Incentivizing Communication through Recommendation:
Signaling, Screening, and Congestion in an Online Labor Market

Extended Abstract

Anaïs Galdin and Jesse Silbert[†]

October 26, 2023

Agents in two-sided matching markets use a mix of platform-guided recommendation and decentralized communication to learn about potential matches before matching. This paper studies whether, and how, market designers can use recommendation systems to incentivize effective communication between potential matches in the context of an online freelance labor market for tasks. With data from a major international online labor market, we build and estimate a novel model of freelancing supply and demand, in which firms post single-task online jobs they wish to outsource, freelancers submit monetary bids and cover letters for those jobs, the platform algorithmically ranks the bids on each job, and firms, seeing these recommendations, choose which freelancer to hire. In our model, freelancers pay up-front effort costs when writing cover letters, and the more writing effort they expend, the more informative the resulting signal is. However, when freelancers face congestion, they lower their equilibrium writing effort, hampering firms' abilities to find the worker with the highest match quality. We use our estimated model to investigate how the platform could redesign the recommender system to lower congestion and thus incentivize more communication effort and improve the quality of realized matches. Specifically we study the equilibrium effects of a recommender system that promotes bids that exhibit more effort.

To answer our research question in our setting, we need a model of communication in a two-sided matching market with transfers. We develop a model that embeds a career concerns model of signaling (Holmstrom, ReStud 1999) into a scoring procurement auction in which firms care both about match quality and price. Importantly, freelancers and firms do not know more about their match quality than what is predicted by their observables. The key idea here is that we model signals as being produced such that effort magnifies the contribution of match quality to the overall signal output relative to the random error. Consequently, all else equal, if we can shift the equilibrium in a counterfactual such that each freelancer increases effort, then it would be possible to make signals more informative about match quality across freelancers. Thus, to increase learning about match quality for any one potential match, a market designer could find a way to increase equilibrium effort from signal senders. This mechanism will be the target of counterfactuals.

[†]Department of Economics, Princeton University, agaldin@princeton.edu and jesseas@princeton.edu

We also incorporate the notion of recommendation by modeling how firms form their consideration sets. Specifically, we know that in our context, and in many contexts with significant congestion, it is not rational for firms to consider all possible applications since doing so can be prohibitively costly. As such, firms rely on the platform's recommendations to form consideration sets á la simultaneous search looking at a subset of bids corresponding to the top ranked bids up to a randomly drawn size of the consideration set. We model the recommendation algorithm using our data source's actual implemented algorithm, and this is the algorithm we change under counterfactuals. The algorithm used by our data source takes as inputs the observable reputation of the freelancers bidding, rather than the attributes of the bids themselves. Freelancers without a strong reputation, namely freelancers who have not yet gotten many matches on the platform, thus face a very low chance of being considered, which incentivizes them not to expend a lot of effort to signal their match quality. Thus, there are possibly many potential good matches that have a very low chance of forming. Our counterfactuals target this market failure by changing the recommendation algorithm to respond to the effort put into the bid: thus, to be considered, freelancers must put effort into signaling her match quality.

Our data comes from Freelancer.com, a major global online labor market. We will have access to all data recording activity on the platform, including posted tasks, freelancer identities, bids, the text of cover letters, observable characteristics about the freelancers and the firms, timestamps measuring how long each freelancer spends writing their cover letter, click data from firms to measure consideration sets, the inputs, outputs, and formula for the recommendation/ranking algorithm, and of course who was chosen and the outcome of each task, including ratings of the freelancers and task completion status. We measure writing effort (assuming we're doing so with classical measurement error) by how long each freelancer spends writing their cover letter. We also construct a novel measure of writing quality, employing a large language model (LLM). We feed each task description to the LLM and prompt it to role-play as the firm posting the task. We then feed each cover letter (without any other information) to the LLM and prompt it to score the writing based on a judgement of how well the freelancer writing the cover letter will do on the task described. We perform various validation checks that our measure of writing quality is stable across multiple independent calls of the LLM and that it corresponds to more conventional, but less tailored, measures of writing quality.

We construct an identification argument that splits our model into three parts, which form the basis for our estimation strategy. First, data on writing quality and writing effort identify the mean of match quality and a linear function of the variances of match quality and writing error. These are identified through slight variants of the usual moments from a linear regression model.

Second, following Guerre, Perrigne, and Vuong (ECMA, 2000), we use first order conditions for bids to identify cost parameters. Specifically, we write down the first order conditions of the freelancers' problem in terms of the reduced-form probability of winning given the choice of bid and writing effort. Given the monotonicity conjectures for the equilibrium strategies, the first order condition with respect to bid price identifies the distribution of opportunity costs, and the first order condition with respect to writing effort identifies the marginal cost of writing effort. The main challenge is to show that the first order conditions themselves are identified, since they involve derivatives of the win probability function which involve off-equilibrium, and thus unobserved, pairs of bids and writing efforts. We overcome this challenge by exploiting the fact that firms only see the resulting writing quality, which is produced randomly once we condition on writing effort, which we as econometricians observe. Thus, we can non-parametrically identify the reduced-form win probability function for off-equilibrium strategies by carefully reweighting the reduced-form win probability function of bids and resulting writing qualities — rather than efforts. Doing so allows us to non-parametrically identify the first order conditions, and thus the desired cost parameters.

Third, we turn to the moments generated by the firm's discrete choice of winning freelancer. The choice between inside options (freelancer vs. freelancer) identifies the firm's relative preference between match quality and price, and the choice between inside options and outside options (freelancer vs. outside option) identifies the ratio of the variances of match quality and writing error. Finally, we exploit random discounts given on the platform to firms to identify firms' price sensitivity relative to the normalized value of the outside option. Putting this all together, we have identified what we need in our model to perform counterfactuals.

Our paper makes three distinct contributions. First, we draw conclusions about how online matching platforms can use recommender systems to incentivize effective communication between potential matches, improving the efficiency of the market. Second, we provide a tractable model of a scoring auction that embeds a signaling game, which can be used in other settings such as procurement auctions with uncertain quality. Finally, we are the first in the economics literature to employ LLMs to transform text data into quantified signals that allow us to estimate a model of strategic communication.