from "less than one week" to "more than six months," to create a dummy variable called *Long_Term*. This information is self-reported regardless of the job type. For any job that reports an expected duration, *Long_Term* equals 1 when that duration is over six months and 0 when it is six months or less.

As Model (3) of Table 7 shows, long-term jobs indeed have a significantly higher disintermediation score for treated clients and high-quality freelancers as compared to short-term jobs.

Client rating. Having established that the disintermediation tendency varies with the business reputations of freelancers, we investigate whether it also varies with clients' reputations. We expect that, given the same SSs, freelancers tend to trust highly-rated clients more and are, therefore, more willing to take the job off-platform.

Using a client's past transactions and corresponding star-rating feedback, we compute the number of five-star jobs in each client's job history. We create the dummy variable *Client_Rating_High*, which equals 1 when that fraction is higher than the median for all clients, and 0 otherwise. Model (4) of Table 7 reports the results. We find that, when both the freelancer

[11] We use category information to define job divisibility because, as we show in Section 5.7, clients may strategically choose their job types after their first jobs. As a robustness check, we use job type information from jobs posted within one month prior to the RCT and obtain the same sets of categories in each of the three groups.

and the client are trustworthy, revealing more information about the freelancer's business reputation boosts their willingness to work off-platform.

We also examine client size as a possible factor that could lead to heterogeneous tendencies in disintermediation on the platform, using self-reported client-size data. As large companies are usually less concerned about cost and more about quality, they tend to place more value on the platform's role in facilitating transactions and, thus, have less incentive to disintermediate, even with trustworthy freelancers. We expect individual clients or clients who post jobs for smaller companies to be more sensitive to job cost than clients who post jobs for large companies. We do not find significant results. The lack of significance could be due to insufficient data, since the information on firm size is self-reported and is missing for 97.4% of the jobs in our sample. It could also be that, since the service fee is a fixed percentage of the transaction value and can become substantial for large jobs, large clients may find the savings from disintermediation just as attractive as small clients do.

5. Robustness Checks

5.1 Contamination Between the Treatment and Control Groups

As with any experiment conducted in a real marketplace, one might be concerned about violations of the stable unit treatment value assumption (Blake and Coey 2014). This concern of possible contamination is mitigated by two factors. First, the total number of clients in the trial constitutes only 3% of all clients on the platform. Second, the platform has substantially more freelancers than clients. Indeed, the job fill rates for the two groups are not different.

Nevertheless, we consider two possible sources of contamination. First, in our setting, treated clients may be better able to target high-quality freelancers, thereby leaving the control clients a pool of lower quality prospective job applicants. This may result in an exaggeration of the treatment effect when we compare the quality of freelancers hired by clients in the two groups. We conduct a t-test to compare the mean SSs of freelancers who apply to job posts in each group and find that there is no significant difference between the SSs of applicants in the treatment group and those in the control group ($p = 0.49$).

Second, it is possible that some freelancers, after the disclosure of their SSs, are approached more often by clients in the treatment group to disintermediate the platform, and may in turn suggest to clients in the control group to conduct transactions off the platform. This possibility may result in an underestimation of the treatment effect. We find that 11.1% jobs in the control group have been assigned to freelancers who have been matched to clients in treatment groups. After excluding these jobs, as expected, our results become slightly stronger (Appendix Table A3).

5.2 Selection of Freelancers

Prior to the availability of SSs, clients may have used price as a quality signal and selected freelancers who charge more. However, with the availability of SSs, clients may select freelancers who have a high SS but charge less because they are more efficient and can complete jobs faster. While this scenario does not explain our findings based on *Disintermediation_Score,* it is consistent with our findings that jobs done by high-SS freelancers in the treatment group take less time and cost less.

To test this alternative explanation, for each freelancer, we calculate the average number of job hours and average total charge for all assignments completed in the six months prior to the study timeframe; we use *Past_Hours* and *Past_Total_Charge* to denote these variables. We then repeat our analysis in Models (1) through (4) of Table 6 with the logarithms of *Past_Hours* and *Past_Total_Charge* as dependent variables. If high-SS freelancers were hired by clients in the treatment group because they could work more efficiently and, therefore, charged less than freelancers in the control group, this difference should be evident in their work completed prior to the study.

Appendix Table A4 reports the results. We find that high-SS freelancers hired by clients in the treatment group during the study do not appear to work faster or charge lesser than high-SS freelancers in the control group prior to the study. The results also suggest that clients in the treatment group are not selecting freelancers with greater tendencies to disintermediate.

As another robustness check, we compare the actual job duration—measured by the number of days between the start date of a job to the date on which the client marks the job as completed—to the expected job duration when the client posts the job. A client is asked to indicate estimated job duration when posting a job by selecting one of the five following options: “More than six months," "Three to six months," "One to three months," "Less than one month," and "Less than one week." First, we compare the distribution between the treatment and control groups and do not observe significant differences in expected job durations (p = 0.239), thereby suggesting that these jobs are comparable. If the fewer hours for jobs in the treatment group are indeed caused by clients' selection of more efficient freelancers, we would expect the actual duration of these jobs to be shorter. However, if they are caused by disintermediation instead, the actual job duration may not be shorter. By not marking the jobs as completed before their actual completion, the clients could leverage instrumental trust (Bapna et al. 2017) by reserving the option to report to the platform or leave negative ratings if the freelancers do not complete the jobs satisfactorily.[12] We compare (the logarithm of) the actual job duration between the two groups, and find that jobs in the treatment group take 0.75% longer than those in the control group (p = 0.260).

These results boost our confidence that the differences in the hours and total charges are caused by greater disintermediation.

5.3 Robustness of the Disintermediation Score Measure

Because the disintermediation score is constructed from a dictionary of company-developed keywords, one might be concerned about whether the keywords accurately capture user intention to disintermediate. To address this concern, we create a more robust disintermediation score based only on keywords that explicitly indicate an intention to disintermediate, such as "avoid fees," "save 10%," and "wire me/you/us," and excluding ambiguous keywords such as "your phone number" and "Skype."

Appendix Table A5 reports the results after repeating our analysis with this new disintermediation score (*Disintermediation_Score_Robust*). The results obtained are qualitatively the same. The coefficient of the interaction term is smaller because the new dictionary includes fewer keywords.

[12] The platform allows clients and freelancers to leave each other reviews within 14 days after their jobs are marked as completed.

5.4 Is it All About Speeding up the Inevitable Outcome?

Tryouts between clients and freelancers may also help build trust (e.g., Gulati 1995). After tryouts, clients and freelancers who had positive experiences with one another may decide to disintermediate. If disintermediation is inevitable because of tryouts, introducing a better reputation score may only help speed up this outcome by reducing the duration of the tryout period.

However, tryouts have their limitations and are not perfect substitutes for a better reputation system. First, for a durational job, the experience can vary within the same job—it is not uncommon to have a good start and a terrible ending. In other words, the initial experience in a job does not necessarily indicate its final success.

Second, experiences with the same freelancer can vary across jobs. For example, it is possible for freelancers to build their reputation first and later exploit customers who trust them the most (e.g., Liu 2011). Further, clients typically use online freelance marketplaces to fulfill ad-hoc needs. On the platform that we studied, although the platform has been in the business for more than ten years, the median number of successfully filled jobs per client is two.[13] Thus, for most clients, their limited number of past interactions would not sufficiently help reduce uncertainty.[14] These clients would value an accurate reputation system—which includes freelancers' past performance for many clients—because these ratings capture the overall satisfaction for all jobs completed, and the rating variation represents truthful uncertainty.

Consistent with these arguments, Kim et al. (2012) find that trust exerts a stronger effect than perceived price on purchase intentions for *both* new and repeat customers of an online store. In a setting like ours, where tasks are heterogeneous across jobs and over time within a job, we expect that the efficiency of tryouts in building trust is even lower.

In fact, one might expect that SSs have a greater impact on clients who have had some past interactions with the same freelancers. A good SS and clients' personal positive experiences can work in combination to build sufficient trust for clients to be willing to disintermediate. Furthermore, after having some interactions with freelancers, clients may feel more comfortable suggesting taking transactions off the platform.[15]

---

[13] A similar statistic is also reported by Barach et al. (2018) in their study of a similar freelance platform.

[14] Because our sample collection requires a client to post a job during the 26-day trial, the RCT selects clients that post jobs more frequently.

[15] As in most marketplaces, disintermediation is a violation of the terms of services on this platform, and the platform encourages users to report such behavior.

In 21.1% of the job assignments in our data set, the clients have already had past interactions with the same freelancers. These clients most likely had positive experiences with the freelancers and thus decided to work with them again. If we restrict the analysis to these repeated interactions, we find that SSs continue to have significant effects (Appendix Table A6). The magnitude of the effects is actually greater for repeated hires than for first-time hires, thereby suggesting that even with tryouts, certain clients still choose not to disintermediate and that the introduction of SSs, in addition to their own experiences, motivates them to disintermediate.

5.5 Client Satisfaction with High-SS Freelancers

Another potential explanation for the lower fees and hours is that high SSs may have raised treated clients' expectations of the freelancers. Treated clients may therefore end jobs prematurely and pay less as a result of dissatisfaction rather than disintermediation.

To test this alternative explanation, we collect data on clients' feedback on freelancers after each job assignment in our sample. The number of jobs for which the client did not leave any feedback is similar for both the treatment and control groups (30.1% and 29.8%, respectively). We replicate the analysis in Table 6, replacing the dependent variable with the client rating for each job assignment. Appendix Table A7 shows that neither the coefficient for being in the treatment group in Model (1) nor the coefficient of the interaction term in Model (2) are statistically significant. Thus, our findings are not driven by reduced client satisfaction with the work of high-SS freelancers.

5.6 Who Initiated the Disintermediation?

As SSs were only revealed to the clients in the treatment group, if revealing SSs indeed resulted in more disintermediation, we expect that clients in the treatment group are more likely to initiate disintermediation. Among the job assignments in which disintermediation is detected, a t-test comparison shows that clients in the treatment group indeed initiate disintermediation 3% more frequently as compared to clients in the control group ($p = 0.042$).

5.7 *Clients' Strategic Choice of Job Type*

After observing SSs, clients interested in disintermediation might strategically select hourly jobs, as these are more conducive to disintermediation. We test this using data on clients' job openings and running several t-tests to compare the number of hourly versus fixed-price jobs that clients posted over time.

Overall, the distribution of job types is different between the treatment and control groups ($p = 0.029$), with clients in the treatment group posting a slightly larger proportion of hourly jobs than clients in the control group (1.2% higher). We split the job-opening sample into subsamples of clients' first job posts and subsequent job posts and then repeat the comparison for each subsample. We find no difference in the distributions of job types between the treatment and control groups in the subsample of clients' first jobs ($p = 0.986$). However, there is a significant difference for the subsample of subsequent job posts: the percentage of hourly jobs posted is 3.5% higher in the treatment group ($p = 0.0004$).

This result suggests that while clients in the treatment group do not appear to take SSs into account when they post their first job (as most become aware of SSs only when they review proposals for their first job posts), they are more likely to post hourly jobs subsequently. This result supports the explanation that clients in the treatment group are, on average, more inclined to disintermediate.

6. Discussion and Conclusion

We provide empirical evidence on disintermediation and show that disintermediation can sometimes render less effective an intermediary's strategy to improve its profitability through enhancing trust. An important challenge in studying disintermediation is the observation of disintermediated transactions. Our study shows that digitization can help overcome this challenge by capturing detailed data on user communication and activities.

We examine only one type of disintermediation. In reality, there are other means by which that users can disintermediate an online platform. For example, a user can use a platform to find an ideal match and then directly contact the other party without ever initiating a transaction on the platform. In this study, the fill rate does not significantly decrease for clients in the treatment group, but this type of disintermediation could occur more frequently in other settings. A user may also complete one transaction on a platform and then take all future transactions with that party off the platform.

In addition, our study examines only the short-term effect of building trust. As more users on the platform realize the benefits of disintermediation, there could be an increase in the negative effects of enhanced trust on platform revenue. Our result that clients change their job format after posting their first job suggests that they do learn from their experiences. Since not all clients hired freelancers twice during our experiment, we would expect to have more disintermediation as a result of this learning if the experiment had run for a longer period of time. Overall, the strategic behavior may increase the negative effects of trust building in the long term. At the same time, a better trust building mechanism may attract more users to the platform, thereby resulting in greater revenue. Thus, the long-term effect in this regard is ambiguous.

Notwithstanding these limitations, it is important to note that the main objective of our research is not to conduct a net benefit analysis to determine whether it is optimal for platforms to implement SSs. Rather, because trust building is important for platform growth, it is of vital importance for a platform to build as much trust as possible. At the same time, our research suggests that as a platform builds more trust to facilitate transactions in its marketplace, it needs to adopt appropriate strategies to counter increased disintermediation.

Platforms could use a variety of strategies to reduce disintermediation as they enhance trust. Airbnb, for example, enhances trust and safety through host ID verification and background checks. At the same time, Airbnb reduces disintermediation by withholding host data, such as listing address or phone number, until the payment is made. Thumbtack, a marketplace that connects consumers with local service providers such as house cleaners, captures value pre-transaction: when customers post job requests on Thumbtack, service providers can send quotes to the customers; service providers pay fees to Thumbtack only if customers respond. Disintermediation affects Thumbtack less strongly because its model captures value before two parties agree to work together.

Other platforms recognize that the motivation to disintermediate comes from the service fees they charge and adopt different value-capture strategies to prevent disintermediation while still enhancing trust. For example, Chinese outsourcing marketplace ZBJ, which launched in 2006 with a 20% commission model, began pursuing other revenue sources after calculating that it could lose as much as 90% of its business to disintermediation. In 2014, ZBJ leveraged big data analytics to find that new business owners often used ZBJ to outsource logo design. However, after logo design, many of these clients would also need business and trademark registration. Thus, ZBJ began offering this service and has now become the largest provider of trademark registration in China. Replicating this experience, ZBJ began providing several other services to its marketplace participants. With these revenue streams, the company decided to significantly reduce its commission to 2% and shifted its resources from fighting disintermediation to growing its user base and building trust (for example, by encouraging clients and freelancers to communicate) (for more details, see Zhu et al. 2018). Because of these changes, the company obtained a valuation over $1.5 billion in 2018. Future research could examine the effectiveness of various strategies that platforms use to mitigate disintermediation.

REFERENCES

Agrawal, Ajay, Nicola Lacetera, and Elizabeth Lyons. 2016. "Does Standardized Information in Online Markets Disproportionately Benefit Job Applicants from Less Developed Countries?" *Journal of International Economics,* 103(1): 1–12.

Armstrong, Mark. 2006. "Competition in Two-Sided Markets." *The RAND Journal of Economics*, 37(3): 668–691.

Arya, Anil, Brian Mittendorf, and David E. M. Sappington. 2007. "The Bright Side of Supplier Encroachment." *Marketing Science*, 26(5): 651–659.

Ba, Sulin, and Paul A. Pavlou. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premium and Buyer Behavior." *MIS Quarterly*, 26(3): 243–268.

Baldwin, Carliss Y., and Kim B. Clark. 2000. *Design Rules: The Power of Modularity*. Vol. 1. Cambridge, MA: MIT Press.

Bapna, Ravi, Liangfei Qiu, and Sarah Rice. 2017. "Repeated Interactions Versus Social Ties: Quantifying the Economic Value of Trust, Forgiveness, and Reputation Using a Field Experiment." *MIS Quarterly*, 41(3): 841–866.

Barach, Moshe, Aseem Kaul, Ming Leung, and Sibo Lu. 2018. "Small Numbers Bargaining in the Age of Big Data: Evidence From a Two-Sided Labor Matching Platform." Available at SSRN: https://ssrn.com/abstract=3277455.

Blake, Thomas, and Dominic Coey. 2014. "Why Marketplace Experimentation is Harder than It Seems: The Role of Test-Control Interference." *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, 567–582.

Bohnet, Iris, and Steffen Huck. 2004. "Repetition and Reputation: Implications for Trust and Trustworthiness When Institutions Change." *American Economic Review*, 94(2): 362–366.

Bolton, Gary, Ben Greiner, and Axel Ockenfels. 2013. "Engineering Trust: Reciprocity in the Production of Reputation Information." *Management Science*, 59(2): 265–285.

Cai, Hongbin, Ginger Z. Jin, Chong Liu, and Li-An Zhou. 2013. "More Trusting, Less Trust? An Investigation of Early E-Commerce in China." *National Bureau of Economic Research Working Paper No. w18961*.

Cabral Luís, and Ali Hortaçsu. 2010. "The Dynamics of Seller Reputation: Evidence from eBay." *Journal of Industrial Economics*, 58(1): 54–78.

Cabral, Luís, Erkut Y. Ozbay, and Andrew Schotter. 2014. “Intrinsic and Instrumental Reciprocity: An Experimental Study.” *Games and Economic Behavior*, 87: 100–121.

Charness, Gary, Ninghua Du, and Chun-Lei Yang. 2011. “Trust and Trustworthiness Reputations in an Investment Game.” *Games and Economic Behavior*. 72: 361–375.

Dai, Weijia, Ginger Z. Jin, Jungmin Lee, and Michael Luca. 2018. “Aggregation of Consumer Ratings: An Application to Yelp.com.” *Quantitative Marketing and Economics*, 16(3): 289–339.

Dellarocas, Chrysanthos. 2003. “The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms.” *Management Science*, 49(10): 1407–1424.

Edelman, Benjamin, and Philip Hu. 2016. “Disintermediation in Two-Sided Marketplaces.” *Harvard Business School Technical Note 917–004*.

Edelman, Benjamin, and Julian Wright. 2015. “Price Coherence and Excessive Intermediation.” *Quarterly Journal of Economics*, 130(3): 1283–1328.

Einav, Liran, Chiara Farronato, and Jonathan Levin. 2016. “Peer-to-Peer Markets.” *Annual Review of Economics* 8: 615–635.

Ert, Eyal, Aliza Fleischer, and Nathan Magen. 2015. “Trust and Reputation in the Sharing Economy: the Role of Personal Photos in Airbnb.” in *Advances in Consumer Research.* Vol. 43, eds. Kristin Diehl and Carolyn Yoon, Duluth, MN: Association for Consumer Research, 518–519.

Gulati, Ranjay. 1995. “Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances.” *Academy of Management Journal*, 38(1): 85–112.

Hagiu, Andrei, and Julian Wright. 2015. “Multi-Sided Platforms.” *International Journal of Industrial Organization*, 43: 162–174.

Horton, John, and Joseph Golden. 2015. “Reputation Inflation: Evidence from an Online Labor Market.” *Working Paper, New York University*.

Horton, John, David Rand, and Richard Zeckhauser. 2011. “The Online Laboratory: Conducting Experiments in a Real Labor Market.” *Experimental Economics*, 14(3): 399–425.

Hu, Nan, Paul A. Pavlou, and Jie (Jennifer) Zhang. 2009. “Overcoming the J-Shaped Distribution of Product Reviews.” *Communications of the ACM*, 52(10).

Hui, Xiang, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan. 2014. "From Lemon Markets to Managed Markets: The Evolution of eBay's Reputation System." *Ohio State University Working Paper*.

Jin, Ginger Z., and Andrew Kato. 2006. "Price, Quality, and Reputation: Evidence from an Online Field Experiment." *The RAND Journal of Economics*, 37(4): 983–1005.

Kapoor, Anuj, and Catherine E. Tucker. 2017. "How Do Platform Participants Respond to an Unfair Rating? An Analysis of a Ride-Sharing Platform Using a Quasi-Experiment." Available at SSRN: https://ssrn.com/abstract=2970772.

Kim, Hee-Woong, Yunjie Xu, and Sumeet Gupta. 2012. "Which is More Important in Internet Shopping, Perceived Price or Trust?" *Electronic Commerce Research and Applications*, 11: 241–252.

Liu, Qingmin. 2011. "Information Acquisition and Reputation Dynamics." *Review of Economic Studies*, 78(4): 1400–1425.

Luca, Michael. 2017. "Designing Online Marketplaces: Trust and Reputation Mechanisms." *NBER Innovation Policy and the Economy*, 77–93.

Masterov, Dimitriy, Uwe Meyer, and Steven Tadelis. 2015. "Canary in the e-Commerce Coal Mine: Detecting and Predicting Poor Experiences Using Buyer-to-Seller Messages." *Proceedings of the Sixteenth ACM Conference on Economics and Computation*.

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

Parker, Geoffrey, and Mashall W. Van Alstyne. 2005. "Two-Sided Network Effects: A Theory of Information Product Design." *Management Science*, 51(10): 1494–1504.

Pavlou, Paul A., and Angelika Dimoka. 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.

Pavlou, Paul A., Huigang Liang, and Yajiong Xue. 2007. "Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective." *MIS Quarterly*, 31(1): 105–136.

Peukert, Christian, and Imke Reimers. 2018. "Digital Disintermediation and Efficiency in the Market for Ideas." *CESifo Working Paper Series No. 6880*.

Resnick, Paul, and Richard Zeckhauser. 2002. "Trust among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." *Advances in Applied Microeconomics*, 11: 127–157.

Rochet, Jean-Charles, and Jean Tirole. 2006. "Two-Sided Markets: A Progress Report." *The RAND Journal of Economics*, 37(3): 645–667.

Sobel, Joel. 2005. "Interdependent Preferences and Reciprocity." *Journal of Economic Literature*, 43(2): 392-436.

Spulber, Daniel F. 2011. "Should Business Method Inventions be Patentable?" *Journal of Legal Analysis*, 3(1): 265–340.

Stanton, Christopher T., and Catherine Thomas. 2016. "Landing the First Job: The Value of Intermediaries in Online Hiring." *Review of Economic Studies*, 83(2): 810–854.

Strader, Troy, and Sridhar Ramaswami. (2002). "The Value of Seller Trustworthiness in C2C Online Markets." *Communications of the ACM*, 45(12): 45–49.

ter Huurne, Maarten, Amber Ronteltap, Rense Corten, and Vincent Buskens. 2017. "Antecedents of Trust in The Sharing Economy: A Systematic Review." *Journal of Consumer Behavior*, 16(6): 485–498.

Utz, Sonja, Uwe Matzat, and Chris Snijders. 2009. "On-Line Reputation Systems: The Effects of Feedback Comments and Reactions on Building and Rebuilding Trust in On-Line Auctions." *International Journal of Electronic Commerce*, 13(3): 95–118.

Waldfogel, Joel. 2012. "And the Bands Played on: Digital Disintermediation and the Quality of New Recorded Music." Available at SSRN: https://ssrn.com/abstract=2117372.

Waldfogel, Joel, and Imke Reimers. 2015. "Storming the Gatekeepers: Digital Disintermediation in the Market for Books." *Information Economics and Policy*, 31(1): 47–58.

Zhu, Feng, Weiru Chen, and Shirley Sun. 2018. "ZBJ: Building a Global Outsourcing Platform for Knowledge Workers (A)." *Harvard Business School Case 618-044*.

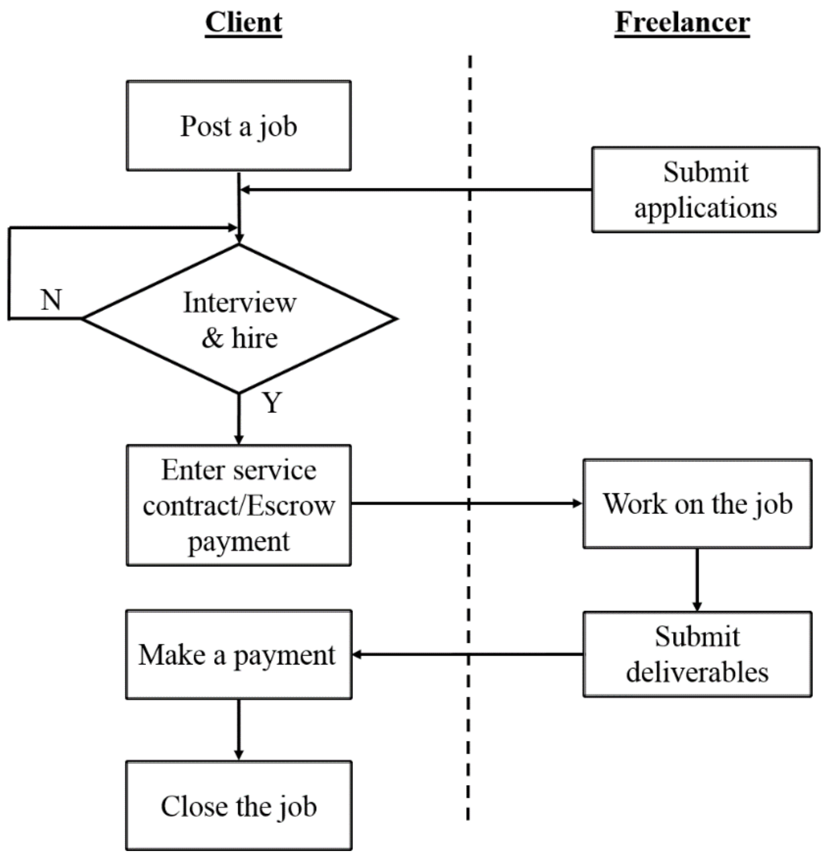

Figure 1. Job Process Flow

Figure 2. User Information Shown to the Treatment and Control Groups

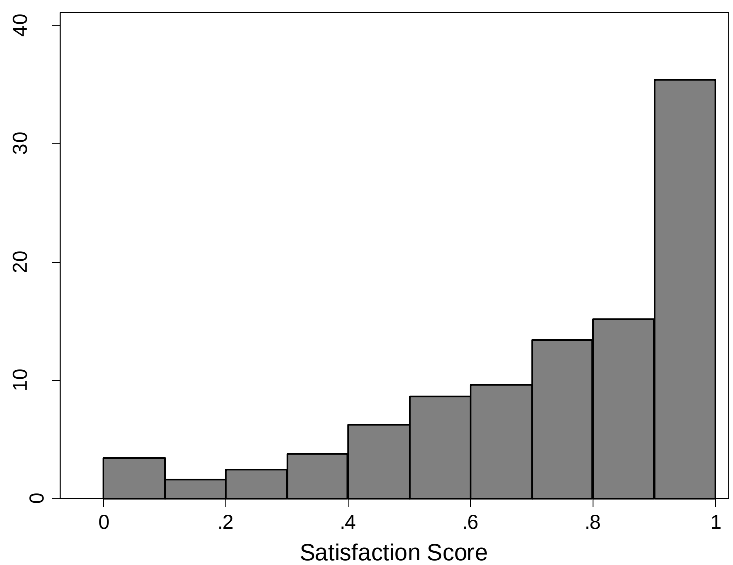

Figure 3. Distribution of Freelancer Satisfaction Scores

Table 1: Examples of Sensitive Words/Phrases Indicating Disintermediation in Messages

off (the platform's name)	Paypal	Venmo
wire me/you/us	avoid fees	apply at/here
outside (the platform's name) / outside of (the platform's name)		
save 10% (or 5%) / save 10 (or 5) percent / save ten (or five) percent		
(my/your/our) (phone / number / phone number / cell phone)		

Table 2: Summary Statistics and Correlations

Variable	Observations	Mean	Std. dev.	Min	Max	(1)	(2)	(3)	(4)	(5)
(1) Treated	33,561	0.504	0.500	0	1	1				
(2) SS	33,561	0.739	0.258	0	1	0.041	1			
(3) SS_High	33,561	0.364	0.481	0	1	0.034	0.679	1		
(4) Hours	14,593	109.00	349.0	0.183	11365.38	-0.003	0.058	0.081	1	
(5) Total_Charge	33,561	679.507	4198.7	0	195731.5	0.005	0.049	0.062	0.683	1
(6) Disintermediation_Score	29,690	7.528	5.642	0	34.1	0.027	-0.059	-0.051	-0.021	0.003

Notes. The number of observations for the main analysis sample is 33,561, except for regressions with *Disintermediation_Score*, which has a non-missing value for 29,690 observations. *Hours* has values only for hourly jobs.

Table 3: Comparison of Clients in the Treatment and Control Groups before the Study

Outcome Variable	Treatment		Control		Paired t-test
	Mean	Standard Error	Mean	Standard Error	t-stats
# of days on the platform	611.17	4.62	605.03	4.63	-0.94
# of jobs in the past 6 months	2.60	0.09	2.73	0.15	0.76
Avg. past job feedback	4.81	0.01	4.81	0.01	0.55
Avg. past job hours	12.62	0.41	13.05	0.44	0.71
Avg. past job total charge	197.00	7.77	185.55	5.31	-1.21

Notes. The unit of analysis is a client in the treatment/control group. Variables are calculated using a past assignment sample including participating clients' job outcomes in the 6 months before the study. None of the above paired t-test results is significant. We also checked the percentage of hourly jobs across the two groups and found no significant difference; the numbers are unreported due to protection of confidentiality.

Table 4: Comparing Treatment and Control Group Observations after Treatment

Outcome Variable	Treatment		Control		Paired t-test
	Mean	Standard Error	Mean	Standard Error	t-stats
Assignments with high freelancer SSs:					
Log(Hours)	2.95	0.04	3.09	0.04	2.80***
Log(Total_Charge)	4.67	0.02	4.76	0.03	2.61***
Log(Disintermediation_Score)	1.76	0.01	1.68	0.01	-4.52***
Assignments with low freelancer SSs:					
Log(Hours)	2.66	0.03	2.67	0.03	0.09
Log(Total_Charge)	4.26	0.02	4.26	0.02	0.10
Log(Disintermediation_Score)	1.83	0.01	1.81	0.01	-1.47

Notes. The unit of analysis is a job assignment from a treatment/control group client during the study. Variables are calculated using the assignment sample for our main analysis. All the paired t-test results in the high-SS group are significant. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Logistic Regressions of the Treatment Effect on Freelancers' Probability of Being Hired

Model	(1)	(2)
Dependent variable	Hired	Hired
Treated	0.0004	-0.0012***
	[0.0004]	[0.0004]
SS_High	0.0079***	0.0049***
	[0.0004]	[0.0006]
Treated x SS_High		0.0060***
		[0.008]
Observations	895,882	895,882
R-squared	0.0004	0.0005

Notes. The unit of analysis is an application to a job posted by the treatment/control group client during the study. The mean for the dummy variable *Hired* is 0.029; the standard deviation is 0.169. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: OLS Regressions on the Treatment Effect of High SSs on Disintermediation

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Log(Hours)	Log(Hours)	Log(Total_Charge)	Log(Total_Charge)	Log(Disintermediation _Score)	Log(Disintermediation _Score)
Treated	-0.058*	-0.003	-0.034*	-0.002	0.038***	0.018
	[0.030]	[0.037]	[0.019]	[0.024]	[0.010]	[0.012]
SS_High	0.348***	0.422***	0.455***	0.499***	-0.098***	-0.128***
	[0.032]	[0.046]	[0.021]	[0.030]	[0.010]	[0.015]
Treated x SS_High		-0.143**		-0.086**		0.058***
		[0.064]		[0.041]		[0.021]
Observations	14,593	14,593	33,561	33,561	29,690	29,690
R-squared	0.009	0.009	0.015	0.015	0.003	0.004

Notes. Observations are all job assignments created during the study period. The sample in Column (1) contains only hourly jobs. The sample in Column (5) and (6) contains only observations with non-missing disintermediation scores. *SS_High* is defined as the freelancer having a Satisfaction Score greater than or equal to 0.9 at the time of the study. *Treated*, a dummy variable for treatment at the client level, equals 1 if the client is in the treatment group. *Log(Hours)* is the logarithm of the number of hours the freelancer worked on the assignment. *Log(Total_Charge)* is the logarithm of the total amount of money charged at the end of the assignment plus 1. *Log(Disintermediation_Score)* is the logarithm of the disintermediation score computed from all messages associated with the assignment plus 1. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Heterogeneity in Disintermediation Tendencies

Model	(1)	(2)	(3)	(4)
Treated	0.019	0.006	0.025	0.063***
	[0.013]	[0.023]	[0.023]	[0.017]
SS_High	-0.102***	-0.126	-0.131***	-0.124***
	[0.016]	[0.028]	[0.029]	[0.021]
Treated x SS_High	0.041*	-0.026	0.028	0.021
	[0.022]	[0.039]	[0.040]	[0.030]
Same_Country	-0.037			
	[0.032]			
Treated x Same_Country	-0.013			
	[0.046]			
SS_High x Same_Country	-0.242***			
	[0.051]			
Treated x SS_High x Same_Country	0.154**			
	[0.074]			
Divisible_Med		0.079***		
		[0.019]		
Divisible_High		0.243***		
		[0.030]		
Treated x Divisible_Med		0.011		
		[0.028]		
Treated x Divisible_High		0.022		
		[0.042]		
SS_High x Divisible_Med		-0.000		
		[0.034]		
SS_High x Divisible_High		0.048		
		[0.053]		
Treated x SS_High x Divisible_Med		0.107**		
		[0.047]		
Treated x SS_High x Divisible_High		0.144**		
		[0.072]		
Long_Term			-0.114***	
			[0.026]	
Treated x Long_Term			-0.038	
			[0.040]	
SS_High x Long_Term			-0.021	
			[0.046]	
Treated x SS_High x Long_Term			0.112*	
			[0.063]	
Client_Rating_High				-0.132***
				[0.017]
Treated x Client_Rating_High				-0.089***
				[0.024]
SS_High x Client_Rating_High				-0.012
				[0.030]
Treated x SS_High x Client_Rating_High				0.071*
				[0.041]
Observations	29,690	29,690	12,118	29,690

R-squared	0.006	0.014	0.010	0.014

Notes. The dependent variable in this table is *Log(Disintermediation_Score)*. Observations are the job assignments created during the study with non-missing disintermediation scores. Model (3) includes only jobs with information on the expected duration. Robust standard errors in brackets.

Appendix

Table A1: Summary Statistics and Correlations, Moderators

Variables	Obs.	Mean	Std. dev.	Min	Max	(1)	(2)	(3)	(4)
(1) Same_Country	29,690	0.080	0.272	0	1	1			
(2) Divisible_High	29,690	0.110	0.312	0	1	0.004	1		
(3) Divisible_Med	29,690	0.617	0.486	0	1	0.040	-0.445	1	
(4) Long_Term	12,118	0.405	0.491	0	1	-0.006	-0.194	-0.020	1
(5) Client_Rating_High	29,690	0.534	0.499	0	1	-0.057	-0.076	-0.019	0.195

Notes: The sample of this table is the main analysis sample, the number of observations in this table is 29,690, except for *Long_Term* which includes only jobs with information on their expected duration.

Table A2: Regression of High SSs on Disintermediation with Discrepant vs. Consistent SS and Five-Star Rating

Panel A – High SS, High Five Star Rating

Model	(1)	(2)	(3)
Dependent variable	Log(Hours)	Log(Total_Charge)	Log(Disintermediation_Score)
Treated	-0.331***	-0.166**	0.112***
	[0.118]	[0.081]	[0.037]
Observations	1,299	2,703	2,199
R-squared	0.006	0.002	0.004

Panel B – High SS, Low Five Star Rating

Model	(1)	(2)	(3)
Dependent variable	Log(Hours)	Log(Total_Charge)	Log(Disintermediation_Score)
Treated	-0.094*	-0.068*	0.066***
	[0.057]	[0.037]	[0.019]
Observations	4,270	9,512	8,435
R-squared	0.001	0.000	0.001

Notes: The sample in this table is the assignment sample, split into cases where the freelancer's five-star rating is above the 90th percentile. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A3: Regression of High SSs on Disintermediation, Without Overlapping Freelancers

Model	(1)	(2)	(3)
Dependent variable	Log(Hours)	Log(Total_Charge)	Log(Disintermediation_ Score)
Treated	-0.028	-0.020	0.026*
	[0.040]	[0.027]	[0.014]
SS_High	0.504***	0.541***	-0.141***
	[0.052]	[0.034]	[0.017]
Treated x SS_High	-0.242***	-0.134***	0.063***
	[0.069]	[0.046]	[0.023]
Observations	12,606	28,003	24,427
R-squared	0.011	0.016	0.005

Notes: The analysis sample is the assignment sample for our main analysis, after dropping freelancers who are matched with control group clients after working with treated clients. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: OLS Regressions on the Treatment Effect of High SSs on Disintermediation, Using Pre-treatment Data

Model	(1)	(2)	(3)	(4)
Dependent variable	Log(Past_Hours)	Log(Past_Hours)	Log(Past_Total_Charge)	Log(Past_Total_Charge)
Treated	0.012	0.013	-0.007	0.006
	[0.020]	[0.022]	[0.017]	[0.020]
SS_High	0.404***	0.406***	0.607***	0.624***
	[0.022]	[0.032]	[0.019]	[0.026]
Treated x SS_High		-0.005		-0.035
		[0.044]		[0.037]
Observations	24,848	24,848	31,825	31,825
R-squared	0.015	0.015	0.035	0.035

Notes: Observations are the job assignments in the main analysis with job outcomes replaced by each freelancer's average past job outcome in the 6 months before the study. There are 33,561 assignments from the study; 31,825 involve a freelancer with at least one previous job in the past 6 months and 24,848 involve a freelancer who worked on an hourly job in the past 6 months. *SS_High* and Treated are defined as in Table IV. *Log(Past_Hours)* is the logarithm of the average number of hours the freelancer worked on each assignment in the past 6 months. *Log(Past_Total_Charge)* is the logarithm of the freelancer's average total charge per assignment in the past 6 months plus 1. Robust standard errors in brackets. *** $p<0.01$.

.

Table A5: Repeating the OLS Regressions on the Treatment Effect of High SSs on Disintermediation using a More Robust Disintermediation Score

Model	(1)	(2)
Dependent variable	Log(Disintermediation_Score_Robust)	Log(Disintermediation_Score_Robust)
Treated	-0.002	-0.005
	[0.002]	[0.003]
SS_High	-0.005**	-0.009***
	[0.002]	[0.003]
Treated x SS_High		0.008**
		[0.004]
Observations	29,690	29,690
R-squared	0.0002	0.0003

Notes: Observations are all the job assignments created during the study period. The sample and independent variable definitions are the same as in Table 5. *Log(Disintermediation_Score_Robust)* is the logarithm of the robust disintermediation score computed based on the subset of the more explicit keywords plus 1. Robust standard errors in brackets. *** p<0.01, ** p<0.05.

Table A6: First-Time Hires vs. Repeated Hires

Panel A: Regression of High SSs on Disintermediation, First-Time Hires Only

Model	(1)	(2)	(3)
Dependent variable	Log(Hours)	Log(Total_Charge)	Log(Disintermediation_Score)
Treated	-0.0003	-0.018	0.027*
	[0.040]	[0.027]	[0.014]
SS_High	0.439***	0.578***	-0.087***
	[0.050]	[0.035]	[0.017]
Treated x SS_High	-0.124*	-0.084*	0.047**
	[0.069]	[0.048]	[0.023]
Observations	12,694	26,469	23,577
R-squared	0.010	0.020	0.002

Panel B: Regression of High SSs on Disintermediation, Repeated Hires Only

Model	(1)	(2)	(3)
Dependent variable	Log(Hours)	Log(Total_Charge)	Log(Disintermediation_Score)
Treated	-0.024	0.059	-0.018
	[0.099]	[0.049]	[0.027]
SS_High	0.318***	0.302***	-0.265***
	[0.121]	[0.056]	[0.030]
Treated x SS_High	-0.282*	-0.138*	0.096**
	[0.166]	[0.079]	[0.044]
Observations	1,899	7,092	6,113
R-squared	0.005	0.006	0.017

Notes: The sample in the above table is the main analysis sample. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: OLS Regressions on the Treatment Effect of High SSs on Client Feedback of Freelancers

Model	(1)	(2)
Dependent variable	Client_Feedback	Client_Feedback
Treated	-0.001	-0.006
	[0.008]	[0.011]
SS_High	0.102***	0.094***
	[0.007]	[0.011]
Treated x SS_High		0.016
		[0.015]
Observations	23,501	23,501
R-squared	0.007	0.007

Notes: Observations are the job assignments in the main analysis with non-missing client feedback. Independent variable definitions are the same as in Table 5. *Client_Feedback* is the five-star feedback rating that the client left when the job assignment was closed. Robust standard errors in brackets. *** p<0.01.