Inferring Private Valuations from Behavioral Data in Bilateral Sequential Bargaining

Lvye Cui , **Haoran Yu***
School of Computer Science & Technology, Beijing Institute of Technology
cui_lvye@outlook.com, yhrhawk@gmail.com

Abstract

Inferring bargainers' private valuations on items from their decisions is crucial for analyzing their strategic behaviors in bilateral sequential bargaining. Most existing approaches that infer agents' private information from observable data either rely on strong equilibrium assumptions or require a careful design of agents' behavior models. To overcome these weaknesses, we propose a **B**ayesian **L**earning-based Val**u**ation Inf**e**rence (BLUE) framework. Our key idea is to derive *feasible intervals* of bargainers' private valuations from their behavior data, using the fact that most bargainers do not choose strictly dominated strategies. We leverage these feasible intervals to guide our inference. Specifically, we first model each bargainer's behavior function (which maps his valuation and bargaining history to decisions) via a recurrent neural network. Second, we learn these behavior functions by utilizing a novel loss function defined based on feasible intervals. Third, we derive the posterior distributions of bargainers' valuations according to their behavior data and learned behavior functions. Moreover, we account for the heterogeneity of bargainer behaviors, and propose a clustering algorithm (K-Loss) to improve the efficiency of learning these behaviors. Experiments on both synthetic and real bargaining data show that our inference approach outperforms baselines.

1 Introduction

With the growing popularity of e-commerce, the bilateral sequential bargaining mechanisms that allow for haggling between sellers and potential buyers have emerged on online platforms [Gujral, 2016; Huang *et al.*, 2013]. One prerequisite for designing and optimizing these bargaining mechanisms is analyzing bargainers' strategic behaviors. To achieve this target, it is important to infer bargainers' private valuations on the negotiated items. For example, if a seller's valuation on an item is known, it is easy to analyze whether the seller deceptively asks a high price during bargaining. Moreover, knowing the seller's valuation helps quantify the efficiency and equality of the bargaining outcome.

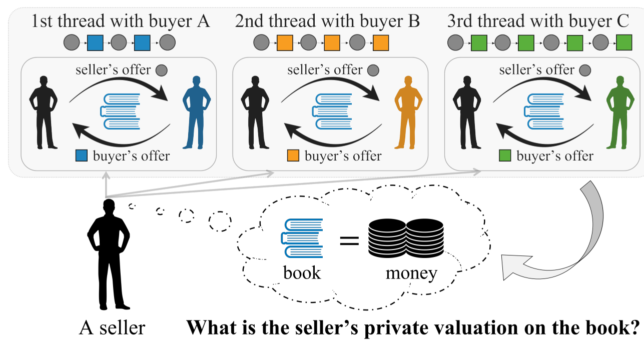


Figure 1: An Example of Our Inference Problem.

Many prior works applied an equilibrium-based approach to infer agents' private information (e.g., their valuations on items) [Athey and Haile, 2002; Jiang and Leyton-Brown, 2007; Athey and Nekipelov, 2010; Bajari *et al.*, 2013]. The approach assumes that all agents are fully rational, each agent's action is the best response to others' actions, and their interactions always converge to a Nash equilibrium. It infers each agent's private information by computing the inverse of his best response function given his observed behavior. However, *this equilibrium-based approach cannot solve the valuation inference problem for real-world bargaining platforms.* This is because the strong assumption that agents are fully rational and play equilibrium strategies may not hold in practice [Goeree and Holt, 2001; Kagel and Roth, 2020; Noti, 2021]. Bargaining platforms usually involve millions of bargainers, who have different capabilities to analyze incomplete information extensive-form games and distinct beliefs about their opponents' information.

We focus on inferring bargainers' private valuations based on their strategic behaviors on bilateral bargaining platforms, without relying on equilibrium assumptions. An example of these platforms is eBay's Best Offer [Backus *et al.*, 2020], where millions of sellers negotiate with buyers over their listed items. We use a *thread* to denote the price sequence generated from a seller-buyer pair's negotiation over an item, and a thread may consist of multiple *rounds*. In each round,

*Corresponding author.

given the opponent's last offer, the seller and buyer can *accept* it, *decline* it, or *make a counteroffer*.

A seller may negotiate with multiple buyers over the price of an item (or different copies of the same item, e.g., coupon cards, commemorative stamps, artworks, and sports trading cards). The seller's valuation on the item across these negotiations can usually be approximated as a constant, and we attempt to infer it by jointly analyzing the threads of these negotiations. Figure 1 shows an example, where a seller has three different threads after negotiating a book's price with three buyers. Given the three thread records, we intend to infer the seller's valuation on this book.

The main challenge of solving our problem is that sellers' private valuations cannot be directly observed in the bargaining data. This implies that we do not have labels to guide the valuation inference process. We address the challenge and propose a **B**ayesian **L**earning-based Val**u**ation Inf**e**rence (BLUE) scheme. Our key idea is to define and derive *feasible intervals* of sellers' private valuations according to their observed behaviors and leverage these feasible intervals to guide the inference of private valuations. Specifically, we make a mild rationality assumption that sellers do not choose *strictly dominated strategies* [Luce and Raiffa, 1989; Gopalakrishnan *et al.*, 2018], e.g., sellers never propose or agree on prices lower than their valuations. This assumption enables us to utilize sellers' behavior data to derive *posterior feasible intervals* of their private valuations, which will be used to guide the inference of these valuations.

Next, we briefly introduce our inference framework. First, we model each seller's behavior function (which maps his bargaining history and private valuation to his decision) via a recurrent neural network. An accurate behavior function enables us to infer the seller's valuation from his observed decision. Second, we learn these behavior functions (i.e., the parameters of the neural networks) using sellers' behavior data. We derive feasible intervals of valuations according to the rationality assumption, and design a novel loss function based on the feasible intervals to train the neural networks. Our loss function ensures that the behaviors characterized by trained neural networks satisfy the rationality assumption and well fit the bargaining data. Third, we apply Bayes' rule to derive the posterior distributions of sellers' private valuations based on sellers' behavior data and trained neural networks. In addition to above three steps, we consider the heterogeneity among sellers' behaviors, and propose a clustering algorithm to enhance the learning of heterogeneous behavior functions.

We evaluate our method on two synthetic datasets and one large real-world dataset [Backus *et al.*, 2020]. Compared with baselines, our method achieves more accurate and effective inference. Our contributions are two-fold:

- On the application side, we develop an effective valuation inference scheme for bilateral bargaining platforms. With inferred valuations of bargainers, these platforms can analyze their behaviors more accurately, and further optimize bargaining mechanisms to induce more bargainers to reach agreements.
- On the methodology side, we propose a novel framework for private information inference, through combining techniques from both economics and machine learning. Our framework is especially applicable to practical environments where agents cannot play equilibrium strategies due to their limited reasoning capabilities.

2 Related Work

2.1 Unknown Parameter Estimation in Games

Our work is related to the research that utilizes players' observed actions to estimate unknown parameters of games (e.g., the payoff matrices in normal-form games). Some works tackled the estimation problem by assuming that players play according to Nash equilibria [Varian, 2007; Vorobeychik *et al.*, 2007; Waugh *et al.*, 2011; Blum *et al.*, 2014; Honorio and Ortiz, 2015]. Some recent studies considered other equilibrium assumptions. For example, [Ling *et al.*, 2018; Ling *et al.*, 2019; Noti, 2021; Wu *et al.*, 2022] utilized the quantal-response equilibria (QRE) or Nested Logit QRE to estimate the parameters in games (such as normal-form games and Stackelberg games). These equilibrium-based estimation methods require strong cognitive and informational assumptions, e.g., players have full information of others' private preferences. Such requirements are usually unsatisfied in practical bilateral bargaining, where bargainers have limited knowledge about others' private information. In contrast, our inference scheme only requires bargainers not to choose strictly dominated strategies, and hence can handle more general human behavior in reality.

2.2 Private Information Inference in Auctions

Our work is also closely related to the literature that utilizes agents' bidding data to infer their private information in auctions. [Athey and Haile, 2002; Jiang and Leyton-Brown, 2007; Athey and Nekipelov, 2010; Bajari *et al.*, 2013] inferred bidders' valuations on items according to their bids, by assuming that they are fully rational and play equilibrium strategies. An emerging line of research relaxes the full rationality assumption, and assumes that agents' strategies can be captured by specific bounded rational behavior models. For example, [Noti and Syrgkanis, 2021; Nekipelov *et al.*, 2015; Nisan and Noti, 2017] inferred the information (e.g., the value-per-click) of each bidder in sponsored search auctions, by assuming that each bidder is a no-regret learner and bids to minimize the regret. These inference schemes require a careful manual design of agents' behavior models, and do not consider the heterogeneity of behaviors across agents. Moreover, they are mainly applicable to repeated auctions, which are distinct from online bargainings. By contrast, our scheme does not impose assumptions on the forms of agents' behavior models. It can automatically characterize bargainers' heterogeneous behaviors using neural networks, instead of manually designing specific unified models for them.

Other related studies include opponent modeling in automated negotiation [Hindriks and Tykhonov, 2008; Williams *et al.*, 2011; Chen and Weiss, 2012; Baarslag *et al.*, 2016; de Jonge, 2022] and preference elicitation [Boutilier *et al.*, 2006; Guo and Sanner, 2010; Viappiani and Boutilier, 2010; Vendrov *et al.*, 2020]. Different from these studies, our work focuses on price negotiation instead of multi-issue negotia-

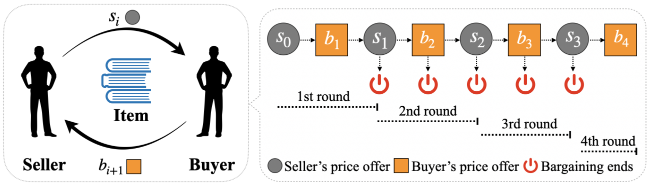


Figure 2: A bargaining thread with at most four rounds: s_0 is the seller's initial price, and the bargaining ends if (i) the seller accepts the buyer's offer, or (ii) the buyer accepts the seller's offer, or (iii) the buyer declines the seller's offer.

tion, and learns private information without querying individuals.

3 Problem Formulation

In this section, we introduce the bilateral sequential bargaining, and describe the valuation inference problem.

3.1 Bilateral Sequential Bargaining

We use a *bargaining thread* to refer to the sequence of price offers made by a seller-buyer pair bargaining over an item. The thread consists of at most L bargaining rounds. Let s_0 denote the seller's initial price offer. Given s_0, the buyer gives a counteroffer b_1 in the first round. Then, the seller and the buyer would make decisions in the following way:

- In round $i = 1, 2, \dots, L-1$, given the buyer's price b_i, the seller may (1) accept the offer, (2) make a counteroffer, or (3) decline the offer (i.e., insisting on his last offer s_{i-1}). If the seller accepts b_i, the bargaining ends in round i; otherwise, bargaining goes to the next round.
- In round $i = 2, 3, \dots, L-1$, the buyer can also accept, decline, or counter in response to the seller's price s_{i-1}. If the buyer accepts or declines s_{i-1}, the bargaining terminates in round i; otherwise, it comes to the next round.
- In the final round L, the buyer accepts or declines the seller's offer s_{L-1}, and the bargaining ends.

This is known as the alternating offers protocol [Raiffa, 1982], and we show an example of $L = 4$ in Figure 2.

3.2 Private Valuation Inference Problem

We use $v_m^{(q)}$ to denote the *private valuation* of seller q for item m. It is defined as the lowest price that seller q can accept for selling item m. We consider a discrete $v_m^{(q)}$, since this lowest price is normally measured in dollars. This private value $v_m^{(q)}$ directly affects seller q's bargaining behavior, as he seeks to sell item m with a price higher than $v_m^{(q)}$.

Let $\{(\boldsymbol{x}_{ih}^{(qm)}, y_{ih}^{(qm)})\}_{i=1}^{I_h}$ denote the set of labeled data points about thread h of seller q for item m, where I_h is the total number of rounds in thread h. Here, $y_{ih}^{(qm)}$ denotes seller q's decision in round i of thread h for item m, and $\boldsymbol{x}_{ih}^{(qm)}$ represents the sequence of all prior offers made before $y_{ih}^{(qm)}$ in thread h. For example, considering the thread in Figure 2, when $i = 1$, there are two offers before the seller's decision of s_1, which indicates that $\boldsymbol{x}_{ih}^{(qm)} = (s_0, b_1)$ and $y_{ih}^{(qm)} = s_1$. When $i = 2$, $\boldsymbol{x}_{ih}^{(qm)} = (s_0, b_1, s_1, b_2)$ and $y_{ih}^{(qm)} = s_2$.

We denote all threads that seller q has participated in to sell item m by $\mathcal{H}_m^{(q)} = \{\{(\boldsymbol{x}_{ih}^{(qm)}, y_{ih}^{(qm)})\}_{i=1}^{I_h}\}_{h=1}^{H_m}$, where H_m is the number of threads related to item m. Then, we characterize all sellers' behaviors by *thread dataset* $\mathcal{D}_{\mathcal{H}} = \{\{\mathcal{H}_m^{(q)}\}_{m=1}^{M_q}\}_{q=1}^{Q}$, where Q is the number of sellers and M_q is the number of items that seller q bargains over. We describe our inference problem as follows:

Problem 1. *Given thread dataset $\mathcal{D}_{\mathcal{H}}$, we attempt to infer all sellers' private valuations on items, i.e., $\{\{v_m^{(q)}\}_{m=1}^{M_q}\}_{q=1}^{Q}$.*

4 Private Valuation Inference Solution

We first explain our idea of solving Problem 1. Since a seller's valuation is usually correlated with his bargaining behavior, we propose to model his behavior via a function $f_{\boldsymbol{\theta}}^{(q)}$: $(v_m^{(q)}, \boldsymbol{x}_{ih}^{(qm)}) \to y_{ih}^{(qm)}$. The function (characterized by parameters $\boldsymbol{\theta}$) maps seller q's valuation $v_m^{(q)}$ and prior offers $\boldsymbol{x}_{ih}^{(qm)}$ to his decision $y_{ih}^{(qm)}$. In particular, $y_{ih}^{(qm)}$ belongs to the set $\{\text{accept}, \text{decline}, \text{counter1}, \dots, \text{counterN}\}$, where set $\{\text{counter1}, \dots, \text{counterN}\}$ includes all possible counter prices (e.g., measured in dollars). Once we learn this function, we can utilize Bayes' rule to derive the posterior distribution $\Pr(v_m^{(q)} | \mathcal{Y}_m^{(q)}; \mathcal{X}_m^{(q)}, \boldsymbol{\theta})$ for $v_m^{(q)}$:

$$\Pr\left(v_m^{(q)}\middle|\mathcal{Y}_m^{(q)};\mathcal{X}_m^{(q)},\boldsymbol{\theta}\right) = \frac{\Pr\left(\mathcal{Y}_m^{(q)}\middle|v_m^{(q)};\mathcal{X}_m^{(q)},\boldsymbol{\theta}\right)\Pr\left(v_m^{(q)}\right)}{\sum_{\tilde{v}_m^{(q)}}\Pr\left(\mathcal{Y}_m^{(q)}\middle|\tilde{v}_m^{(q)};\mathcal{X}_m^{(q)},\boldsymbol{\theta}\right)\Pr\left(\tilde{v}_m^{(q)}\right)}$$

$$= \frac{\prod_{h=1}^{H_m}\prod_{i=1}^{I_h}\Pr\left(y_{ih}^{(qm)}\middle|v_m^{(q)};\boldsymbol{x}_{ih}^{(qm)},\boldsymbol{\theta}\right)\Pr\left(v_m^{(q)}\right)}{\sum_{\tilde{v}_m^{(q)}}\left(\prod_{h=1}^{H_m}\prod_{i=1}^{I_h}\Pr\left(y_{ih}^{(qm)}\middle|\tilde{v}_m^{(q)};\boldsymbol{x}_{ih}^{(qm)},\boldsymbol{\theta}\right)\Pr\left(\tilde{v}_m^{(q)}\right)\right)}, \tag{1}$$

where $\Pr(v_m^{(q)})$ is the prior probability distribution of $v_m^{(q)}$, and $\mathcal{X}_m^{(q)}$ and $\mathcal{Y}_m^{(q)}$ include the prior offers and seller behaviors of all threads related to item m and seller q, respectively. Here, we assume independence among threads to derive the second equality of (1), which can reduce the computational complexity of our scheme. With $\Pr(v_m^{(q)}|\mathcal{Y}_m^{(q)}; \mathcal{X}_m^{(q)}, \boldsymbol{\theta})$, we can infer each $v_m^{(q)}$ by:

$$v_m^{(q)} = \arg\max_{\tilde{v}_m^{(q)}} \Pr\left(\tilde{v}_m^{(q)}\middle|\mathcal{Y}_m^{(q)};\mathcal{X}_m^{(q)},\boldsymbol{\theta}\right). \tag{2}$$

Derivation of (1) requires an accurate behavior model $f_{\boldsymbol{\theta}}^{(q)}$ to compute $\Pr(y_{ih}^{(qm)}|v_m^{(q)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta})$. To achieve this, we design a novel loss function to optimize $\boldsymbol{\theta}$ given thread dataset $\mathcal{D}_{\mathcal{H}}$. Our loss function is defined based on a *posterior feasible interval* of $v_m^{(q)}$. Next, we introduce this feasible interval in Section 4.1, and study the optimization of $\boldsymbol{\theta}$ in Section 4.2. To simplify the explanation, we first assume all sellers have the same behavior pattern, i.e., $f_{\boldsymbol{\theta}}^{(q)} = f_{\boldsymbol{\theta}}$ for all $q \in \{1, \dots, Q\}$. In Section 4.3, we extend our solution to the case where sellers have distinct behavior patterns.

4.1 Feasible Interval of Valuation

In the economics and game theory literature, strictly dominated strategies refer to the strategies that always result in lower payoffs than alternative strategies regardless of other players' strategies [Luce and Raiffa, 1989; Gopalakrishnan *et al.*, 2018; Qin *et al.*, 2019; Maschler *et al.*, 2020]. We make a rationality assumption that sellers never choose strictly dominated strategies. Specifically, if the bargaining ends with an agreement, the seller's payoff equals the agreed price minus his valuation; otherwise, his payoff is zero. Under our rationality assumption, accepting the buyer's offer implies that the seller's valuation is no greater than the agreed price. This is because once the seller's valuation is greater than the price, accepting becomes a strictly dominated strategy and the seller will not choose this decision. Formally, our rationality assumption is as follows.

Assumption 1. *In the bargaining process:(i) if a seller accepts a buyer's offer, the seller's valuation is no greater than the price; (ii) if a seller declines the buyer's offer in the last round, the seller's valuation is no less than the price; (iii) a seller never proposes a price less than his valuation.*

We define the feasible interval of a seller's valuation as the set of all possible valuation values satisfying Assumption 1. Given thread dataset $\mathcal{D}_{\mathcal{H}}$, we could derive the bounds $c_m^{(q)}$ and $d_m^{(q)}$ of the feasible interval of valuation $v_m^{(q)}$ as follows:

$$\begin{cases} c_m^{(q)} &= \max\left\{c_{m1}^{(q)}, c_{m2}^{(q)}, \ldots, c_{mh}^{(q)}, \ldots, c_{mH_m}^{(q)}\right\}, \\ d_m^{(q)} &= \min\left\{d_{m1}^{(q)}, d_{m2}^{(q)}, \ldots, d_{mh}^{(q)}, \ldots, d_{mH_m}^{(q)}\right\}. \end{cases}$$

Here, we calculate $[c_m^{(q)}, d_m^{(q)}]$ as the intersections of the feasible intervals of all H_m threads. For thread h, we calculate its bounds $c_{mh}^{(q)}$ and $d_{mh}^{(q)}$ according to seller q's decision $y_{I_h h}^{(qm)}$ in last round I_h: (i) When $y_{I_h h}^{(qm)} \neq$ accept, we have $c_{mh}^{(q)} = b_{I_h h}^{(qm)}$ and $d_{mh}^{(q)} = \min\{s_{1h}^{(qm)}, s_{2h}^{(qm)}, \ldots, s_{I_h h}^{(qm)}\}$; (ii) Otherwise, we have $c_{mh}^{(q)} = 0$ and $d_{mh}^{(q)} = \min\{b_{I_h h}^{(qm)}, s_{1h}^{(qm)}, \ldots, s_{(I_h-1)h}^{(qm)}\}$.

4.2 Learning of Homogeneous Behavior

Given thread dataset $\mathcal{D}_{\mathcal{H}}$ and the set of our derived bounds $\{\{c_m^{(q)}, d_m^{(q)}\}_{m=1}^{M_q}\}_{q=1}^{Q}$, we design a novel loss function to learn $\boldsymbol{\theta}$. Our loss function consists of the following two parts:

- The first part is related to the rationality assumption. Given $\boldsymbol{\theta}$, we can infer the seller's valuation $v_m^{(q)}$. With accurate $\boldsymbol{\theta}$, the inferred $v_m^{(q)}$ will fall into its feasible interval $[c_m^{(q)}, d_m^{(q)}]$. Thus, our loss function's first part is the following negative log-likelihood function:

$$-\sum_{q=1}^{Q}\sum_{m=1}^{M_q}\sum_{h=1}^{H_m}\sum_{i=1}^{I_h} \log \Pr\left(c_m^{(q)} \le v_m^{(q)} \le d_m^{(q)} \middle| y_{ih}^{(qm)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}\right).$$

  Here, $\Pr(c_m^{(q)} \le v_m^{(q)} \le d_m^{(q)} | y_{ih}^{(qm)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta})$ is the probability that the inferred valuation $v_m^{(q)}$ belongs to the feasible interval $[c_m^{(q)}, d_m^{(q)}]$, given the prior offers $\boldsymbol{x}_{ih}^{(qm)}$, the seller decision $y_{ih}^{(qm)}$, and the parameters $\boldsymbol{\theta}$.

Algorithm 1 Homogeneous Behavior Learning Algorithm

Input: Thread dataset $\mathcal{D}_{\mathcal{H}}$, prior distribution $p(v_m^{(q)})$, learning rate η, maximum epoch times T, weight factor α.
Output: Model $f_{\boldsymbol{\theta}_t}$.

1: Compute the interval bounds $\{\{c_m^{(q)}, d_m^{(q)}\}_{m=1}^{M_q}\}_{q=1}^{Q}$.
2: Initialize $\boldsymbol{\theta}$ with a random $\boldsymbol{\theta}_0$, and set $t = 0$.
3: **while** $t \le T$ **and** $\boldsymbol{\theta}$ does not converge **do**
4: Compute the gradient $\nabla_{\boldsymbol{\theta}_t}^{(1)}$ of the first part of the loss function according to Equation (4).
5: Estimate $v_m^{(q)}$ given $\boldsymbol{\theta}_t$ based on Equations (1)-(2).
6: Compute the gradient $\nabla_{\boldsymbol{\theta}_t}^{(2)}$ of the second part of the loss function (shown in Equation (3)).
7: Update $\boldsymbol{\theta}$: $\boldsymbol{\theta}_{t+1} \leftarrow \boldsymbol{\theta}_t - \eta\alpha\nabla_{\boldsymbol{\theta}_t}^{(1)} - \eta(1-\alpha)\nabla_{\boldsymbol{\theta}_t}^{(2)}$.
8: $t \leftarrow t + 1$.
9: **end while**
10: **return** $\boldsymbol{\theta}_t$

- The second part is associated with the accuracy of the behavior model. With accurate $\boldsymbol{\theta}$, we can predict the seller decision under its valuation $v_m^{(q)}$ and prior offers $\boldsymbol{x}_{ih}^{(qm)}$ as $f_{\boldsymbol{\theta}}(v_m^{(q)}, \boldsymbol{x}_{ih}^{(qm)})$, and it will be close to the seller's actual decision $y_{ih}^{(qm)}$. Hence, the second part of our loss function is the following function:

$$\sum_{q=1}^{Q}\sum_{m=1}^{M_q}\sum_{h=1}^{H_m}\sum_{i=1}^{I_h} \mathcal{L}\left(f_{\boldsymbol{\theta}}\left(v_m^{(q)}, \boldsymbol{x}_{ih}^{(qm)}\right), y_{ih}^{(qm)}\right), \quad (3)$$

  where $\mathcal{L}$ is the cross-entropy loss function. In the implementation, we do not know the actual value of $v_m^{(q)}$, which is our inference target. Instead, we use its estimated value, where the estimation is based on (1)-(2).

Our loss function is the weighted sum of the above two parts. Let $\alpha \in (0, 1)$ and $1-\alpha$ be the weights of the first and second parts, respectively. By minimizing the loss function over $\boldsymbol{\theta}$, we can learn a function $f_{\boldsymbol{\theta}}$ that both satisfies the rationality assumption and well fits the seller bargaining behavior.

We design a homogeneous behavior learning algorithm (Algorithm 1) to learn seller behavior via minimizing the loss function. In lines 4 to 6 of Algorithm 1, we compute the gradient of our loss function. Specifically, in line 4, we compute the gradient of the first part of the loss, which is shown in (4). After getting $\boldsymbol{\theta}$ via Algorithm 1, we utilize the Bayes' rule to infer each valuation $v_m^{(q)}$ with Equations (1) and (2).

4.3 Extension for Heterogeneous Behavior

In the presence of heterogeneous behavior, a natural approach is to learn a behavior model for each seller utilizing only his thread data. However, each seller may only have a small number of bargaining threads, which may be insufficient for behavior learning. Considering that many sellers have similar behavior patterns, we propose an algorithm to cluster sellers, as shown in Algorithm 2.

Specifically, we partition sellers into K clusters (denoted by $\{C_k\}_{k=1}^{K}$), and the sellers in C_k are assumed to have the same behavior pattern. Our algorithm iterates between two

$$
\begin{aligned}
\nabla_{\boldsymbol{\theta}_t}^{(1)} &= -\sum_{q=1}^{Q}\sum_{m=1}^{M_q}\sum_{h=1}^{H_m}\sum_{i=1}^{I_h} \frac{\partial \log \Pr\left(c_m^{(q)} \le v_m^{(q)} \le d_m^{(q)} \middle| y_{ih}^{(qm)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t} \\
&= \sum_{q=1}^{Q}\sum_{m=1}^{M_q}\sum_{h=1}^{H_m}\sum_{i=1}^{I_h} \left\{ \frac{\sum\limits_{v_m^{(q)}} \frac{\partial \Pr\left(y_{ih}^{(qm)} \middle| v_m^{(q)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t} \Pr\left(v_m^{(q)}\right)}{\sum\limits_{v_m^{(q)}} \Pr\left(y_{ih}^{(qm)} \middle| v_m^{(q)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}_t\right) \Pr\left(v_m^{(q)}\right)} - \frac{\sum\limits_{c_m^{(q)} \le v_m^{(q)} \le d_m^{(q)}} \frac{\partial \Pr\left(y_{ih}^{(qm)} \middle| v_m^{(q)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}_t\right)}{\partial \boldsymbol{\theta}_t} \Pr\left(v_m^{(q)}\right)}{\sum\limits_{c_m^{(q)} \le v_m^{(q)} \le d_m^{(q)}} \Pr\left(y_{ih}^{(qm)} \middle| v_m^{(q)}; \boldsymbol{x}_{ih}^{(qm)}, \boldsymbol{\theta}_t\right) \Pr\left(v_m^{(q)}\right)} \right\}.
\end{aligned} \tag{4}
$$

Algorithm 2 K-Loss Clustering Algorithm

Input: Thread dataset $\mathcal{D}_{\mathcal{H}}$, cluster number K, maximum epoch times E.
Output: $\{C_1, \ldots, C_K\}$.

1: Randomly assign sellers into K clusters $\{C_k\}_{k=1}^{K}$.
2: Let $e = 0$.
3: **while** $e \le E$ **and** the clustering does not converge **do**
4:     **for** $k = 1, \ldots, K$ **do**
5:         Learn behavior function $g_{\boldsymbol{\theta}}^{(C_k)}$, only based on the thread data related to the sellers in cluster C_k.
6:     **end for**
7:     Re-assign all sellers according to the prediction losses of K learned models on the seller thread data, i.e., assign seller q to cluster C_k if $g_{\boldsymbol{\theta}}^{(C_k)}$ has the smallest prediction loss on his bargaining behavior data.
8:     $e \leftarrow e + 1$.
9: **end while**
10: **return** $\{C_1, \ldots, C_K\}$

steps: (i) learning the behavior model of sellers in each cluster, and (ii) updating the assignment of sellers to K clusters based on the fitness between the sellers' behavior data and the learned behavior models. Particularly, we use $g_{\boldsymbol{\theta}}^{(C_k)}$ to denote the behavior pattern of sellers in cluster C_k. This $g_{\boldsymbol{\theta}}^{(C_k)}$ is similar to $f_{\boldsymbol{\theta}}$ and can be learned via Algorithm 1. Since we are mainly concerned with the clustering performance in Algorithm 2, one approach to accelerating the learning of $g_{\boldsymbol{\theta}}^{(C_k)}$ in line 5 is that we assume uniform valuations across sellers and simplify $g_{\boldsymbol{\theta}}^{(C_k)}$ as a mapping from seller's bargaining thread histories to seller's decision.

After clustering sellers into K clusters, we further learn their behavior functions (captured by $f_{\boldsymbol{\theta}_k}$: $(v_m^{(q)}, \boldsymbol{x}_{ih}^{(qm)}) \rightarrow y_{ih}^{(qm)}, \forall q \in C_k$) via Algorithm 1. Then, we infer their private valuations utilizing learned functions based on (1) and (2).

5 Experiments

5.1 Dataset Description

We generate synthetic data by simulating the process of bilateral bargaining (depicted in Figure 2). To model the diversity of sellers' and buyers' behaviors, we consider three kinds of patterns, i.e., *simple*, *random* and *payoff-maximization*.

- For the *simple* pattern, an agent accepts the opponent's offer as long as he can obtain a non-negative payoff.
- For the *random* pattern, an agent first estimates its opponent's private valuation, and then randomly takes actions from those that give it a non-negative expected payoff.
- For the *payoff-maximization* pattern, an agent chooses the action that maximizes his payoff, given his belief about the opponent. Specifically, the agent first updates his belief about the opponent's valuation based on the opponent's prior actions. Then, the agent reasons about the opponent's potential response in the new round, and chooses his action to maximize his expected payoff.

We synthesize two thread datasets (i.e., *synthetic data I* and *synthetic data II*), based on different distributions of agents' valuations. In *synthetic data I* and *synthetic data II*, all agents' valuations follow uniform and categorical distributions, respectively. We let the support of the valuation distribution be $\{10, 14, \cdots, 98\}$. In each of *synthetic data I* and *synthetic data II*, we simulate 900 sellers (with bargaining behaviors uniformly chosen among the three patterns) and there are about $120,000$ bargaining threads.

We also conduct experiments on a large dataset collected from eBay's Best Offer platform [Backus *et al.*, 2020]. It contains concrete bargaining information, e.g., seller ID, item ID, thread ID, and each thread's complete history. If a seller only has a limited number of decisions recorded, it is nearly impossible to infer his valuations on different items. Thus, we filter out these sellers, and our real dataset includes 32,099 sellers with 358,641 threads.

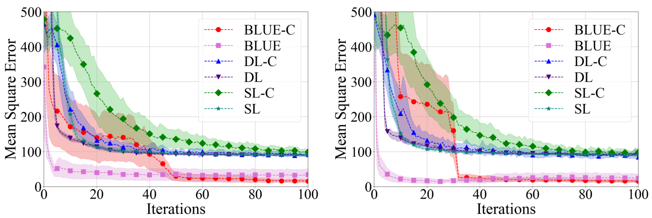
5.2 Experiment Settings

We compare six learning-based valuation inference methods (in Appendix A, we show the results of two sampling-based inference methods):

- **BLUE** (**B**ayesian **L**earning-based Val**u**ation Inf**e**rence): We assume all sellers have homogeneous behaviors, and implement Algorithm 1 for valuation inference.
- **BLUE-C** (**BLUE** with **C**lustering): We first cluster sellers via Algorithm 2, and then make inference for the sellers in each cluster via Algorithm 1.
- **DL** (**D**ual **L**earning): Dual learning [Qin, 2020; He *et al.*, 2016] is a learning framework that utilizes the primal-dual structure of AI tasks to advance the learning

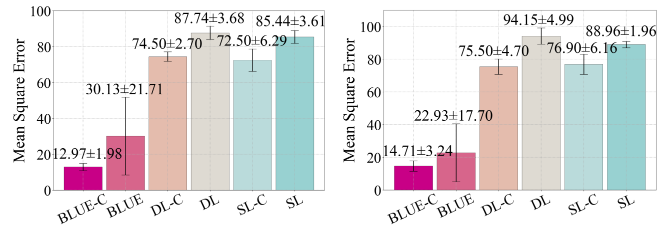
Methods/Datasets	G1		G2		G3	
	Precision	Recall	Precision	Recall	Precision	Recall
K-Means/D-I	0.647 ± 0.088	0.629 ± 0.156	0.480 ± 0.109	0.479 ± 0.067	0.543 ± 0.072	0.809 ± 0.209
AE-Image/D-I	0.698 ± 0.084	0.755 ± 0.118	0.736 ± 0.088	0.666 ± 0.136	0.993 ± 0.008	0.979 ± 0.027
K-Loss/D-I	$\mathbf{0.992 \pm 0.006}$	$\mathbf{0.998 \pm 0.002}$	$\mathbf{0.998 \pm 0.002}$	$\mathbf{0.989 \pm 0.011}$	$\mathbf{1.000 \pm 0.000}$	$\mathbf{1.000 \pm 0.000}$
K-Means/D-II	0.904 ± 0.085	0.300 ± 0.191	0.467 ± 0.077	0.657 ± 0.080	0.730 ± 0.086	0.952 ± 0.028
AE-Image/D-II	0.701 ± 0.042	0.776 ± 0.079	0.763 ± 0.044	0.661 ± 0.084	0.976 ± 0.015	0.993 ± 0.007
K-Loss/D-II	$\mathbf{0.995 \pm 0.003}$	$\mathbf{0.993 \pm 0.002}$	$\mathbf{0.998 \pm 0.002}$	$\mathbf{0.992 \pm 0.003}$	$\mathbf{1.000 \pm 0.000}$	$\mathbf{1.000 \pm 0.000}$

Table 1: Clustering on Two Synthetic Datasets: D-I and D-II represent *synthetic data I* and *synthetic data II*, respectively.



(a) Synthetic Data I. (b) Synthetic Data II.

Figure 3: Convergence of MSEs on Synthetic Validation Data.



(a) Synthetic Data I. (b) Synthetic Data II.

Figure 4: MSEs of Six Methods on Synthetic Testing Data.

process. **DL** applies the dual learning framework to infer valuations, by treating valuation inference as the primal task and behavior prediction as the dual task. In **DL**, all sellers are assumed to have homogeneous behaviors.

- **DL-C** (**D**ual **L**earning with **C**lustering): It extends **DL** by considering sellers' heterogeneous behaviors. When solving the dual task (i.e., behavior prediction), it utilizes our clustering Algorithm 2 to cluster sellers and learns behavior models for different clusters separately.
- **SL** (**S**ingle **L**earning): Unlike **DL**, single learning only considers the task of valuation inference. **SL** learns a function that uses the thread history and seller decision as inputs and the inferred valuation as its output.
- **SL-C** (**S**ingle **L**earning with **C**lustering): It extends **SL** by considering sellers' heterogeneous behaviors. It clusters sellers via Algorithm 2 and learns valuation inference models for different clusters separately.

In addition, we also compare different clustering schemes:

- ***K*-Loss**: It refers to Algorithm 2 in the last section.
- ***K*-Means**: It first learns the latent embedding of each seller's behavior via an auto-encoder [Vincent *et al.*, 2010], and then uses K-means algorithm [Hartigan and Wong, 1979] to cluster these embeddings.
- **AE-Image**: Each seller's behavior is first embedded using an auto-encoder and then converted into an image. This scheme then extracts $4,096$-dimensional features from these behavior images with VGG16 [Simonyan and Zisserman, 2014], and further partitions sellers into K clusters based on these features.

Implementation Details

For both synthetic data and real data, we randomly select 80% of all threads for training, 10% for validation, and 10% for testing. We model each behavior function $f_{\boldsymbol{\theta}}$ via a gated recurrent unit (GRU). The Adam optimizer with a learning rate of 0.001 is applied for our network training. The epoch number T is set to 500 with a batch size of 64, and the weight factor α is set to 0.6. The source code and data are available at: https://github.com/cuilvye/Bargaining-project.

5.3 Experiment Results

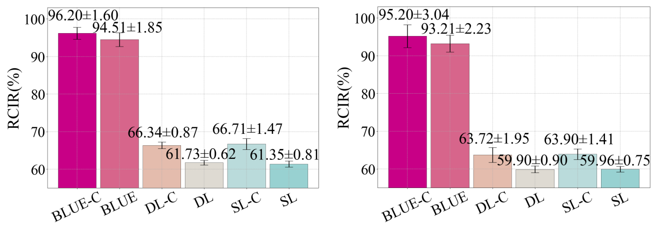
Heterogeneous Behavior Clustering

We evaluate the clustering performance of different methods on two synthetic datasets, where we know the ground truth of different sellers' behavior patterns. Recall that in both *synthetic data I* and *synthetic data II*, we simulate 900 sellers whose behaviors are randomly chosen among three pattern groups. We denote the three pattern groups by **G1**, **G2**, and **G3**. Table 1 summarizes the clustering precisions and recalls of three clustering approaches after five runs of experiments on **G1**, **G2**, and **G3**. Our **K-Loss** achieves the best performance in all cases, and can accurately assign the sellers with the same bargaining pattern into the same cluster.

Valuation Inference on Synthetic Data

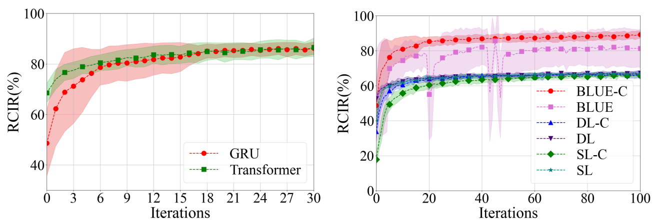
On the synthetic datasets, we characterize the valuation inference performance by the mean squared error (MSE), where the error of an inferred valuation is its squared distance to the actual valuation. Figure 3 shows the convergence performance of different inference schemes. It plots the MSEs achieved on synthetic validation data after running these schemes for different numbers of iterations. We test six random seeds, and the shadows in Figure 3 indicate the standard deviations of MSEs. Our **BLUE-C** and **BLUE** converge to much lower MSEs than other schemes on both datasets.

Figure 4 presents different methods' average MSEs after six runs of experiments on the testing data of two synthetic datasets. **BLUE-C** achieves the lowest MSE (less than 15.0) on both datasets, and other methods have relatively large MSEs (with the smallest MSE being 72.5). **BLUE** also outperforms all comparison methods on both datasets. We also



(a) Synthetic Data I. (b) Synthetic Data II.

Figure 5: RCIRs of Six Methods on Synthetic Testing Data.



(a) Different Neural Networks. (b) Different Inference Methods.

Figure 6: Convergence of RCIRs on Real Validation Data.

notice that clustering-based inference methods perform better than those without clustering, e.g., **DL-C** is better than **DL**.

In addition to MSEs, we compare different inference methods in terms of the fractions of inferred valuations satisfying the rationality assumption. Specifically, we define *Rationality-Compatible Inference Rate* (RCIR) as the percentage of inferred valuations that belong to the feasible intervals derived based on rationality assumption. Intuitively, a method achieving a higher RCIR can infer valuations more accurately. Figure 5 shows the average RCIRs of different methods after six runs of experiments on the testing data of synthetic datasets. The RCIRs of our schemes (**BLUE-C** and **BLUE**) are above 90%, and those of other methods are below 67%. This result is consistent with Figure 4, and implies that RCIR is a good indicator of the inference performance.

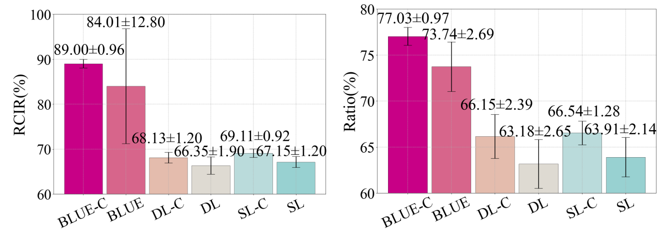
Valuation Inference on Real Data

Our real data from eBay does not record sellers' actual valuations. Hence, we cannot directly compare the MSEs of different methods on the real data. Instead, we compare different methods using two criteria: the RCIR and the ratio of satisfying the secret bounds (as defined later).

In order to justify our choice of the recurrent neural network (GRU), we investigate the performance of Transformer under our **BLUE-C** scheme. Figure 6a presents the RCIRs of the two neural networks at different iterations on the real validation data. They converge to similar RCIRs. Since Transformer has more learning parameters, its training is much slower than GRU. Hence, we employ GRU to implement all inference schemes when performing other experiments.

Figure 6b shows the RCIRs of different inference schemes at different iterations on the real validation data. We can observe that our **BLUE-C** and **BLUE** converge to much higher RCIRs than other schemes.

We show the performance of different methods on real testing data. Figure 7a presents different methods' RCIRs. Our **BLUE-C** achieves the highest RCIR (i.e., 89.0%). In Figure 7b, we compare different methods using a new criterion. To enable auto-declining and auto-accepting of offers, each seller can inform eBay of an upper bound and a lower bound of its private valuation. For example, if a buyer's offer exceeds the upper bound, eBay will accept the offer on behalf of the seller. We call these *secret bounds*. They are also recored in eBay's dataset, and are closely related to the sellers' actual private valuations. Figure 7b shows the ratio of inferred valuations satisfying these secret bounds. Our **BLUE-C** still achieves the highest ratio (i.e., 77.0%), while the ratios under all baselines are no greater than 66.5%.



(a) RCIR. (b) Secret Bounds.

Figure 7: Comparison Between Methods on Real Testing Data.

	Random	Mean	**BLUE-C**
D-I	66.64 ± 2.81	41.22 ± 2.17	$\mathbf{12.97 \pm 1.98}$
D-II	69.89 ± 4.43	42.89 ± 1.73	$\mathbf{14.71 \pm 3.24}$
D-R	$63.54\% \pm 0.33\%$	$65.49\% \pm 0.60\%$	$\mathbf{77.03\% \pm 0.97\%}$

Table 2: Comparison with Sampling-Based Methods.

6 Conclusion

In this paper, we proposed a novel framework for inferring bargainers' private valuations on items in bilateral sequential bargaining. Our inference framework works in general scenarios where bargainers may not play equilibrium strategies due to their limited reasoning capabilities and knowledge. It can automatically characterize bargainers' behaviors through neural networks rather than manually designed models. Our future research will explore the theoretical foundations of our methods and extend them to account for time-discounted valuations. Another extension is generalizing our framework to multilateral bargaining, where an agent negotiates the price of a homogeneous product with different opponents simultaneously to elicit a more favorable offer [Thomas and Wilson, 2002; Thomas and Wilson, 2014].

A Comparison with Sampling-Based Methods

We consider two sampling-based methods: **Random**: selecting random values from the feasible intervals; **Mean**: using the means of the bounds of feasible intervals. We evaluate them based on the mean squared errors on two synthetic datasets (D-I, D-II) and the ratios of satisfying the secret bounds on real data (D-R). Table 2 shows that our scheme outperforms these two methods.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62202050) and the Beijing Institute of Technology Research Fund Program for Young Scholars.

References

[Athey and Haile, 2002] Susan Athey and Philip A Haile. Identification of standard auction models. *Econometrica*, 70(6):2107–2140, 2002.

[Athey and Nekipelov, 2010] Susan Athey and Denis Nekipelov. A structural model of sponsored search advertising auctions. In *Proceedings of the Sixth Ad Auctions Workshop*, volume 15, 2010.

[Baarslag *et al.*, 2016] Tim Baarslag, Mark JC Hendrikx, Koen V Hindriks, and Catholijn M Jonker. Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, 30:849–898, 2016.

[Backus *et al.*, 2020] Matthew Backus, Thomas Blake, Brad Larsen, and Steven Tadelis. Sequential bargaining in the field: Evidence from millions of online bargaining interactions. *The Quarterly Journal of Economics*, 135(3):1319–1361, 2020.

[Bajari *et al.*, 2013] Patrick Bajari, Han Hong, and Denis Nekipelov. Game theory and econometrics: A survey of some recent research. In *Advances in Economics and Econometrics, 10th World Congress*, volume 3, pages 3–52, 2013.

[Blum *et al.*, 2014] Avrim Blum, Nika Haghtalab, and Ariel D Procaccia. Learning optimal commitment to overcome insecurity. *Advances in Neural Information Processing Systems*, 27, 2014.

[Boutilier *et al.*, 2006] Craig Boutilier, Relu Patrascu, Pascal Poupart, and Dale Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9):686–713, 2006.

[Chen and Weiss, 2012] Siqi Chen and Gerhard Weiss. An efficient and adaptive approach to negotiation in complex environments. In *Proceedings of the 20th European Conference on Artificial Intelligence*, volume 242, pages 228–233, 2012.

[de Jonge, 2022] Dave de Jonge. An analysis of the linear bilateral anac domains using the micro benchmark strategy. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*, pages 23–29, 2022.

[Goeree and Holt, 2001] Jacob K Goeree and Charles A Holt. Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, 91(5):1402–1422, 2001.

[Gopalakrishnan *et al.*, 2018] Raga Gopalakrishnan, Theja Tulabandhula, and Koyel Mukherjee. Sequential individual rationality in dynamic ridesharing. *Available at SSRN 3288319*, 2018.

[Gujral, 2016] Kritee Gujral. *Bargaining vs. Posted Prices: An Analysis Using the eBay Automobile Market*. PhD thesis, University of Florida, 2016.

[Guo and Sanner, 2010] Shengbo Guo and Scott Sanner. Real-time multiattribute bayesian preference elicitation with pairwise comparison queries. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pages 289–296, 2010.

[Hartigan and Wong, 1979] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of The Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[He *et al.*, 2016] Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. Dual learning for machine translation. *Advances in Neural Information Processing Systems*, 29, 2016.

[Hindriks and Tykhonov, 2008] Koen Hindriks and Dmytro Tykhonov. Opponent modelling in automated multi-issue negotiation using bayesian learning. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent systems - Volume 1*, pages 331–338, 2008.

[Honorio and Ortiz, 2015] Jean Honorio and Luis E Ortiz. Learning the structure and parameters of large-population graphical games from behavioral data. *J. Mach. Learn. Res.*, 16(1):1157–1210, 2015.

[Huang *et al.*, 2013] He Huang, Robert J Kauffman, Hongyan Xu, and Lan Zhao. A hybrid mechanism for heterogeneous e-procurement involving a combinatorial auction and bargaining. *Electronic Commerce Research and Applications*, 12(3):181–194, 2013.

[Jiang and Leyton-Brown, 2007] Albert Xin Jiang and Kevin Leyton-Brown. Bidding agents for online auctions with hidden bids. *Machine Learning*, 67(1):117–143, 2007.

[Kagel and Roth, 2020] John H Kagel and Alvin E Roth. *The handbook of experimental economics, volume 2*. Princeton university press, 2020.

[Ling *et al.*, 2018] Chun Kai Ling, Fei Fang, and J Zico Kolter. What game are we playing? end-to-end learning in normal and extensive form games. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 396–402, 2018.

[Ling *et al.*, 2019] Chun Kai Ling, Fei Fang, and J Zico Kolter. Large scale learning of agent rationality in two-player zero-sum games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 6104–6111, 2019.

[Luce and Raiffa, 1989] R Duncan Luce and Howard Raiffa. *Games and decisions: Introduction and critical survey*. Courier Corporation, 1989.

[Maschler *et al.*, 2020] Michael Maschler, Shmuel Zamir, and Eilon Solan. *Game theory*. Cambridge University Press, 2020.

[Nekipelov *et al.*, 2015] Denis Nekipelov, Vasilis Syrgkanis, and Eva Tardos. Econometrics for learning agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 1–18, 2015.

[Nisan and Noti, 2017] Noam Nisan and Gali Noti. An experimental evaluation of regret-based econometrics. In *Proceedings of the 26th International Conference on World Wide Web*, pages 73–81, 2017.

[Noti and Syrgkanis, 2021] Gali Noti and Vasilis Syrgkanis. Bid prediction in repeated auctions with learning. In *Proceedings of the Web Conference*, pages 3953–3964, 2021.

[Noti, 2021] Gali Noti. From behavioral theories to econometrics: Inferring preferences of human agents from data on repeated interactions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 5637–5646, 2021.

[Qin *et al.*, 2019] Xinghong Qin, Qiang Su, Samuel H Huang, Uco Jillert Wiersma, and Ming Liu. Service quality coordination contracts for online shopping service supply chain with competing service providers: integrating fairness and individual rationality. *Operational Research*, 19(1):269–296, 2019.

[Qin, 2020] Tao Qin. *Dual Learning*. Springer, 2020.

[Raiffa, 1982] Howard Raiffa. *The art and science of negotiation*. Harvard University Press, 1982.

[Simonyan and Zisserman, 2014] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[Thomas and Wilson, 2002] Charles J Thomas and Bart J Wilson. A comparison of auctions and multilateral negotiations. *RAND Journal of Economics*, pages 140–155, 2002.

[Thomas and Wilson, 2014] Charles J Thomas and Bart J Wilson. Horizontal product differentiation in auctions and multilateral negotiations. *Economica*, 81(324):768–787, 2014.

[Varian, 2007] Hal R Varian. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, 2007.

[Vendrov *et al.*, 2020] Ivan Vendrov, Tyler Lu, Qingqing Huang, and Craig Boutilier. Gradient-based optimization for bayesian preference elicitation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 10292–10301, 2020.

[Viappiani and Boutilier, 2010] Paolo Viappiani and Craig Boutilier. Optimal bayesian recommendation sets and myopically optimal choice query sets. *Advances in neural information processing systems*, 23, 2010.

[Vincent *et al.*, 2010] Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, Pierre-Antoine Manzagol, and Léon Bottou. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 11(12), 2010.

[Vorobeychik *et al.*, 2007] Yevgeniy Vorobeychik, Michael P Wellman, and Satinder Singh. Learning payoff functions in infinite games. *Machine Learning*, 67(1):145–168, 2007.

[Waugh *et al.*, 2011] Kevin Waugh, Brian D Ziebart, and J Andrew Bagnell. Computational rationalization: The inverse equilibrium problem. In *Proceedings of the 28th International Conference on Machine Learning*, pages 1169–1176, 2011.

[Williams *et al.*, 2011] Colin R Williams, Valentin Robu, Enrico H Gerding, and Nicholas R Jennings. Using gaussian processes to optimise concession in complex negotiations against unknown opponents. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 432–438, 2011.

[Wu *et al.*, 2022] Jibang Wu, Weiran Shen, Fei Fang, and Haifeng Xu. Inverse game theory for stackelberg games: the blessing of bounded rationality. *Advances in Neural Information Processing Systems*, 2022.