# Data-Driven Knowledge-Aware Inference of Private Information in Continuous Double Auctions

**Lvye Cui, Haoran Yu***

School of Computer Science & Technology, Beijing Institute of Technology
cui_lvye@outlook.com, yhrhawk@gmail.com

**Abstract**

Inferring the private information of humans from their strategic behavioral data is crucial and challenging. The main approach is first obtaining human behavior functions (which map public information and human private information to behavior), enabling subsequent inference of private information from observed behavior. Most existing studies rely on strong equilibrium assumptions to obtain behavior functions. Our work focuses on continuous double auctions, where multiple traders with heterogeneous rationalities and beliefs dynamically trade commodities and deriving equilibria is generally intractable. We develop a knowledge-aware machine learning-based framework to infer each trader's private cost vectors for producing different units of its commodity. Our key idea is to learn behavior functions by incorporating the *statistical knowledge about private costs given the observed trader asking behavior across the population*. Specifically, we first use a neural network to characterize each trader's behavior function. Second, we leverage the statistical knowledge to derive the posterior distribution of each trader's private costs given its observed asks. Third, through designing a novel loss function, we utilize the knowledge-based posterior distributions to guide the learning of the neural network. We conduct extensive experiments on a large experimental dataset, and demonstrate the superior performance of our framework over baselines in inferring the private information of humans.

## 1 Introduction

Many recent studies in economics and applied machine learning have explored the inference of humans' private information (e.g., a seller's private cost of producing its product) based on their observed decisions in strategic environments. These environments encompass a wide range of applications, such as normal-form games (Ling, Fang, and Kolter 2018; Noti 2021), Stackelberg games (Wu et al. 2022), alternating-offer bargaining (Larsen and Zhang 2018; Cui and Yu 2023), and sponsored-search auctions (Nekipelov, Syrgkanis, and Tardos 2015; Noti and Syrgkanis 2021). Accurate private information inference enables designers to better understand humans' decision-making processes, compute their payoffs, and quantify the efficiency and equity of the designed mechanisms or policies.

The main inference idea is first obtaining humans' behavior functions (which map public information and humans' private information to their decisions), and then inferring private information based on the behavior functions and the observed decisions. Prior studies obtained behavior functions mainly by assuming that humans possess sufficient rationality and knowledge to make decisions according to equilibria (Athey and Nekipelov 2010; Ling, Fang, and Kolter 2019; Wu et al. 2022). This inference approach heavily relies on assumptions regarding humans' rationality and knowledge about each other (Nisan and Noti 2017; Kagel and Roth 2020), and can hardly be applied to complex strategic environments where deriving equilibria is infeasible and information about human rationality and knowledge is unavailable. To address this problem, some recent studies proposed a machine learning-based inference approach (Cui and Yu 2023), which applies machine learning models to learn behavior functions from data.

**Knowledge-Aware Information Inference** The key challenge of applying the machine learning-based inference approach lies in accurately learning behavior functions. Because the private information of each individual human is typically unobservable and not recorded in the data, learning behavior functions from the data (which only record public information and human decisions) cannot be accomplished using standard supervised learning techniques. In this work, we propose a novel knowledge-aware method for learning behavior functions to achieve accurate inference. Specifically, although the private information of each individual human is unknown, the statistical knowledge about private information given observed decisions across the population may be obtained in many applications. We study how to derive this statistical knowledge and incorporate it into behavior function learning.

We focus on inferring private information in continuous double auctions (CDAs). In a CDA, multiple sellers and buyers engage in trading multiple units of a commodity. They can submit asks or bids at any time throughout the trading period, and transactions take place as soon as asks and bids are matched. For ease of exposition, we study inferring the private cost vectors of sellers for producing differ-

*Corresponding author.

ent units of the commodity (our method can be straightforwardly extended to infer buyer private information). Due to the complex rule and dynamic nature, deriving equilibria for CDAs is generally intractable, except in some restricted settings (e.g., single-unit CDAs (Ruijgrok 2012)). Therefore, the equilibrium-based inference approach is not applicable.

The platforms operating CDAs typically have access to the distribution of the ratio between a seller's ask and its private cost across the entire seller population. By investigating a large experimental dataset, we demonstrate that this distribution is well-behaved, and hence is easy to be estimated. Our work proposes a machine learning-based cost inference method that incorporates the knowledge of ask-cost ratios.

Our key idea is utilizing the knowledge to guide the training of the machine learning model that characterizes the seller behavior function. First, with the statistical knowledge of ask-cost ratios and observed seller asks, we can derive the posterior distributions of seller costs. Second, with the behavior function characterized by the machine learning model and other observed information, we can also derive the posterior distributions of seller costs. When the behavior function is accurate, the posterior distributions derived through these two different ways will be consistent. Building upon this insight, we train the machine learning model to get an accurate behavior function, and then use it for inference.

We summarize our contributions as follows:

- To the best of our knowledge, our work is the first attemp to improve machine learning-based information inference by incorporating the statistical knowledge about private information. In particular, we theoretically derive the knowledge-based posterior distribution of private information, and propose a novel algorithm to train the machine learning model using the knowledge-based posterior distribution.
- Our work is also the first study investigating private information inference for CDAs, which is more challenging than the inference for games with simpler strategic interactions (e.g., static or two-player games). Unlike most prior studies on information inference, the private information of each seller in CDAs is represented by a vector rather than a scalar, introducing a unique challenge to our inference method design.
- We conduct extensive experiments using a large experimental dataset, and compare our inference method with other machine learning-based methods and knowledge-aware methods (e.g., a method that directly infers private information solely according to the knowledge-based posterior distribution). The results demonstrate that the integration of machine learning and knowledge improves the accuracy of private information inference.

## 2 Related Work

**Inverse Game Theory** Our work is related to the field of *inverse game theory* (Kuleshov and Schrijvers 2015), which focuses on inferring the utilities of humans by assuming that they adopt equilibrium strategies. Previous studies have addressed this problem using different equilibrium concepts. For instance, some studies leveraged Nash equilibria to estimate player utility functions in static games (Athey and Nekipelov 2010; Bertsimas, Gupta, and Paschalidis 2015) or noncooperative dynamic games (Tsai, Molloy, and Perez 2016; Molloy et al. 2022). Some studies utilized quantal-response equilibria (McKelvey and Palfrey 1995) to infer human utility functions in Stackelberg games (Haghtalab et al. 2016; Wu et al. 2022) or payoff matrices in normal-form games (Ling, Fang, and Kolter 2018; Noti 2021). Moreover, a few studies inferred bidders' valuations on items using their bids in auctions, assuming that they play Bayesian-Nash equilibrium strategies (Jiang and Leyton-Brown 2007; Bajari, Hong, and Nekipelov 2013).

Multi-unit CDAs involve complex trading rules, and the information about traders' heterogeneous rationalities and beliefs is typically unavailable. Therefore, computing equilibria for CDAs is generally intractable (Vytelingum, Cliff, and Jennings 2006, 2008; Friedman 2018), making the inverse game theoretic approach inapplicable.

**Learning-Based Information Inference** To relax the assumptions in inverse game theory, some studies applied learning techniques to account for the bounded rationality of real-world behavior. For example, (Nekipelov, Syrgkanis, and Tardos 2015; Nisan and Noti 2017; Noti and Syrgkanis 2021) estimated the information (e.g., value-per-click) of bidders in sponsored-search auctions, assuming that they employ no-regret learning strategies. Unlike these repeated ad auctions, a CDA cannot be decomposed into multiple repetitions of the same game, making it difficult to define the regret and apply the no-regret learning. (Cui and Yu 2023) inferred the valuation of each bargainer by capturing its behavior function using machine learning. Compared with this study, it is more challenging to infer private information in CDAs, which is represented by a vector rather than a scalar. Moreover, we propose to incorporate the statistical knowledge about private information into the inference process.

Our work is also related to inverse reinforcement learning (IRL), which aims to recover an agent's reward function from its observed behavior by assuming that it applies a reinforcement learning policy (NG 2000; Yu, Song, and Ermon 2019). One major limitation of IRL lies in its demand for a substantial amount of the agent's behavioral data to learn its reward function (which maps the environment state and the agent behavior to its reward) (Rothkopf and Dimitrakakis 2011; Evans, Stuhlmüller, and Goodman 2016; Skalse and Abate 2023). In CDAs, a seller typically only has a limited number of asking records (e.g., on average, each seller in our dataset has around 5 records), rendering IRL inapplicable.

## 3 Problem Formulation

### 3.1 Continuous Double Auctions

We refer to a continuous double auction (CDA) as a *centralized market* which consists of multiple traders. Each trader is either a buyer or a seller, and is not allowed to change his role (e.g., buy a unit and then resell it). In this market, there is *a single commodity* which is available in *integer amounts*. Multiple units of this commodity can be traded during a fixed-duration trading period. Let $K$ denote the maximum number of units which a buyer may buy or a seller may sell. A CDA has the following detailed settings:

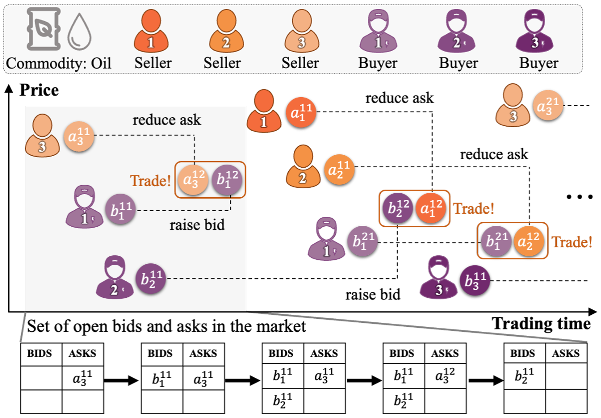



Figure 1: A CDA involving 3 buyers and 3 sellers: each trader can buy or sell up to $K = 3$ units. Three transactions have already taken place. Here, $a_j^{kl}$ denotes the $l$-th ask submitted by seller $j$ for unit $k$, while $b_j^{kl}$ is the $l$-th bid proposed by buyer $j$ for unit $k$.

- **Bidding rules**: At any time during the trading period, both sellers and buyers can make offers (known as *asks* and *bids*, respectively) or change their existing offers to sell and buy a single unit of the commodity.
- **Information revelation**: All the offers submitted and the transaction prices achieved by traders are publicly announced to all the participants in the market.
- **Clearing policy**: A transaction takes place when the bid of a buyer matches or exceeds the ask of a seller. The matching of bids and asks are conducted based on their values and the orders in which they are submitted.

Figure 1 shows an example of a CDA, where traders make offers and update their offers after observing others' offers.[1] On the bottom, we show the set of open bids and asks in the market. Initially, seller 3 submits his first ask for his first unit (i.e., $a_3^{11}$). Then, buyers 1 and 2 submit their first bids for their first units (i.e., $b_1^{11}, b_2^{11}$). Next, seller 3 reduces his ask for his first unit (i.e., submits $a_3^{12}$). After buyer 1 raises his bid to match $a_3^{12}$, a trade takes place, and only buyer 2's bid $b_2^{11}$ is open in the market. Due to space limit, we only show the change of the set up to the first trade in the market.

### 3.2 Private Information Inference Problem

**Private Cost Vector** In the CDA, each seller $j$ has a private cost vector $\boldsymbol{c}_j = \left(c_j^1, \cdots, c_j^K\right)$. Here, $c_j^k>0$ represents the cost of producing the $k$-th unit of the commodity, and hence equals the minimum price that the seller can accept for selling the unit. In particular, the marginal costs are increasing, i.e., $c_j^1 \leqslant c_j^2 \leqslant \ldots \leqslant c_j^K$. The cost vector $\boldsymbol{c}_j$ directly influences seller $j$'s ask behavior, e.g., he will never sell the $k$-th unit at a price lower than $c_j^k$.

[1]One real-world example is the agricultural commodity market, where farmers trade with food processors. Commodities can vary across markets, including wheat, corn, rice, and soybeans. The market operates continuously during specific trading hours.

**Observable Data** For market $m$, let $\left\{\left(\boldsymbol{h}_{mj}^{kl}, a_{mj}^{kl}\right)\right\}_{l=1}^{n_{mj}^k}$ represent the set of labeled data points associated with the $k$-th unit of seller $j$, where $n_{mj}^k$ is the number of asks that seller $j$ submits for selling its $k$-th unit. Here, $a_{mj}^{kl}$ denotes the $l$-th ask among these $n_{mj}^k$ asks, and $\boldsymbol{h}_{mj}^{kl}$ represents the sequence of prior offers that affect seller $j$'s choice of $a_{mj}^{kl}$.

Sequence $\boldsymbol{h}_{mj}^{kl}$ mainly consists of two types of offer information: (i) seller $j$'s prior asks for selling the commodity, and (ii) other traders' *open* bids and asks when seller $j$ chooses $a_{mj}^{kl}$. In the market in Figure 1, when $j = 1$ and $k = 1$, there are $n_{mj}^k = 2$ asks associated with the $k$-th unit of seller $j$. When $l = 2$, $a_{mj}^{kl}$ is the second ask and is $a_1^{12}$. In this case, sequence $\boldsymbol{h}_{mj}^{kl}$ includes the information of $a_1^{11}$, $a_2^{11}$, $b_1^{21}$, and $b_2^{12}$.

We denote the set of data related to all units which seller $j$ sells in market $m$ by $\mathcal{D}_{mj} = \{\{(\boldsymbol{h}_{mj}^{kl}, a_{mj}^{kl})\}_{l=1}^{n_{mj}^k}\}_{k=1}^{K_m}$, where $K_m$ is the maximum number of units that a seller can sell in this market. Seller $j$ may participate in multiple markets to sell different commodities. It is common that only $\mathcal{D}_{mj}$ is observable while $\boldsymbol{c}_{mj}$ remains undisclosed. We describe our inference problem as follows:

**Problem 1.** *Given $\mathcal{D}_{mj}$ for all seller $j$ and market $m$, we attempt to infer private cost vector $\boldsymbol{c}_{mj} = \left(c_{mj}^1, c_{mj}^2, \cdots, c_{mj}^{K_m}\right)$ for all seller $j$ and market $m$.*

We consider discrete private costs, seller asks, and buyer bids, e.g., they can be measured by dollars.

## 4 Private Information Inference Solution

We first describe our idea of solving Problem 1. Considering the fact that a seller's private cost vector directly influences his asking behavior, we propose to model his behavior via a neural network with weights $\boldsymbol{\theta}$. Then, we can denote the behavior function by $F_{\boldsymbol{\theta}}\left(c_{mj}^k, \boldsymbol{h}_{mj}^{kl}\right)$, which maps seller $j$'s $k$-th private cost $c_{mj}^k$ and the prior offers $\boldsymbol{h}_{mj}^{kl}$ to a probability distribution over all possible values of his submitted ask $a_{mj}^{kl}$. Once we learn $F_{\boldsymbol{\theta}}$, we utilize Bayes' rule to derive the posterior distribution $\Pr\left(\boldsymbol{c}_{mj}\middle|\mathcal{A}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right)$ for $\boldsymbol{c}_{mj}$:

$$\begin{aligned}
&\Pr\left(\boldsymbol{c}_{mj}\middle|\mathcal{A}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right)\\
&=\frac{\Pr\left(\mathcal{A}_{mj}\middle|\boldsymbol{c}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right)\Pr\left(\boldsymbol{c}_{mj}\right)}{\sum\limits_{\tilde{c}_{mj}^1}\sum\limits_{\tilde{c}_{mj}^2}\cdots\sum\limits_{\tilde{c}_{mj}^{K_m}}\Pr\left(\mathcal{A}_{mj}\middle|\tilde{\boldsymbol{c}}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right)\Pr\left(\tilde{\boldsymbol{c}}_{mj}\right)}\\
&=\frac{\left(\prod\limits_{k=1}^{K_m}\prod\limits_{l=1}^{n_{mj}^k}\Pr\left(a_{mj}^{kl}\middle|c_{mj}^k;\boldsymbol{h}_{mj}^{kl},\boldsymbol{\theta}\right)\right)\Pr\left(\boldsymbol{c}_{mj}\right)}{\sum\limits_{\tilde{c}_{mj}^1}\sum\limits_{\tilde{c}_{mj}^2}\cdots\sum\limits_{\tilde{c}_{mj}^{K_m}}\left(\prod\limits_{k=1}^{K_m}\prod\limits_{l=1}^{n_{mj}^k}\Pr\left(a_{mj}^{kl}\middle|\tilde{c}_{mj}^k;\boldsymbol{h}_{mj}^{kl},\boldsymbol{\theta}\right)\right)\Pr\left(\tilde{\boldsymbol{c}}_{mj}\right)},
\end{aligned} \tag{1}$$

where $\Pr\left(\boldsymbol{c}_{mj}\right)$ is the prior probability of $\boldsymbol{c}_{mj}$. Sets $\mathcal{A}_{mj}$ and $\mathcal{H}_{mj}$ include the ask behavior and corresponding prior offers associated with all units sold by seller $j$ in market $m$, respectively, i.e., $\mathcal{A}_{mj} = \{\{a_{mj}^{kl}\}_{l=1}^{n_{mj}^k}\}_{k=1}^{K_m}$ and $\mathcal{H}_{mj} = \{\{\boldsymbol{h}_{mj}^{kl}\}_{l=1}^{n_{mj}^k}\}_{k=1}^{K_m}$. In the second step of (1), we split the probability into products of $\Pr(a_{mj}^{kl}|c_{mj}^k;\boldsymbol{h}_{mj}^{kl},\boldsymbol{\theta})$, which establishes the relation between the posterior probability $\Pr(\boldsymbol{c}_{mj}|\mathcal{A}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta})$ and behavior function $F_{\boldsymbol{\theta}}$ (it gives $\Pr(a_{mj}^{kl}|c_{mj}^k;\boldsymbol{h}_{mj}^{kl},\boldsymbol{\theta})$). We refer to $\Pr\left(\boldsymbol{c}_{mj}|\mathcal{A}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right)$ as the *model-based posterior distribution*, since it is derived using the behavior model characterized by $\boldsymbol{\theta}$. With this distribution, we can solve Problem 1 and infer $\boldsymbol{c}_{mj}$ by using maximum a posteriori (MAP) estimation:

$$\boldsymbol{c}_{mj} = \arg\max_{\tilde{\boldsymbol{c}}_{mj}} \Pr\left(\tilde{\boldsymbol{c}}_{mj}\middle|\mathcal{A}_{mj};\mathcal{H}_{mj},\boldsymbol{\theta}\right). \quad (2)$$

Derivation of (1) requires an accurate $F_{\boldsymbol{\theta}}$ to calculate $\Pr\left(a_{mj}^{kl}\middle|c_{mj}^k;\boldsymbol{h}_{mj}^{kl},\boldsymbol{\theta}\right)$. To achieve this, we propose a novel loss function to optimize $\boldsymbol{\theta}$ given the observable data $\mathcal{D}_{mj}$. Our loss function is defined based on the *feasible set* and *knowledge-based posterior distribution* of $c_{mj}^k$. We present these two concepts in Sections 4.1 and 4.2, respectively, and propose the loss function for optimizing $\boldsymbol{\theta}$ in Section 4.3.

### 4.1 Feasible Set of Private Cost

Since $c_{mj}^k$ is the minimum price that seller $j$ accepts for selling the $k$-th unit, it should be no greater than any of seller $j$'s asks for the $k$-th unit, i.e., $c_{mj}^k \leqslant \min\left\{a_{mj}^{kl}\right\}_{l=1}^{n_{mj}^k}$. Hence, given the ask history associated with the $k$-th unit of seller $j$, we can deduce an upper bound $u_{mj}^k$ for each cost $c_{mj}^k$ (i.e., $u_{mj}^k = \min\left\{a_{mj}^{kl}\right\}_{l=1}^{n_{mj}^k}$). Consequently, we can determine the feasible set of each $c_{mj}^k$ as $\left\{1,\ldots,u_{mj}^k\right\}$.

### 4.2 Knowledge-Based Posterior Distribution

We normally have the knowledge about the distribution of the ratio between an ask and the corresponding cost across the entire seller population. We can utilize this knowledge to derive the posterior distribution of $c_{mj}^k$, and compare it with the *model-based posterior distribution*, which guides our learning of the behavior model. Next, we first introduce the ask-cost ratio, and then derive the *knowledge-based posterior distribution* of $c_{mj}^k$.

**Ask-Cost Ratio** We define $\gamma_{mj}^{kl} \triangleq \frac{a_{mj}^{kl}}{c_{mj}^k}$, and refer to it as the ask-cost ratio. Apparently, $\gamma_{mj}^{kl}$ is no less than 1. It is a good indicator of the deviation of $a_{mj}^{kl}$ from $c_{mj}^k$. Compared with other measures of the deviation (e.g., $a_{mj}^{kl} - c_{mj}^k$), $\gamma_{mj}^{kl}$ effectively mitigates the influence of the scale of $c_{mj}^k$.

We normally know the distribution of $\gamma_{mj}^{kl}$, and one common scenario is that $\gamma_{mj}^{kl}$ follows a shifted exponential distribution. Specifically, the ask-cost ratio $\gamma_{mj}^{kl}$ is closely associated with the corresponding seller's expectation about the waiting time of a transaction. A higher ratio indicates a longer duration that the seller is willing to wait for a trade. The waiting time in double auctions has been extensively modeled using the exponential distribution (Raberto and Cincotti 2005; Cincotti et al. 2006; Cartea and Jaimungal 2013). The probability density function of $\gamma_{mj}^{kl}$ can be characterized as follows:

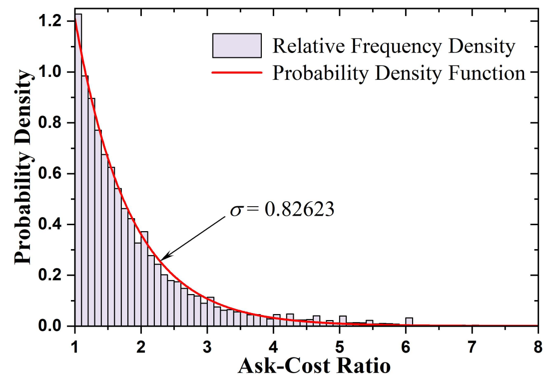



Figure 2: The Fitted Shifted Exponential Distribution Function of The Ask-Cost Ratios in Our Experimental Data.

$$g\left(\gamma_{mj}^{kl};\sigma\right) = \frac{1}{\sigma}\exp\left(-\frac{\gamma_{mj}^{kl}-1}{\sigma}\right), \gamma_{mj}^{kl} \geqslant 1, \sigma>0, \quad (3)$$

where $\sigma$ is the scale parameter. To validate this model, we plot the normalized histogram of $\gamma_{mj}^{kl}$ in our dataset containing $295,245$ ask-cost pairs. As Figure 2 shows, it can be well approximated by a shifted exponential distribution.

Note that our private information inference framework is not limited to the shifted exponential distribution of $\gamma_{mj}^{kl}$. Our framework can be applied as long as we have the knowledge of $g\left(\gamma_{mj}^{kl};\boldsymbol{\sigma}\right)$, where $\boldsymbol{\sigma}$ includes all parameters that characterize the probability density function.

**Knowledge-Based Posterior Distribution of $c_{mj}^k$** Using the knowledge of $g\left(\gamma_{mj}^{kl};\boldsymbol{\sigma}\right)$, we can derive the posterior distribution of $c_{mj}^k$ given the corresponding seller asks. We first consider a continuous $c_{mj}^k$ and introduce the following proposition.

**Proposition 1.** *Suppose that $c_{mj}^k>0$ is a continuous variable. Given $a_{mj}^{kl}$, the posterior cumulative distribution function of $c_{mj}^k$ can be computed as*

$$\Psi\left(c\middle|a_{mj}^{kl};\boldsymbol{\sigma}\right) = 1 - \Pr\left(\gamma_{mj}^{kl} < \frac{a_{mj}^{kl}}{c}\middle|\boldsymbol{\sigma}\right). \quad (4)$$

For example, if $g\left(\gamma_{mj}^{kl};\boldsymbol{\sigma}\right)$ takes the form in (3), we can derive the concrete form of $\Psi\left(c\middle|a_{mj}^{kl};\boldsymbol{\sigma}\right)$ as

$$\Psi\left(c\middle|a_{mj}^{kl};\sigma\right) = \exp\left(-\frac{\frac{a_{mj}^{kl}}{c}-1}{\sigma}\right), c \in \left(0, a_{mj}^{kl}\right]. \quad (5)$$

Since $c_{mj}^k$ actually takes a discrete value, we use the survival function discretization method (Kemp 2004;

Chakraborty 2015; Tovissodé et al. 2021) to get the following posterior probability mass function of $c_{mj}^k$:

$$\Pr\left(c_{mj}^k | a_{mj}^{kl}; \boldsymbol{\sigma}\right) = \Psi\left(c_{mj}^k + 1 | a_{mj}^{kl}; \boldsymbol{\sigma}\right) - \Psi\left(c_{mj}^k | a_{mj}^{kl}; \boldsymbol{\sigma}\right). \tag{6}$$

The above $\Pr\left(c_{mj}^k | a_{mj}^{kl}; \boldsymbol{\sigma}\right)$ characterizes the distribution of $c_{mj}^k$ given one $a_{mj}^{kl}$. Overall, seller $j$ submits $n_{mj}^k$ asks for the $k$-th unit in market $m$. We compute the distribution of $c_{mj}^k$ given all these asks in the following proposition (the computation is challenging due to the dependence between the $n_{mj}^k$ asks).

**Proposition 2.** *With the knowledge of $g\left(\gamma_{mj}^{kl}; \boldsymbol{\sigma}\right)$, the distribution of $c_{mj}^k$ given $\left\{a_{mj}^{kl}\right\}_{l=1}^{n_{mj}^k}$ can be characterized by*

$$\Pr\left(c_{mj}^k \middle| \left\{a_{mj}^{kl}\right\}_{l=1}^{n_{mj}^k}; \boldsymbol{\sigma}\right) = \frac{\prod_{l=1}^{n_{mj}^k} \Pr\left(c_{mj}^k \middle| a_{mj}^{kl}; \boldsymbol{\sigma}\right)}{\left(\prod_{l=1}^{n_{mj}^k - 1} \Pr\left(c_{mj}^k\right)\right) \cdot R}, \tag{7}$$

*where* $R = \sum_{\tilde{c}_{mj}^k} \left( \frac{\prod_{l=1}^{n_{mj}^k} \Pr\left(\tilde{c}_{mj}^k \middle| a_{mj}^{kl}; \boldsymbol{\sigma}\right)}{\prod_{l=1}^{n_{mj}^k - 1} \Pr\left(\tilde{c}_{mj}^k\right)} \right)$ *is a normalization factor, and* $\Pr\left(c_{mj}^k\right)$ *is the prior probability of* $c_{mj}^k$.

### 4.3 Learning of Behavior Model

Given the observable data $\mathcal{D}_{mj}$, the feasible set and the knowledge-based posterior distribution of each $c_{mj}^k$, we design a novel loss function to learn the behavior model $F_{\boldsymbol{\theta}}$. Our loss function consists of the following three parts:

- The first part is a cross-entropy loss $\mathcal{L}_{mj}^{\text{CE}}$, which is associated with the accuracy of the behavior model. With precise $\boldsymbol{\theta}$, we can predict a seller's ask utilizing its cost $c_{mj}^k$ and prior offers $\boldsymbol{h}_{mj}^{kl}$ as $F_{\boldsymbol{\theta}}\left(c_{mj}^k, \boldsymbol{h}_{mj}^{kl}\right)$, and it will be close to the seller's true ask $a_{mj}^{kl}$. Hence, the first part of our loss function is the following function:

$$\mathcal{L}_{mj}^{\text{CE}} = -\sum_{k=1}^{K_m} \sum_{l=1}^{n_{mj}^k} \log \Pr\left(a_{mj}^{kl} \middle| c_{mj}^k; \boldsymbol{h}_{mj}^{kl}, \boldsymbol{\theta}\right). \tag{8}$$

  During the implementation, we do not know the actual value of $c_{mj}^k$, which is our inference target. Hence, we utilize an estimated value instead, e.g., estimating $c_{mj}^k$ by randomly sampling a cost value from its feasible set $\left\{1, \ldots, u_{mj}^k\right\}$.

- The second part is a relative entropy loss $\mathcal{L}_{mj}^{\text{KL}}$, which is related to the knowledge-based posterior distribution of $c_{mj}^k$. Given $\boldsymbol{\theta}$, we can compute the model-based posterior distribution of $c_{mj}^k$ as in (1). With accurate $\boldsymbol{\theta}$, the model-based posterior distribution will be close to the knowledge-based posterior distribution computed in Proposition 2. We use the Kullback-Leibler divergence (Csiszar 1975) to quantify the closeness between the two distributions. The second part of our loss function is

$$\begin{aligned}\mathcal{L}_{mj}^{\text{KL}} &= \sum_{k=1}^{K_m} \sum_{l=1}^{n_{mj}^k} D_{KL}\left(\Pr\left(c_{mj}^k \middle| \mathcal{A}_{mj}^k; \boldsymbol{\sigma}\right) \middle\| \Pr\left(c_{mj}^k \middle| a_{mj}^{kl}; \boldsymbol{h}_{mj}^{kl}, \boldsymbol{\theta}\right)\right) \\ &= \sum_{k=1}^{K_m} \sum_{l=1}^{n_{mj}^k} \sum_{c_{mj}^k} \Pr\left(c_{mj}^k \middle| \mathcal{A}_{mj}^k; \boldsymbol{\sigma}\right) \log\left(\frac{\Pr\left(c_{mj}^k \middle| \mathcal{A}_{mj}^k; \boldsymbol{\sigma}\right)}{\Pr\left(c_{mj}^k \middle| a_{mj}^{kl}; \boldsymbol{h}_{mj}^{kl}, \boldsymbol{\theta}\right)}\right).\end{aligned} \tag{9}$$

  Here, $\mathcal{A}_{mj}^k$ includes the asks of seller $j$ for its $k$-th unit in market $m$, i.e., $\mathcal{A}_{mj}^k = \left\{a_{mj}^{kl}\right\}_{l=1}^{n_{mj}^k}$.

- The third part is a loss $\mathcal{L}_{mj}^{\text{MC}}$, which is associated with the monotonicity constraint of marginal costs. With precise $\boldsymbol{\theta}$, we can infer vector $\boldsymbol{c}_{mj}$ according to (1), and it will satisfy the constraint $c_{mj}^1 \leqslant c_{mj}^2 \leqslant \ldots \leqslant c_{mj}^{K_m}$. Hence, the last part of our loss function is the following function:

$$\mathcal{L}_{mj}^{\text{MC}} = -\log \Pr\left(\mathcal{COS}\left(\boldsymbol{c}_{mj}\right) \middle| \mathcal{A}_{mj}; \mathcal{H}_{mj}, \boldsymbol{\theta}\right). \tag{10}$$

  Here, $\mathcal{COS}\left(\boldsymbol{c}_{mj}\right)$ refers to the event that the inferred vector $\boldsymbol{c}_{mj}$ satisfies the monotonicity constraint.

Our loss function is the weighted sum of the above three parts considering all markets and sellers, i.e., $\sum_{(m,j)}\left(w_1 \mathcal{L}_{mj}^{\text{CE}} + w_2 \mathcal{L}_{mj}^{\text{KL}} + w_3 \mathcal{L}_{mj}^{\text{MC}}\right)$. Here, weights $w_1$, $w_2$, and $w_3$ belong to $(0, 1)$ and satisfy $w_1 + w_2 + w_3 = 1$. By minimizing the loss function over $\boldsymbol{\theta}$, we can learn a function $F_{\boldsymbol{\theta}}$ that both (i) fits the seller ask behavior and (ii) can

---

Algorithm 1: Behavior Learning Algorithm

**Input**: Observable data $\mathcal{D}_{mj}$ for all $m$ and $j$, probability density function $g\left(\gamma_{mj}^{kl}; \boldsymbol{\sigma}\right)$, loss weights $w_1, w_2, w_3$, learning rate $\eta$, and maximum epoch times $T$.
**Output**: Behavior model $F_{\boldsymbol{\theta}_t}$.

1: Compute the feasible set $\left\{1, \cdots, u_{mj}^k\right\}$ for each $c_{mj}^k$;
2: Compute the conditional probability $\Pr\left(c_{mj}^k | a_{mj}^{kl}; \boldsymbol{\sigma}\right)$ according to (4) and (6);
3: Calculate the knowledge-based posterior distribution $\Pr\left(c_{mj}^k \middle| \mathcal{A}_{mj}^k; \boldsymbol{\sigma}\right)$ according to (7);
4: Initialize $\boldsymbol{\theta}$ with $\boldsymbol{\theta}_0$, and set $t = 0$;
5: **while** $t \leqslant T$ **and** $\boldsymbol{\theta}$ does not converge **do**
6:   Randomly divide all data into training batches;
7:   **for** each training batch $\mathcal{B}$ **do**
8:     Estimate each $c_{mj}^k$ in batch $\mathcal{B}$ to compute $\mathcal{L}_{mj}^{\text{CE}}$;
9:     Compute the gradient $\nabla_{\boldsymbol{\theta}_t}^{(\mathcal{B})}$ utilizing batch $\mathcal{B}$:
$$\nabla_{\boldsymbol{\theta}_t}^{(\mathcal{B})} = \sum_{(m,j) \text{in batch} \mathcal{B}} \frac{\partial\left(w_1 \mathcal{L}_{mj}^{\text{CE}} + w_2 \mathcal{L}_{mj}^{\text{KL}} + w_3 \mathcal{L}_{mj}^{\text{MC}}\right)}{\partial \boldsymbol{\theta}_t};$$
10:     Update model parameters $\boldsymbol{\theta}$: $\boldsymbol{\theta}_t \leftarrow \boldsymbol{\theta}_t - \eta \nabla_{\boldsymbol{\theta}_t}^{(\mathcal{B})}$;
11:   **end for**
12:   $\boldsymbol{\theta}_{t+1} \leftarrow \boldsymbol{\theta}_t$; $t \leftarrow t + 1$.
13: **end while**
14: **return** $F_{\boldsymbol{\theta}_t}$

be used to infer private costs which match the knowledge of ask-cost ratios and satisfy the monotonicity constraint.

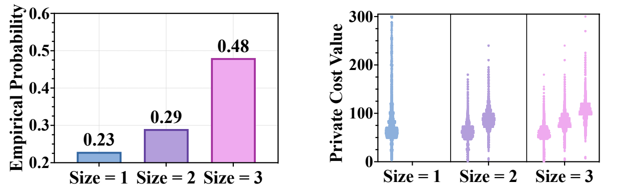
We design a behavior learning algorithm (Algorithm 1) to learn seller ask behavior via minimizing the loss function. In lines 2 to 3 of Algorithm 1, we compute the knowledge-based posterior distribution for each $c_{mj}^k$. In lines 4 to 13, we train the behavior model with the proposed loss function. Specifically, we randomly partition the training data into training batches. For each batch, we compute the gradient of the loss function, and update $\boldsymbol{\theta}$ accordingly. After getting $\boldsymbol{\theta}$ via Algorithm 1, we utilize Bayes' rule to infer each cost vector $\boldsymbol{c}_{mj}$ with (1) and (2).

# 5 Experiments

## 5.1 Dataset Description

We conduct all experiments on a large experimental CDA dataset (Lin et al. 2020). The data were collected from over $9,000$ markets during classroom experiments conducted in various countries using the MobLab platform.[2] The dataset contains concrete auction information, e.g., market IDs, seller IDs, seller private cost vectors, seller asks, and the corresponding prior offers. It includes $55,507$ market-seller pairs with $295,245$ data points.

In real-world CDAs, it is infeasible to collect data on the private cost vectors of sellers. On the MobLab platform, experiment designers can predefine the cost vector for each participant acting as a seller. This enables us to use the ground truth of cost vectors to evaluate the accuracy of our inference framework. Our dataset consists of cost vectors with sizes $K$ ranging from 1 to 3. Figure 3a displays the empirical probability distribution of the vector size $K$. Additionally, each element $c_{mj}^k$ of a cost vector $\boldsymbol{c}_{mj}$ ranges from 1 to 300. Figure 3b plots the distribution of the element values across cost vectors of different sizes. For example, when the vector size $K$ is three, we can see the distributions of $c_{mj}^1$, $c_{mj}^2$, and $c_{mj}^3$.




(a) Vector Size Distribution. (b) Elements of Cost Vectors.

Figure 3: Private Cost Vectors in The Dataset.

## 5.2 Experimental Settings

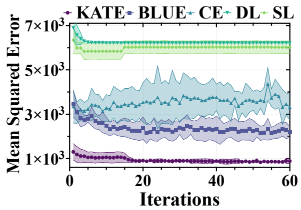
**Comparison Methods** We compare our proposed method (**KATE**: **K**nowledge-**A**ware Cos**t** Inf**e**rence) with the following three categories of inference methods:

- Learning-based inference methods that do not use knowledge of ask-cost ratios:
  - **CE** (**C**ross**E**ntropy): It only uses the cross-entropy loss (i.e., $\mathcal{L}_{mj}^{\text{CE}}$) to learn the behavior model $F_{\boldsymbol{\theta}}$, and then uses the model for cost vector inference.
  - **BLUE** (**B**ayesian **L**earning-based Val**u**ation Inf**e**rence) (Cui and Yu 2023): It was primarily developed for inferring bargainer private valuations in bilateral sequential bargaining. It cannot utilize the knowledge of the ratios between offers and valuations.
  - **DL** (**D**ual **L**earning): We treat private cost inference as the primal task and ask behavior prediction as the dual task, and learn to jointly solve the two tasks through the dual learning framework proposed in (Qin 2020; He et al. 2016).
  - **SL** (**S**ingle **L**earning): We only consider the task of cost inference, and directly learn a mapping from prior offers and seller asks to seller costs.
- Inference methods that use knowledge of ask-cost ratios:
  - **DL-KL** (**D**ual **L**earning with **K**ullback-**L**eibler Divergence Loss): We modify **DL** to consider a KL divergence loss function for the primal task (i.e., cost inference). This loss function quantifies the closeness between each cost's inferred distribution and its knowledge-based posterior distribution.
  - **SL-KL** (**S**ingle **L**earning with **K**ullback-**L**eibler Divergence Loss): We modify **SL** by considering a KL divergence loss, similar to **DL-KL**.
  - **DI-KP** (**D**irect **I**nference with **K**nowledge-based **P**osterior Distribution): This inference method takes the value with the highest knowledge-based posterior probability as the estimated cost.

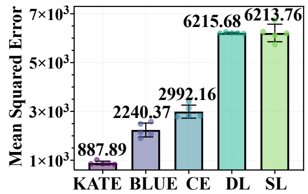
  Note that the design of the above three methods incorporates the results derived in this paper. Specifically, all three methods utilize the knowledge-based posterior distribution derived in Proposition 2. Moreover, **DL-KL** and **SL-KL** apply the KL-divergence term defined in (9).
- Other inference methods:
  - **Median**: It takes the median of all values in each cost's feasible set as the estimation.
  - **Mean**: It takes the mean of all values in each cost's feasible set as the estimation.
  - **Uniform**: It randomly takes one value uniformly from each cost's feasible set as the estimated cost.
  - **Truthfulness-Based Inference**: It performs inference by assuming that whenever a transaction occurs, the transaction price (i.e., the seller's last ask) reveals the seller's true cost.

**Implementation Details** We randomly select $80\%$ of all market-seller pairs for training, $10\%$ for validation, and $10\%$ for testing. The behavior function $F_{\boldsymbol{\theta}}$ is modeled via a gated recurrent unit (GRU) network (we have also explored the Transformer architecture, which achieves a similar inference accuracy to GRU but has a much slower training process). For the hyperparameters $w_1$, $w_2$, and $w_3$, we use the grid search method to optimize them in a discretized space, minimizing the MSE on the validation data. Their ultimate values are 0.2, 0.6, and 0.2, respectively, and we apply the Adam optimizer with a learning rate of 0.001 for network training. Our data and codes are available at: https://github.com/cuilvye/Inference-CDAs.

[2]MobLab (moblab.com) is a company that provides a platform for carrying out in-class experiments in economics courses.

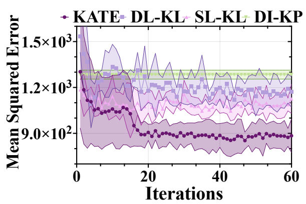



(a) Convergence of MSEs.

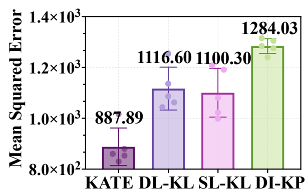



(b) MSEs on Testing Data.

Figure 4: Comparison with Knowledge-Unaware Methods.




(a) Convergence of MSEs.




(b) MSEs on Testing Data.

Figure 5: Comparison with Knowledge-Aware Methods.

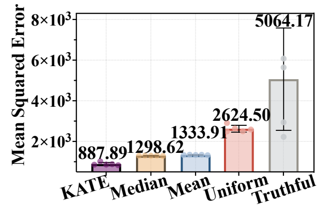
### 5.3 Experimental Results

We use the mean squared error (MSE) to evaluate the performance of cost vector inference methods. Given each actual cost vector $\boldsymbol{c}_{mj}$ recorded in the experimental dataset and the inferred cost vector $\hat{\boldsymbol{c}}_{mj}$, the MSE is calculated as $\frac{1}{K_m}\sum_{k=1}^{K_m}\left(c_{mj}^k - \hat{c}_{mj}^k\right)^2$. We focus on the average value of this metric over all market $m$ and seller $j$.

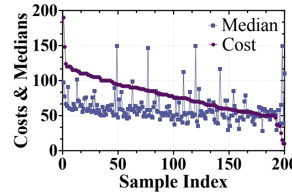
**Comparison with Knowledge-Unaware Learning-Based Methods** In Figure 4, we compare our **KATE** with learning-based inference methods that do not utilize knowledge of ask-cost ratios. Specifically, Figure 4a presents the MSE achieved on the validation data against the number of iterations for running each method. We conduct five experiments using different random seeds for dataset splitting, and the shadows in Figure 4a indicate the standard deviations of MSEs. It can be observed that our **KATE** converges to a much lower MSE on the validation data, compared with the other knowledge-unaware methods.

Figure 4b presents the MSEs of different inference methods on the testing data. The five circles around each bar indicate the MSEs achieved in the five experimental runs. Our **KATE** achieves the lowest MSE (around $887.89$), and the other knowledge-unaware methods have relatively large MSEs (with the lowest one being $2240.37$). This demonstrates that our **KATE** can effectively leverage the knowledge of ask-cost ratios to infer private cost vectors.

**Comparison with Knowledge-Aware Methods** Figure 5 compares our **KATE** with inference methods that utilize knowledge of ask-cost ratios. Figure 5a illustrates that our **KATE** converges to the lowest MSE on the validation data. In particular, we can observe that the performance of **DI-KP** (which is not a learning-based method) remains unchanged across different iterations.




(a) MSEs on Testing Data.




(b) Actual Costs vs. Medians.

Figure 6: Comparison with Other Inference Methods.

Figure 5b shows the average MSEs of different methods on the testing data (over five runs of experiments). **KATE** exhibits the most accurate inference. Recall that the design of the three knowledge-aware methods incorporates the results that we derive in this paper. These knowledge-aware methods significantly outperform the knowledge-unaware methods shown in Figure 4b. This demonstrates the importance of leveraging the knowledge of ask-cost ratios and also validates the results in this paper (such as the knowledge-based posterior distribution in Proposition 2).

Furthermore, **DI-KP** is the least accurate knowledge-aware method. This is because it solely relies on the knowledge-based posterior distribution. It does not learn the behavior function (i.e., the mapping from private costs and prior offers to asks), and hence fails to utilize the information of prior offers in inference.

**Comparison with Other Inference Methods** Figure 6a compares our **KATE** with the methods that only utilize the derived feasible sets or the truthfulness assumption to infer private costs. We can observe that **KATE** reduces the MSE by approximately $410$, compared with the best feasible set-based method, i.e., **Median**.

**Median** and **Mean** achieve better performance than **Uniform**. This is mainly due to a feature of our dataset, where most costs are close to the medians or means of their feasible intervals. To illustrate this, we randomly choose 200 samples of costs, and sort them in descending order of values. Figure 6b compares them with the medians of their feasible intervals. When applied to the datasets without this feature, the performance of **Median** and **Mean** may deteriorate.

## 6 Conclusion

In this paper, we studied information inference for multi-unit CDAs, which involve sophisticated strategic interactions and multi-dimensional private information. We leveraged neural networks to learn behavior functions, and proposed a framework to incorporate statistical knowledge about private information into neural network training. Experimental results demonstrate the superiority of our framework over three categories of inference baselines. Our idea of performing knowledge-aware machine learning-based inference is general, and may be applied to infer information from strategic behavior in other applications.

## 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. We acknowledge the authors of (Lin et al. 2020) for releasing the CDA dataset publicly.

## References

Athey, S.; and Nekipelov, D. 2010. A structural model of sponsored search advertising auctions. In *Proceedings of the Sixth Ad Auctions Workshop*, volume 15.

Bajari, P.; Hong, H.; and Nekipelov, D. 2013. Game theory and econometrics: A survey of some recent research. In *Advances in Economics and Econometrics, 10th World Congress*, volume 3, 3–52.

Bertsimas, D.; Gupta, V.; and Paschalidis, I. C. 2015. Data-driven estimation in equilibrium using inverse optimization. *Mathematical Programming*, 153: 595–633.

Cartea, A.; and Jaimungal, S. 2013. Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, 20(6): 512–547.

Chakraborty, S. 2015. Generating discrete analogues of continuous probability distributions-A survey of methods and constructions. *Journal of Statistical Distributions and Applications*, 2: 1–30.

Cincotti, S.; Focardi, S. M.; Ponta, L.; Raberto, M.; and Scalas, E. 2006. The waiting-time distribution of trading activity in a double auction artificial financial market. In *The Complex Networks of Economic Interactions: Essays in Agent-Based Economics and Econophysics*, 239–247. Springer.

Csiszar, I. 1975. *I*-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1): 146 – 158.

Cui, L.; and Yu, H. 2023. Inferring private valuations from behavioral data in bilateral sequential bargaining. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 2624–2632.

Evans, O.; Stuhlmüller, A.; and Goodman, N. 2016. Learning the preferences of ignorant, inconsistent agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Friedman, D. 2018. *The double auction market: institutions, theories, and evidence*. Routledge.

Haghtalab, N.; Fang, F.; Nguyen, T. H.; Sinha, A.; Procaccia, A. D.; and Tambe, M. 2016. Three strategies to success: learning adversary models in security games. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, 308–314.

He, D.; Xia, Y.; Qin, T.; Wang, L.; Yu, N.; Liu, T.-Y.; and Ma, W.-Y. 2016. Dual learning for machine translation. *Advances in Neural Information Processing Systems*, 29.

Jiang, A. X.; and Leyton-Brown, K. 2007. Bidding agents for online auctions with hidden bids. *Machine Learning*, 67(1): 117–143.

Kagel, J. H.; and Roth, A. E. 2020. *The handbook of experimental economics, volume 2*. Princeton university press.

Kemp, A. W. 2004. Classes of discrete lifetime distributions. *Communications in statistics. Theory and methods*, 33(11-12): 3069–3093.

Kuleshov, V.; and Schrijvers, O. 2015. Inverse game theory: Learning utilities in succinct games. In *Proceedings of the 11th International Web and Internet Economics Conference*, 413–427. Springer.

Larsen, B.; and Zhang, A. L. 2018. A mechanism design approach to identification and estimation. Technical report, National Bureau of Economic Research.

Lin, P.-H.; Brown, A. L.; Imai, T.; Wang, J. T.-y.; Wang, S. W.; and Camerer, C. F. 2020. Evidence of general economic principles of bargaining and trade from 2,000 classroom experiments. *Nature Human Behaviour*, 4(9): 917–927.

Ling, C. K.; Fang, F.; and Kolter, J. Z. 2018. 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*, 396–402.

Ling, C. K.; Fang, F.; and Kolter, J. Z. 2019. Large scale learning of agent rationality in two-player zero-sum games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 6104–6111.

McKelvey, R. D.; and Palfrey, T. R. 1995. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1): 6–38.

Molloy, T. L.; Inga Charaja, J.; Hohmann, S.; and Perez, T. 2022. Inverse noncooperative dynamic games. In *Inverse Optimal Control and Inverse Noncooperative Dynamic Game Theory: A Minimum-Principle Approach*, 143–187. Springer.

Nekipelov, D.; Syrgkanis, V.; and Tardos, E. 2015. Econometrics for learning agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 1–18.

NG, A. 2000. Algorithms for inverse reinforcement learning. In *Proc. of 17th International Conference on Machine Learning, 2000*, 663–670.

Nisan, N.; and Noti, G. 2017. An experimental evaluation of regret-based econometrics. In *Proceedings of the 26th International Conference on World Wide Web*, 73–81.

Noti, G. 2021. 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, 5637–5646.

Noti, G.; and Syrgkanis, V. 2021. Bid prediction in repeated auctions with learning. In *Proceedings of the Web Conference*, 3953–3964.

Qin, T. 2020. *Dual Learning*. Springer.

Raberto, M.; and Cincotti, S. 2005. Modeling and simulation of a double auction artificial financial market. *Physica A: Statistical Mechanics and its applications*, 355(1): 34–45.

Rothkopf, C. A.; and Dimitrakakis, C. 2011. Preference elicitation and inverse reinforcement learning. In *Proceedings* *of the 2011 European Conference on Machine Learning and Knowledge Discovery in Databases-Volume Part III*, 34–48. Springer.

Ruijgrok, M. 2012. A single-item continuous double auction game. MPRA Paper 42086, University Library of Munich, Germany.

Skalse, J.; and Abate, A. 2023. Misspecification in inverse reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 15136–15143.

Tovissodé, C. F.; Honfo, S. H.; Doumatè, J. T.; and Glèlè Kakaï, R. 2021. On the discretization of continuous probability distributions using a probabilistic rounding mechanism. *Mathematics*, 9(5): 555.

Tsai, D.; Molloy, T. L.; and Perez, T. 2016. Inverse two-player zero-sum dynamic games. In *2016 Australian Control Conference (AuCC)*, 192–196. IEEE.

Vytelingum, P.; Cliff, D.; and Jennings, N. R. 2006. Evolutionary stability of behavioural types in the continuous double auction. In *Workshop on Trading Agent Design and Analysis*, 103–117. Springer.

Vytelingum, P.; Cliff, D.; and Jennings, N. R. 2008. Strategic bidding in continuous double auctions. *Artificial Intelligence*, 172(14): 1700–1729.

Wu, J.; Shen, W.; Fang, F.; and Xu, H. 2022. Inverse game theory for Stackelberg games: the blessing of bounded Rationality. *Advances in Neural Information Processing Systems*.

Yu, L.; Song, J.; and Ermon, S. 2019. Multi-agent adversarial inverse reinforcement learning. In *International Conference on Machine Learning*, 7194–7201. PMLR.