Bidder Behavior in Complex Trading Environments: Modeling, Simulations, and Agent-Enabled Experiments

A Thesis
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Ali Mahdavi Adeli

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Gediminas Adomavicius, Alok Gupta

January 2018

Acknowledgements

Special thanks to my Advisor Gediminas (Gedas) Adomavicius for his continued support throughout my Ph.D. and all he has taught me. His contagious enthusiasm for research has been a source of motivation. I am grateful to Alok Gupta (my co-advisor) for his insightful feedbacks and guidance and to Shawn Curley for his support and advice. I would also like to thank Maria Gini, the external member of my defense committee. I appreciate all their contributions of time and ideas. I would like to express my heartfelt gratitude to my fellow colleagues and friends who made this otherwise lonely journey enjoyable and enriched my time in various ways. From intriguing discussions, random coffee breaks, and walks around campus, to occasional trips, get-togethers, and personal conversations. You know who you are, thank you all for being there during some period of this journey.

Dedication

To Mom & Dad.

Abstract

Combinatorial auctions represent sophisticated market mechanisms that are becoming increasingly important in various business applications due to their ability to improve economic efficiency and auction revenue, especially in settings where participants tend to exhibit more complex user preferences and valuations. While recent studies on such auctions have found heterogeneity in bidder behavior and its varying effect on auction outcomes, the area of bidder behavior and its impact on economic outcomes in combinatorial auctions is still largely underexplored. One of the main reasons is that it is nearly impossible to control for the type of bidder behavior in real world or experimental auction setups. In my dissertation I propose two data-driven approaches (heuristic-based in the first part and machine-learning-based in the second part) to design and develop software agents that replicate several canonical types of human behavior observed in this complex trading mechanism. Leveraging these agents in an agent-based simulation framework, I examine the effect of different bidder compositions (i.e., competing against bidders with different bidding strategies) on auction outcomes and bidder behavior. I use the case of continuous combinatorial auctions to demonstrate both approaches and provide insights that facilitate the implementation of this combinatorial design for online marketplaces. In the third part of my thesis, I conduct human vs. machine style experiments by integrating the bidding agents into an experimental combinatorial auction platform, where participants play against (human-like) agents with certain pre-determined bidding strategies. This part investigates the impact of different competitive environments on bidder behavior and auction outcomes, the underlying reasons for different behaviors, and how bidders learn under different competitive environments.

Table of Contents

List of Tables

List of Figures

1. Introduction

Combinatorial auctions are important market mechanisms that allow bidders to express interdependencies in their valuations for combinations of items, which leads to more efficient allocation of resources in complex market environments. While there has been research on a number of topics in this area – e.g., winner determination problem in combinatorial auctions (Cramton et al. 2006; De Vries and Vohra 2003; Demir and Gini 2007; Mueller 2006; Rothkopf et al. 1998; Sandholm 2002), different combinatorial auction designs (Ausubel et al. 2006; Ausubel and Milgrom 2006a; Bichler et al. 2013b; Kwasnica et al. 2005; Parkes 2006; Porter et al. 2003), practicality of these designs (Adomavicius and Gupta 2005; Bichler et al. 2014; Hoffman et al. 2006; Pekeč and Rothkopf 2003) for online marketplaces (Adomavicius et al. 2007; Bichler 2000), and comparison of different auction mechanisms (Bapna et al. 2001; Bichler et al. 2009; Brunner et al. 2010; Kwasnica et al. 2005) – the issues related to bidder behavior in these auctions have been largely underexplored. The main challenge has been the complexity of decision space for bidders, and how different bidders handle this complexity. Prior research has shown significant heterogeneity in bidder behavior (Adomavicius et al. 2012; Bapna et al. 2000; Bichler et al. 2013a; Scheffel et al. 2011), however, since it is not possible to control for bidder behavior in real-world settings or experimental auction setups, it is hard to address a number of important and interesting research questions, such as understanding how bidder behavior changes in the presence of different types of competition and how the auction outcomes are affected as the result.

In the first and second parts of my dissertation, I propose two data-driven approaches (heuristic-based in the first and machine learning in the second part) to design and develop software agents that replicate several canonical types of human behavior observed in this complex trading mechanism. Leveraging these agents in an agent-based simulation framework, I examine the effect of different bidder compositions (i.e., competing against bidders with different bidding strategies) on auction outcomes and bidder behavior. I use the case of continuous combinatorial auctions to demonstrate both approaches and provide insights that facilitate the implementation of this combinatorial design for online marketplaces. In the third part of my thesis, I conduct human vs. machine style experiments by integrating the bidding agents into an experimental continuous combinatorial auction platform; i.e., participants play against (human-like) agents with certain pre-determined bidding strategies. The goal is to investigate the underlying reasons for different bidder behaviors and to understand the effect of competition type on participants' acceptance of CCAs as a viable trading mechanism.

1.1 Combinatorial auctions

In combinatorial auctions (CA), bidders can bid not only on individual items but also on their combinations (bundles), which allows for more expressive bidding. This often leads to increased efficiency in resource allocation and higher auction revenue in complex economic environments, especially when there are complementarities or positive externalities among items (Banks et al. 2003; Porter et al. 2003), i.e., when a set of items as a whole has greater value to the bidder than the sum of the values for the individual items. Various combinatorial auction designs have been proposed for important applications, such as FCC spectrum auctions (Bichler et al. 2014; McAfee and McMillan 1996), allocation of airport takeoff and landing slots to competing airlines (Rassenti et al. 1982), allocation and pricing of payload resources for the international space station (Banks et al. 1989), procurement of transportation services (Sheffi 2004), task scheduling (Collins and Gini 2009), network routing (Hershberger and Suri 2001), industrial procurements (Bichler et al. 2006), and allocation of airspace system resources (Ball et al. 2006).

Different combinatorial auction designs differ in their rules (e.g., ascending versus sealed-bid, single round vs. multi-round) and the auction environment, which is determined by the number of items being traded, the number of participants in each auction, the preferences of buyers and sellers, and the private information available to each of the parties (Cramton et al. 2006). Combinatorial auction designs include: single-round, first-price sealed bidding; Vickrey-Clarke-Groves (VCG) mechanisms; and iterative combinatorial auction mechanisms (Pekeč and Rothkopf 2003) such as simultaneous multi-round auctions, the combinatorial clock auction (Porter et al. 2003) and its extended version CC+ (Bichler et al. 2013b), ascending proxy auctions (Ausubel and Milgrom 2006a) and clock proxy auctions (Ausubel et al. 2006), and the Resource Allocation Design auction (Kwasnica et al. 2005). In the combinatorial generalization of Vickrey auctions, known as Vickrey-Clarke-Groves (VCG) auctions, bidders report their valuations for all packages without knowing the bids of other bidders, and items are then allocated efficiently to maximize total revenue. One of the shortcomings in VCG auctions is that bidders have to express their true values for all bundles without any information about prices, which limits its usefulness and practical application (Ausubel and Milgrom 2006b). Therefore, *iterative* combinatorial auctions are favored over sealed-bid auctions in complex economic environments for several reasons. Specifically, they help bidders express their preferences by providing pricing and allocation information at each bidding iteration (Parkes 2006), enhance bidders' ability to detect intensity of competition and learn when and how high to bid (Porter et al. 2003), and do not require bidders to reveal their preferences for all possible bundles. Iterative combinatorial auction designs can have multi-round or continuous timing rules (Porter et al. 2003). *Continuous* versions of iterative mechanisms strive to make the auction participation as simple as possible by enabling free-flowing, asynchronous participation, i.e., allowing bidders to join the auction and submit bids at any time without being subjected to activity-based rules and restrictions, and by providing relevant real-time information to facilitate the bidding process (such as information on current bundle prices). Continuous designs are generally more desirable in consumer-oriented (i.e., B2C or C2C) online marketplaces, and it is the type of mechanism we are focusing on in this study.

For combinatorial auction mechanisms to become practical for online marketplaces it is important to be able to implement and empirically test proposed theoretical mechanisms. This also helps to reduce the disparity between academic discussions of combinatorial auctions and their online marketplace implementations (Pekeč and Rothkopf 2003). To this end, Hoffman et al. (2006) consider how to best implement the ascending proxy auction; Bichler (2000) investigates multi-attribute auctions using web-based experiments that are designed based on game theoretical simulation results, and identifies implementation issues such as the need for decision support tools in these auctions; and Bichler et al. (2014) address implementation issues for combinatorial clock auctions and experimentally analyze designs for spectrum auctions; the aforementioned studies look at combinatorial designs with multi-round timing rules. Adomavicius and Gupta (2005) present a continuous bidder support scheme for iterative combinatorial auctions and test the effectiveness of their proposed design using lab experiments (Adomavicius et al. 2007). In this paper, we use data from this latter continuous combinatorial auction design that provides real-time bidder feedback (see Section 2 for details).

A stream of research has compared different iterative combinatorial auction designs in terms of auction outcomes, such as revenue and allocative efficiency, and identified auction parameters that affect these outcomes. Bapna et al. (2001) use online lab experiments to compare efficiency between a variant of English auctions and Vickrey auctions for multiple identical units. In another study, they use an analytical decision-theoretic approach complemented by an empirical investigation to study multi-unit online auctions and find that bid increment is an important parameter for online auctioneers that influences auction revenue and bidders' bidding strategies (Bapna et al. 2003a). Kwasnica et al. (2005) use lab experiments to compare their proposed Resource Allocation Design (RAD) auction (an iterative multi-item design that allows package bidding) with simultaneous multi-round (SMR) and Adaptive User Selection Mechanism (AUSM) auctions (Banks et al. 1989) in terms of efficiency, overall revenue, and auction duration. Bichler et al. (2009) use computer simulations to compare different iterative combinatorial auction designs in terms of allocative efficiency and revenue distribution, among other parameters. Brunner et al. (2010) use lab experiments to compare the simultaneous multi-round (SMR) auction (used by the FCC in 1994) in terms of seller revenue and efficiency against simultaneous multi-round package bidding (SMRPB), combinatorial clock (Porter et al. 2003), and RAD (Kwasnica et al. 2005) auctions.

1.2 Understanding bidder behavior in combinatorial auctions

An integral aspect for the design of effective and feasible mechanisms is the study of bidder behavior in combinatorial auctions since user preferences and behaviors are intricately linked to the design of information systems (Bapna et al. 2004). Understanding and modelling bidder behavior in auctions is important and can reveal interesting managerial insights or identify important issues that may have been overlooked in theoretical studies based upon normative models of bidder behavior (Shen and Su 2007), given that previous studies have observed that bidders do not consistently follow rational bidding (or best response) strategies that are expected in theory (Adomavicius et al. 2012; Kagel et al. 2010; Scheffel et al. 2011). Recent studies have analyzed bidder behavior in combinatorial auctions (Adomavicius et al. 2012; Brunner et al. 2010) and observed its effect on auction performance metrics (e.g., efficiency, bidder surplus, and auctioneer revenue). Park and Bradlow (2005) are among the first to provide a stochastic dynamic model of bidder behavior for online single-item auctions (ascending first-price auctions) and acknowledge the need for further research on this topic. Bapna et al. (2000) use a game-theoretical model to study multi-item online auctions, empirically compare them with multi-item Vickrey auctions, and find three different bidder types. In another study, they use simulations to study bidders' and auctioneers' strategies in online multi-unit auctions with multiple bidders (also called Yankee auctions), where the simulations are designed based on theoretically derived properties, and simulation parameters are instantiated based on empirical data from online auctions (Bapna et al. 2003b). They empirically identify five different bidding strategies in these auctions (Bapna et al. 2004). Schneider et al. (2010) use numerical simulations to compare the performance of several combinatorial auction mechanisms (in terms of allocative efficiency, revenue distribution, and auction duration) with respect to different bidding strategies and valuation models. Scheffel et al. (2011) use lab experiments to compare several iterative combinatorial auction mechanisms (ALPS, combinatorial clock, and iBundle) with VCG auctions in terms of efficiency and revenue distribution, and observe heterogeneous bidder behavior in all auction formats. Bichler et al. (2013a) use lab experiments to compare efficiency and revenue between simultaneous multi-round and combinatorial clock auctions, examine bidder behavior in both auction formats, and point out the need for further research to better understand bidders' bidding strategies.

Three main methodological approaches have been used to study the different aspects of combinatorial auctions discussed above (e.g., different combinatorial auction designs, practicality and implementation, auction outcomes and bidder behavior). *Game theoretical (analytical) modeling* is prevalently used to derive theoretical properties within simplified models (Krishna and Rosenthal 1996; Rothkopf et al. 1998). However, the multi-unit and discrete nature of iterative combinatorial auctions make the derivation of equilibrium bidding strategies intractable (Nautz and Wolfstetter 1997). Moreover, the traditional assumptions of homogeneous, risk-neutral bidders who adopt Bayesian Nash equilibrium strategies do not hold for most combinatorial auctions (Bapna et al. 2003b). *Experiments (both online and in-lab)* have been used to study auction outcomes and bidder behavior under different combinatorial designs. Experiments are usually costly and the number of treatment variables that can be analyzed is limited. More importantly, some research questions are not readily amenable to experimental methodologies due to the fact that it is not possible to a priori control for some important aspects of auctions in realistic settings, e.g., the bidding strategies employed by different participants in an auction experiment. Therefore, *computer simulation* has become an important research method that complements analytical, empirical, and experimental approaches (Bapna et al. 2003b; Bichler et al. 2009; Schneider et al. 2010). Simulations make it possible to investigate scenarios and study phenomena that are difficult to analyze analytically and/or are difficult (sometimes impossible) to test in experimental settings. The design of simulation models can be based on or informed by, theoretically derived properties and/or empirical data from experiments or real-world cases. For example, Adomavicius et al. (2009) use simulations that are designed based on analytical models of a multi-unit, weighted-average ascending price auction mechanism to design intelligent bidding strategies. Simulation methodologies have also been used in combination with each other, for example, analytical models can inform the better design of experiments (Bichler 2000), and computational simulations can be based on theoretical models (Adomavicius et al. 2009; Bapna et al. 2003b; Bichler et al. 2013b).

1.3 Overview of the three parts

To address the limitations in prior studies (i.e., understanding bidder behavior, interaction of different bidding strategies, and the effect of these two on auction outcomes), we adopt an agent-based modelling approach (Bonabeau 2002; Gilbert 2008; Jennings et al. 1998; Jennings and Wooldridge 1998; Macal and North 2005) that enables us to create, analyze, and experiment with simulation models composed of bidding agents that interact within the auction environment. Agent-based models are computational simulations that are similar to mathematical modeling in terms of rigor, but better suited for situations when agents are autonomous and heterogeneous, when there are complex interactions between agents, and when lower-level actions and interactions can lead to the emergence of system-level structures (Ren and Kraut 2014). In an auction scenario, bidder behaviors translate into lower-level actions and auction outcomes translate into system-level structures. We develop and validate novel software agents that replicate human bidder behavior, based on experimental data from continuous combinatorial auctions (henceforth called CCAs).

In the **first part,** I derive decision heuristics to model different aspects of bidder behavior and implement them in the bidding agents. The heuristic-based approach, while useful, lacks generalizability and does not provide much insight into the modeled aspects of behavior. In the **second part,** I use machine learning techniques to provide a more generalizable approach to model bidder behavior, which also outperforms the heuristic-based approach in terms of replicating the observed bidding strategies. These bidding agents are used in (agent-based) auction simulations to explore a wide variety of possible auction scenarios and different bidder type compositions (i.e., different competition types), many of which may not be easily encountered organically when running combinatorial auctions in experimental lab environments. My study draws upon, but differs from, existing research on automated bidding agents (Collins et al. 2010; He et al. 2003; Wellman et al. 2007) in that my agents are designed to *replicate* human behavior, i.e., not to outperform human participants, to compete against other agents, or to optimize a given task. Findings from the computational agent-based simulations also allow for bottom-up theorizing (Miller and Page 2009) and deeper understanding of how individual bidders' behavior interact and lead to emergent auction outcomes. In the **third part,** I use the agents developed in Part II in experimental auctions to control for the type of competition (i.e., bidding strategies) human participants face, and analyze the dynamics of bidder behavior and emerging auction outcomes such as revenue and efficiency. These human vs machine style experiments intend to examine the underlying reasons for different bidder behaviors and to understand the effect of competition type on participants' acceptance of CCAs as a viable trading mechanism. The experimental design, implementation and tuning of agents for these experiments, and findings are presented and discussed.

2. Background

2.1 Combinatorial auction fundamentals

In combinatorial auctions, bidders can bid on a single item or a combination of items (i.e., a bundle or package). At any time during the auction, any bid that has been submitted by an auction participant can be in one of three states: (a) *winning*, (b) *dead*, i.e., no chance of winning in the future, or (c) *live*, i.e., not currently dead or winning but may change to one these states depending on future bids. This is substantially different from single-item auctions where a bid can only be either winning or dead. Winning bids are bids on non-overlapping bundles that create the highest revenue (finding such bids is known as the winner determination problem in combinatorial auctions), which are updated upon any new bid. Given the three possible states of a bid, there are naturally two important bidding levels for any bundle *b* at any given time in an auction (Adomavicius and Gupta 2005): *deadness level* (DL) and *winning level* (WL), where $DL(b) \leq WL(b)$. A bid amount above the WL will make a bid *winning*, below the DL will result in a *dead* bid, and a bid amount in-between DL and WL will result in a *live* bid. Adomavicius and Gupta (2005) developed a bidder support scheme for real-time iterative combinatorial auctions that incorporates these two constructs (i.e., DL and WL) and allows for comprehensive real-time bidder feedback. Auction revenue (or auctioneer revenue) is the amount obtained by the auctioneer from winning bids, which equals to the sum of winning bid amounts in a first-price auction mechanism. Another important auction outcome is allocative efficiency, which measures how close the allocation of items is to the optimal allocation at the end of an auction. Allocative efficiency is defined as:

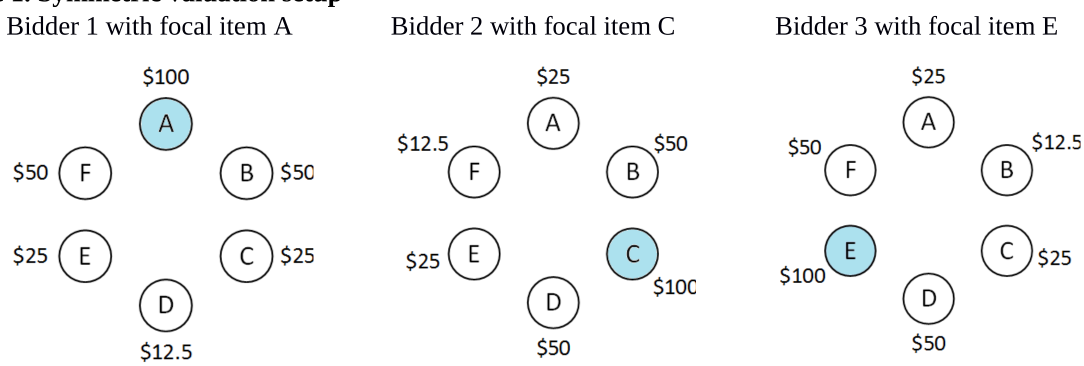
$$\frac{\text{All Bidders' Surplus} + \text{Auction Revenue } (= \text{Total Surplus})}{\text{Maximum Possible Surplus}} = \frac{\text{Bidders' Valuation for Won Bundles}}{\text{Maximum Possible Surplus}}$$

Higher allocative efficiency is often stated as a desirable goal in auction literature, because it leads to greater social welfare (Banks et al. 2003; Banks et al. 1989; Cramton et al. 2006; Porter et al. 2003), since allocative efficiency is maximized when items are acquired by bidders who value them the most.

2.2 Continuous combinatorial auction setup

To illustrate the proposed agent-based approach, we use the following experimental continuous combinatorial auction (CCA) setup first introduced by Adomavicius and Gupta (2005) as well as the dataset collected in their study (Adomavicius et al. 2012), referred to as the "baseline experimental data" in the rest of the paper. Three bidders compete to acquire six items, representing six real-estate properties around a lake. A systematic valuation scheme is used where each bidder is designated a *focal item*, which has the highest value for that bidder among all items. The remaining items' value decreases by 50% the farther they are from the focal item (see Figure 1). For example, with six lots around a lake, the right and left neighbors of the focal lot are both half as valuable as the focal item, and the lot on the opposite side of the focal item is valued at 12.5% (= 0.5 × 0.5 × 0.5) of the focal lot.

Figure 1. Symmetric valuation setup



Complementarities among items are defined by adding a super-additive valuation of 10% for each adjoining item in a bundle. For example, if a bundle consists of focal item A ($100 value), its neighbor item B ($50 value), and B's neighbor C ($25 value), then the valuation for bundle ABC is ($100 + $50 + $25)×(1 + 0.1 + 0.1) = $210 since there are two adjoining items in this bundle. Another example for bidder 1, if a bundle consists of item B ($50 value), adjoining item C ($25 value), and non-adjoining item E ($25 value), then the valuation for bundle BCE is ($50 + $25 + $25)×(1 + 0.1) = $110. Bidders are provided comprehensive information feedback throughout the auction, i.e., they can see bids placed so far in the auction, the provisional winning allocation at the current auction state, and the WL and DL for any bundle of interest. Since there are six items in each auction there are 63 ($= 2^6 - 1$) possible bundles, i.e., all possible subsets of 6 distinct items except for the empty set.

2.3 Baseline experimental data

Using cluster analysis of bids and clicks generated by bidders, three stable bidder strategies/types have been identified in prior work (Adomavicius et al. 2012), namely: *Analyzers* (A), *Participators* (P), and *Explorers* (E). These bidder types were shown to exhibit different behaviors in terms of several bidder-specific variables, including:

- **Bids**: the number of bids placed by a bidder throughout the auction;
- **Spans**: the number of distinct bundles a bidder bids on throughout the auction;
- **Surplus**: the bidder's valuation for the bundles won at the end of the auction minus the amount the bidder has to pay for them (i.e., the winning bid amount);
- **Effort**: total number of clicks made by a bidder divided by the total number of bids placed by that bidder (i.e., mean number of clicks per bid) during an auction, representing the level of information-seeking effort by the bidder prior to submitting a bid (e.g., in terms of looking at DLs and WLs of various potential bundles of interest).

Table 1 shows cluster means and standard deviations for bidder-specific variables based on cluster analysis of baseline experimental data.

Table 1. Summary statistics of bidder-specific variables for different bidder types in baseline experimental data (means are in regular font, standard deviations in italics)

Bidder Type	Number of	Bids		Spans		Surplus		Effort	
Analyzer	4	15.50	(6.61)	8.00	(3.56)	77.50	(50.08)	32.97	(26.20)
Participator	34	23.06	(8.84)	9.88	(3.67)	47.26	(44.65)	17.45	(8.61)
Explorer	7	52.86	(16.88)	21.57	(4.89)	31.79	(28.45)	11.15	(3.86)

Analyzers (A) represent the most rational bidders who spend more effort on analyzing the auction state and dynamics. Compared to the other two bidder types, a typical Analyzer places fewer bids on a smaller set of bundles for which s/he has higher valuation, and derives higher surplus as a result. *Participators* (P) demonstrate only moderate effort investigating the auction activity and maintain a steady participatory behavior. A typical Participator places more bids than Analyzers on a wider variety of bundles and derives a lower surplus than Analyzers, but higher surplus than Explorers. *Explorers* (E) spent the least effort on analyzing the auction environment and can display very random, exploratory behavior. Compared to the other two types, a typical Explorer places the most bids on a wider variety of bundles and derives the lowest surplus as a result.

Since three bidders can participate in each auction and there are three different bidder types (A, P, and E), there can be 10 different possible bidder type combinations in a single auction: AAA, AAP, AAE, APP, APE, AEE, PPP, PPE, PEE, and EEE. We refer to these combinations of competing bidders as *competition types*. In general, if there are *n* possible bidder types and *k* bidders participate in an auction, there are C possible competition types, where $C = \binom{n+k-1}{k}$, i.e., the k-combination *with repetitions* from a set of size n.

Only 3 out of these 10 possible competition types were observed frequently enough in the baseline experimental CCAs (Adomavicius et al. 2012) to be included in statistical analysis. Out of 15 auctions, there were 4 auctions with APP competition (i.e., one Analyzer and two Participators in an auction), 5 auctions with PPP competition, 5 auctions with EPP competition, and a single instance with EEP competition. This further highlights the issue that it may be very difficult to collect data about all possible combinations of behaviors in experimental setups organically, i.e., without the explicit ability to control for the type of competition that given bidder may face. The few types of competition observed in this dataset do not allow us to directly study the dynamics of bidder behavior and its effect on auction outcomes. For example, we do not know how a typical Analyzer would behave under the other 5 possible types of competition it could face, and how auction outcomes would look like in those cases. This limitation also applies to other previous studies that have used data from real world or experimental combinatorial auctions, where only a limited number of all possible competition types were observed, and/or where the number of observations was not sufficient for analysis.

In summary, the limitations of other methodological approaches motivate our data-driven agent-based modeling approach. Developing software agents that replicate the different human bidding behaviors observed in experimental CCAs allows us to run any number of auction simulations for *all* possible competition types and, hence, enables us to study the effects of different competition types on bidder behavior and auction outcomes. Comprehensive auction simulations also let us discover other possible existing relationships among competition type, bidder behavior, and auction parameters.

2.4 Difference among bidder types

The different bidder types were identified based on the cluster differences in several variables as presented in Table 1. To model bidder behavior in software agents, we only need to use Bids, Spans, and Surplus data from baseline experiments. We do not explicitly need to consider Effort because the number of clicks indicates effort that has been exerted by a *human* bidder in order to explore the bidding environment; the bidding environment is directly available to software agents simply by analyzing the available information. Table 2 compares these three main variables among different bidder types using t-tests to identify significant differences that guide our agents' design. We also run non-parametric Mann-Whitney-U tests to make sure our comparisons are consistent if the normality assumptions do not hold for some of the compared samples; both tests indicate consistent results.

Table 2. Comparison of outcome variables between three bidder types in baseline experimental data

		Mann Whitney U test		Welch t-test		
Hypothesis		**W**	**p-value**	**T**	**p-value**	**Significance**
Bids	A < P	34.5	0.0580	-2.0794	0.0500	marginally significant
	P < E	8	0.0001	-4.545	0.0014	significant
	A < E	0	0.0030	-5.2004	0.0003	significant
Spans	A < P	55.5	0.2832	-0.9971	0.1890	not significant
	P < E	6.5	0.0000	-5.9815	0.0002	significant
	A < E	0	0.0053	-5.2873	0.0003	significant
Surplus	A > P	95	0.1033	1.1549	0.1596	not significant
	P > E	137	0.2714	1.1721	0.1311	not significant
	A > E	22	0.0777	1.6775	0.0832	marginally significant

significant: below or at 0.01, marginally significant: below or at 0.1

There is a significant difference in Bids and Spans among the three bidder types. The difference in Surplus, even though not always statistically significant, is still considered in our agent design, because the observed higher surplus of Analyzers is theoretically meaningful, and the limitation of the experimental data (e.g. few data points for Analyzers) can be a reason for this statistical insignificance.

Part I: Modeling Bidder Behavior: A Data-Driven Heuristic Approach

3. Characterizing bidder behaviors

In a typical CCA, at any time in the auction, a bidder can select a bundle s/he is interested in, check the DL and WL for the selected bundle, and decide to either place a bid of a certain monetary amount on the selected bundle or to not bid at this time. By analyzing the experimental bid-level and clickstream-level data, we observe different temporal bidding patterns for the three bidder types and characterize these behaviors in terms of "how often they bid at any auction state" (**bidding frequency**), "what they bid on" (**bundle selection**), and "how much they bid on a selected bundle" (**bid amount**). These three aspects capture the dynamics of bidder behavior and, when implemented in our software agents, determine bidding agent behavior at any time in the auction. It is important to note that we model for *the three distinct behaviors and the differences in temporal bidding patterns,* not merely the bidder-specific outcome variables (i.e., Bids, Spans, and Surplus) that were used to identify, a posteriori, different bidding strategies (discussed in Section 2.2). The bidding agents are expected to generate comparable bidder-specific outcomes, which will be verified as part of our validation. Moreover, we are not modeling for outlier type of behavior that might be occasionally observed in experimental CCAs, such as the premature closing of auctions due to bidders' unexpectedly stopping to participate (potentially due to fatigue, boredom, etc.). Introducing such aspects into our model would not be useful for the aim of understanding dynamics of bidder behavior, but the simulation model does have the flexibility to incorporate such rare suboptimal behaviors if it would be useful for future studies.

To model each bidder type (Analyzer, Participator, and Explorer) in terms of the behavioral aspects discussed above, we aggregate bidding data for bidders of the same type across all auctions. Also, since bidder behavior is likely dependent on the *current state* of the auction, we want to parameterize auction progress. Time is one possible continuous indicator of an auction state (e.g., represented as early, mid, and late portions of the auction), but it has certain limitations. A point of time is not a consistent indicator of auction progress (or auction state) across different auctions. This is because the duration of auctions varies significantly and mere passing of time does not always translate into bidder activity; e.g., bidders may not place any bids for some period of time during the auction. We propose to use current *auction revenue* as a proxy for auction progress and discretize the temporal bidding data based on revenue (note that revenue is monotonically increasing in any auction). By taking auction progression and mapping it to auction revenue, revenue becomes a discrete, time-independent parameter that represents an auction state in the CCA context. Another advantage of a time-independent auction progression model is that it allows to straightforwardly scale the bidding dynamics simulation to any desired auction duration (in terms of auction running time), which is especially important when deploying agent-based bidders in auction experiments together with human bidders. Note that information about the DL and WL of any bundle is updated upon any new bid, and is available from the auction framework to every bidder. The discretization step of our approach draws upon the unsupervised binning with equal-width bins method discussed in machine learning and data mining literature (Chmielewski and Grzymala-Busse 1996; Dougherty et al. 1995; Holte 1993; Liu et al. 2002). The discretized auction state is then used to model the dynamic aspects of bidder behavior, as discussed below.

3.1 Bidding frequency

Figure 2 shows the high-level pseudocode for modeling bidding frequency as a function of auction state. To derive a dataset from which we can extract the dynamic bidding frequency aspect, we pre-process the data by discretizing and aggregating the experimental baseline data based on revenue (lines 3-8 of the pseudocode). Given a certain revenue binwidth, we calculate the number of bids each bidder has placed within each bin's revenue range; this is the bidder's bidding frequency at each revenue bin (discretization steps in lines 3-4 of the pseudocode). For example, if a bidder has placed four bids when auction revenue is between \$25 and \$50 (i.e., within the 2nd bin's revenue range when using binwidth of \$25), his bidding frequency equals 4 for the 2nd bin. For each bidder type (Analyzer, Participator, and Explorer), we calculate the average bidding frequency and standard deviation at each bin by aggregating the bidding frequency values of all same-type bidders for that bin (aggregation step in line 5 of the pseudocode). E.g., if three Analyzers have the bidding frequency values of 4, 5, and 6 for the 2nd bin, the average bidding frequency value for Analyzers at the 2nd bin equals 5, with a (bidding frequency) standard deviation of 1; similar average and standard deviation values are calculated for all bins (line 6 of the pseudocode). The result is average bidding frequency and standard deviation series for each bidder type (lines 7-8 of the pseudocode), which represent how bidding frequency changes with auction state.

Figure 2. Pseudocode for deriving bidding frequency functions

Deriving the bidding frequency function: f:Auction State → Bidding Frequency

Inputs:
Bid-level data D
Bidder type b, where $b \in \{Analyzer\ (A), Participator\ (P), Explorer\ (E)\}$
Class of candidate fitting functions $\mathbb{F}$

Outputs:
f_b // the mean bidding frequency function
f_b' // the corresponding standard deviation function for type b

Procedure:
1 D_b = bid level data for bidders of type b ($D_b \subset$ D)
2 Foreach $bw \in \{\$10, \$15, \ldots, \$35, \$40\}$ // *bw* – revenue binwidth, bin i represents auction state *i*
3 Foreach bidder j in D_b and all bins i = 1 to n, where n = $\lceil max\ Revenue\ /\ bw \rceil$
4 c_{ij} = number of bids placed by bidder j in bin i
5 M_i = mean of c_{ij} across all j (for each *i*)
6 S_i = standard deviation of c_{ij} across all j (for each *i*)
7 $M_{b,bw}$= $[M_1, M_2, \ldots, M_n]$ // series of all M_i derived for type b bidders with binwidth bw
8 $S_{b,bw}$= $[S_1, S_2, \ldots, S_n]$ // series of all S_i
9 Foreach ∈ $\mathbb{F}$ **:** fit function $f_{b,bw}$ on $M_{b,bw}$ and calculate goodness of fit measures
10 $f_b = f_{b,bw}$ that has highest robust adjusted R^2 AND does not produce negative values at any state
11 Fit linear function f'_b on $S_{b,bw}$ with bw of above selected $f_{b,bw}$
12 Return $f_b\ and\ f_b'$

Since the initial binwidth choice affects the resulting series, the discretization and aggregation procedure is repeated for several different binwidth values to avoid missing meaningful patterns in bidding frequency in our pre-processing steps (as indicated in line 2 of the pseudocode). The reason we do this is that there is no generally applicable approach to select the best number of bins and/or binwidth for discretizing the data; this is comparable to the bandwidth selection problem in kernel density estimation (Fan and Gijbels 1995; Jones et al. 1996; Ruppert et al. 1995). Here, the aforementioned bidding frequency series are derived for the following values of the binwidth parameter: \$10, \$15, \$20, \$25, \$30, \$35, and \$40 (seven different binwidths). We use maximum binwidth of \$40 since the auctions usually close at a revenue of about \$400 in the baseline experimental data, which gives us about 10 data points after aggregation. Fewer data points would be unreliable and could lead to over-smoothing and/or underfit the data. We use a minimum binwidth of \$10 to avoid overfitting the data in the subsequent modeling steps; binwidths smaller than \$10 are too granular and result in too many empty bins (too many missing data points after discretization).

To determine the best frequency model for each bidder type, we fit polynomials of degree 1 (linear fit) to 4 on the mean bidding frequency series (all seven series derived using different binwidths) using the robust linear least squares method with bisquare weights to account for possible noisiness of the derived series by down-weighting outliers (line 9 of the pseudocode). We derive 28 (= 7 × 4) functions for each bidder type. The most suitable function is selected (line 10 in pseudocode) based on goodness of fit (robust R-square and adjusted R-square) and the expert-based understanding of bidder behavior (e.g., a function that can produce substantial negative values is ruled out even if it is a good fit, since bidding frequency cannot be negative). The following functions are selected for each bidder type and determine the average bidding frequency as a function of state (plotted in Figure 3):

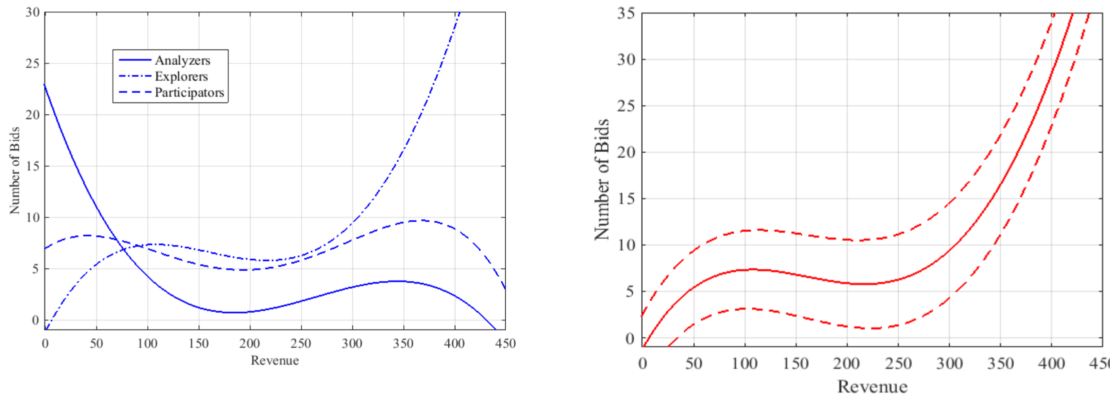
- Analyzers: 3rd degree polynomial fitted on data that is discretized using \$35 revenue binwidths (robust $R^2 = 0.7801$, Adjusted $R^2 = 0.6858$).
- Participators: 4th degree polynomial on data that is discretized using \$30 revenue binwidths (robust $R^2 = 0.6313$, Adjusted $R^2 = 0.4838$).
- Explorers: 4th degree polynomial on data that is discretized using \$35 revenue binwidths (robust $R^2 = 0.7593$, Adjusted $R^2 = 0.6218$).

To model the variation in bidding frequency among bidders of the same type as a function of auction state, we fit a linear function on the standard deviation series that is discretized using the same binwidth that is used for the selected mean function (line 11 in pseudocode). E.g., for

Analyzers, a linear fit on the standard deviation series discretized using the $35 revenue binwidth is used. As an illustration, Figure 3b shows the average bidding frequency function for Explorer types plus/minus the linear standard deviation function which represents one standard deviation (rough 68% confidence interval) around the mean.

Figure 3. Comparing average bidding frequency functions

(a) Comparing Average Bidding Frequency Functions for Different Bidder Types (b) Average Bidding Frequency Function with Confidence Interval Approximation for Explorers



Consistent with prior literature (Adomavicius et al. 2012), our frequency models (depicted in Figure 3) indicate that Analyzers place consistently fewer bids compared to Explorers and Participators. They bid more frequently early in the auction and place slightly fewer bids as the auction proceeds. Their variation in bidding frequency is low and does not fluctuate significantly throughout the auction, as indicated by the very small slope and relatively small intercept in the linear standard deviation modeling function for Analyzers, i.e., $0.03765\,x + 2.328$. Participators and Explorers bid more frequently throughout the auction and have higher variation in their bidding frequencies, as indicated by the steeper slope and larger intercept of their standard deviation modeling functions (for Participators, $0.2226\,x + 3.585$; for Explorers, $0.1615\,x + 3.738$). Participators' and Explorers' bidding frequency is similar until about the middle of the auction. From the midst of the auction onwards, Explorers keep increasing their bidding frequency whereas Participators keep a relatively steady bidding frequency and reduce their bidding behavior only towards the very end of the auction. Bidding agents determine their bidding frequency at any auction state by making a random pick from a normal distribution whose mean and standard deviation is determined by the two corresponding mean and standard deviation functions.

3.2 Bundle selection

Given the same valuation scheme and identical focal items, the three bidder types vary in their propensity to bid on different bundles. For example, as observed in baseline experimental data, Analyzers are much more likely to realize that it may be advantageous to bid on a bundle that consists of all six items compared to Participators and Explorers. Bidders' interest in different bundles has both a static aspect, their inherent interest for different bundles, and a dynamic/contextual aspect, which depends on how a bidder's preference for different bundles changes at different auction states (over time). For example, a bidder who has the same likelihood of bidding on any bundle in a set of equally valuable bundles at the beginning of an auction (his static preference for those bundles), becomes more likely to bid on bundles with less competition as the auction proceeds (his dynamic preference for those bundles); i.e., a bidder would prefer bundles with less competition among equally valuable bundles. We model the bundle selection aspect of bidder behavior as static, to represent bidders' inherent interest in different bundles. The dynamic aspect is captured when the bundle selection aspect is combined with the two other aspects of bidder behavior (i.e., bidding frequency and bid amount). Referring to our earlier example, higher competition on certain bundles is reflected in the higher DL and WL for those bundles, which is captured by the bid amount aspect (see Section 3.3), and a possible increase in bidding frequency by other bidders trying to acquire those bundles, which is captured by the bidding frequency aspect (see Section 3.1).

To model the bundle selection aspect in a data-driven manner, we use the aggregated bidding data across all auctions for each bidder type and derive discrete probability distributions that specify the likelihood of a bundle being selected; the 63 possible bundles are the possible values for the categorical random variable (on the x-axis). Figure 4 shows the pseudocode for modeling the bundle selection aspect.

We do not always observe bids on all 63 bundles in our baseline experimental data due to the limited number of observations (e.g., for Analyzers). However, given the valuation scheme of the items, we know that the value of some of these unobserved bundles are equal or similar to bundles on which bidders have placed bids. These unobserved bundles are considered as likely to be selected as other equally valuable bundles that we have observed in the data (line 5 in pseudocode). For example, two bundles that both consist of the focal item and its left or right neighboring item (let us name these bundles FL and FR correspondingly), are both equally valuable. If we do not observe any bids on the FR bundle in the baseline experimental data but observe bids on the FL bundle, we use the same likelihood we have estimated for FL being selected, for FR's selection as well. As an illustration, Figure 5 shows the probability distribution we use to model the bundle selection aspect of Participators' behavior. Note that bidders' likelihood of selecting a bundle, as modeled above, and their valuation for that bundle are not correlated (correlation coefficients close to zero for all three bidder types).

Figure 4. Pseudo code for modeling bundle selection

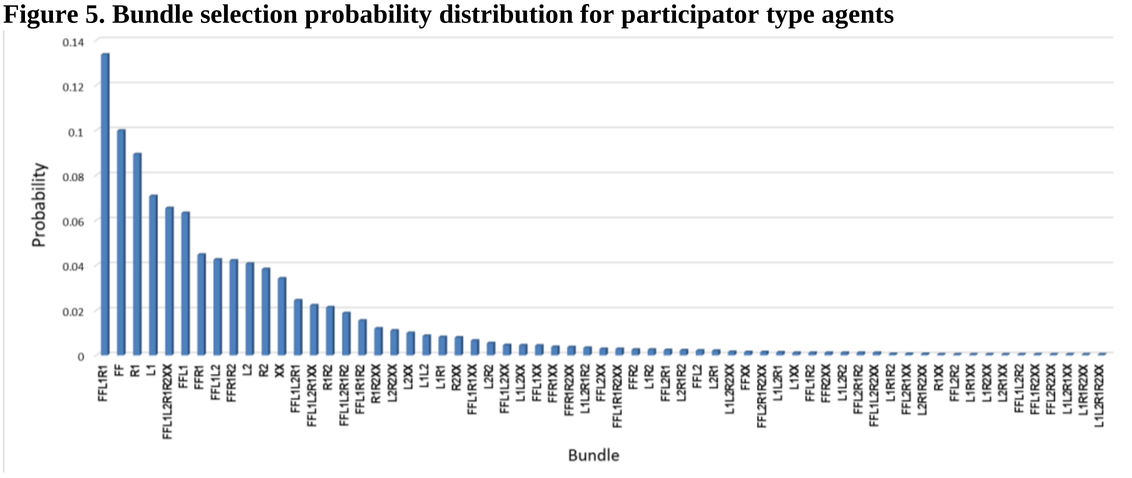
Bundle selection function: $Probability\ Mass\ Function\ (PMF): Valuation \rightarrow Bundle$

Inputs:
Bid-level data D
Bidder type b, where $b \in \{Analyzer\ (A), Participator\ (P), Explorer\ (E)\}$

Outputs:
P_b // Probability Mass Function for bidder type b

Procedure:
D_b = bid level data for bidders of type b ($D_b \subset$ D)
Foreach bundle x ∈ $\{1,2,3,\ldots,63\}$ // all possible bundles
  Y = set of bundles with equal valuation as x // bundles equivalent to x
  If bid(s) on x exist in D_b Then c_x = count of bids on x in D_b // count this bundle
  ElseIf bids on y ∈ Y exist in D_b Then c_x = mean of c_y for all y ∈ Y //count equivalent bundles
  Else c_x = 0.5 × $\min(c_{x'})$, where x' = any bundle except x //assign min selection probability
$C_b = [c_1, c_2, \ldots, c_{63}]$ // series of bundle frequencies for type *b*
Derive probability mass function P_b from C_b
Return P_b

Figure 5. Bundle selection probability distribution for participator type agents



Note: the labels represent bundle names with respect to the focal item of a bidder (FF – focal item, XX – item on the opposite side of the focal item, R1 – first neighboring item to the right, etc.). For example, FFL1 represents a bundle that consists of the focal item and its first left neighbor.

In our auction simulations, a bidding agent determines which bundle to bid on by making a random pick from its bundle selection distribution given its focal item. For example, if the pick from the distribution returns the bundle that consists of the focal item, and the 1st and 2nd neighboring items on the right (labeled as bundle FFR1R2 in Figure 5), the agent for bidder 1 will select bundle ABC when the setup in Figure 1 (from Section 2) is used. Our approach for modeling the bundle selection aspect of bidder behavior can be easily generalized to auction setups with more than 6 items as well.

3.3 Bid amount

To properly model how much different bidder types bid on bundles across different auctions, we need to take into account the changing WL and DL of bundles as the auctions proceed. We introduce a **bid position** variable that represents the bid amount *relative* to the bundle's WL and DL, and calculate it for every bid as follows:

- If WL = DL : bid position = bid amount / WL;
- If WL > DL : bid position = (bid amount – DL) / (WL – DL);
- If WL and/or DL are not yet set for the bundle (i.e., in the beginning of the auction), the initial bid amounts are used to establish separate starting bid distributions which we use to determine an amount for first bids placed by bidding agents.

In the $WL > DL$ case, a bid position above 1 means the bid amount is above the WL for that bundle (hence a winning bid), a bid position below 0 means the bid amount is below the DL for that bundle (hence a dead bid), and a bid position greater than 0 and lower than 1 means the bid amount is between the DL and WL for that bundle (hence a live bid). In the case where DL = WL, a bid position above 1 means the bid amount is above the WL (which equals the DL) for that bundle and a bid position below 1 means the bid amount is below the WL. This bid position variable allows us to meaningfully analyze the bid amount aspect of the three bidder types across different auctions. Figure 6 shows the pseudocode for modeling bid-position as a function of auction state.

As with bidding frequency modeling, we discretize and aggregate the bid positions derived for each individual bidder in the baseline experimental data using the different revenue bins ($10 to $40 in $5 steps). The result is seven series of average bid positions (and seven series of corresponding standard deviations) for each of the three bidder types. To model the distinct patterns we observe for the three different bidder types, we fit different functions (polynomials of degree 1 to 4 using the robust linear least squares method with bisquare weights) on each of these average bid positions series and select the most appropriate one, considering goodness of fit (based on R-square and adjusted R-square of each fitted function), complexity of the function, and the interpretability of the function given our understanding of bidders' behavior (similarly to what was done for bidding frequency).

Figure 6. Pseudo code for deriving bid-position function

Deriving the bid-position function: $f\colon Auction\ State \rightarrow Bid\ position$

Inputs:

Bid-level data D

Bidder type b, where $b \in \{Analyzer\ (A), Participator\ (P), Explorer\ (E)\}$

Class of candidate fitting functions $\mathbb{F}$

Outputs:

f_b // the mean bid-position function

f_b' // the corresponding standard deviation function for type b

Procedure:

D_b = bid level data for bidders of type b ($D_b \subset \mathrm{D}$)

Calculate bid-position variable for each observation (i.e., bid) in D_b //as explained in text

Foreach $bw \in \{\$10, \$15, \ldots, \$35, \$40\}$

    For bin i = 1 to n, where n = $\lceil max\ Revenue\ /\ bw \rceil$

        M_i = mean bid-position value of all bids placed in bin i

        S_i = bid-positions' standard deviation for all bids placed in bin i

    $M_{b,bw}$= series of all M_i derived for type b bidders with binwidth bw

    $S_{b,bw}$= series of all S_i

    Foreach $\in \mathbb{F}$**:** fit function $f_{b,bw}$ on $M_{b,bw}$ and calculate goodness of fit measures

$f_b = f_{b,bw}$ that has highest robust adjusted R^2

Fit linear function f'_b on $S_{b,bw}$ with bw of above selected $f_{b,bw}$

Return f_b and f_b'

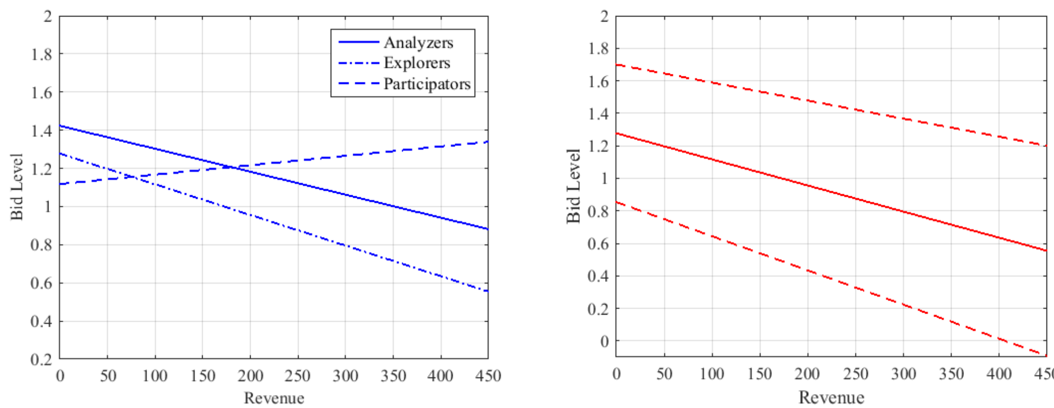
We select the following functions to model the average bid positions for each of the bidder types as a function of auction revenue; see Figure 7a:

- Analyzers: linear fit on bid position data discretized using \$35 revenue binwidths (R^2 = 0.7789, Adjusted R^2 = 0.6841).
- Participators: linear fit on bid position data discretized using \$15 revenue binwidths (R^2 = 0.8161, Adjusted R^2 = 0.8093).
- Explorers: linear fit on bid position data discretized using \$10 revenue binwidths (R^2 = 0.7735, Adjusted R^2 = 0.7676).

These bid position functions are plotted in Figure 7a. In other words, to model the difference among bidders of the same type in terms of bid position, we use a linear function that is fit on the standard deviation series discretized with the same revenue binwidth we selected for the average bid positions function; i.e., linear fit on the standard deviation series discretized using revenue bins of width \$35 for Analyzers, \$15 for Participators, and \$10 for Explorers, corresponding to the above-selected functions for average bid positions. As an example, Figure 7b shows the average bid position function for Explorer type bidders, plus/minus the linear standard deviation function which gives a rough 68% confidence interval (one standard deviation) around the mean.

Figure 7. Comparing average bid-levels functions for different bidder types

(a) Comparing Average Bid-Levels Functions for Different Bidder Types (b) Average Bid-Level Function with Confidence Intervals for Explorer Type



Based on Figure 7a, all three bidder types start the auction with comparably aggressive bid positions; i.e., the amount they bid is similar, relative to the DL and WL of bundles they bid on, with the mean bid above WL. Analyzers begin by placing bids above the WL (i.e., bid positions above 1) that are winning or live; towards the end of the auction they are likely to place more strategic bids that are very close to the WL, which makes them more likely to derive higher surplus, as compared to Participators who are more likely to bid more aggressively above the WL (i.e., to submit jump bids) in order to maintain the winning status throughout the auction. Explorers, because of their largely non-systematic, exploratory nature, are more likely to place dead bids (bid positions close to and below 0) as the auction proceeds.

A bidding agent determines the amount it bids on a selected bundle by specifying a bid position and knowing the WL and DL for the bundle at any given auction state. The bid position is specified by making a random pick from a normal distribution whose mean and standard deviation are determined by the corresponding average and standard deviation functions explained above.

3.4 Agent-based auction simulations

Each of the three main aspects of bidder behavior (bidding frequency, bid amount, and bundle selection) are implemented in our agents as separate modules. A fourth module determines the agent's final decision on whether or not to bid, after combining values returned by the three above modules and considering the agent's bidding history so far. Figure 8 provides an overview of the auction simulation with three bidding agents and the agent's decision-making procedure.

Figure 8. Auction simulation pseudocode

```
Auction simulation:
1   NonActivity = 0
2   While (1)
3       Shuffle {                          // execute the following statements in random order
4           bA = pingAgent1(state)         // check if agent wants to place a bid at current state
5           bB = pingAgent2(state)
6           bC = pingAgent3(state)     }
7       If (bA = NULL and bB = NULL and bC = NULL)       // if none of the agents want to bid
8           NonActivity = NonActivity + 1                // increase non-activity count
9       Else
10          Update state                          // update auction state with newly submitted bids
11      If (NonActivity ≥ Threshold)       // if sufficiently long inactivity
12          Break                          // end auction simulation

Bidding agent:
13      If (number of bids placed in current state < BiddingFrequency(state)) // BiddingFrequency module
14          Select bundle                                      // using BundleSelection module
15          Determine bid amount for bundle given current state // using BidPosition module
16          If (NOT underbidding itself AND bid-amount + margin ≤ valuation)
17              Return bid(bundle, bid amount)
18          Else
19              Return NULL        // not bidding
20      Else
21          Return NULL            // not bidding
```

An auction is simulated as a sequence of virtual rounds (the While loop in Figure 8). In each round, all participating bidding agents are called in random order (lines 3-6 in Figure 8), and each agent decides to either place a bid at the given auction state or not to bid. The bidding decision is determined by the agent's three internal modules (i.e., bidding frequency, bundle selection, and bid amount) and taking into account the bids it has placed so far in the auction. Each bidding agent first checks whether it has surpassed its expected number of bids at the current auction state (line 13 in Figure 8), as determined by its bidding frequency module. If not, it decides which bundle and how much it wants to bid (lines 14-15 in Figure 8), using its bundle selection and bid amount modules. Finally, if the agent is not underbidding itself and if the determined bid amount (plus a margin) is not higher than the agent's valuation, it places the bid (lines 16-17 in Figure 8). The aforementioned margin parameter is modeled differently for different bidder types representing an inherent level of aggressiveness observed in the baseline experimental data. If none of the bidding agents place a bid in a round (line 7 in Figure 8), it is considered as a count of non-activity, which models the inactivity of bidders in experimental auctions. When the non-activity count reaches a specified stopping threshold the auction stops (line 11-12 in Figure 8); this threshold, models the extended inactivity time that ends experimental auctions, mimicking the so-called “soft stopping” rule. For auction simulations, we choose a threshold value of 150, which is calibrated to have auction simulations stop at a revenue similar to the average revenues observed in experimental auctions. We run simulations with different threshold values in our sensitivity analysis in Section 4.5. In summary, the developed simulation platform allows us to run auction simulations with any desired combination of bidding agent types.

4. Data-driven validation of agent-based simulations

Once our bidding agents are developed and run correctly from a programming perspective (i.e., no errors or bugs in the code), we need to verify that they replicate the behavior of human bidders observed in experimental CCAs. Agent-based simulations are part of the more general family of computational simulations. The literature on simulation validation provides various frameworks for verification and validation of computer simulations ranging from technical approaches (e.g., spectral analysis) to behaviorally oriented ones such as the Turing test (Sargent 2005; Van Horn 1971). Naylor and Finger (1967) propose a multistage approach where the researcher selects a set of postulates that describe the behavior of a system, empirically verifies these selected postulates, and tests the model’s ability to predict the behavior of the system. Van Horn (1971) proposed a similar approach of using simulation models with face validity, statistically testing assumptions, and comparing input-output transformations generated by the model to those generated in the real world. Model validation techniques generally address the fit between a theory and the model of that theory, or the fit between the model and the real-world phenomenon that the model is supposed to simulate (Gilbert 2008). The validation techniques we use for our simulation model are more of the latter type (i.e., fit between the model and the real-world phenomenon), therefore, we focus on assessing the *outcome validity* of our models. Outcome validity can be assessed at four levels: pattern validity, where the pattern of results generated by the computational model matches patterns of real data; point validity, where the behavior of the model on each dependent variable, taken one at a time, has the same mean as the real data; distributional validity, where the distribution of results generated by the computational model has the same distributional characteristics as the real data; and value validity, where the specific results from the computational model match on a point by point basis the real data (Carley 1996). Which level to choose depends on the model’s purpose and is at the researcher’s discretion. It is understood that there can be multiple correct answers for the simulation validation problem, and researchers' decision of which methods and techniques to use depends on the modeled phenomena and the intent of the model (Kleindorfer et al. 1998), among other criteria.

In designing our bidding agents, we proposed that three main aspects (namely, bidding frequency, bundle selection, and bid amount) characterize different bidding behaviors (Sections 3.1-3.3). In this section, we verify whether our bidding agents, modeled based on these three agent-level parameters, correctly replicate human bidders' behavior and auction-level variables via their emergent behaviors under similar conditions (i.e., competition types). We assess the validity of our agent-based simulations in terms of matched bidder-specific variables (Bids, Spans, and Surplus) and auction outcome (revenue), by statistically comparing outcome variables generated by our simulation model with outcome variables from experimental CCAs (for competitions observed in experimental CCAs). The level of assessment for different matched variables varies from pattern validity to distributional validity, as we will elaborate in the following subsections.

4.1 Replicating competition types observed in experimental auctions

To verify that our bidding agents replicate the behavior of human bidders under similar competition types, we run auction simulations with the same types of competition observed in experimental auctions, namely APP, EPP, and PPP. The number of auction simulations for each of the competition types is proportional to their occurrence in experimental auctions (five times as many simulated auctions as observed in experimental ones) so that the data from our simulations is comparable to data from human bidders. Table 3 presents summary statistics of bidder-specific variables per each competition type for auction simulations and baseline experimental data.

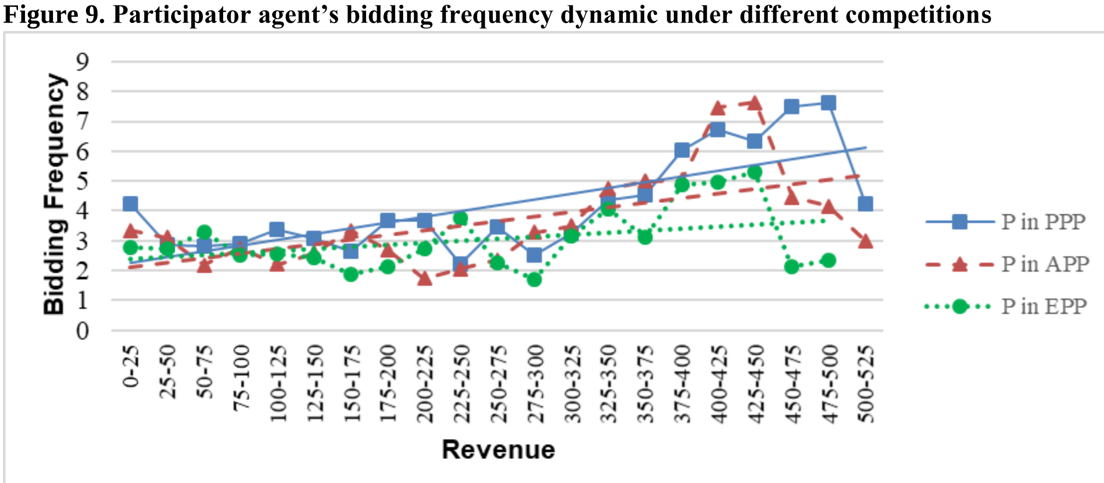
In designing the bidding agents, we implicitly model context awareness via the dynamic aspects of bidding behavior (i.e., bidding frequency and bid position), arguing that auction state is an emergent property of interacting bidders. For instance, the difference in Participators' bidder-specific outcomes under different competitions (upper part of Table 3) demonstrates that a Participator-type agent, modeled to behave according to a typical Participator, is context-aware and responds differently to different competitions. Our complete simulation results in Section 5 further support this argument by providing evidence for agents of all three bidder types.

Figure 9 shows the temporal differences of the dynamic bidding frequency aspect in Participator agents' behavior under different competition types. This further illustrates that the proposed simulation approach can capture the bidder's context-awareness in terms of its dynamic behavioral aspects.

Table 3. Summary statistics of bidder-specific variables for human bidders and bidding agents (standard deviations in *italics*)

Competition Type	Human Bidders in Baseline Experimental Data					Auction Agents in Our Simulations				
	Bidder Type	# of Auctions	Bids	Span	Surplus	Agent Type	# of Auctions	Bids	Span	Surplus
APP	P	4	23.62 *11.07*	10.38 *5.04*	72.25 *59.12*	P	20	27.05 *5.62*	13.125 *2.09*	4.013 *11.82*
EPP	P	5 (+1 EEP)	28.82 *8.94*	9.91 *2.91*	58.91 *32.84*	P	30	17.517 *5.61*	10.7 *2.79*	16.25 *18.05*
PPP	P	5	18.53 *4.19*	9.6 *3.58*	25.38 *34.72*	P	25	20.64 *5.89*	11.36 *2.52*	5.4 *12.66*
All Auctions Pooled Together	A	15	15.5 *6.61*	8 *3.56*	77.5 *50.08*	A	75	20.2 *6.51*	9.35 *2.2*	17.1 *15.83*
	P		23.06 *8.84*	9.88 *3.67*	47.26 *44.65*	P		20.64 *5.89*	11.36 *2.52*	5.4 *12.66*
	E		52.86 *16.88*	21.57 *4.89*	31.79 *28.45*	E		49.67 *6.49*	22.03 *2.39*	-2.03 *11.116*

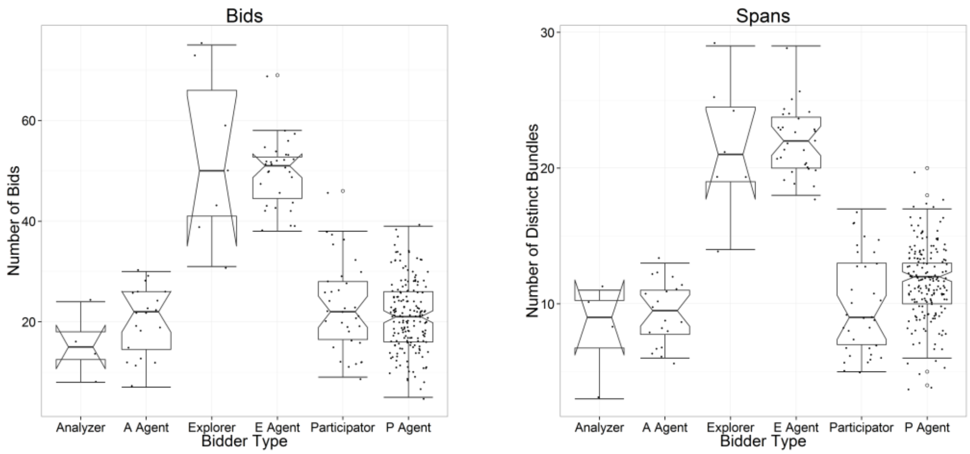
Figure 9. Participator agent's bidding frequency dynamic under different competitions



4.2 Replicating bundle diversity and number of bids

Figure 10 shows the side-by-side comparison of Bids (number of bids) and Spans (diversity of bundles) variables between human bidders and bidding agents. The horizontal axis indicates the human bidder or bidding agent type (e.g., "E Agent" stands for Explorer type bidding agents). For all boxplots presented in this paper, the middle line is the median and the upper and lower "hinges" correspond to the third and first quartiles. The upper (lower) whiskers extend from the hinges to the highest (lowest) value that is within $1.5 \times IQR$ of the hinge, where IQR is the inter-quartile range or distance between the first and third quartiles. The notches extend $\pm(1.58 \times IQR)/sqrt(n)$, which gives a roughly 95% confidence interval for comparing medians; for further details see (Chambers 1983; McGill et al. 1978). Overlaid jitter plots indicate the plotted variable's distribution.

Figure 10. Comparison of bids and spans variables between human bidders and bidding agents



We use t-tests, Wilcoxon rank-sum (WRS aka. Mann-Whitney U) tests, and Kolmogorov-Smirnov (KS) tests to compare the data generated by our bidding agents in simulated auctions with data from experimental auctions generated by human bidders. Mann-Whitney U tests are considered non-parametric counterparts for t-tests, and it is typical to assume that data generated from human subjects (the baseline experimental data) do not meet the normality assumptions required for a t-test (see for example Kwasnica et al. 2005; Porter 1999). When the t-tests and WRS tests indicate no significant difference, we can infer point validity; i.e., the means and medians of the compared variables are not significantly different. We also run two-sample Kolmogorov-Smirnov (KS) tests, to verify whether variables generated by humans and bidding agents have the same distribution. When the KS test does not show a significant difference we can infer distributional validity for the compared variable. Table 4 shows the comparison of the Bids and Spans variables between human bidders and bidding agents of the same type. Based on the results, all three bidding agent types generate an equal number of bids (i.e., Bids variable) as human bidders. Since none of the tests show a significant difference in Bids for any of the three bidder types, our simulation model has distributional validity in terms of Bids.

The bundle diversities (i.e., Spans variable) generated by Analyzer and Explorer type bidding agents have the same distributions as Spans generated by human bidders of the same type. Participator type bidding agents seem to slightly differ from human bidders; i.e., Spans difference is significant at the 1% level (p-values close to 0.01 for all three tests). However, by looking at Participator data broken down by competition we see that the difference in Spans is only marginally significant when Participators compete against other Participators (i.e., in auctions with PPP type competition); there is no significant difference in Spans between Participator type human bidders and bidding agents in auctions with EPP and APP competitions. Nevertheless, the *relative* difference in bundle diversity among different bidding agent types always matches those generated by human bidders; i.e., for both bidding agents and human bidders we have: Spans (Analyzers) < Spans (Participators) < Spans (Explorers). In terms of bundle diversity, our simulation model has (in the worst case) pattern validity.

Table 4. Comparing bidder-specific variables between human bidders and bidding agents

H0: Human = Agent H: Human ≠ Agent			Wilcox Rank Sum		t-test		Kolmogorov Smirnov	
	Comparing Human vs. Bidding Agent of Type		W	p-value	T	p-value	D	p-value
Bids	Analyzer		25.5	0.277	-1.296	0.26	0.45	0.5095
	Explorer		109.5	0.876	0.49	0.6394	0.395	0.3379
	Participator		3325	0.2782	1.266	0.2128	0.125	0.7635
Spans	Analyzer		31.5	0.532	-0.73	0.5113	0.25	0.9853
	Explorer		98.5	0.814	-0.243	0.815	0.295	0.7056
	Participator		2096.5**	0.0062	-2.499*	0.0166	0.342**	0.0026
	Participators broken down by competition type	P in APP	123.5	0.3146	-1.517	0.1703	0.5+	0.0713
		P in EPP	269.5	0.3368	-0.832	0.4197	0.212	0.797
		P in PPP	348.5*	0.0198	-1.814+	0.087	0.387*	0.0476

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

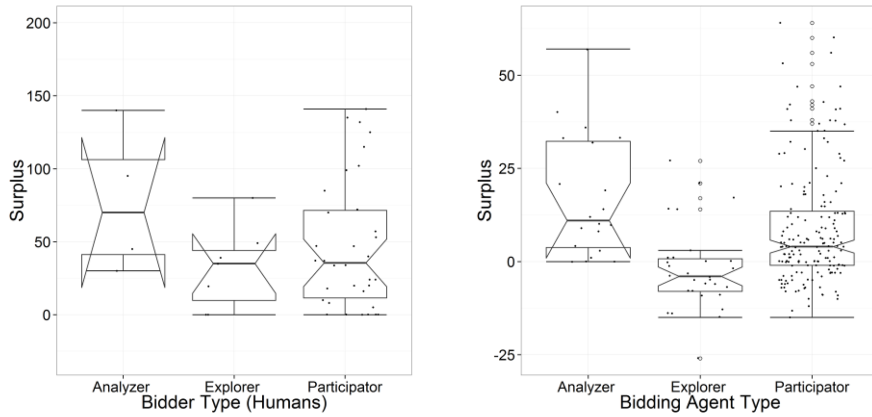
4.3 Replicating surplus

With respect to bidders' Surplus, our results indicate pattern validity. Figure 11 shows that the relative difference in surplus among bidding agent types is similar to the relative difference among human bidders; i.e., for both bidding agents and human bidders we have: Surplus (Analyzers) > Surplus (Participators) > Surplus (Explorers). The overall lower surplus in simulations (as indicated by the scale on the Y-axis) is the result of excluding outlier-type behavior in modeling the canonical behavior of bidders, whereas in real auctions (baseline experimental data) there are instances where some human bidders unexpectedly stop bidding prematurely, which generally leads to a higher total surplus for other bidders in the auction (and lower auction revenue).

4.4 Replicating auction outcomes

We also compare auction revenue between simulated and experimental auctions with identical competition types to verify comparable auction-specific outcomes. Auction revenue has a similar pattern in both experimental and simulated auctions as shown in Table 5. The significance of these

Figure 11. Comparison of surplus between human bidders and bidding agents



revenue differences among different competition types is not always the same, based on the Wilcoxon rank-sum (WRS) tests shown in Table 5 (since revenue is not normally distributed). As mentioned earlier, the reason for higher revenues in simulated auctions (compared to experimental auctions) is that our agents do not model outlier type of behavior, which leads to a lower average surplus for bidding agents (as explained in Section 4.3) and higher auction revenues.

Table 5. Comparing auction revenue across competition types observed in experimental auctions

	Revenue in Competition of Type:				
Experimental Auctions	APP	≈	EPP	<*	PPP
Mean Revenue (standard deviation)	258 (141)		333.7 (66.6)		381.6 (49.9)
Auction Simulations	APP	≈	EPP	<***	PPP
Mean Revenue (standard deviation)	468 (22.8)		463.4 (21.7)		482 (22.1)

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Auction revenue is lower in experimental auctions with EPP type competition compared to those with PPP competition. We cannot make statistically significant revenue comparisons for experimental auctions with APP competition type due to the limited number of observations in the Baseline Experimental Data, even though these auctions seem to have lower median (and mean) revenue. In our simulations, auctions with APP and EPP competition generate significantly less revenue than auctions with PPP competition, but there is no statistically significant difference in revenue between auctions with APP and EPP competition types (we will elaborate on why this is the case in Section 5.3 after our comprehensive auction simulations). Our results show that our simulation model has pattern validity in terms of auction revenue.

Based on our analyses, we conclude that our agent-based simulation model has (in the worst case) pattern validity when we assess our model in terms of bidder-specific outcome variables (i.e., Bids, Spans, and Surplus) and auction outcome (i.e., revenue). Pattern validity is sufficient to leverage our model for studying dynamics of bidder behavior and consequent auction outcomes under different competition types. The validation results indicate that characterizing bidder behavior in the complex decision environments of CCAs in terms of bidding frequency, bid amount, and bundle selection allows us to successfully build bidding agents that replicate different canonical real-world bidding behaviors.

4.5 Sensitivity to auction-level parameter: stopping threshold

In this section, we test the sensitivity and robustness of our simulation model to changes in an important macro-level (i.e., environment-level) simulation parameter, i.e., the auction stopping criterion. The stopping threshold models the extended inactivity time (i.e., the soft stopping rule) that ends experimental auctions, as noted in Section 4.3. In all the main simulations (used in Sections 4 and 5), a stopping threshold of 150 is used, which stops auction simulations at a revenue similar to the average revenues observed in experimental auctions. We run separate sets of auction simulations with much lower (i.e., 50) and much higher (i.e., 250) stopping thresholds to verify our simulation model's consistent behavior in terms of relative differences in bidder-specific variables between the three bidder types, and the model's sensitivity in terms of auction revenue. A low threshold in auction simulations models premature end of auctions. As a result, we would expect less opportunity for bidder types to differentiate themselves from other types in terms of bidder-specific variables as well as lower average revenues since bidders are prematurely stopped from exploring the possibilities to place further advantageous bids on more diverse bundles. Conversely, a high threshold, models allowing much longer inactivity time before stopping experimental auctions. Thus, we would expect bidder types to fully differentiate themselves in terms of bidder-specific variables and would expect higher average revenues, since there is a higher likelihood for bidders to place higher bids on more diverse bundles.

For this sensitivity analysis, we run 90 auction simulations for each stopping threshold, 30 per each competition type observed in experimental CCAs (APP, EPP, and PPP). Table 6 shows average values for bidder-specific variables and auction revenue across all three sets of auction simulations. The horizontal comparisons in Table 6 show that number of bids (Bids) and bundle diversity (Spans) generally increase with stopping threshold for all three bidding agent types, whereas average surplus decreases, resulting in an increase in auction revenue (also see Table 7).

In auction simulations with a low stopping threshold of 50, the similar surplus between Participators and Explorers and the marginally lower Bids for Analyzers, indicate that auction simulations are stopping too early, which does not allow for bidding agents to fully express their preferences for different bundles by placing further bids and to differentiate themselves in terms of Surplus. This also explains why the difference in revenue between replicated competition types across auction simulations changes from insignificant to significant under higher stopping thresholds (last row of Table 6). These results are consistent with what we would expect from human bidders. Stopping thresholds higher than 250 do not result in further significant changes in any of the studied outcomes since allowing for longer inactivity time would not result in further bids once all bidder types have fully expressed their valuations. Table 7 shows the insignificant change in revenue for auction simulations with a higher stopping threshold of 300; other bidder-specific variables were also observed to have similar insignificant changes.

Table 6. Analyzing sensitivity of bidder-specific variables and auction revenue across different stopping thresholds; WRS tests are used for comparisons and "mean, median" values are shown

Variable	Bidding Agent Type	Stopping Threshold				
		50		**150**		**250**
Bids	Analyzer	14.03 , 14	< $^{+}$	16.83 , 17	< ***	24.43 , 23.5
		< $^{+}$		< ***		< $^{+}$
	Participator	15.9 , 16	< ***	21.52 , 21	< ***	26.28 , 25
		< ***		< ***		< ***
	Explorer	28.97 , 28	< ***	52.87 , 50.5	< **	63.5 , 61
Spans	Analyzer	8.33 , 8.5	≈	8.87 , 9	< $^{+}$	10.07 , 10
		< ***		< ***		< ***
	Participator	10.28 , 10	< ***	11.75 , 12	< ***	12.78 , 13
		< ***		< ***		< ***
	Explorer	16.77 , 17	< ***	22.3 , 23	< **	24.93 , 25
Surplus	Analyzer	27.97 , 19	≈	23.13 , 15	** >	12.93 , 5
		* >		*** >		*** >
	Participator	15.93 , 9	*** >	5.99 , 3	*** >	1.49 , -1
		≈		*** >		*** >
	Explorer	13.7 , 13	*** >	-3.37 , -4.5	≈	-6.3 , -7
Auction Revenue		424.42 , 423.5	<***	474.34 , 475.5	<***	494.84 , 499.5
Comparing Auction Revenue between Competition Types		EPP ≈ APP ≈ PPP		EPP ≈ APP <*** PPP		EPP <* APP <*** PPP

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

These results indicate our bidding agents' consistent and realistic behavior under different stopping thresholds; i.e., relative differences in bidder-specific variables are similar between bidding agent types in simulations. This indicates that our agent-level modeling choices are appropriately sensitive to changes in the auction-level stopping threshold parameter.

Table 7. Comparing revenue across auction simulations with different stopping thresholds for competition types observed in experimental CCAs

		Stopping Threshold						
		50		150		250		300
Auction Revenue	Mean	424.42		474.34		494.84		496.08
	(*std. dev*)	*(30.29)*	<***	*(19.28)*	<***	*(17.81)*	≈	*(19.45)*
	Median	423.5		475.5		499.5		499

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

5. Leveraging bidding agents

Our agent-based auction simulations allow us to explore the effect of competition on the dynamics of bidder behavior and auction outcomes. Having validated the agents and knowing that they replicate human bidder behavior reasonably closely, we are interested in finding out whether the bidding dynamics that result from different competitions, while the bidder agents bid based on the coded canonical behaviors, provide interesting insights into potential auction outcomes through "emergent" behaviors of the agents under different competitive environments. We run 100 auction simulations (with 3 bidders competing to acquire 6 items in each auction) for each of the 10 possible competition types (total of 1000 auction simulations) in this part of the study using the same valuation setup used in experimental CCAs (and a stopping threshold of 150). Table 8 shows mean and standard deviation values for auction specific and bidder-specific variables grouped by bidder type and competition type. This dataset is used for all analyses in this section.

5.1 Effect of competition on bidder behavior

Each bidder can face 6 types of competition based on the composition of bidding strategies it encounters in the auction, namely: AA (i.e., competing against two Analyzers), AP (i.e., competing against an Analyzer and a Participator), AE, PP, EP, and EE. Even if we assume that bidder's behavior is endowed, we can clearly hypothesize that the type of competition a bidder faces affects his/her behavior, including both the number of placed bids (Bids) and the diversity of bundles bid on (Spans). Based on the three bidder types' characteristics, we conjecture that Analyzers generally make for tougher competition, Explorers make for easier competition, and Participators make for medium competition. The comprehensive simulations will allow us to further characterize the 6 aforementioned competition types.

We use ANOVA to study how bidders' own type, the competition type they face, and the possible interaction of these two factors affect key bidder-specific outcome variables (Bids and Spans). Table 9 shows ANOVA results for the two models (one model for each dependent variable)

Table 8. Summary statistics of auction simulations for all possible competition types

Competition Type	Auction Specific Variables		Bidder Specific Variables				
	Revenue	Allocative Efficiency	Agent Type	n	Bids	Spans	Surplus
AAA	456.26 *(16.84)*	0.9801 (0.0289)	Analyzer	300	27.35 *(6.765)*	11.01 *(2.164)*	12.897 *(13.225)*
APE	451.59 *(19.933)*	0.9427 (0.0593)	Analyzer	100	17.45 *(6.306)*	9.52 *(2.681)*	17.835 *(20.058)*
			Explorer	100	56.64 *(10.712)*	22.87 *(2.185)*	-2.13 *(11.848)*
			Participator	100	23.39 *(8.238)*	12.44 *(2.994)*	8.785 *(14.398)*
APP	468.86 *(20.009)*	0.9752 (0.0413)	Analyzer	100	18.71 *(5.93)*	9.32 *(2.222)*	17.44 *(17.118)*
			Participator	200	26.19 *(7.045)*	13.205 *(2.503)*	3.088 *(11.08)*
EAA	432.19 *(18.759)*	0.9446 (0.0522)	Analyzer	200	18.31 *(4.687)*	9.77 *(2.144)*	22.803 *(21.83)*
			Explorer	100	62.04 *(10.601)*	24.5 *(2.941)*	-0.795 *(14.322)*
EEA	449.19 *(20.422)*	0.9325 (0.0568)	Analyzer	100	20.63 *(6.313)*	10.59 *(2.454)*	15.69 *(19.835)*
			Explorer	200	51.37 *(10.306)*	21.26 *(2.71)*	3.015 *(13.806)*
EEE	466.08 *(23.19)*	0.9705 (0.0401)	Explorer	300	45.303 *(8.474)*	20.377 *(2.836)*	8.015 *(15.645)*
EEP	470.64 *(22.05)*	0.9714 (0.0402)	Explorer	200	45.655 *(7.693)*	20.225 *(2.69)*	3.968 *(14.331)*
			Participator	100	20.77 *(6.119)*	11.91 *(2.786)*	12 *(16.162)*
EPP	469.78 *(18.642)*	0.9666 (0.0409)	Explorer	100	49.98 *(8.112)*	21.82 *(2.67)*	-3.385 *(8.809)*
			Participator	200	18.07 *(5.39)*	10.845 *(2.65)*	10.878 *(15.74)*
PAA	453.38 *(20.605)*	0.9719 (0.0397)	Analyzer	200	20.09 *(5.817)*	9.645 *(2.184)*	18.555 *(16.738)*
			Participator	100	32.72 *(7.979)*	14.72 *(2.324)*	0.335 *(8.301)*
PPP	487.34 *(21.753)*	0.9882 (0.0268)	Participator	300	21.457 *(6.731)*	11.44 *(2.513)*	3.903 *(10.812)*

Standard deviations in parentheses

using type III sum of squares to account for unbalanced samples (i.e., controlling for the other two factors when testing each of the main effects and the interaction effect). The models are highly significant (p-value ≤ 0.001) with a high explanation of variance (adjusted $R^2 > 75\%$). This tells us that bidding agents' own strategy and the competition type they face both significantly affect the number of bids and the diversity of bundles they bid on. The significant interactions indicate that the effect of competition type on agents' behavior depends upon the agent's bidding strategy.

We use ANOVA to study how bidders' own type, the competition type they face, and the possible interaction of these two factors affect key bidder-specific outcome variables (Bids and Spans). Table 9 shows ANOVA results for the two models (one model for each dependent variable) using type III sum of squares to account for unbalanced samples (i.e., controlling for the other two factors when testing each of the main effects and the interaction effect). The models are highly significant (p-value ≤ 0.001) with a high explanation of variance (adjusted $R^2 > 75\%$). This tells us that bidding agents' own strategy and the competition type they face both significantly affect the number of bids and the diversity of bundles they bid on. The significant interactions indicate that the effect of competition type on agents' behavior depends upon the agent's bidding strategy.

Table 9. ANOVA of factors influencing bidding agents' behavior and surplus

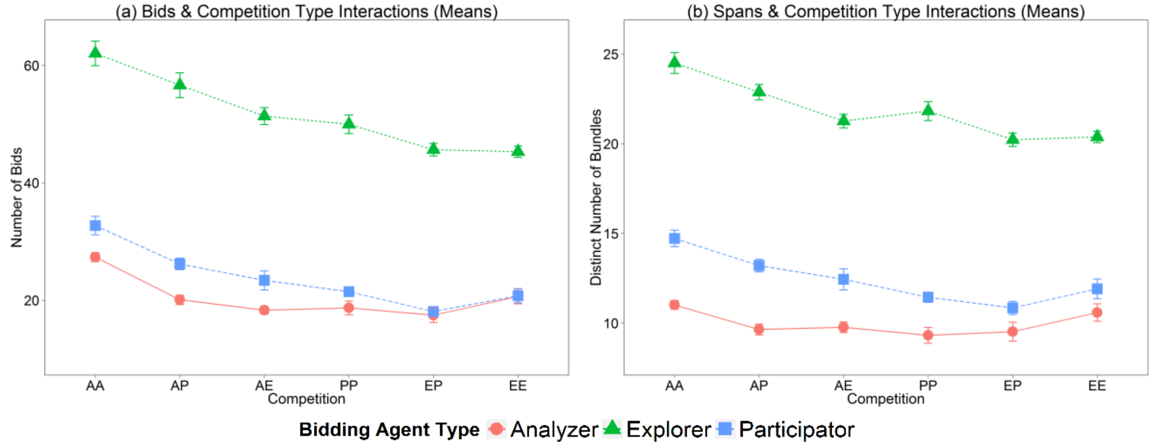
	Degrees of Freedom	Bids	Spans	Surplus
		F	F	F
Bidder Type	2	816.011***	1058.414***	46.119***
Competition Type Bidder Faces	5	54.408***	13.306***	11.0939***
Bidder Type × Competition Type	10	13.302***	14.843***	5.4314***
R^2		0.7742	0.7956	0.1867
Adj. R^2		0.7729	0.7944	0.1821
		F(17,2982) = 601.5 ***	F(17,2982) = 682.6 ***	F(17,2982) = 40.28 ***

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

We use interaction plots to uncover the patterns of these interactions in Figure 12. Each point shows the group means and the error bars indicate 95% confidence intervals (1.96 × standard errors); each group consists of a certain agent type facing a specific competition, e.g., Analyzers facing AP competition. These plots also provide a guideline for further post-hoc tests to verify significant differences. Distribution of Bids and Spans variables do not violate normality assumptions for most groups.

Under any competition type, differences between the three bidding agent types (Participator, Analyzer, and Explorer) are consistent with the differences among the three bidding strategies observed in experimental CCAs; i.e., Analyzers place the least number of bids and focus on a narrower set of (lower lines with circles in Figure 12); Explorers place the most bids, bid on a broad variety of bundles (upper lines with triangles in Figure 12); and Participators' bidding behavior is

Figure 12. Effect of competition on bidding agent behavior for different bidding strategies



the middle ground between Analyzers and Explorers in terms of both number of overall bids and the variety of bundles they bid on (middle lines with squares in Figure 12). For each bidder type, we use pairwise t-tests to compare mean Bids and Spans between any two competition types. When we suspect non-normal distributions, we also use nonparametric WRS tests for comparisons (results of both tests are consistent in such cases). Tables 10-12 report the mean and median values for these variables under different competition types as well as significant differences (significance level of the less significant test is considered to make conclusions about significant differences when the two tests have different significance levels).

Table 10. Comparing explorers' behavior across different competition types using WRS tests

		Explorer Facing Competition of Type:										
Bids		AA	>***	AP	>***	PP	≈	AE	>***	EP	≈	EE
	Mean	62.04		56.64		49.98		51.37		45.655		45.303
	Median	61		57		49.5		51		44		45
Spans		AA	>***	AP	>***	PP	>+	AE	>***	EP	≈	EE
	Mean	24.5		22.87		21.82		21.26		20.225		20.377
	Median	25		23		22		21		20		21

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

For Explorers, as the competition changes from AA to EE, there is a significant and consistent drop in both the overall number of bids they place (Bids) and the overall variety of bundles they bid on (Spans); see Table 10. Fewer Bids and Spans generally imply a less intense competition that demands less exploration. Once there are two Explorers in the competition, the third bidder's strategy does not significantly affect Explorers' bidding behavior (EP vs. EE competition types) unless the third bidder is an Analyzer (AE competition type), where Explorers place slightly more bids on more various bundles.

Participators' change in Bids and Spans as the competition changes from AA to EE is different from what we observed for Explorers; see Table 11. This implies that some competition types are perceived differently by Participators compared to Explorers in terms of intensity. Competing against two Analyzers results in the highest Bids and Spans for Participators, and both these variables decrease once one of the Analyzers is replaced by a Participator. Competitions of type PP, AE, and EE are equally demanding in terms of Bids and Spans, and make for less challenging competition compared to AA and AP competitions. Participators place the least number of bids on the fewest number of distinct bundles when they bid against EP competition type, implying that this is the least demanding competition. Interestingly, bidding against two Explorers – which we initially expected to be the least demanding competition – makes Participators place more bids on more various bundles, likely in response to the frequent random bids Explorers place on various bundles.

Table 11. Comparing participator' behavior across different competition types

		Participator Facing Competition of Type:										
Bids		AA	>***	AP	>**	AE	>*	EE	≈	PP	>***	EP
	Mean	32.72		26.19		23.39		20.77		21.457		18.07
	Median	32		25.5		21		21		20.5		17.5
Spans		AA	>***	AP	>*	AE	≈	EE	≈	PP	>*	EP
	Mean	14.72		13.205		12.44		11.91		11.44		10.845
	Median	15		13		12		12		11		11

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Table 12. Comparing analyzers' behavior across different competition types

		Analyzer Facing Competition of Type:										
Bids		AA	>***	EE	≈	AP	>***	AE	≈	PP	≈	EP
	Mean	27.35		20.63		20.09		18.31		18.71		17.45
	Median	27		20		20		18		19		17
Spans		AA	>**	EE	>**	AP	≈	AE	$>^{+}$	PP	≈	EP
	Mean	11.01		10.59		9.645		9.77		9.32		9.52
	Median	11		11		10		10		9		9

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Analyzers are characterized as the most rational among the three bidder types who make the most effort to place smarter bids that are expected to maximize their surplus. Competing against two Analyzers is, therefore, considered to be the toughest competition other bidders (including other Analyzers) can face. The bidding agents' behavior supports this hypothesis, as implied by the highest Bids and Spans for all three agent types when they face competition of type AA (Figure 12 and Tables 10-12). The competition becomes less demanding when one of the Analyzers is substituted with a Participator (facing AP competition), where Analyzers' Bids and Spans both significantly decrease. AE, PP, or EP are all (equally demanding) less intense competitions for Analyzers as compared to auctions with AP type of competition, implied by Analyzers' lower Bids and Spans. Interestingly, these three different types of competition are all equally demanding for Analyzers, which is not the case for Participators and Explorers who, for example, perceive EP competition as the least intense.

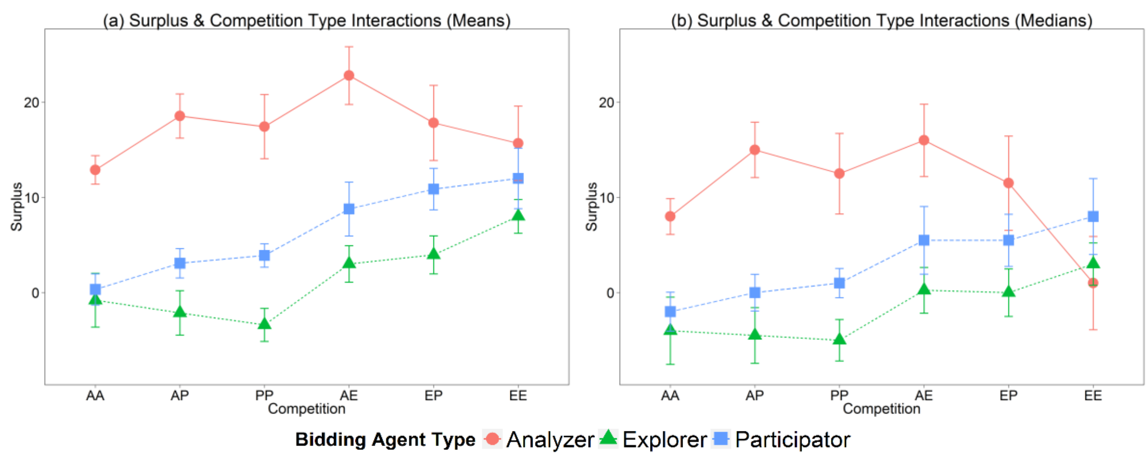
Competing against two Explorers seems to provide a disadvantage to Analyzer agents since they behave similarly to when they face AA and AP types of competition, even though the EE competition type is composed of two rather randomly bidding Explorers (expected to be less demanding) as compared to AA and AP competitions that are composed of more rational bidders (expected to be more demanding). A typical Analyzer agent has to place more bids on more various bundles trying to take advantage of, and respond to, the rather random behavior of Explorers.

Our analysis of agents' emerging behaviors provides insights into human bidders' probable behaviors under similar competitions. Participation of **Analyzers** in an auction generally makes for a tougher competition. Other bidder types, including another Analyzer, place more bids (higher Bids) on a wider variety of bundles (higher Spans) in the presence of Analyzers. **Explorers** generally make for a weaker competition. Other bidder types, including another Explorer, generally bid less frequently (lower Bids) on less various bundles (lower Spans) when competing with Explorers. However, Explorers' rather unexpected behavior in terms of placing random bids on various bundles can also confuse other bidders and makes for more uncertain outcomes. When the majority of bidders are Explorer types, other bidders (i.e., bidders facing EE competition type) can get confused if they overanalyze Explorers' random behavior and may, therefore, place more bids on more diverse bundle. Analyzers are seemingly more likely to be affected in this way than Participators. **Participators** are the middle ground in terms of bidding frequency and bundle diversity. They are less competitive than Analyzers but more competitive than Explorers, without Explorers' random bidding behavior. Their participatory behavior can neutralize the competitiveness induced by Analyzers and the effect of Explorers' random behavior on other bidders (e.g., in auctions with APE competition).

5.2 Competition and bidders' economic welfare

We are also interested in the effect of competition on bidders' economic welfare, as measured by the surplus of winning bidders after the auction ends. A bidder's surplus is her/his valuation for the won bundles minus the amount s/he has to pay for them. We use ANOVA to study how bidding agents' own type, the competition type they face, and the possible interaction of these two factors, affect the surplus retained by bidders. Our ANOVA results (Table 9) show that both factors, as well as their interaction, are highly significant (p-value ≤ 0.001), with a reasonable explanation of variance (adjusted $R^2 > 18\%$). We use interaction plots to uncover the patterns of these interactions in Figure 13. Since the distribution of surplus is positively skewed for some groups, we draw interaction plots for both group means and group medians to take into account possible differences between mean and median patterns. The error bars in Figure 13 indicate roughly 95% confidence intervals (equal to 1.96 × standard error); median standard errors equal regular standard errors multiplied by 1.25 (1.2533 × standard error). Differences between the mean and median plots indicate under which competition types we have unusual surplus distributions that we take into account for pairwise comparisons, e.g., Analyzers' surplus when facing competition type EE.

Figure 13. Effect of competition on bidder surplus for different bidding strategies



Under any competition type, Analyzers generally derive the highest surplus, Explorers derive the least (and sometimes negative) surplus, and Participators generally derive less surplus than Analyzers but more than Explorers (this pattern is consistent with the differences in surplus between the three bidding strategies observed in the baseline experimental data). For each bidding agent type, we make pairwise comparisons between competition types to explore significant differences in surplus.

The mean and median surplus for **Explorer** agents follows the same pattern under all competition types; lower lines with triangles in Figure 13. Table 13 shows Explorers' mean and median surplus and indicates significant differences, using t-tests as well as nonparametric WRS tests (if they have different significance levels, the less significant one is considered). When there is only one Explorer type agent participating in an auction (i.e. Explorers facing PP, AP, or AA competition), the other two bidders (Participators or Analyzers) can exploit his/her random behavior, and the Explorer ends up making zero or negative surplus (i.e., Explorers tend to overpay their valuations when trying to win a bundle). Having two Explorer type agents in the auction makes both of them derive significantly more (and positive) surplus as compared to any other type of competition with only a single Explorer (AE and EP competitions). Explorers benefit from having more bidders of their type in the auction since it becomes less likely for their random behavior to be exploited by other bidding strategies (i.e., Analyzers and Participators). Explorer agents make the highest surplus when they face EE competition type since there is no other bidder type who would exploit their random behavior (EE was the least demanding type of competition for Explorers in terms of Bids and Spans; see Section 5.1). The significant increase in Explorers' surplus, as the competition changes from AE or EP to EE, supports the argument that Explorers' random behavior is exploited by other bidder types who place more calculated bids.

Table 13. Explorers' surplus across different competition types

		Explorer Facing Competition of Type:										
Surplus		PP	≈	AP	≈	AA	<*	AE	≈	EP	<**	EE
		PP		≈		AA						
	Mean	-3.385		-2.13		-0.795		3.015		3.968		8.015
	Median	-5		-4.5		-4		0.25		0		3

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Table 14. Participators' surplus across different competition types

		Participator Facing Competition of Type:										
Surplus		AA	<*	AP	≈	PP	<**	AE	≈	EP	≈	EE
		AA		<**		PP		AE		≈		EE
	Mean	0.335		3.0875		3.903		8.785		10.8775		12
	Median	-2		0		1		5.5		5.5		8

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Table 15. Analyzers' surplus across different competition types

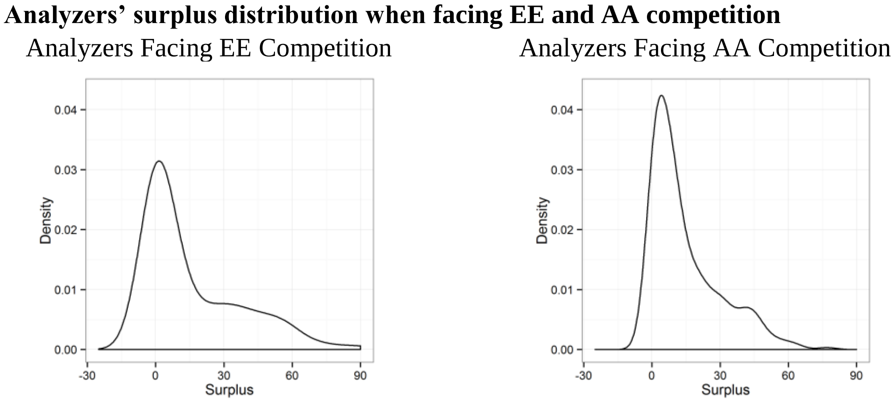
		Analyzer Facing Competition of Type:										
Surplus		AA	< >	EE	<***	PP	≈	EP	≈	AP	<*	AE
	Mean	12.897		15.69		17.44		17.835		18.555		22.803
	Median	8		1		12.5		11.5		15		16

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Participator agents derive higher surplus as the competition becomes less demanding in terms of Bids and Spans; increase in both mean and median surplus can be observed in Figure 13. Participators make their highest surplus in competitions where Explorers participate (i.e., AE, EP, and EE), see Table 14. Participators derive less surplus when facing AP or PP competitions, without a significant difference in surplus between these two competition types, and make the least surplus (close to zero, on average) when competing against two Analyzers (AA competition type is also the most demanding in terms of Bids and Spans). The pattern in Participators' surplus implies that they are capable of deriving higher surplus by exploiting Explorer's random behavior, but can themselves be exploited by Analyzers.

Analyzer agents, like the other two agent types, make the least surplus when facing AA type of competition. Analyzers derive the highest average surplus when facing AE type of competition where both Analyzers can exploit the single Explorer (AE is one of the least demanding competition types in terms of Bids and Spans for Analyzers). However, they are left with uncertain and lower surplus when competing with two Explorers (facing EE competition type); implied by the large difference between mean and median surplus (see column EE in Table 15) and the shape of the surplus distribution, shown in Figure 14 with an average low surplus of 1 but a fat tail (skew = 0.94, kurtosis = -0.39, standard error = 1.99). When comparing Analyzers' surplus between AA and EE competitions, the results of the t-test and WRS test are inconsistent and we can only conclude unequal surplus; the skewed surplus distributions under AA and EE competitions (Figure 14) are the reason behind these inconsistent test results. Having at least one Participator agent in the competition makes Analyzers derive medium surplus as compared to competitions without a Participator. The pattern in Analyzers' surplus implies that they can exploit both Explorers' and Participators' behavior, but have a disadvantage when facing a majority of Explorers.

Figure 14. Analyzers' surplus distribution when facing EE and AA competition



Based on our analysis of bidder behavior and economic welfare under different types of competition, we see some general patterns that hold true for all bidder types (e.g., AA competition is the toughest competition for all bidder types); however, we find strong and consistent evidence for the context-dependence of the competition impact, i.e., that different competition types manifest themselves differently depending on the focal bidder's type (i.e., bidding strategy).

5.3 Effect of competition on auction outcomes

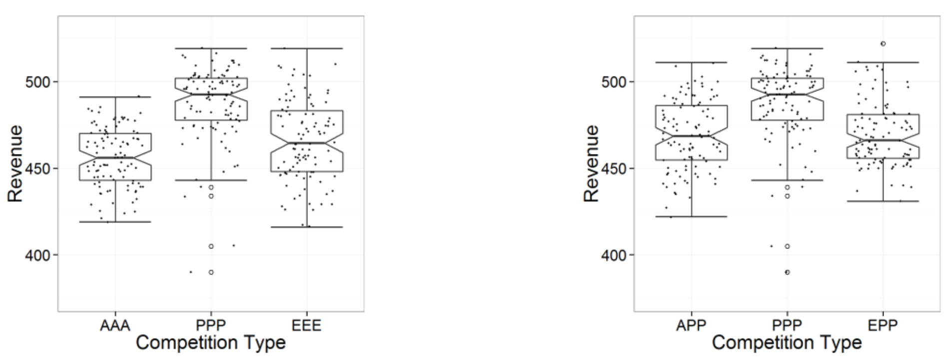
In this subsection, we study how bidding agents' emergent behaviors under different competitions lead to differences in auction revenue and allocative efficiency. Competition type of an entire auction refers to the combination of three bidder types in this subsection (different from previous subsections where competition type referred to the combination of two bidding strategies faced by a bidder). ANOVA results (Table 16) indicate that competition has a significant effect on both revenue and allocative efficiency.

Table 16. ANOVA for effect of competition types on auction outcomes

	Degrees of Freedom	Auction Revenue	Allocative Efficiency
		F	F
Competition Type	9	168.37***	51.764***
R^2		0.3363	0.1348
Adj. R^2		0.3343	0.1322
		F(9,2990) = 168.4 ***	F(9,2990) = 51.76 ***

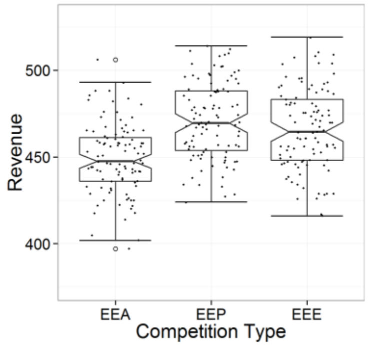
Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

Figure 15. Auction revenue under different competition types (sorted by mean)

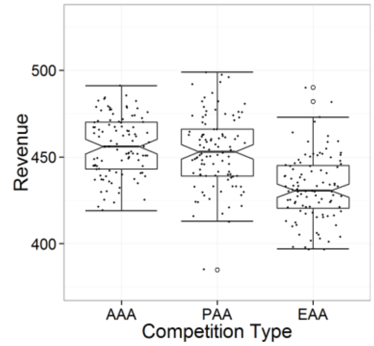


a) Homogenous Bidders

b) Participator Majority



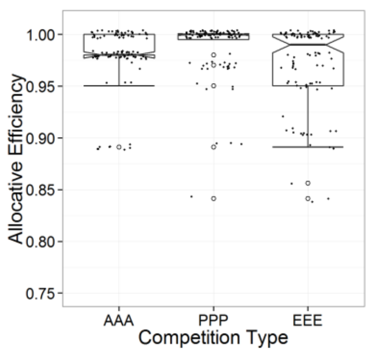
c) Explorer Majority



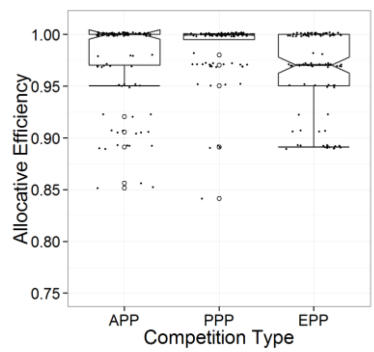
d) Analyzer Majority

To better understand the effect of competition on auction revenue we group auctions into the following groups: where all bidding agents are of the same type (Figures 15a, 16a), where the majority are Participator types (Figures 15b, 16b), where the majority are Explorer types (Figures 15c, 16c), and where the majority are Analyzer types (Figures 15d, 16d). See Table 8 for mean revenue and allocative efficiency values under different competition types.

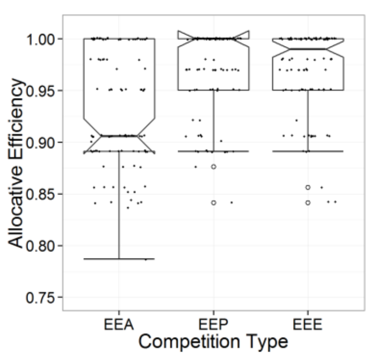
Figure 16. Allocative efficiency under different competition types (sorted by mean)



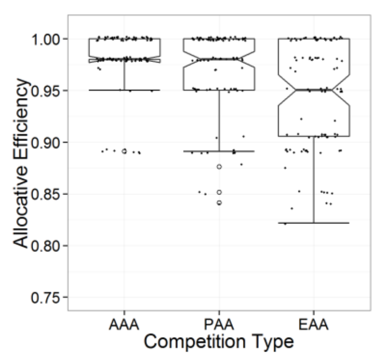
a) Homogenous Bidders



b) Participator Majority



c) Explorer Majority



d) Analyzer Majority

When all bidders are of the same type (Figures 15a, 16a), the auctioneer makes the least revenue when there are only Analyzers participating in an auction (AAA competition). Analyzers are the least likely to overbid or place uncalculated bids that would create opportunities for other bidders to exploit. The auctioneer derives the highest revenue when there are only Participators in an auction because all three Participators maintain their moderate behavior under this competition and derive a relatively medium surplus when there are neither Analyzers who could occasionally exploit them nor Explorers whose random behavior would allow Participators to derive higher surplus (see Table 14). This also leads to consistently high allocative efficiency under PPP competitions (see Figure 16a). When there are only Explorers participating in an auction, revenue is in between the revenue of auctions with PPP and AAA competition. Interestingly, Explorers derive higher average surplus than Participators in auctions with homogenous bidder types; their random overbidding yields higher surplus when there are no other bidder types that would exploit their behavior (as compared to other competitions they can face; see Table 13). However, Explorers' randomness also leads to the lowest allocative efficiency among auctions with homogeneous bidders (see Figure 16a).

In auctions where the majority of bidders are Participators (Figure 15b, 16b), auction revenue significantly decreases when we substitute a Participator with an Analyzer or Explorer. This is because the average surplus of individual bidding agents increases in both of these cases, leading to lower revenue for the auctioneer. Participators' average surplus increases when an Explorer substitutes a Participator (competition changing from PPP to EPP) while the Explorer derives significantly less surplus, leading to an overall higher average bidder surplus. When an Analyzer substitutes a Participator (competition changing from PPP to APP), Participators average surplus decreases but the Analyzer derives much higher surplus compared to Participators under the APP competition, which leads to an overall higher average bidder surplus. Auctions with APP competition have lower allocative efficiency compared to PPP auctions, due to the Analyzer's attempt to exploit other bidders. Average allocative efficiency further decreases in auctions with EPP competition as a result of Explorers' random behavior.

In auctions where the majority of bidders are Explorers (Figure 15c, 16c), substituting an Explorer with a Participator (competition changing from EEE to EEP) slightly increases auction revenue as well as average allocative efficiency. This is because Explorers' average surplus decreases significantly in the presence of a Participator (see Table 13) while the new Participator's surplus is only slightly higher than the Explorer it has replaced under the EEP competition, which leads to an overall lower average bidder surplus (i.e., higher auction revenue). When the substituted agent is an Analyzer (competition changing from EEE to EEA), Explorers' average surplus decreases and the introduced Analyzer agent derives higher surplus by exploiting the random behavior of Explorers, which leads to lower auction revenue as well as significantly lower (and highly varying) allocative efficiency.

In auctions where the majority of bidders are Analyzers (Figure 15d, 16d), substituting an Analyzer with a Participator (competition changing from AAA to PAA) slightly decreases auction revenue and does not significantly change average allocative efficiency. Analyzers derive significantly higher surplus under PAA competition compared to AAA, however the Participator agent derives a much lower average surplus compared to the Analyzer it has replaced (0.335 compared to 12.89), which results to an overall slightly lower average bidder surplus in auctions with PAA competition and, hence, a slightly higher auction revenue. When an Analyzer is replaced by an Explorer (competition changing from AAA to EAA), the two Analyzers derive an even higher surplus under the new EAA competition (compared to the previous PAA competition) leaving the Explorer with an average negative surplus (see Tables 13 and 15 for surplus comparisons), which leads to an overall higher average bidder surplus that results in lower average auction revenue. Average allocative efficiency also decreases under EAA competition, compared to auctions with AAA and PAA competition (see Figure 17b), due to Explorers acquiring suboptimal items as implied by their negative average surplus.

Figure 17. Comparing auction outcomes under different competition types

a) Revenue Comparison (Sorted by Mean) b) Allocative Efficiency Comparison (by mean)

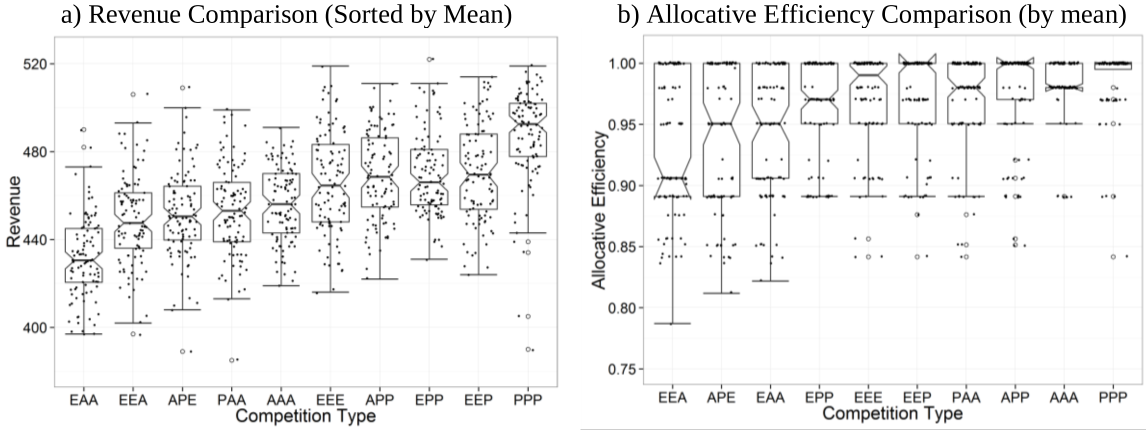


Figure 17a shows the overall trend in auction revenue across all different competitions. Table 17 shows results of pairwise comparisons to verify significant differences in auction revenue using t-tests; revenue is not normally distributed for all groups, but the sample size is large enough (100 auctions per each competition type) to use t-tests.

Table 17. Comparison of revenue across different competition types

Revenue under Competition Type																		
EAA	<***	EEA	≈	APE	≈	PAA	<*	AAA	<***	EEE	<*	EPP	≈	APP	≈	EEP	<***	PPP
		EEA		<*		PAA						EPP		≈		EEP		

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

In general, Explorer agents' random behavior provides other bidder types a better opportunity to derive higher surplus. The auctioneer derives the lowest average revenue in auctions with EAA competition, where two Analyzers derive the highest surplus by exploiting the random behavior of a single Explorer (Analyzers derive the highest surplus when facing EA competition, see Table 15). Under EEA competition type, there is only one Analyzer who exploits the randomness of Explorers, leaving the auctioneer with relatively higher revenue (as compared to auctions with EAA competition). Participators can also exploit opportunities created by Explorers (i.e., EPP and EEP competition types) but cannot exploit them as well as Analyzers do (i.e., EAA, EEA, and APE competition types). Analyzers not only exploit Explorers' random behavior but can also exploit opportunities created by Participators (i.e., APE, PAA, and APP competition types). There is no significant difference in Participators' surplus when facing EP or EE types of competition (see Table 14) which makes for a similar revenue in auctions with EPP or EEP competition.

Figure 17b shows the overall trend in allocative efficiency across all different competition types, and Table 18 shows significant differences in allocative efficiency between different competition types using t-tests. Note that allocative efficiency is not normally distributed under any of the competition types but the large sample size (100 observations) allows for valid t-tests.

Table 18. Comparing allocative efficiency across different competition types

Allocative Efficiency under Competition Type																		
EEA	<*	APE	≈	EAA	<***	EPP	≈	EEE	≈	EEP	≈	PAA	≈	APP	<+	AAA	<***	PPP
						EPP				<*				APP				

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Average allocative efficiency is always above 0.9 and close to 1 under all competition types (see Table 8 for allocative efficiency values). Given the high variation in allocative efficiency under some competition types (i.e., EEA, APE, and EAA) and the relatively small magnitude of difference in cases where there is statistical significance (e.g., allocative efficiency in EPP auctions is significantly less than allocative efficiency in APP auctions by only 0.0085, less than 1%), we group auctions that yield similar allocative efficiencies into three groups for more meaningful comparisons:

- Varying medium efficiency: auctions with EEA, EAA, or APE competitions (average allocative efficiency of 0.94 with a mean standard deviation of 0.056);

- High efficiency: auctions with EPP, AAA, PAA, EEE, APP, or EEP competitions. (average allocative efficiency of 0.97 with an average standard deviation of 0.038);
- Consistently high efficiency: auctions with PPP competition (average allocative efficiency of 0.99 with a mean standard deviation of 0.027).

The first group consists of auctions with both Explorer and Analyzer bidders, which lead to lower average allocative efficiencies that are more varying (average standard deviations of 0.056). Auctions with these competitions also result in lower auction revenue (see Figure 17a). As mentioned earlier, this is because Explorers' random behavior can generally be exploited by Analyzers, and to a lesser extent by Participators, leading to suboptimal allocation of items in these competitions. Auctions with PPP competition type, which produce the highest auction revenue, also make for consistently higher allocative efficiency, due to Participators' consistent moderate behavior in absence of other bidder types.

6. Limitations

The heuristic-based approach, presented in Part I, results in reasonably well-performing bidding agents. However, it involves many subjective modeling decisions, e.g., the choice of binwidth for the aggregation and discretization procedures and, subsequently, the choice of functions fit on the discretized data (see Section 3). More subjective modeling decisions imply that the chosen heuristics are context dependent and less generalizable. Another limitation is that the heuristic-based models do not provide insights about determinants that lead to differences in the modeled behavioral aspects (i.e., bundle choice, bid amount, and bidding frequency) between the three bidder types. For example, the heuristic-based approach might not be the best way to simultaneously capture both static and dynamic factors that lead to differences in bundle choice. In Part II we intend to address such limitations and better connect the observed behaviors (human bidders) with the modeled behaviors (bidding agents).

Part II: A Machine Learning Approach to Model Bidder Behaviors

In this part, we propose a machine learning (ML) approach to model different bidder behaviors observed in continuous CAs and implement these models in bidding agents that mimic these bidding behaviors. We validate the agents' human-like behavior in agent-based simulations by comparing auction-level and bidder-level outcomes between simulated auctions (purely using agents) and experimental auctions (with human participants). Further simulation studies allow us to explore various possible auction scenarios and different bidder type compositions (i.e., different competition types), many of which may not be easily encountered organically when running combinatorial auctions in experimental lab environments.

7. Characterizing bidder behaviors

In a typical continuous CA, at any time in the auction, a bidder can select a bundle s/he is interested in, check the DL and WL for the selected bundle, and decide to either place a bid of a certain monetary amount on the selected bundle or to not bid at that time.

Characterizing bidder behavior in terms of "what bidders bid on" (**bundle selection**), "how much they bid on a selected bundle" (**bid amount**), and "how long they wait in between subsequent bids" (**waiting time**), allows us to capture the dynamics of different bidding strategies and, when modeled correctly, to design bidding agents with human-like behavior. It is worth noting that what is presented here is the working result after experimenting with various ways to characterize bidder behavior (addressing the question of "what to model?") and different approaches to model each of these aspects (addressing the question of "how to model it?"); unsuccessful attempts such as, including "bidder effort" in modeling different behaviors, are not reported.

We formulate the bundle selection aspect as a discrete choice problem and the bid amount decision as a numeric prediction problem. The waiting time in-between subsequent bid attempts (in seconds) would be ideally modeled as numeric prediction; however, when it is not possible to derive a good model (i.e., when the models do not perform significantly better than random chance or model accuracy is not significantly larger than the no information rate), we derive appropriate probability distributions to model this behavioral aspect.

For each aspect, we use machine learning techniques to select the best parsimonious model based on the model's stand-alone predictive accuracy and its explanatory performance and, more importantly, based on how it contributes to the bidding agents' overall performance (in terms of similarity to human bidders). For example, to model the bid amount aspect, we may derive a linear regression model (where bid amount is the dependent variable), let us name it LR1, which has a high adjusted R-squared (apparently good explanatory performance) but does not perform well in the ensemble alongside other derived models for bundle selection and waiting time when implemented in the bidding agents (e.g., as indicated by the frequent erroneous bid amount outputs). An alternative model, LM2, that would perform better in the ensemble but has a slightly lower stand-alone accuracy (and/or explanatory performance) would be the superior choice. The following sections will elaborate on how each of the behavioral aspects is modeled and implemented in the agents.

All model building procedures were conducted in R using the *nnet* (https://cran.r-project.org/web/packages/nnet/) and *caret* (https://cran.r-project.org/web/packages/caret/) packages/libraries, and the *fitdisrplus* (https://cran.r-project.org/web/packages/fitdistrplus/) library was used for distribution fitting (along with any other required libraries). The derived models were then implemented in the agents by exporting all model parameters from R and coding them in ASP.net utilizing MathNet libraries (https://numerics.mathdotnet.com/) and Accord machine learning framework (http://accord-framework.net/). All models are derived from the baseline experimental data. The following candidate predictor variables are considered for model building. This information is also available to a bidder before selecting a bundle and placing a bid in the auction (i.e., agents are provided the same information as human bidders):

- Valuation: valuation for the bundle bid on;
- WL: winning level of the bundle bid on (as well as its normalized value);
- WL.norm: WL divided by valuation;
- DL: deadness level of the bundle bid on (as well as its normalized value);
- DL.norm: DL divided by valuation;
- Size: number of items in the bundle;
- ExpValue: expected surplus of the bundle bid on relative to the other 62 bundles at current auction state, min-max normalized (i.e., a value of 1 indicates the bundle has the highest expected surplus);
- Peak: whether the focal item is included in the bundle (binary variable);
- LiveBids: number of live bids before placing the bid;
- Surplus_pre: bidder's surplus right before placing the bid;
- Revenue: auction revenue right before bid was placed (represents auction state);
- Diversity: number of unique bundles that have been bid on so far (bidder's personal information based on her/his history).

To present models we use the notation used in R to define models. For example, $Y \sim X_1 + X_2$ means Y is a function of X_1 and X_2, i.e., $Y = f(X_1, X_2)$; $Y \sim X_1 * X_2$ means Y is a function of X_1, X_2 and the interaction of X_1 and X_2, i.e., $Y = f(X_1, X_2, X_1{:}X_2)$.

7.1 Bundle selection

Given the same valuation scheme and identical focal items, different bidder types vary in their propensity to bid on different bundles. Bundle selection can be formulated as a discrete choice problem, where the bidder can choose one of the 63 possible bundles upon any bid attempt. Given the nature of this problem, we derive a multinomial logit model to predict the likelihood of selecting any of the 63 bundles at any time during the auction. The most appropriate model is built by selecting predictor variables from a set of candidate predictors (i.e. features), considering both explanatory and predictive performance. Models are trained and evaluated using the processed bid-level baseline experimental data, where each observation is a bid (together with bidder and auction related variables).

We use both stepwise forward selection and recursive feature elimination to select the best parsimonious model. Predictive performance in terms of balanced accuracy is the main model selection criteria in either approach but explanatory power in terms of AIC and/or residual deviance is also evaluated (the model with the best explanatory power does not always provide superior predictive performance). For example, if there was a model with high predictive accuracy but very high residual deviance it will not be chosen.

Predictive power is assessed in terms of balanced accuracy (i.e., average accuracy across all outcome classes) since the prediction for all classes (i.e., all bundles) is important. We use 5-fold (or 10-fold) repeated cross-validation (CV) with 20 repeats for the model building procedure. To consistently compare different models (models with different input features), the random seed for the CV procedure is set at the same constant value when training models for each of the three bidder types. We also check models' predictive accuracy relative to the "no information rate" (NIR) so that the calculated model accuracy is indeed significantly better than random chance. If we take all known classes in the testing sample and randomly guess which records belong to which class, the NIR is the proportion of records we would guess correctly just because of luck. This is important because we have class imbalance for the outcome variable, e.g., most observations could be on bundles number 24 and 16, with few or single observations on the other bundles. A one-sided binomial test is used (as part of the model checking procedure using repeated k-fold CV in the caret package) to see if the derived model accuracy is significantly better than the "no information rate" (NIR), which is taken to be the largest class percentage in the data. A low p-value for this test (i.e., Accuracy > NIR) indicates that the model's predictive accuracy is significantly better than random chance (i.e., assigning any new input to the majority class). The final multinomial logit models for each of the three bidder types (A, P, and E) have the following formula:

- For Analyzer and Explorer types: $Bundle \sim valuation + Size + ExpValue$
- For Participator types: $Bundle \sim valuation + Size$

The model fitting procedure uses a single-hidden-layer feed forward neural network with the number of outputs equal to the number of classes and a softmax output stage (optimization is done using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm when using the *multinom* function); for details see *multinom* documentation https://cran.r-project.org/web/packages/nnet/.

Valuation and *Size* are both static variables as these values do not change for a bundle throughout the auction. *ExpValue* of a bundle is a dynamic variable since it changes as the auction proceeds. The derived bundle selection models for Analyzers and Explorers have both static and dynamic elements. For Participators, the *ExpValue* variable does not significantly improve model accuracy (or its explanatory power). This was also further tested, without improvement, by deriving and implementing both models for the Participator type agents and comparing their performance in simulations (as we will explain in later sections). More complex models that would include, besides the above variables, some of the reasonable interaction terms were tested as well but interestingly did not provide a significant improvement over the above-selected models. When implemented in the agents, the model determines the likelihood of selecting each of the bundles at any time during the auction. This basically allows creating the probability mass function (PMF) for a discrete "bundles" variable with 63 possible values.

7.1.1 Dealing with unobserved outcomes

In the baseline experimental data, there were observations on only 55 out of the 63 possible bundles (i.e., the 63 possible classes for the outcome). However, we still want to allow the possibility of selecting those unobserved bundles when the models are implemented in the agents, even if it very unlikely. At run-time, the likelihood for these unobserved bundles' is set equal to the least probable bundle, times a constant (after the likelihood for all observed bundles are derived using the multinomial logit models). The value of this constant is one of the tuning parameters (it is set to 0.25 for the agent-based simulations).

7.1.2 Bidding history

Bidders (may) also consider their bidding history when placing new bids and, depending on the auction state, decide to select a bundle they have bid on in the past or select a new bundle. We code this by defining a binary variable "*New*" for every observation (i.e., bid) in the baseline experimental data, which takes the value 0 if the bundle has been bid on before by that bidder and

1 otherwise. We then model the likelihood of selecting a new bundle using a logistic regression, where "*New*" is the output variable. The procedure for selecting the best model (i.e., best input features) uses stepwise forward selection. The potential features are the same above introduced variables. Models are assessed in terms of balanced accuracy as well as precision and recall for the 1 class. The following is the final model used by the agents:

$$New \sim (Revenue + Diversity) * BidderType$$

In this model, "Bidder Type" is included as a variable (a categorical variable that takes the 3 values A, P, or E) instead of creating separate models for each of the three bidder types. Building separate models for each of the three bidder types using the same forward selection procedure resulted in models with the same features, e.g., for Participator types the resulting model was $New \sim Revenue + Diversity$. Other, more complex models were tested but did not significantly improve balanced accuracy.

7.1.3 Implementation

The two above models are combined to determine agent's bundle selection decision and, once implemented in the agents, the bundle selection decision works as follows:

- Likelihood of bidding on a new bundle is calculated using the logistic regression model; likelihood of selecting a new bundle = N.
- Probability b_i of selecting each bundle i ($i \in [1, 63]$) is calculated using the multinomial logit model presented above.
- For unobserved bundles j in the training data:

$$b_j = \min\{\frac{s \times cumulative\ prob(b_i\ for\ all\ i)}{63-(no\ \ of\ observed\ bundles\ in\ training\ data)}, \frac{\min(b_i)}{d}\}$$

    - ***s*** represents what percentage of the cumulative probability of the observed bundles we want to assign to unobserved bundles; s = 0.1 after tuning agents.
    - ***d*** specifies what fraction of the least probable observed bundle we want to assign to unobserved bundles; d = 4 after tuning agents.
- Final bundle selection likelihood for any bundle i (or j) is:
    - $p_i = b_i \times N$ if the bundle has not been bid on before (is a new bundle), and
    - $p_i = b_i \times (1 - N)$ otherwise
- Create a cumulative density function (CDF) for all 63 bundles using p_i values.
- Select corresponding bundle based on a random pick from this CDF.

7.2 Bid amount

The amount different bidder types place on selected bundles varies as the auction proceeds. Initially, we tried to formulate it as a numeric prediction problem and modeled it using a linear regression model that includes time-varying variables (such as WL and DL). The resulting models did not perform well in the agents after implementation. Upon further examination, it turns out that the WL and DL for a selected bundle are two important decision anchors (recall that this information is provided as feedback in the auction environment), and the bid-amount decision is made relative to the WL and DL of the bundle. Therefore, we formulate the bid-amount decision as following: a bidder first decides whether to bid above or below the WL, since this information is more consequential than the DL (i.e., whether to place a winning bid). If a winning bid is placed, an amount above the WL is determined (we will elaborate on this later). Otherwise, the bidder places a live bid below the WL or there is a chance that a dead bid (i.e., below the DL) is placed. If a live bid is placed, a bid amount relative to (and in between) the DL and WL is determined (we will elaborate on this). We model this decision as a 2-level nested logit, where the first logit model determines the likelihood of bidding above the WL (placing a winning bid), $\boldsymbol{P_W}$, and the second logit (nested within the first) determines the likelihood of bidding above the DL (placing a live bid), $\boldsymbol{P_L}$. Based on this model, at any time in the auction, the likelihood of placing a winning bid is P_W, the likelihood of placing a live bid is $(1 - P_W) \times P_L$, and the likelihood of placing a dead bid is $(1 - P_W) \times (1 - P_L)$.

The two logit models that determine P_W and P_L are derived using the stepwise forward selection procedure where model performance is assessed in terms of balanced accuracy, and precision and recall for the 1 class. Following are the final logit models for each bidder type to determine P_W and P_L.

- Analyzer type:
  - $Win \sim DL.norm + WL.norm + Size + ExpValue$
  - $Live \sim DL + Revenue$
- Participator type:
  - $Win \sim WL + WL.norm + valuation + ExpValue + LiveBids$
  - $Live \sim WL + DL.norm + valuation + Peak + ExpVlaue + LiveBids$
- Explorer type:
  - $Win \sim DL.norm + WL.norm + Peak + ExpValue + Surplus\_pre$
  - $Live \sim DL + WL + Size + ExpValue + Surplus\_pre + Revenue * (WL.norm +$

$LiveBids + Diversity)$

Win and *Live* are binary variables that are added to the bid-level baseline experimental data to derive the above models. The value of *Win* equals 1 if the bid is winning and 0 otherwise; the value of *Live* equals 1 if the bid is live (but not winning) and 0 otherwise; both variables equal 0 for a dead bid.

The bid-amount in each of the three states (i.e., winning, live, and dead) is then modeled separately depending on the bid state. For *winning* bids, we derive separate linear regression models for each bidder type. The best model is derived using a stepwise forward selection procedure where model performance is assessed in term of RMSE (R-squared is also checked to ensure the model's explanatory performance). The selected linear regression models for each of the bidder types are as follows:

- Analyzer type: *BidAmount* ~ *DL* + *WL* + *valuation* + *Size*
- Participator type: *BidAmount* ~ *WL* + *valuation* + *LiveBids* + *Diversity*
- Explorer type: *BidAmount* ~ *WL* + *DL.norm* + *Peak* + *LiveBids*

For *live* bids, where bid-amount is between the DL and WL (and DL < WL), we derive the bid-amount's relative position to the DL and WL as $(BidAmount - DL)/(WL - DL)$, which is always a ratio between 0 and 1, and fit a beta distribution on this ratio. The resulting distributions are: Analyzer ~ Beta(0.3124502, 0.1900098), Participator ~ Beta(0.2609996, 0.2572557), Explorer ~ Beta(0.3095985, 0.4367436); Beta(α, β), see https://en.wikipedia.org/wiki/Beta_distribution/ for details. A possible alternative to modeling this ratio using a beta regression was not successful; none of the beta regression models have significant explanatory or predictive performance. For *dead* bids, bid-amount is set equal to a random value slightly less than the DL (e.g., DL – RandomUniform(1,5)), since it is practically irrelevant how much below the DL the bid-amount is; dead bids do not change auction state and are not shown to bidders in the combinatorial auction platform.

7.2.1 Endgame behavior

Towards the end of the auction, e.g. when auction revenue has gone above $300, the above-defined linear regression models to determine winning bid amounts do not perform well upon implementation. A probable explanation is that the endgame behavior, for which we have less data, is not well captured with the linear models. We use the auctions' average closing revenue in the baseline experimental data ($330) as the threshold for switching to endgame behavior. After experimenting with several choices, we use triangular distributions that are fitted on the endgame partition of the baseline data (i.e., bids where the revenue variable is equal or above $330) to determine how much above the WL each of the bidder types bid. The resulting distributions are: Analyzer ~ Tri(1, 4, 1), Participator ~ Tri(1, 8, 1), Explorer ~ Tri(1, 19, 1); Triangular(*lower limit, upper limit, mode*), see https://en.wikipedia.org/wiki/Triangular_distribution/ for details. Upon implementation, when auction revenue exceeds the endgame threshold (e.g., $330) agents would use these triangular distributions instead of the linear regression models to determine winning bid amounts. Note that the amount of live or dead bids, where we did not use models, is still determined as explained in the previous sub-section.

Figure 18 shows the pseudo code for how agents determine their bid amount for a selected bundle once the above-derived models are implemented.

Figure 18. Bid Amount determination procedure

```
Determine P_W and P_L using logit models
If RandomUniform(0,1) > P_W
        place winning[1] bid
Else
        If RandomUniform(0,1) > P_L
                If  WL = DL  //Live bids do not exist under this condition
                        place winning[1] bid
                Else
                        place live[2] bid
        Else
                place dead[3] bid

[1]Winning:
        If Revenue > endgame_threshold  //$330 used
                bid-amount = WL + Random pick from triangular distribution
        Else
                bid-amount determined using linear regression model

[2]Live: bid-amount = Random pick from beta distribution × (WL – DL) + DL
[3]Dead: bid-amount = DL – Random pick from Uniform(1,5)
```

7.3 Waiting time (when to bid)

Ideally, we would like to derive a model to predict “when a bidder bids” during the auction. The idea was to derive a logit model that would determine the likelihood of bidding vs. other types of possible actions on the platform (i.e., refresh, select/deselect items), given some auction state input variables. Each observation in the clickstream level data indicates whether the bidder refreshed the page, selected/deselected an item, or placed a bid, however it does not tell us anything about what people think or actions they intend but do not perform (which could play a role in the timing decision). None of the derived logit models performed significantly better than random chance in predicting “whether to bid at the current state”; all logit models were trained on the clickstream data but included different set of features (the outcome variable in the training data is 1 if the type of action is “placing a bid” and 0 otherwise).

As the next best alternative, we defined waiting time as the period of time (in seconds) in between consequent bids (we added a waiting time variable to the baseline experimental data) and fitted a distribution on waiting time for each of the three bidder types. The best fit for all three types turned out to be a gamma distribution with different shape and rate parameters; Analyzer ~ Gamma(1.377, 0.0181), Participator ~ Gamma(1.19, 0.0185), Explorer ~ Gamma(1.357, 0.0373); Gamma(*shape*, *rate*), see https://en.wikipedia.org/wiki/Gamma_distribution/ for details. The derived gamma distributions for waiting time are statistically different between the three bidder types, as verified using the KS test.

Once implemented, each agent waits for x seconds after placing a bid before re-evaluating and deciding to place the next bid, where x is a random pick from the waiting time distribution (derived as above). For tuning purposes and to allow agents to be faster (which contributes to shorter overall auction durations), x is multiplied by a constant value, alpha, between 0.1 and 1. Alpha is set to 0.5 after tuning the agents, which results in an average auction duration of about 20 minutes (average auction duration is close to 40 minutes in the baseline experimental data). This feature facilitates multiple shorter auctions while maintaining agent’s general bidding strategy, which will be important when agents are used in experiments with human participants (Part III of the thesis).

7.4 Agent-based auction simulations

Each of the three main aspects of bidder behavior (bundle selection, bid amount, and waiting time) are implemented in our agents as separate modules. A fourth module determines the agent’s final decision on whether or not to bid, after combining values returned by the three above modules and considering the agent’s bidding history to avoid overbidding itself on the exact same bundle. Figure 19 provides an overview of the auction simulation with three bidding agents and the agent’s decision-making procedure.

An auction is simulated as a sequence of virtual rounds (the While loop in Figure 19). In each round, all participating bidding agents are called in random order (lines 3-6 in Figure 19), and each agent decides to either place a bid at the given auction state or not to bid. The bidding decision is determined by the agent’s three internal modules and considering the bids it has placed so far in the auction. Each bidding agent decides which bundle and how much it wants to bid (lines 10-11 in Figure 19), using its bundle selection and bid amount models. If the agent is not underbidding itself and if the determined bid amount (plus a margin) is not higher than the agent's valuation, it places the bid and then waits for *x* seconds as determined by its waiting time model (lines 12-14 in Figure 19). The margin parameter is modeled differently for different bidder types representing an inherent level of aggressiveness observed in the baseline experimental data. When the inactivity time reaches a specified threshold the auction stops (line 8-9 in Figure 19); this threshold, models the so-called "soft stopping" rule and is set to 1 minute which is the same value used in the baseline experimental data.

Figure 19. Auction simulation pseudocode (ML-based agents)

```
Auction simulation:
1  While (1)
2        Shuffle {                          // execute the following statements in random order
3              bA = pingAgent1(state)       // checks whether bidding agent wants to place a bid at
current state
4              bB = pingAgent2(state)
5              bC = pingAgent3(state)       }
6        If (bA ≠ NULL OR bB ≠ NULL OR bC ≠ NULL)       // if any of the agents want to bid
7              Update state                 // update auction state with newly submitted bids
8        If (Inactivity Time ≥ threshold)   // if there are no Live or Wining bids placed for 1 minute
9              Break                        // end auction simulation

Bidding agent:
10       Select bundle                                               // using Bundle Selection model
11       Determine bid amount for bundle given current state         // using Bid Amount model
12       If (NOT underbidding itself AND bid-amount + margin ≤ valuation)
13             Return bid(bundle, bid amount)
14             Wait x seconds // x determined using waiting time model
15       Else
16             Return NULL      // not bidding
```

8. Validating bidding agents

Once the agents are implemented and run correctly without errors/bugs, we need to verify that they replicate the behavior of human bidders observed in experimental CCAs. In designing the bidding agents, we proposed that three main aspects (namely, bundle selection, bid amount, and waiting time) characterize different bidding behaviors (Sections 7.1 – 7.3). In this section, we verify whether our bidding agents correctly replicate (human) bidder-specific variables via their emergent behaviors under similar conditions (i.e., competition types). We assess the validity of our agent-based simulations in terms of matched bidder-specific variables (Bids, Spans, and Surplus) by statistically comparing outcome variables generated by our simulation model with outcome variables from experimental CCAs (for competitions observed in experimental CCAs). The level of assessment for different matched variables varies from pattern validity to distributional validity, as we will elaborate in the following subsections.

We run auction simulations with the same types of competition observed in experimental auctions, namely APP, EPP, and PPP. The number of auction simulations for each of the competition types is proportional to their occurrence in experimental auctions. Table 19 shows summary statistics of bidder-specific variables per each competition type for auction simulations and baseline experimental data. Figure 20 shows the side-by-side comparison of Bids (total number of bids) and Spans (diversity of bundles) variables between human bidders and bidding agents. The horizontal axis indicates the human bidder or bidding agent type (e.g., "A Agent" stands for Analyzer type bidding agents).

Table 19. Summary statistics of bidder-specific variables for human bidders and ML-based bidding agents (standard deviations in *italics*)

	Human Bidders in Baseline Experimental Data					Bidding Agents in Simulations				
Competition Type	Bidder Type	# of Auctions	Bids	Span	Surplus	Agent Type	# of Auctions	Bids	Span	Surplus
APP	P	4	23.62	10.38	74.12	P	24	22.29	11.9	15.62
			11.07	*5.04*	*60.42*			*8.17*	*2.68*	*22.89*
APP	A		15.5	8	77.5	A		18.42	8.96	67.18
			6.61	*3.56*	*50.08*			*6.31*	*1.99*	*37.81*
EPP	P	5 (+1 EEP)	28.82	9.91	63.91	P	30 (+6 EEP)	24.86	12.03	18.12
			8.94	*2.91*	*31.98*			*6.92*	*2.96*	*27.48*
EPP	E		52.86	21.57	31.79	E		43.79	20.14	16.3
			16.88	*4.89*	*28.45*			*9.84*	*3.41*	*23.26*
PPP	P	5	18.53	9.6	26.38	P	20	19.87	10.83	32.45
			4.19	*3.58*	*34.77*			*6.25*	*2.97*	*38.92*
All auctions pooled	P	15	23.06	9.88	49.76	P	80	22.43	11.58	22.37
			8.84	*3.67*	*45.34*			*7.35*	*2.93*	*31.58*

The Bids and Spans generated by all three bidding agent types are comparable to those generated by human bidders. Mann-Whitney tests to compare median Bids and Spans (and t-tests for mean comparisons) between bidding agents and human bidders of each type would result in no statistically significant difference when the boxplot notches overlap. This is the case for all 5 comparisons except for the comparison of Spans between Participator type bidding agents, which generate slightly higher Spans than Participator type human bidders. But the desired relative difference between the three bidding agent types holds; i.e., Analyzer < Participator < Explorer in terms of Spans (as well as Bids).

Figure 20. Comparing Bids and Spans variables between human bidders and bidding agents

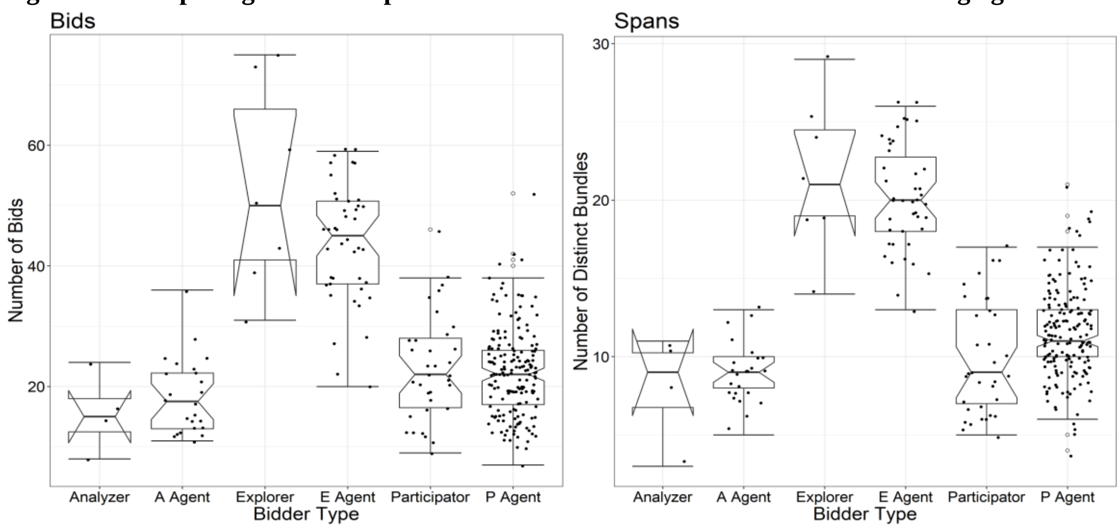
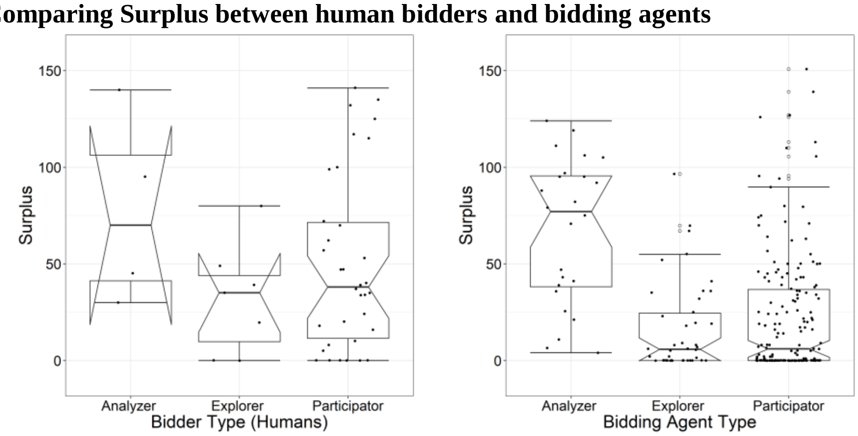


Figure 21 compares surplus between human bidders and bidding agents. The relative difference in surplus among the 3 types is similar between bidding agents and human bidders (i.e., pattern validity). We run 6 times as many auction simulations as there are experimental auctions in the baseline data, which makes for a higher total number of bidding agents who make 0 surplus since one or two (out of the three) bidders in each auction may not win any items. For the type of competitions we simulate (APP, EPP, and PPP), usually, either Explorer or Participator type agents do not win any items at the end of the auction. This leads to an overall higher frequency of Participator and Explorer type agents with 0 surplus and a lower average surplus for these agent types compared to the same type of human bidders in the baseline experimental data.

Figure 21. Comparing Surplus between human bidders and bidding agents



Overall, our bidding agents' behavior in terms of Bids, Spans, and Surplus replicates the three canonical bidding strategies observed in experimental combinatorial auctions.

9. Comparing ML-based and heuristic-based agents

The ML-based approach involves less subjective modeling decisions compared to the heuristic-based approach. For example, there is no aggregation or discretization that would depend on the binwidth choice (Section 3). Secondly, the ML-based agents' behavior better connects with the observed human behaviors. I.e., instead of the bidding frequency aspect (in the heuristic-based approach) we model *waiting time* (in the ML-based approach), which allows for more natural auction simulations that do not require the introduction of additional modeling constructs such as "non-activity count". Moreover, the ML-based models for bundle selection (Section 7.1) and bid amount (Section 7.2) better connect with how human bidders behave in the auction environment and provide insights about static and dynamic factors that determine differences in these two behavioral aspects between the three bidder types.

Figure 22. Comparing Bids between heuristic-based and ML-based agents

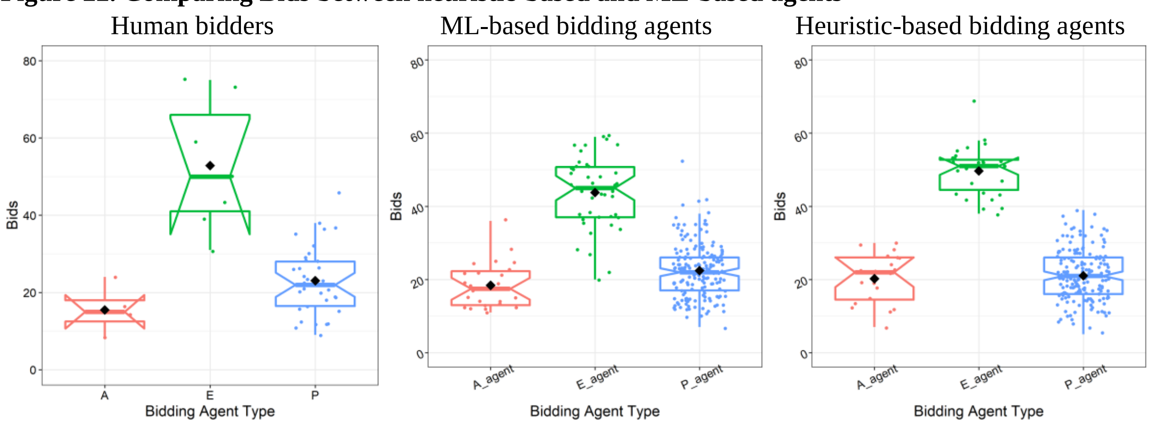


Figure 23. Comparing Spans between heuristic-based and ML-based agents

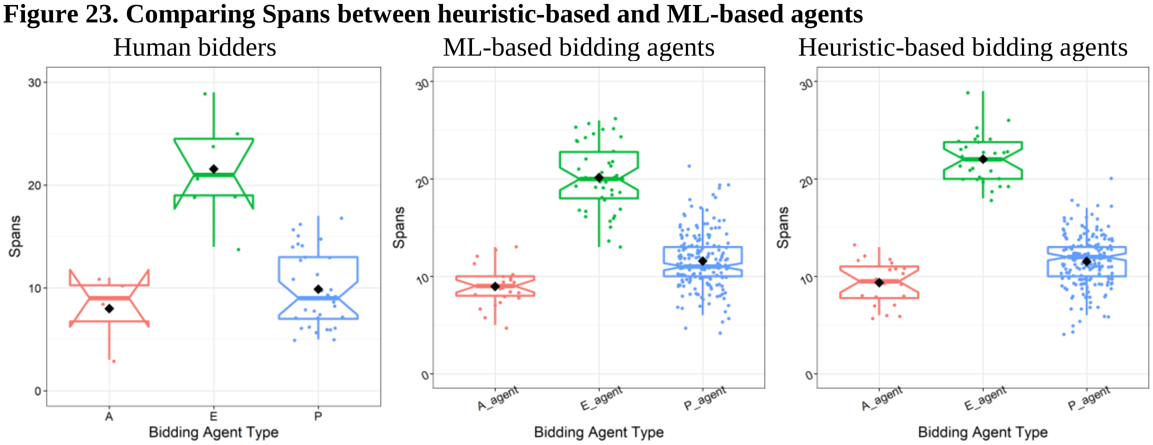
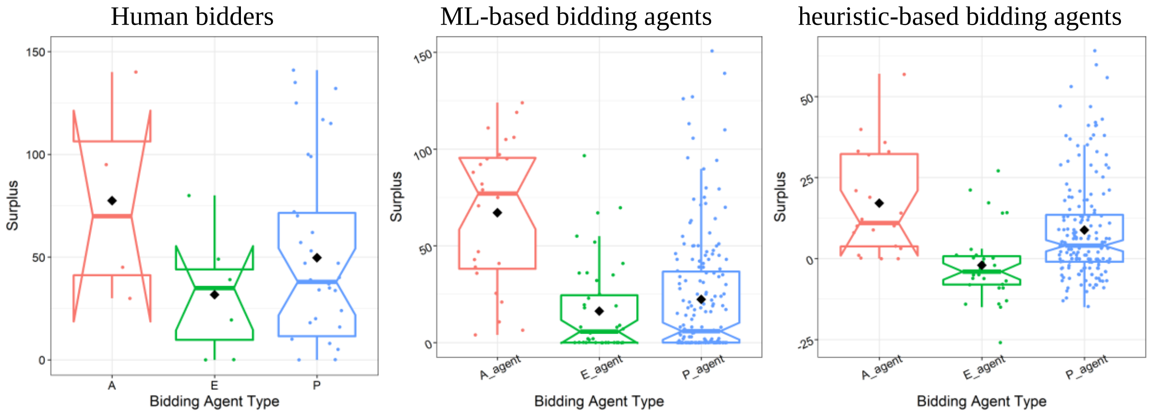
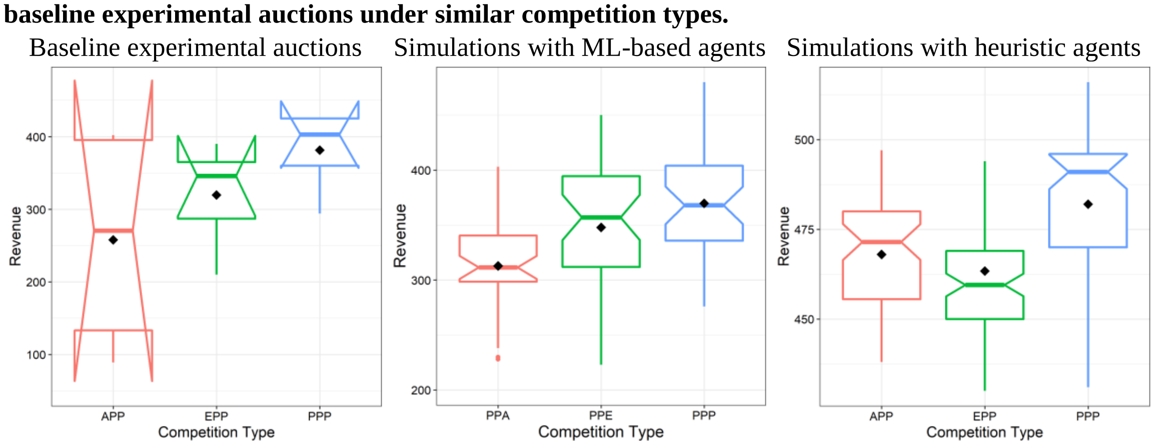


Figure 24. Comparing Surplus between heuristic-based and ML-based agents



In terms of bidder-specific outcomes both ML-based and heuristic based agents perform similarly in terms of Bids and Spans variables, as shown in Figures 22 and 23. However the ML-based agents perform much better in term of Surplus compared to the heuristic-based agents as seen in Figure 24; i.e., the pattern of surplus between the three bidder types (Analyzer, Participator, and Explore) is similar in all three plots (recall pattern-validity as discussed in the validation sections), however, the scale on the vertical axis is the same for ML-based agents and human bidders but much smaller for the heuristic-based agents.

Figure 25. Comparing auction revenue between auction simulations with ML-based agents and baseline experimental auctions under similar competition types.



In terms of auction revenue, auction simulations using the ML-based agents produce better results (compared to auction simulations using the heuristic-based agents) as seen in Figure 25; the patterns in revenue across competition types observed in the baseline data are better replicated in auction simulations using the ML-based agents, both in terms of patterns and values (note the different scaled on the vertical axis).

In summary, ML-based agents perform better in terms of bidder-specific and auction outcomes and better connect with and represent the observed human bidding strategies.

10.Leveraging ML-based agents

Agent-based auction simulations allow us to explore the effect of competition on the dynamics of bidder behavior and auction outcomes. Having validated the agents and knowing that they reasonably replicate human bidding strategies, we are interested in finding out whether the bidding dynamics that result from different competitions, while the bidder agents bid based on the coded canonical behaviors, provide interesting insights into potential auction outcomes through "emergent" behaviors of the agents under different competitive environments. We run 40 auction simulations (with 3 bidders competing to acquire 6 items in each auction) for each of the 10 possible competition types (total of 400 auction simulations) in this part of the study using the same valuation setup used in experimental CCAs. Table 20 shows mean and standard deviation values for auction specific and bidder-specific variables grouped by bidder type and competition type. This dataset is used for the analysis in the forthcoming sub-sections.

Table 20. Summary statistics for auction simulations grouped by competition types

Competition Type	Auction Specific Variables		Bidder Specific Variables				
	Revenue	Allocative Efficiency	Agent Type	n	Bids	Spans	Surplus
AAA	339.40 *(66.12)*	0.87 (0.11)	Analyzer	120	13.72 *(7.30)*	7.31 *(2.14)*	33.27 *(37.69)*
APE	292.57 *(58.39)*	0.85 (0.09)	Analyzer	40	11.93 *(4.84)*	6.45 *(1.71)*	60.12 *(50.26)*
			Explorer	40	25.52 *(8.60)*	16.68 *(3.42)*	45.48 *(47.88)*
			Participator	40	14.12 *(5.22)*	9.60 *(2.49)*	29.17 *(40.48)*
APP	312.95 *(37.55)*	0.87 (0.08)	Analyzer	40	10.68 *(3.74)*	6.53 *(1.62)*	65.02 *(42.86)*
			Participator	80	13.80 *(3.95)*	9.31 *(2.53)*	30.42 *(39.64)*
EAA	318.25 *(61.71)*	0.86 (0.08)	Analyzer	80	12.78 *(6.03)*	6.97 *(1.83)*	45.14 *(43.42)*
			Explorer	40	27.32 *(8.71)*	16.88 *(3.16)*	26.99 *(40.09)*
EEA	322.55 *(50.40)*	0.87 (0.06)	Analyzer	40	12.32 *(4.13)*	7.28 *(1.91)*	60.27 *(45.64)*
			Explorer	80	26.98 *(7.00)*	15.99 *(3.18)*	29.17 *(37.98)*

EEE	429.20 *(30.16)*	0.91 (0.05)	Explorer	120	43.24 *(9.35)*	19.41 *(3.70)*	10.21 *(17.68)*
EEP	412.10 *(40.44)*	0.90 (0.04)	Explorer	80	42.84 *(8.03)*	19.44 *(3.33)*	14.29 *(20.42)*
			Participator	40	24.07 *(6.86)*	11.82 *(2.95)*	15.17 *(29.97)*
EPP	347.95 *(61.33)*	0.84 (0.10)	Explorer	40	32.90 *(10.13)*	17.40 *(3.93)*	33.67 *(37.05)*
			Participator	80	19.35 *(8.45)*	10.60 *(3.25)*	20.63 *(34.62)*
PAA	344.65 *(47.98)*	0.88 (0.07)	Analyzer	80	15.64 *(7.28)*	7.59 *(1.99)*	36.40 *(41.04)*
			Participator	40	18.50 *(6.42)*	12.03 *(3.05)*	25.43 *(40.64)*
PPP	369.82 *(50.40)*	0.89 (0.05)	Participator	120	25.25 *(8.23)*	12.34 *(3.0)*	25.96 *(35.46)*

Standard deviations in parentheses

10.1 Competition and bidder behavior

Each bidder can face 6 types of competition based on the composition of bidding strategies it encounters in the auction, namely: AA (i.e., competing against two Analyzers), AP (i.e., competing against an Analyzer and a Participator), AE, PP, EP, and EE. Even if we assume that bidder's behavior is endowed, we can clearly hypothesize that the type of competition faced by a bidder affects his/her behavior, including both the number of placed bids (Bids) and the diversity of bundles bid on (Spans). The comprehensive simulations will allow us to further characterize these 6 competition types.

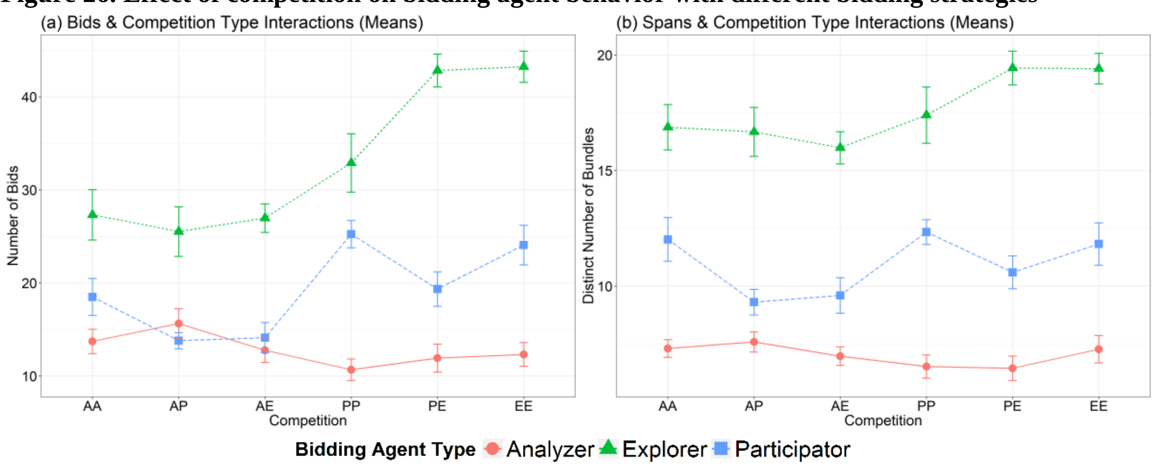
Table 21. ANOVA of factors influencing bidding agents' behavior and surplus

	Degrees of Freedom	Bids	Spans	Surplus
		F	F	F
Bidder Type	2	959.3***	1496.57***	43.94***
Competition Type Bidder Faces	5	59.72***	17.12***	4.59***
Bidder Type × Competition Type	10	32.112***	10.223***	6.4***
R^2		0.6823	0.7291	0.1289
Adj. R^2		0.6777	0.7252	0.1163
		$F(17,1182)$ = 149.3 ***	$F(17,1182)$ = 187.1 ***	$F(17,1182)$ = 10.28 ***

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

The ANOVA results in Table 21 show that bidders' own type, the competition type they face, and the possible interaction of these two factors affect key bidder-specific outcome variables (Bids and Spans). All three factors are highly significant (p-value ≤ 0.001) and both models have a high explanation of variance. This tells us that bidding agents' own strategy and the competition type they face both significantly affect the number of bids and the diversity of bundles they bid on. The significant interactions indicate that the effect of competition type on agents' behavior depends upon the agent's bidding strategy. We use interaction plots to uncover the patterns of these interactions in Figure 26. Each point shows a group's mean and the error bar indicates 95% confidence intervals (1.96 × standard errors); each group consists of a certain agent type facing a specific competition, e.g., Analyzers facing AP competition. These plots also provide a guideline for further post-hoc tests to verify significant differences. The distribution of Bids and Spans variables do not violate normality assumptions for most groups.

Figure 26. Effect of competition on bidding agent behavior with different bidding strategies

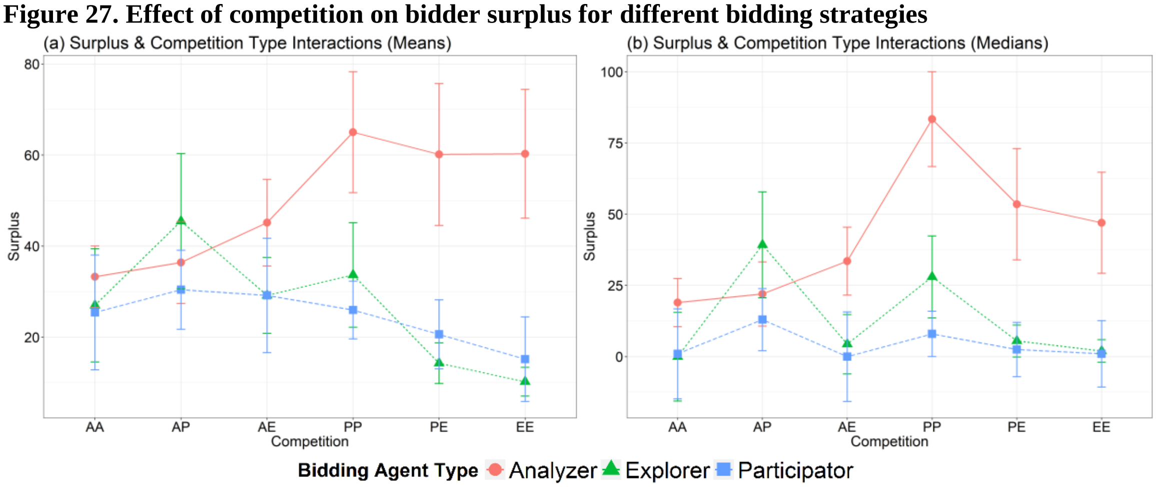


Under most competition type, differences between the three bidding agent types (Participator, Analyzer, and Explorer) are consistent with the differences among the three bidding strategies observed in experimental CCAs; i.e., Analyzers place the least number of bids and focus on a narrower set of (lower lines with circles in Figure 26); Explorers place the most bids, bid on a broad variety of bundles (upper lines with triangles in Figure 26); and Participators' bidding behavior is the middle ground between Analyzers and Explorers in terms of both number of overall bids and the variety of bundles they bid on (middle lines with squares in Figure 26).

10.2 Competition and bidders' economic welfare

We are also interested in the effect of competition on bidders' economic welfare, as measured by the surplus of winning bidders after the auction ends. A bidder's surplus is her/his valuation for the won bundles minus the amount s/he has to pay for them. We use ANOVA to analyze how bidding agents' own type, the competition type they face, and the possible interaction of these two factors, affect the surplus retained by bidders. Our ANOVA results (last column for Surplus in Table 21) show that both factors, as well as their interaction, are highly significant (p-value ≤ 0.001), with a reasonable explanation of variance (adjusted $R^2 > 18\%$). We use interaction plots to uncover the patterns of these interactions in Figure 27. Since the distribution of surplus is positively skewed for some groups, we draw interaction plots for both group means and group medians to take into account possible differences between mean and median patterns. The error bars in Figure 27 indicate roughly 95% confidence intervals (equal to 1.96 × standard error); median standard errors equal regular standard errors multiplied by 1.25 (1.2533 × standard error). Differences between the mean and median plots indicate under which competition types we have non-normal surplus distributions that we take into account for pairwise comparisons. Under most competition types, Analyzers generally derive the highest surplus compared to Explorers and Participators, who derive approximately the same surplus under most competitions.

Figure 27. Effect of competition on bidder surplus for different bidding strategies



Analyzers derive a higher surplus when there are no other Analyzers in the auction (facing PP, PE, and EE competitions), even though their number of bids (Bids) decreases only slightly and the variety of bundles (Spans) does not significantly change (as compared to competitions with Analyzer types present). There is also a slight increase in Analyzers' surplus when the competition they face changes from AA to AP to AE. These patterns indicate that Analyzers can exploit both Participator and Explorer type bidding strategies without placing more bids on more diverse bundles. The difference in **Participator**s' surplus under different competitions is not significant. However, they place more bids (higher Bids) on more diverse bundles (higher Spans) when facing homogenous AA, PP, and EE competitions, or when they face competitions without Analyzers (i.e., PP). **Explorers** derive significantly higher surplus when facing AP and PP competitions. The number of bids they place and the variety of bundles they bid on increases significantly in the absence of Analyzers and is highest when facing PE and EE competitions where there are other Explorers in the auction.

10.3 Competition and auction outcomes

In this subsection, we study how bidding agents' emergent behaviors under different competitions lead to differences in auction revenue and allocative efficiency. In this subsection, competition type of an entire auction refers to the combination of three bidder types (different from previous subsections where competition type referred to the combination of two bidding strategies faced by a bidder). ANOVA results (Table 22) indicate that competition has a significant effect on both revenue and allocative efficiency. Figures 26a and 26b show the overall trend in auction revenue and allocative efficiency across all different competitions.

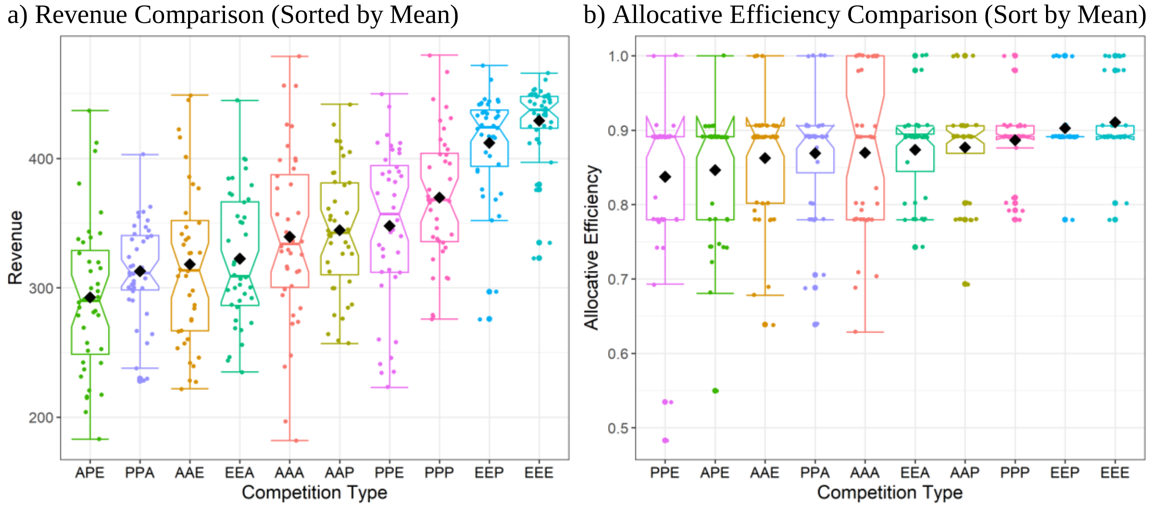
Table 22. ANOVA for effect of competition types on auction outcomes

	Degrees of Freedom	Auction Revenue	Allocative Efficiency
		F	F
Competition Type	9	27.98 ***	3.584 ***
R^2		0.3923	0.076
Adj. R^2		0.3783	0.055
		F(9,390) = 27.98 ***	F(9,390) = 3.584 ***

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

Figure 28. Comparing auction outcomes under different competition types

a) Revenue Comparison (Sorted by Mean) b) Allocative Efficiency Comparison (Sort by Mean)



Under homogenous competition types (AAA, PPP, and EEE competition types in Figure 28a), the auctioneer makes the least average revenue when there are only Analyzers participating in the auction, however this type of competition leads to the highest variance in revenue (in comparison to not only the other two homogenous types of competition but also to any of the 7 other types). The auctioneer derives significantly higher revenue under the PPP competition type and the highest revenue with the lowest variance under EEE competition (see Table 23 comparisons using t-tests).

Table 23. Comparing revenue between different competition types using t-tests

	Revenue between different competition types				
	Heterogenous				Homogenous
Analyzer majority	AA**E**	<*	AA**P**	≈	AAA
					<*
Participator majority	PP**A**	<**	PP**E**	<*	PPP
					<***
Explorer majority	EE**A**	<***	EE**P**	<*	EEE

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1, ≈ no significant difference

Under heterogenous competition types, the auctioneer derives the lowest revenue under the APE competition type. When we have a majority of a bidding strategy in the auction (i.e., two bidders of the same type) an interesting pattern in revenue based on the third bidder's bidding strategy can be observed (see Table 23). Revenue decreases when the third bidder type changes from Participator to Explorer (for Analyzer majority competitions), from Explorer to Analyzer (for Participator majority competitions), or from Participator to Analyzer (for Explorer majority competitions); i.e., revenue increases when the third bidder's strategy changes from Analyzer to Explorer to Participator, but there is no significant increase in revenue between the heterogenous competition type with higher revenue and the homogenous competition for each bidder type.

Figure 28b shows that the homogenous EEE competition type leads to the highest average (0.91) and the AAA competition type has the highest variance (0.11) in allocative efficiency; average efficiency ranges between 0.84 to 0.91. The differences in average allocative efficiency under other competition types are not significant enough to support a general pattern/explanation.

Part III: Human vs. Machine Auction Experiments

In this part, we integrate the bidding agents that were developed in Part II into an experimental combinatorial auction platform. This allows for fine-grained control of the competitive environment human bidders are going to face, i.e., participants join each auction and play against agents with certain pre-determined bidding strategies. The goal of this study is to investigate the underlying reasons for different bidder behaviors and to understand the effect of competition on auction outcomes and participants' acceptance of continuous CAs as a viable trading mechanism. The comprehensive simulations from the previous part will guide the design of experiments by suggesting which competition types to include. The goal is to address the following questions:

- How are auction outcomes (e.g., efficiency, auction revenue) influenced by the types of competition bidders face?
- Is the behavior that bidders exhibit an endowed trait that does not change over time? Or is there a learning effect, in terms of change in sophistication level, from observing the auction dynamics?
- If there is learning, how is it affected by the type of competition a bidder faces when participating in multiple auctions? i.e., how does the auction dynamics experienced by different bidders impact their learning?
- How does the competitiveness experienced by participants affect their perception of continuous CAs as a trading mechanism? How will such behavioral variables affect the sustainability of continuous CAs as a viable market mechanism?

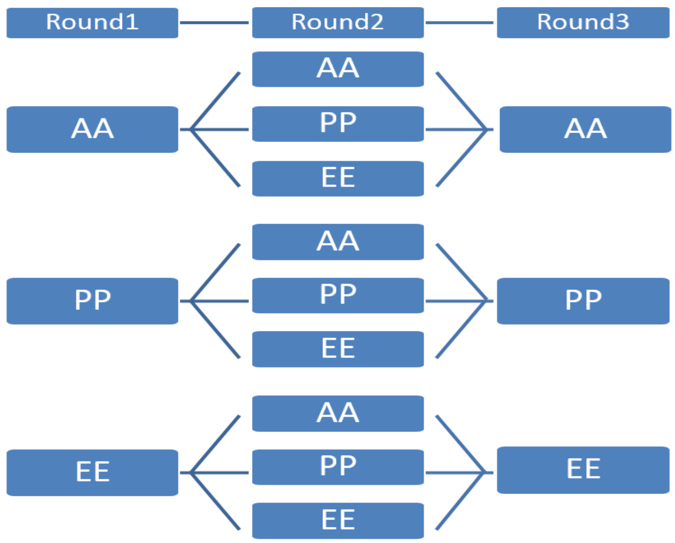
The findings are expected to also support and enhance the results from simulation experiments in the second part of my thesis.

11.Experiment design

For the main part of the experiment, participants take part in three subsequent auctions (same auction mechanism presented in Section 2.2) in which they face different types of competition as shown in Figure 29; the sequence of 3 auctions is a treatment and participants are randomly assigned to one of the 9 possible treatments. Three homogeneous types of competition, AA, PP, and EE (among the 6 possible competitions) are selected for this experiment to allow for better identification of competition effects. The agents developed using the ML approach (presented in Part II) are used for this experiment. For example, a participant who is given the second treatment, AA-PP-AA, will compete against two Analyzer type agents in the first and third rounds, and two Participator type agents in the second round. Analyzer type agents are placing fewer bids, bid on more valuable bundles, and derive higher surplus as compared to Participator and Explorer type agents. Participator and Explorer type agents do not significantly differ in the amount of surplus they derive but Explorer agents place more bids on a wider variety of bundles (i.e., have exploratory and random behavior) compared to Participator types (see Sections 8 and 10 for agent behavior).

At the beginning of each session, participants take a short demographics survey, followed by detailed instructions, a practice test, and a demo auction to make sure they understand the experiment scenario. Participants do not know that they are playing against software agents. They are not told how many other bidders are in each auction and only know about their own valuation setup. Participants take a post-experiment survey once they have completed all three auctions. Upon completing the survey, they are shown their earned amount in each of the three auctions and their total payoff.

Figure 29. Experiment design



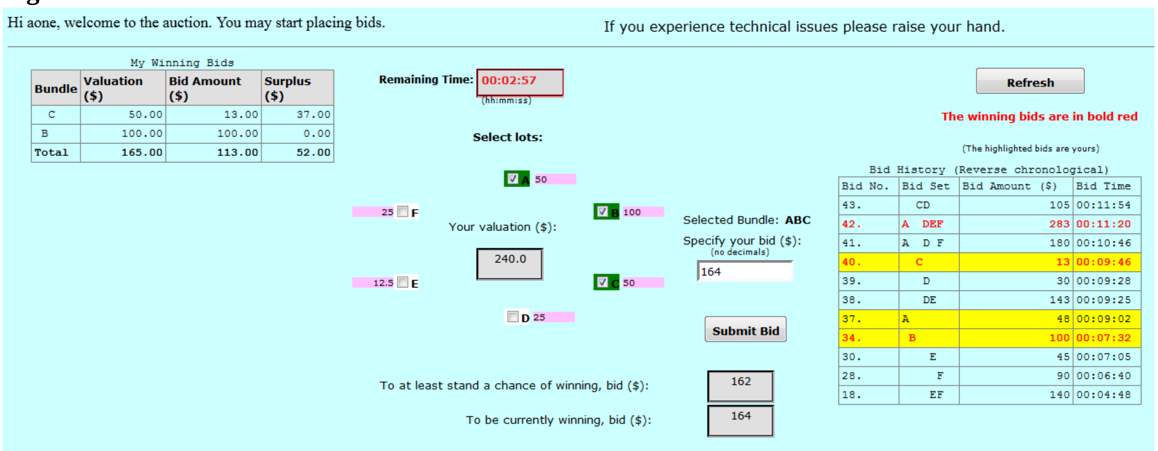
11.1 Auction procedure

The auction setup is the same described in Section 2.2. Each participant is assigned one of the valuation setups shown in Figure 1 when s/he enters each auction, and each of the two bidding agents receives one of the other two valuation setups. None of the players know about other participants' valuation (i.e., bidder's valuation is private information for both human participants and agents) or how many other players there are in the auction.

The initial auction duration is set to 12 minutes, but it is extended to have 1 full minute remaining if there is any bidder activity within the last minute. This so-called soft stopping rule prevents auction sniping (i.e., bidders placing last second winning bids in limited-time auctions). Participants are provided comprehensive information feedback throughout the auction, i.e., they can see bids placed so far in the auction, the provisional winning allocation at the current auction state, and the WL and DL for any bundle of interest. They are shown the amount of surplus they make at the end of each auction. A screenshot of the auction page as observed by participants in the middle of an auction is shown in Figure 30.

Participants are paid $15 as their participation fee and can earn/lose 10 cents for every dollar of positive/negative surplus they make in the auctions (up to $50 and down to $10). A complete experimental session including the pre and post-experiment surveys takes about 80 to 120 minutes.

Figure 30. Auction screenshot



11.2 Participants

We have a total of 135 participants. Table 24 shows the descriptive statistics for participants' demographic variables based on their responses to the pre-experiment survey. The 120 student participants (89%) are majoring in 76 different disciplines.

Table 24. Descriptive statistics of participants (pre-experiment survey)

135 participants across 11 sessions	
Age	Mean = 24, Median = 22, sd = 8.6, min = 18, max = 79
Gender	47% male, 53% female
Education	high school diploma = 10 (7.4%) college freshman year = 13 (9.6%) college sophomore year = 20 (14.8%) college junior year = 28 (20.7%) college degree = 36 (26.7%) graduate degree = 25 (18.5%) doctorate degree = 3 (2.2%)
Currently student?	Yes = 88.9%, No = 11.1%
Experience with online auctions?	Yes = 20%, No = 80%

11.3 User perception

The post-experiment survey is used to assess participants' perception along several dimensions. Each construct is assessed using several questions/items that are answered on a 7-point Likert scale, where 1 stands for "strongly disagree" and 7 stands for "strongly agree". The constructs and questions are as follows.

- Perceived competitiveness of the auction (**Comp**):
    - There was a lot of competition for the property lots that I wanted.
    - Other bidders in the auction seemed to want the same property lots that I wanted.
    - The auctions were competitive.
- Participant's perception about improvements in their decision making, i.e., learning (**LRN**):
    - I was more comfortable making bidding decisions in the last auction compared to the first auction.
    - Making bidding decisions was easier in the first auction compared to the last auction
    - participating in multiple auctions helped me place better bids.
    - I found it harder to make good bidding decisions in the first auction compared to the last auction.
    - Making bidding decision became progressively easier from the first to the last auction.
- Participants' perception of the auction platform as a suitable fit for the proposed market mechanism, derived from the task-technology fit (**TTF**) model (Goodhue 1995; Goodhue and Thompson 1995):
    - The system helped me in determining how much to bid on my selected lot(s).
    - The system provided the information I needed at all stages of the auction.
    - The system helped me to determine winning bid(s) at all stages of the auction.
    - The system helped me in determining the lot(s) to bid on.
- I would encourage my friends to participate in this type of auction (**RCM**).

The four following constructs are derived from the technology acceptance model (TAM) (Davis 1989). The goal is to investigate how the change in the experienced competition would impact users' acceptance of the auction platform.

- Perceived ease of use (**EOU**):
    - I felt very comfortable using the system.
    - I found the system easy to use.
    - I found this mechanism understandable for buying or selling multiple items.
    - I found the system to be user-friendly.
- Intention to use (**ITU**):
    - Assuming I had access to the system for participating in a combinatorial auction, I would intend to use it.
    - If I had access to the system for participating in a future combinatorial auction, I predict

I would use it.

- If I had access to the system for participating in a combinatorial auction, I would plan to use it.
- I would participate in a combinatorial auction if it were available to me in the future.

- Perceived performance (**PP**):
  - I am pleased with the decisions I made during the auction.
  - I am satisfied with my performance in this auction.
- Perceived usefulness (**PU**):
  - I found this mechanism suitable for buying or selling multiple items.
  - Using combinatorial auctions for buying and selling multiple items seems like a good idea.
  - I found this mechanism effective for buying or selling multiple items.
  - Combinatorial auctions are beneficial for buying or selling multiple items.

Questions/items that measure each construct are randomized in the survey. The score for each construct is the average of the item scores (inverse coded questions are weighted correspondingly). For example, if a participant's responses to the perceived performance (PP) questions are 4 and 5, his score for the PP construct is 4.5.

12.Results

12.1 Descriptive statistics

Table 25 shows the descriptive statistics for auction outcomes across grouped by type of competition used as treatment (i.e., type of bidding agents used as treatment in that auction) and auction sequence (i.e., whether it was the 1st, 2nd, or 3rd auction).

Table 25. Mean values for auction-specific outcome variables (standard deviations in parenthesis)

Sequence	Competition type	Number of auctions	Revenue in $		Duration in minutes		Allocative efficiency	
1	AA	46	406.39	(50.72)	21.09	(6.52)	0.96	(0.07)
	EE	44	421.57	(43.21)	19.78	(4.07)	0.95	(0.06)
	PP	45	381.56	(67.86)	18.9	(4.61)	0.94	(0.06)
2	AA	44	398.7	(41.19)	20.58	(5)	0.96	(0.06)
	EE	47	417.83	(43.52)	19.36	(5.58)	0.95	(0.07)
	PP	44	367.98	(56.64)	19.26	(5.15)	0.92	(0.06)
3	AA	46	403.13	(56.75)	21.41	(6.18)	0.97	(0.06)
	EE	44	431.86	(32.02)	21.23	(4.83)	0.96	(0.05)
	PP	45	360.29	(62.61)	18.49	(4.48)	0.93	(0.06)

Table 26 shows descriptive statistics for participants' total number of bids placed throughout the auction (Bids), number of distinct bundles bid on (Spans), their surplus at the end of each auction (Surplus), and the effort they made in terms of number of total clicks per placed bid (Effort), grouped by competition type and sequence of the auction. Table 27 shows descriptive statistics for user perception variables presented in Section 11.3.

Table 26. Mean values for bidder-specific outcome variables (standard deviations in parenthesis)

Sequence	Competition type	# of bidders	Bids	Spans	Surplus	Effort
1	AA	46	33.35 *(16.4)*	12.85 *(6.68)*	24.46 *(28.71)*	11.82 *(5.48)*
	EE	44	28.02 *(12.93)*	11.36 *(5.83)*	33.62 *(31.59)*	21.21 *(53.77)*
	PP	45	30.16 *(17.44)*	13.42 *(8.26)*	49.18 *(55.08)*	11.7 *(7.15)*
2	AA	44	37.55 *(17.5)*	13.57 *(5.81)*	34.11 *(25.13)*	16.21 *(21.36)*
	EE	47	25.79 *(10.71)*	10.32 *(4.64)*	36.76 *(35.67)*	15.25 *(13.8)*
	PP	44	29.09 *(16.9)*	12.82 *(7.85)*	68.07 *(42.93)*	15.51 *(13.02)*
3	AA	46	33.83 *(20.17)*	11.5 *(7.37)*	27.2 *(32.83)*	15.1 *(11.73)*
	EE	44	31.55 *(19.6)*	11.45 *(5.68)*	39.94 *(27.33)*	21.04 *(28.26)*
	PP	45	31.67 *(20.32)*	12.31 *(5.73)*	66.09 *(41.54)*	12.14 *(6.53)*

Table 27. Means of user perception variables per treatment (standard deviations in parenthesis)

Treatment	Number of participants	EOU	ITU	PP	PU	LRN	COMP	TTF	RCM
PP-EE-PP	16	6.25 *(0.61)*	6.06 *(0.71)*	5.59 *(1.24)*	6.06 *(0.77)*	5.53 *(0.98)*	5.31 *(1.38)*	6.45 *(0.58)*	5.81 *(1.14)*
PP-PP-PP	14	5.84 *(1.03)*	5.79 *(1.16)*	5.39 *(1.46)*	5.66 *(1.13)*	5.27 *(1.06)*	5.95 *(0.78)*	6.18 *(0.78)*	5.57 *(2.05)*
PP-AA-PP	15	5.67 *(1.33)*	5.37 *(1.46)*	5.3 *(1.12)*	5.47 *(1.23)*	4.75 *(1)*	5.78 *(1.21)*	5.67 *(1.22)*	5.8 *(1.49)*
AA-EE-AA	16	5.17 *(1.59)*	4.47 *(1.64)*	4.19 *(1.48)*	4.33 *(1.8)*	4.99 *(1.36)*	5.27 *(0.93)*	5.22 *(1.42)*	4.88 *(1.63)*
AA-PP-AA	15	6.07 *(0.56)*	5.72 *(0.69)*	5.07 *(0.99)*	5.78 *(0.66)*	5.23 *(0.6)*	6.09 *(0.69)*	6.15 *(0.59)*	5.67 *(0.88)*
AA-AA-AA	15	5.62 *(1.11)*	4.82 *(1.64)*	4.87 *(1.19)*	5.18 *(1.22)*	5.05 *(1.15)*	5.2 *(1.05)*	5.72 *(0.87)*	4.6 *(1.72)*
EE-EE-EE	15	5.65 *(0.71)*	5.08 *(1.48)*	5.47 *(0.77)*	5.37 *(1.08)*	5.12 *(0.96)*	5.82 *(1.21)*	6.08 *(0.68)*	5.13 *(1.56)*
EE-PP-EE	15	5.62 *(1.22)*	5.25 *(1.16)*	5.13 *(1.17)*	5.37 *(1.19)*	4.93 *(1.14)*	5.44 *(1.31)*	5.77 *(1.32)*	4.8 *(1.58)*
EE-AA-EE	14	5.54 *(0.91)*	5.18 *(1.03)*	4.82 *(1.22)*	5.23 *(0.88)*	5.01 *(0.77)*	5.81 *(1.06)*	5.59 *(1.09)*	5.29 *(1.35)*

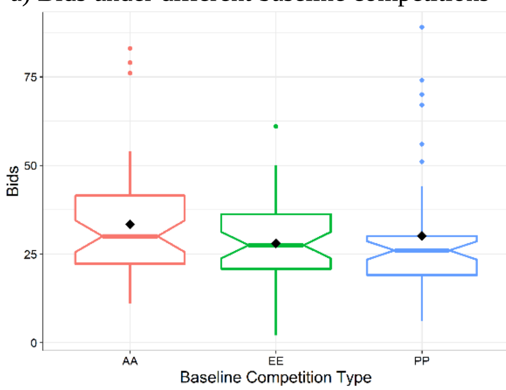
The average value for all 8 variables in Table 27 are significantly higher than 4 (based on t-test results); except for ITU, PP, and PU under the AA-EE-AA treatment; and ITU and RCM under the AA-AA-AA treatment.

12.2 Impact of competition on bidder behavior

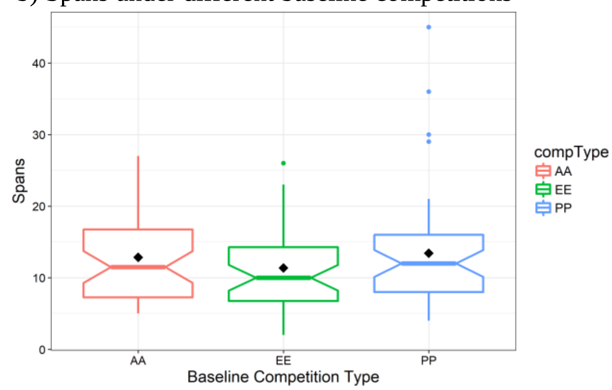
Using the data from only 1st round auctions (see the 3 rows corresponding to sequence 1 in Table 26), we can study the impact of competition on bidder behaviors; participants have not gained experience at this stage and are randomly assigned to one of the baseline treatments (AA, PP, or EE). Figure 31 compares participants' bidder-specific outcome variables (presented in Table 26) across different baseline competition treatments.

Figure 31. Comparing bidder-specific outcome variables across 1st round auctions

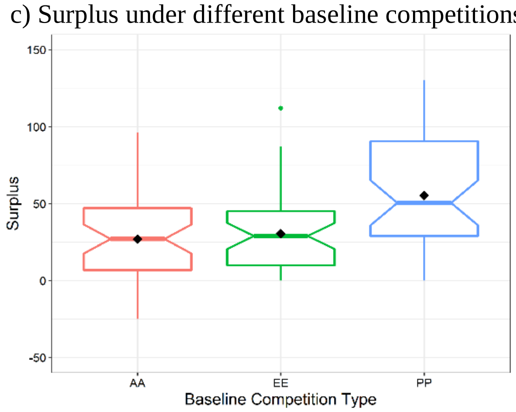
a) Bids under different baseline competitions



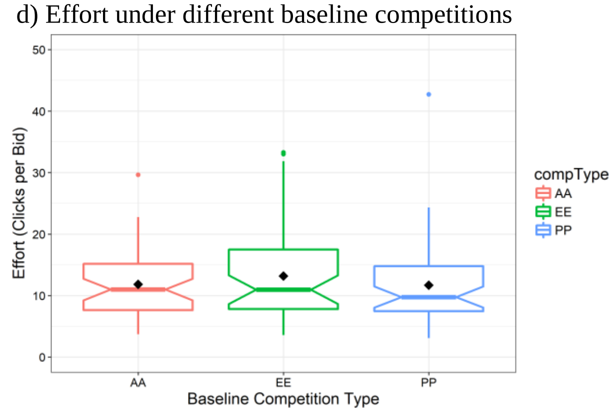
b) Spans under different baseline competitions



c) Surplus under different baseline competitions



d) Effort under different baseline competitions



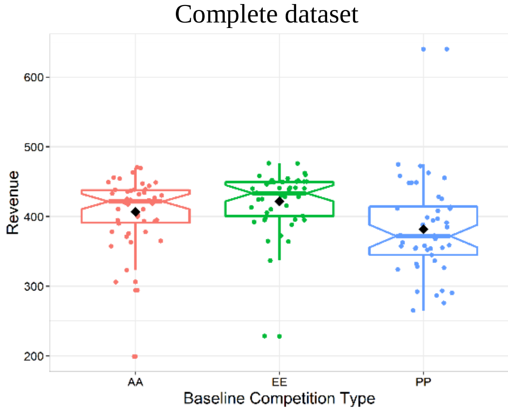
We use t-tests to verify significant differences between means and Wilcoxon-Rank-Sum (WRS) tests to check for significant differences between medians. Participants' average Bids are marginally higher when they receive the AA baseline competition (p-value = 0.045 for t-test of AA > EE), compared to when they receive the EE competition. Average Surplus is significantly higher when participants receive the PP competition (p-value=0.005 for t-test of PP > AA, and p-value=0.05 for t-test of PP > EE), but there is no significant difference in average surplus between the AA and EE competitions. The differences in participants' Spans and Effort are not significant across different baseline competitions. These results suggest that participants place more bids when facing the more competitive AA competition. When given the moderate PP competition (competing against participatory and more predictable bidding strategies) participants derive the highest surplus and both AA (competing against more competitive and rational bidding agents that rarely finish an auction without deriving positive surplus) and EE (competing against more randomly behaving bidding agents that place frequent bids on diverse bundles) competitions result in lower surplus (via different mechanisms).

12.3 Impact of competition on auction outcomes

In this subsection we look at the impact of competition on auction outcomes, using only data from 1st round auctions where participants are not yet exposed to any learning effects. Figure 32 compares auction revenue between different baseline treatments. Note that competition type indicates our bidding agents bidding strategy and the participant's (3rd bidder in each auction) bidding strategy is not identified here. The right boxplot in Figure 32 merely excludes the outlier points seen in the left plot and readjusts the revenue scale on the vertical axis accordingly. Using one-sided t-tests to check significant differences in average revenue between different competitions, we find that auctioneer's revenue is significantly lower under the PP competition (compared to both AA and EE competitions), and marginally higher under the EE competition as compared to the AA competition; p-value=0.06 for t-test of EE > AA, p-value=0.026 for t-test of AA > PP, and p-value=0.000 for t-test of EE > PP.

Figure 32. Comparing auction revenue across 1st round auctions under different baseline competitions

Complete dataset



Excluding outliers from left boxplot

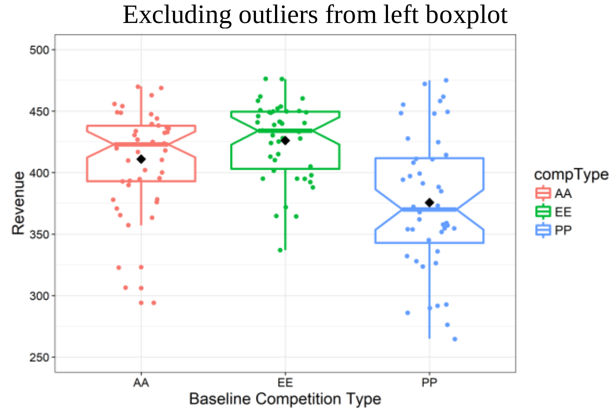
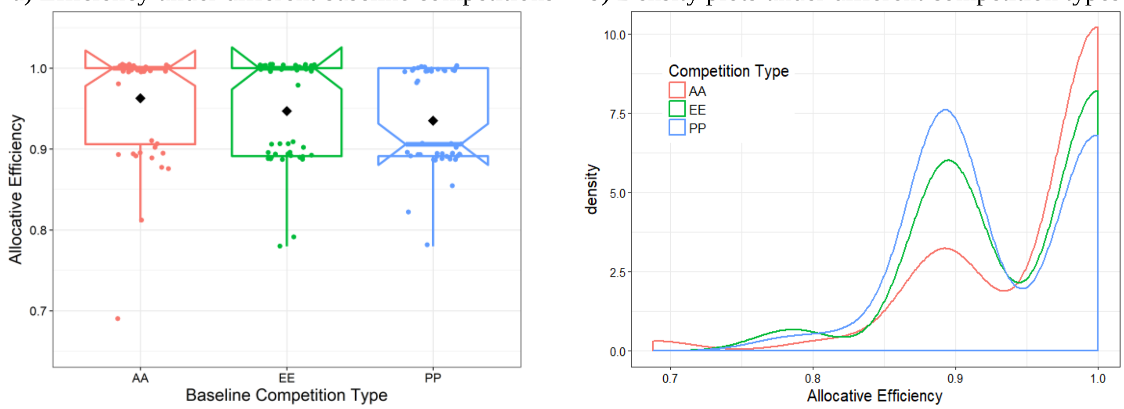


Figure 33 compares allocative efficiency between different baseline treatments. Median efficiency is similar under AA and EE competitions and both are significantly higher than median efficiency under PP competition (based on WRS tests). Average allocative efficiency (shown by the black dots on the boxplots in Figure 33) under AA competition is significantly higher than efficiency under PP competition. However, note that the distribution of allocative efficiency is bimodal as indicated by the density plots in Figure 33; most auctions end with the optimal allocative efficiency of 1 and remaining auctions have an average allocative efficiency of about 0.89.

Figure 33. Comparing allocative efficiency across 1st round auctions

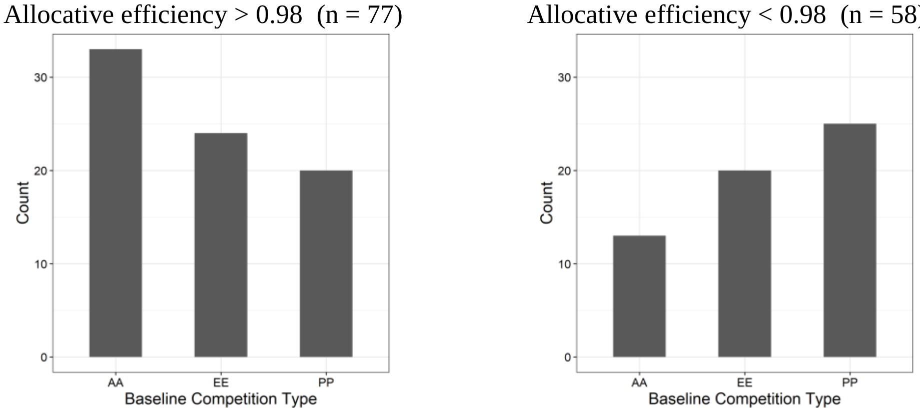
a) Efficiency under different baseline competitions b) Density plots under different competition types



To better understand these observed differences given the bimodal distribution of efficiency, Figure 34 shows the frequency of competition types for auctions where allocative efficiency is above 0.98 (optimal) and auctions where efficiency is below 0.98 (sub-optimal). We can see that auction with baseline AA competition are mostly optimal, auctions with PP baseline competition more frequently lead to sub-optimal auctions (as well as lower revenue), and auctions with EE competition (almost) equally lead to optimal and sub-optimal auctions.

Figure 34. Frequency competition types grouped by optimal and sub-optimal auctions

Allocative efficiency > 0.98 (n = 77) Allocative efficiency < 0.98 (n = 58)



12.4 Bidder behavior and learning

In this subsection, we study the change in participants behavior across different treatments. We calculate the change in participants' bidder-specific variables (Bids, Spans, Surplus, and Effort) between their 1st and 3rd auctions, where they receive the same baseline competitions as treatment, for example, a participant in treatment group PP-AA-PP will receive the PP competition in her 1st and 3rd auctions. For any bidder-specific variable X for participant i, $\Delta X_i = X_i$(in 3rd auction) – X_i(in 1st auction); a positive value indicates an increase and a negative value indicates a decrease in X. Table 28 shows the mean change (mean ΔX) for each of the four bidder-specific outcome variables per treatment. For example, under the PP-EE-PP treatment the average number of bids placed by participants throughout the auction decreases by about 3 bids ($\Delta Bids$=-3.31).

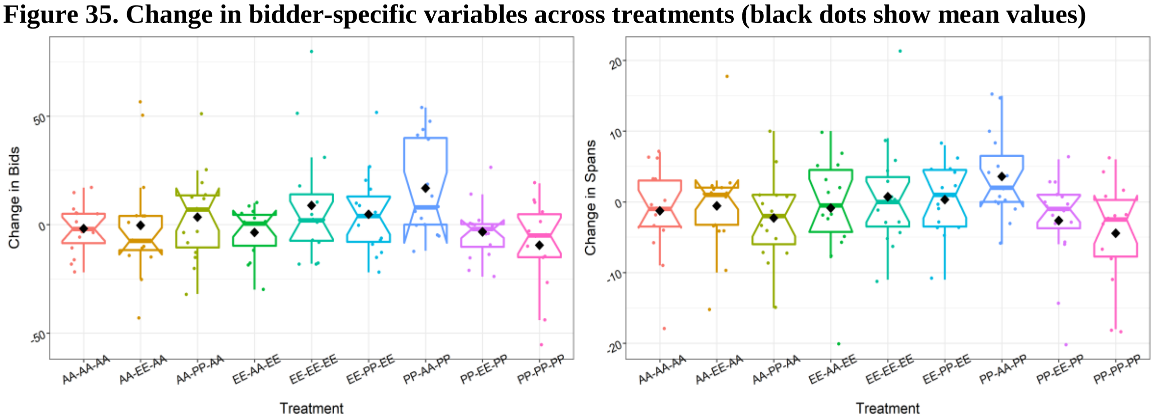
Table 28. Average change in bidder-specific variables across treatments (std. dev. in parenthesis)

Treatment	ΔBids		ΔSpans		ΔSurplus		ΔEffort	
PP-EE-PP	-3.313	(12.279)	-2.625	(6.52)	15.75	(35.815)	0.35	(7.16)
PP-PP-PP	-9.429	(20.891)	-4.429	(7.345)	35.429	(87.026)	3.33	(5.36)
PP-AA-PP	16.867	(22.239)	3.6	(6.139)	0.867	(36.701)	-2.17	(6.44)
AA-EE-AA	-0.25	(25.005)	-0.563	(6.928)	-14.891	(38.843)	1.84	(7.57)
AA-PP-AA	3.533	(20.553)	-2.267	(6.1)	9.3	(30.823)	3.95	(12.48)
AA-AA-AA	-1.8	(11.465)	-1.267	(6.475)	14.983	(41.876)	4.15	(11.38)
EE-EE-EE	8.867	(27.163)	0.733	(7.526)	-4.517	(36.374)	-1.12	(5.9)
EE-PP-EE	4.8	(18.937)	0.333	(5.08)	16.867	(24.121)	8.61	(31.89)
EE-AA-EE	-3.571	(11.837)	-0.857	(7.543)	6.643	(38.218)	10.51	(27.57)

Positive values indicate an increase and negative values indicate a decrease.

Figure 35 shows the change in the four above bidder-specific variables from the 1st to 3rd auctions. The boxplots also suggest heterogeneity in variance across different treatments for all 4 variables. To analyze under which treatment(s) the increase/decrease in these bidder-specific variables is significantly higher/lower than zero, we use one-sided t-tests. The results in Table 29 show that the change in these bidder-specific variables is significant (marginally significant) only under some of the treatments (based on the p-values of the t statistics).

Figure 35. Change in bidder-specific variables across treatments (black dots show mean values)



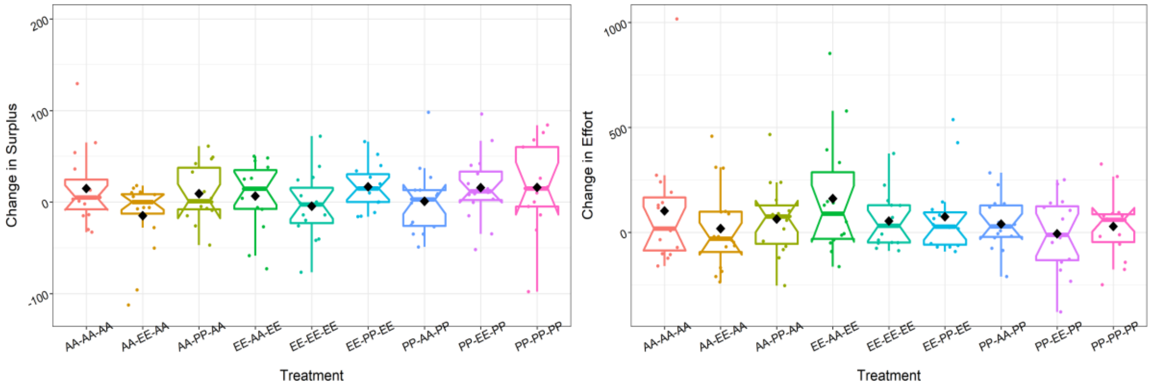


Table 29. Testing increase/decrease in bidder-specific outcome variables using one-sided t-tests (p-values are shown in parenthesis)

Treatment	ΔBids <> 0		ΔSpans <> 0		ΔSurplus <> 0		ΔEffort <> 0	
PP-EE-PP	t = -1.08	(0.1488)	t = -1.61^{+}	(0.0641)	t = 1.76*	(0.0495)	t = -0.15	(0.4412)
PP-PP-PP	t = -1.69^{+}	(0.0575)	t = -2.26*	(0.021)	t = 1.52^{+}	(0.0758)	t = 0.68	(0.2552)
PP-AA-PP	t = 2.94***	(0.0054)	t = 2.27*	(0.0197)	t = 0.09	(0.4642)	t = 1.24	(0.1185)
AA-EE-AA	t = -0.04	(0.4843)	t = -0.33	(0.3749)	t = -1.53^{+}	(0.073)	t = 0.37	(0.3567)
AA-PP-AA	t = 0.67	(0.2582)	t = -1.44^{+}	(0.086)	t = 1.17	(0.131)	t = 1.42^{+}	(0.0888)
AA-AA-AA	t = -0.61	(0.2765)	t = -0.76	(0.2306)	t = 1.39^{+}	(0.0937)	t = 1.38^{+}	(0.0952)
EE-EE-EE	t = 1.26	(0.1134)	t = 0.38	(0.3558)	t = -0.48	(0.319)	t = 1.63^{+}	(0.063)
EE-PP-EE	t = 0.98	(0.1715)	t = 0.25	(0.4015)	t = 2.71**	(0.0085)	t = 1.60^{+}	(0.0658)
EE-AA-EE	t = -1.13	(0.1397)	t = -0.43	(0.3388)	t = 0.65	(0.2634)	t = 2.12*	(0.0271)

The direction of inequalities are based on values in Table 28, i.e., if the change is positive the test is for $\Delta > 0$, and if the change is negative the test is for $\Delta < 0$.

Table 30 shows the correlations between the change in the four variables and auction duration. This gives us a better picture of simultaneous changes before we proceed with regression modeling. Change in duration is highly correlated with change in Bids and Effort, and relatively correlated with change in Spans, which is not unexpected given the soft-stopping rule of the auction. Longer auctions mean that more bids are placed which require more effort and are associated with bids on more diverse bundles (significant positive correlation of ΔSpans and ΔBids). Interestingly change in Surplus is only marginally correlated with change in auction duration, and the correlations suggest that an increase in the total number of bids and diversity of bundles is associated with a

Table 30. Correlation matrix for change variables; significance levels for Pearson correlations

	ΔBids	ΔSpans	ΔSurplus	ΔEffort
ΔSpans	0.743***	-	-	-
ΔSurplus	-0.265**	-0.267**	-	-
ΔEffort	-0.368***	-0.299***	0.09	-
ΔDuration	0.452***	0.297***	-0.175*	0.059

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

significant decrease in surplus (significant negative correlation of ΔSurplus with ΔSpans and ΔBids).

Next, we build regression models to investigate the effect of different treatments (different sequences of competition types) on the change in each of these 4 outcome variables. For each model, the outcome variable is ΔY (the change in variable Y) and the potential predictors are chosen from, treatment (as a categorical variable), the 5 variables presented in Table 24, and interaction terms using the stepwise regression procedure with variable selection in both directions (both forward and backward selection). ΔDuration is not included as an independent variable in the models since auction duration (or the change of it) is by nature the result of bidder behaviors. Model performance is assessed in terms of AIC and adjusted-R^2 for models that have a significant F-test of overall significance (when F-test of overall significance is not significant the model is not a better fit than an intercept-only model). Insignificant interaction terms and controls that do not significantly improve the model are dropped for parsimony. The results in Table 31 show the final model as well as the ANOVA results for factors influencing the change in each of the 4 outcome variables. Unless otherwise mentioned, for all ANOVA tests presented in this manuscript type 2 sum of squares are used when there are no significant interactions, where the presence of each main effect is tested while controlling for other main effects (using Anova(,type=2) function instead the default anova() function which runs type 1 tests). When there are significant interactions, type 3 sums of squares are used (using the Anova(,type=3) function from *car* library in R) where the presence of each main effect is tested after the other main effects and interactions. The default type 1 tests may show higher significance but would be incorrect for the unbalanced sample sizes across treatment groups; i.e., main effects are tested sequentially, and the effect of treatment might be tested without controlling for the effect of other variables and interactions.

The results in Table 29 and 31 provide some evidence that bidders learn as they move from the 1st to the 3rd auction, as indicated by the change in Bids, Spans, and Surplus. However, this change in behavior is influenced by the sequence of competitions participants face and significant under some treatments. Participants learn how to derive higher surplus (marginally significant increase) under PP-EE-PP, PP-PP-PP, AA-EE-AA, AA-AA-AA, and EE-PP-EE treatments as can be seen in the ΔSurplus column in Table 29. This learning seems to be moderated by whether participants have had prior experience with online auctions as indicated by the significant interaction term in the ΔSurplus model in Table 31; under some treatments (e.g., PP-PP-PP) the change in surplus is larger when participants have not had prior experience with online auctions.

The change in Effort is only marginally significant under some treatments (column ΔEffort in Table 29) and this change does not significantly vary across treatments (insignificant F statistics for treatment variable in ΔEffort model in Table 31). The change in Bids varies across treatments (significant F statistics for treatment variable in ΔBids model in Table 31) but it only significantly increases under the PP-AA-PP treatment (ΔBids column in Table 29). The change in bundle diversity from the 1st to the 3rd auctions slightly differs across treatments (treatment variable is marginally significant for the ΔSpans model in Table 31). Furthermore, ΔSpans is only significant (marginally) under some treatment groups (see ΔSpans column in Table 29).

Table 31. Modeling change in bidder-specific outcome variables and ANOVA results for factors influencing the change in bidder-specific outcome variables

ΔBids ~ Treatment + Gender				
ΔSpans ~ Treatment				
ΔSurplus ~ Treatment * AuctionExp(Y/N)				
ΔEffort ~ Treatment * (Student(Y/N) + Gender) + Age + Education				
Outcome:	**ΔBids**	**ΔSpans**	**ΔSurplus**	**ΔEffort**
Variable:	F statistics (p-values shown in parenthesis)			
Treatment	2.16^{*} (0.0348)	1.749^{+} (0.0943)	1.64 (0.122)	2.26^{*} (0.036)
Gender(M/F)	2.28 (0.1334)	-	-	8.49^{**} (0.0051)
Education	-	-	-	2.48^{*} (0.034)
Student(Y/N)	-	-	-	4.89^{*} (0.031)
AuctionExp(Y/N)	-	-	0.328 (0.568)	0.32 (0.57)
Treatment:Gender	-	-	-	3.06^{*} (0.006)
Treatment:Education	-	-	-	2.28^{**} (0.0025)
Treatment:Student	-	-	-	5.98^{***} (0.0000)
Treatment:AucExp	-	-	2.18^{*} (0.034)	2.11^{+} (0.057)
R^2	14.31%	9.97%	21.25%	74.35%
Adjusted-R^2	8.14%	4.26%	9.8%	37.97%
Model F statistic	2.32^{*} (0.0189)	1.749^{+} (0.0943)	1.857^{*} (0.0287)	2.04^{**} (0.0029)

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

Overall, learning is more likely when participants are given the PP competition as their baseline (i.e., playing against two Participator type agents in their 1st and 3rd auctions). This competition type allows for more improvement in terms of surplus (note the lower average revenue in auctions with PP competition in Table 25) and we see more change in both diversity of bundles and number of total placed bids as compared to treatments where participants receive AA or EE competition types as their baseline. When playing against two Analyzer agents (AA competition) in the 1st and 3rd rounds, which are the most competitive treatments, there are only marginally significant changes in Spans, Surplus, and Effort (see Table 29). This suggests that there is not much room for improvement under these treatments. When participants receive the EE competition in their 1st and 3rd auctions, there is a highly significant improvement in Surplus under the EE-PP-EE treatment and a significant increase in Effort under the EE-AA-EE treatment while there is no significant change in any of the outcome variables under the EE-EE-EE treatment. Participants learn how to better play against the random EE competition and derive higher surplus in their 3rd round when they face the moderate PP competition in their 2nd round. But they do not learn how to improve their performance in the 3rd round (in terms of surplus) when they experience the tougher AA competition in the 2nd round.

12.5 Initial bidding strategies and its impact on learning

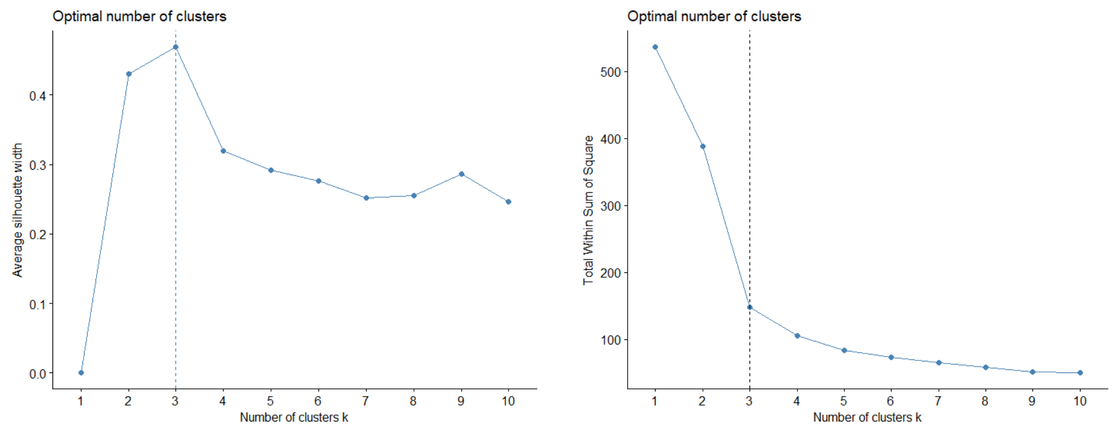
In this subsection, we analyze participants' initial bidding strategies and how it influences bidder behavior as they proceed in the auctions under different treatments. We cluster participants in the 1st round auctions across all 9 treatment groups; at this stage, participants have not yet been exposed to the treatments and their behavior should reflect their bidding strategy before they learn from the auction environment. We want this characterization to be purely based on participants' observed behavior in the auctions and therefore do not use any of the demographic variables from the pre-experiment survey (i.e., age, gender, education, auction experience, and student status) for the cluster analysis. The selected variables for the cluster analysis are: total number of bids placed throughout the auction (Bids), variety of bundles bid on (Spans), amount of effort calculated as the total number of clicks per placed bid (Effort), and when the participant placed the first bid (time of entry in seconds – TOE). These features are selected consistent with (Adomavicius et al. 2012) to allow for comparisons between the derived clusters. Variables are normalized for the cluster analysis and k-means clustering using Euclidean distance is used. To determine the optimal number of clusters (k) we use the following approaches:

1. The average silhouette method (Rousseeuw 1987), where the optimal number of clusters is the one that maximizes the average silhouette of observations, using the *fviz_nbclust* function from the *factoextra* library in R (see https://cran.r-project.org/web/packages/factoextra/ for details). This method suggests 3 clusters as shown in Figure 36a.
2. The elbow method (using the *fviz_nbclust* function), where the optimal number of clusters is determined so that the total intra-cluster variation (or the total within-cluster sum of squares) does not significantly decrease by increasing the number of clusters, i.e., the elbow in Figure 36b. This method suggests 3 clusters.
3. Using the *NbClust* package (Charrad et al. 2014) in R, which provides 30 indices for

determining the number of clusters and proposes the best clustering scheme from the different results obtained by varying all combinations of number of clusters, distance measures, and clustering methods (see https://cran.r-project.org/web/packages/NbClust/ for details). According to the majority rule using this method, the best number of clusters is 4.

Figure 36. Determining optimal number of clusters

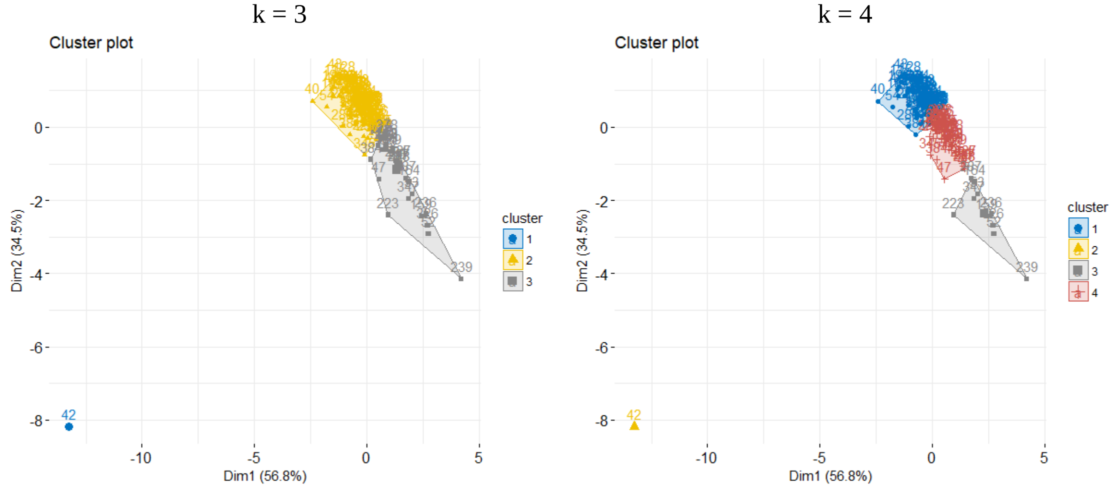
a) Average silhouette method b) Elbow method



We run k-means for both k=3 and k=4 and by inspecting the clustering results we find that in both cases one of the clusters consists of only one observation, which is the same point in both

Figure 37. Cluster plots for k-means results on the 1st auctions dataset

k = 3 k = 4



cases; point 42 in both cluster plots in Figure 37. In the cluster plots in Figure 37 principal component analysis (PCA) is used to reduce the four input features into two dimensions that are shown as the horizontal and vertical axis of the cluster plots; the percentages indicate the amount of overall variability explained by the 1st (horizontal axis) and 2nd (vertical) component. We repeat the analysis after dropping this outlier and all three approaches recommend 4 clusters as the optimal number. Nevertheless, we run the k-means analysis with k values from 3 to 10 and evaluate the resulting cluster centroids (in terms of Bids, Spans, Effort, TOE, as well as Surplus) for each run (i.e., each k value). The goal is to evaluate whether deriving more (or less) than 4 clusters, or further clustering large clusters (i.e., with many data points), would allow us to better characterize the different behaviors; complementing cluster evaluation measures such as within-cluster sum of squares with interpretability of the results based on our domain knowledge. We find that 4 clusters (derived from the dataset without the outlier point) make for most interpretable cluster characteristics as well.

Table 32. Cluster centroids of bidder types (standard deviation in parenthesis) in 1st round auctions

Cluster	Bidding behavior	size	Bids	Spans	Effort	TOE	Surplus
1	Explorer	19	57.00 *(16.97)*	24.89 *(6.72)*	6.29 *(2.12)*	35.11 *(12.65)*	14.75 *(28.86)*
2	Analyzer	25	18.68 *(7.58)*	6.68 *(2.54)*	21.79 *(6.67)*	45.60 *(18.44)*	44.80 *(39.45)*
3	Analyzer	19	26.00 *(14.33)*	10.26 *(5.89)*	11.86 *(6.33)*	111.32 *(26.6)*	54.62 *(40.94)*
4	Participator	71	29.27 *(8.39)*	12.08 *(3.77)*	10.52 *(3.64)*	37.04 *(12.72)*	31.94 *(41.60)*
outlier	Analyzer	1	2	2	366.5	834	112

Table 32 shows cluster centroids (and standard deviations) plus mean Surplus for each cluster (note that Surplus has not been used as an input feature for the k-means clustering); mean values are shown in the original (non-normalized) scale. The centroids in Table 32 are based on results using the *NbClust* function (instead of the *kmeans* function in R) since this function derives more robust clusters by running the k-means algorithm multiple times using various indices and assigning points to clusters using a majority rule. The cluster characteristics are consistent with the bidder clusters identified in earlier studies (Adomavicius et al. 2012) and are named accordingly. However, we will refer to them as *initial bidding strategies/behaviors* since participants' bidding behavior changes in the 2nd and 3rd auctions.

Initial Analyzer types generally start placing bids later than the two other types, spend more effort, and place fewer bids on less diverse (more valuable) bundles, which makes them derive the highest surplus. Initial Participator types place more bids on more diverse bundle than Analyzer types but spend less effort per placed bid and derive less surplus as a result. Initial Explorer types generally start bidding earlier than the other types, place the most bids on a wide variety of bundles while spending the least effort, and end up with the lowest surplus. Both clusters 2 and 3, while different from each other, represent Analyzer type of behavior. The outlier point represents an obvious Analyzer type behavior as well (i.e., very low Bids and Spans, high effort, and late TOE, as well as high Surplus), however, we will drop it for the subsequent analysis. Table 33 shows the variables presented in Table 32 regrouped based on identified initial bidding behavior, excluding the outlier point. Table 34 shows the mean values for bidder specific-outcomes (except for Effort) for the three bidding agent types that were playing against participants in the 1st auctions, for comparison purposes.

Table 33. Participants' initial bidding strategies in 1st round auctions

Initial Behavior	Number of Bidders	Bids	Spans	Effort	TOE	Surplus
Explorer	19	57.00 *(16.97)*	24.89 *(6.72)*	6.29 *(2.12)*	35.11 *(12.65)*	14.75 *(28.86)*
Analyzer	45	21.84 *(11.46)*	8.23 *(4.62)*	17.5 *(8.15)*	73.98 *(39.63)*	49.04 *(39.93)*
Participator	71	29.27 *(8.39)*	12.08 *(3.77)*	10.52 *(3.64)*	37.04 *(12.72)*	31.94 *(41.6)*

Table 34. Bidding agents' bidder-specific outcomes in 1st rounds auctions

Agent Type	Number of agents	Bids		Spans		Surplus	
Analyzer	92	26.88	(8.13)	9.34	(1.98)	27.71	(24.45)
Participator	90	28.44	(8.32)	12.71	(3.07)	20.76	(24.7)
Explorer	88	53.03	(11.93)	20.05	(4.14)	11.51	(16.83)

While it is possible to independently cluster bidders in 2nd and 3rd round auctions, it would not enable a useful interpretation since participants bidding strategy is influenced by their previous experience. Table 35 shows the average change in the four bidder-specific outcome variables based on participant's initial bidding strategy (see Table 28 for average change per treatment group independent of initial bidding strategy) and indicates whether the change is significantly lower/higher than zero using t-tests (i.e., testing null hypothesis that average change is 0).

Table 35. Average change in bidder-specific outcome variables (positive/negative values indicate an increase/decrease) for different initial bidding strategies and its significance based on t-test results

Initial bidding behavior	$\Delta Bids$		ΔSpan$s$		ΔSurplus		ΔEffort	
Analyzer	8.54**	(19.98)	2.02**	(5.48)	-1.07	(33.12)	1.17	(19.44)
Participator	1.21	(17.77)	-0.76	(5.7)	11.29*	(49.57)	4.05*	(14.77)
Explorer	-11.95*	(25.13)	-7.9***	(8.4)	25.46**	(41.96)	4.22**	(6.06)

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1 (standard deviations shown in parenthesis)

As we can see, the change in each of these four bidder-specific variables significantly varies across the three initial bidding strategies. Those with the initial Analyzer bidding strategy increase their number of bids and the diversity of bundles they bid on as they proceed to the 3rd auction, but their derived surplus does not significantly increase; note that this type already derives the highest surplus in the 1st auctions (see Table 33) and there is not much room for improvement. Participants with initial Participator behavior do not significantly change the number of bids and the diversity of bundles they bid on, but they increase their effort and learn how to derive higher surplus. Participants who are initially Explorers significantly reduce the number of bids and variety of bundles they bid on, as they get to the 3rd auctions. They also spent significantly more effort and learn how to derive higher surplus (note that this type derives the lowest surplus in the 1st auctions; see Table 33). Participants with this initial bidding behavior change the most compared to the other two types.

Next, we study how the change in participants behavior is influenced by their initial bidding strategies and the different treatments (different sequences of competition types) they receive. We build separate regression models for the change in Bids, Spans, Surplus, and Effort. Note that with 3 initial types across 9 treatments, there are 27 (unbalanced) groups for the ANOVA analysis, and the limited number of observations can result in models with insignificant explanatory power (in terms of model F-statistic). For each model, the outcome variable is ΔY (the change in variable Y) and the potential predictors and controls are chosen from, treatment (as a categorical variable), participant's initial bidding strategy (A, P, or E), the interaction of these two variables, and the 5 variables presented in Table 24 as controls. We first evaluate straightforward models with only two main effects (i.e., Treatment and Initial bidding strategy as independent variables) and then also use the stepwise regression procedure with variable selection in both directions (both forward and backward selection) to evaluate possible better models in terms of higher adjusted R-square and AIC. Models that have insignificant (or marginally significant) F-statistics are eliminated (when F-test of overall significance is not significant the model is not a better fit than an intercept-only model) and insignificant interaction terms and controls that do not significantly improve the model are dropped. Table 36 shows the final model for each of the four outcome variables and ANOVA results for factors influencing the change in these outcomes.

The models for change in Bids, Spans, and Surplus improve upon the regression models presented in Table 31, however none of the models for change in Effort that include either the Treatment or Initial-Type variable are significant. The ANOVA results in Table 36 show that participants' initial bidding behavior significantly influences their learning. The change in bidders' behavior as they move from the 1st to the 3rd auction is mainly influenced by their initial bidding strategy (significant effect on change in Bids and Spans and marginally significant effect on change in Surplus). The sequence of competitions participants experience (treatment) influences their learning as well, as indicated by its effect on the change in number of bids and change in effort (see ΔEffort model in Table 31).

Table 36. ANOVA for factors influencing the change in bidder-specific outcome variables including participants' initial bidding strategies (p-values are in parenthesis)

ΔBids ~ Treatment + InitialType + Gender + Education				
ΔSpans ~ InitialType				
ΔSurplus ~ Treatment * AuctionExp(Y/N)				
ΔEffort ~ 1				
Outcome:	**ΔBids**	**ΔSpans**	**ΔSurplus**	**ΔEffort**
Variable:	F statistics (p-values shown in parenthesis)			
Treatment	2.09^{*} (0.0418)	-	1.6 (0.132)	None of the models that include Treatment or InitialType are significant
Initial Type (A, P, E)	7.57^{***} (0.0008)	17.68^{***} (0.0000)	2.87^{+} (0.061)	
Gender(M/F)	3.72^{+} (0.0561)	-	-	
Education	2.74 (0.1006)	-	-	
Student(Y/N)	-	-	-	
AuctionExp(Y/N)	-	-	0.029 (0.864)	
Age	-	-	-	
Treatment:InitialType	-	-	-	
Treatment:AuctionExp	-	-	2.28^{*} (0.027)	
R^2	25.04%	21.26%	24.52%	-
Adjusted-R^2	17.61%	20.06%	11.94%	-
Model F statistic	3.368^{***} (0.000)	17.68^{***} (0.0000)	1.949^{*} (0.0166)	-

Significance levels: *** 0.001, ** 0.01, * 0.05, + 0.1

12.6 Initial bidding strategies and auction outcomes

In this subsection, we look at the impact of participants initial bidding strategy together with the baseline treatment competition (agent types) on auction revenue and allocative efficiency. Table 37 shows the mean revenue grouped by competition type and participants' initial bidding strategy as well as average allocative efficiency grouped by competition type across 1st around auctions. Auctioneer's revenue is on average lower when the PP competition type is used as treatment (as compared to AA or EE), particularly when the participant's initial bidding behavior is an Analyzer type. Revenue is generally higher when the EE competition type is used, particularly when the participant's initial bidding behavior is an Explorer type.

Allocative efficiency seems to be highest under the AA competition type and decrease as we use the EE and PP competitions. However, the difference in average allocative efficiency is only significantly different between the AA and PP competition types (based on t-test results); all other pairwise comparisons are not (or only marginally) significant. Under a specific treatment competition type (AA, PP, or EE), allocative efficiency does not significantly vary across different initial bidding strategies.

Table 37. Revenue and allocative efficiency under different competition types for 1st round auctions

Revenue		Initial bidding strategy				Allocative Efficiency
		Analyzer	Participator	Explorer		
Treatment competition type	AA	383.9 (73.5) n = 12	409.9 (39.9) n = 27	431.6 (26.1) n = 7	406.4 (50.7) n = 46	0.963 (0.067)
	PP	346.7 (56.1) n = 14	384.8 (69.3) n = 24	440.1 (41.0) n = 7	381.6 (67.9) n = 45	0.935 (0.061)
	EE	417.8 (55.7) n = 18	429.2 (27.8) n = 20	421.6 (32.9) n = 5	423.5 (41.7) n = 43	0.947 (0.062)

(Standard deviations in parenthesis)

13.Discussion

The experiment results allow us to answer some of the questions we stated at the beginning of this part. Observed bidder behaviors are both an innate trait and can be learned; i.e., different bidder types learn differently. For example, bidders who exhibit an initial Analyzer bidding strategy have on average less room for improvement and learn differently compared to bidders who exhibit an initial Explorer or Participator bidding strategy. The type of competition they experience also influences their learning, but its impact is less than bidder's innate (i.e., initial) bidding strategy. Overall, all types increase the amount of effort they spend and learn how to derive higher surplus.

The type of competition a bidder experiences does not have a negative impact on their perception of the auction platform; responses to all post-experiment survey questions are positive as indicated by the average above 4 values for all variables in Table 27.

Consistent with the findings from the simulation studies in Part II, we see that competition type impacts auction revenue as well as allocative efficiency. Auctioneer's revenue is higher on average when most bidders are of Explorer types as compared to when the majority are of Analyzer or Participator types. Revenue is on average lower when most bidders are Participator types, which also results in lower allocative efficiency. Allocative efficiency is higher when most bidders are Analyzer types as compared to having a majority of the other two types. Knowing that bidders can learn to improve their performance, by guiding bidders toward Analyzer type of behavior we can achieve a win-win scenario where allocative efficiency is maximized; i.e., auctioneers derive high (but not the highest) revenue and bidders earn higher surplus (compared to adopting Participator or Explorer type behavior).

14.Conclusions

The contributions of this work are three-fold. First, we propose and design a data-driven approach for developing software agents that replicate multi-faceted human bidder behavior in complex decision environments of CCAs, which are important, sophisticated market mechanisms that are becoming increasingly used in various business applications. Our study is the first to combine an agent-based modeling approach with machine learning for this purpose; it differs from existing work on automated software agents in that we use machine learning techniques to design agents that *replicate* human behavior, i.e., are not designed to outperform human participants, or optimize a given task. The agent validity is demonstrated by replicating bidder-specific outcome variables that were observed in experimental CCAs with human participants.

Second, we successfully leverage these agents to better understand dynamics of bidder behavior and explore competition types not observed in experimental auctions. The simulations indicate how different competition types affect auction outcomes, such as revenue, and show that different bidder types are affected differently (i.e., in a context-dependent manner) by the type of competition they face in ways that are not always intuitive. We analyzed bidding agents' emergent behaviors under different competition types and discussed the probable underlying mechanisms that lead to these different behaviors and welfare outcomes. In this study, we run simulations with 3 agents in each auction so that the results can be directly compared with results from experimental auctions. However, our simulation platform and knowledge gained from this work provide the groundwork for future agent-based simulation studies with more than 3 bidders (and/or more than 6 items in each auction) to explore how the identified effects of competition can be further generalized to even more complex auction setups.

Third, the capabilities of our modeling approach allowed experimental studies to analyze how human bidder behavior and auction outcomes are affected under different competition scenarios (e.g., when a human bidder competes against Analyzers vs. Participators vs. Explorers vs. a mix of bidding strategies). In these experiments, human participants compete with software agents that exhibit bidding strategies of our choice in a controlled environment, which allowed us to address interesting questions that could not otherwise be answered. The results show that bidder behavior is both innate and influenced by the contextual auction dynamics. Bidders with different initial behaviors perform and learn differently under similar competition scenarios. Moreover, we found that the type of competition participants experience influences their perception of the proposed platforms' usefulness and their tendency to adopt the proposed market mechanism.

Finally, our study has implications for the design and implementation of combinatorial auctions in digital marketplaces and contributes to the better design of smart markets (Bichler et al.

2010) by providing a better understanding of bidder behavior in CCAs and uncovering the effect of competition on bidder behavior and auction outcomes. We show how heterogeneity in auction outcomes is driven by bidders own bidding strategy and the competition they face; for example, certain competition types lead to lower auctioneer revenue as well as suboptimal allocative efficiency. We further show how bidders learn under different competitive environments depending on their initial bidding strategies. Such understanding is necessary for possible customer segmentation and market designs to attract different participant types to certain auctions (i.e., designing incentive mechanisms). Auctioneers can utilize this understanding to incentivize participants to adopt different strategies that would improve the allocative efficiency of auctions and benefit both bidders and auctioneers. This study also contributes to the better design of user-centric artificial bidding agents by developing software agents that demonstrate strategic and human-like behavior in a complex market environment; similar agents can be used to compete with or assist human participants in online auction markets.

Bibliography

Adomavicius, G., Curley, S.P., Gupta, A., and Sanyal, P. 2012. "Effect of Information Feedback on Bidder Behavior in Continuous Combinatorial Auctions," *Management Science* (58:4), pp. 811-830.

Adomavicius, G., and Gupta, A. 2005. "Toward Comprehensive Real-Time Bidder Support in Iterative Combinatorial Auctions," *Information Systems Research* (16:2), p. 169.

Adomavicius, G., Gupta, A., and Zhdanov, D. 2009. "Designing Intelligent Software Agents for Auctions with Limited Information Feedback," *Information Systems Research* (20:4), pp. 507-526.

Adomavicius, G., Sanyal, P., Gupta, A., and Curley, S. 2007. "Design and Effects of Information Feedback in Continuous Combinatorial Auctions," ICIS 2007 Proceedings.

Ausubel, L.M., Cramton, P., and Milgrom, P. 2006. "The Clock-Proxy Auction: A Practical Combinatorial Auction Design,").

Ausubel, L.M., and Milgrom, P. 2006a. "Ascending Proxy Auctions," *Combinatorial Auctions*), pp. 79-98.

Ausubel, L.M., and Milgrom, P. 2006b. "The Lovely but Lonely Vickrey Auction," *Combinatorial Auctions* (17), pp. 22-26.

Ball, M., Donohue, G., and Hoffman, K. 2006. "Auctions for the Safe, Efficient, and Equitable Allocation of Airspace System Resources," *Combinatorial auctions* (1).

Banks, J., Olson, M., Porter, D., Rassenti, S., and Smith, V. 2003. "Theory, Experiment and the Federal Communications Commission Spectrum Auctions," *Journal of Economic Behavior & Organization* (51:3), pp. 303-350.

Banks, J.S., Ledyard, J.O., and Porter, D.P. 1989. "Allocating Uncertain and Unresponsive Resources: An Experimental Approach," *The Rand journal of economics*), pp. 1-25.

Bapna, R., Goes, P., and Gupta, A. 2000. "A Theoretical and Empirical Investigation of Multi-Item on-Line Auctions," *Information Technology and Management* (1:1-2), pp. 1-23.

Bapna, R., Goes, P., and Gupta, A. 2001. "Comparative Analysis of Multi-Item Online Auctions: Evidence from the Laboratory," *Decision Support Systems* (32:2), pp. 135-153.

Bapna, R., Goes, P., and Gupta, A. 2003a. "Analysis and Design of Business-to-Consumer Online Auctions," *Management Science* (49:1), pp. 85-101.

Bapna, R., Goes, P., and Gupta, A. 2003b. "Replicating Online Yankee Auctions to Analyze Auctioneers' and Bidders' Strategies," *Information Systems Research* (14:3), pp. 244-268.

Bapna, R., Goes, P., Gupta, A., and Jin, Y. 2004. "User Heterogeneity and Its Impact on Electronic Auction Market Design: An Empirical Exploration," *MIS Quarterly*), pp. 21-43.

Bichler, M. 2000. "An Experimental Analysis of Multi-Attribute Auctions," *Decision Support Systems* (29:3), pp. 249-268.

Bichler, M., Davenport, A., Hohner, G., and Kalagnanam, J. 2006. "Industrial Procurement Auctions," *Combinatorial auctions*), pp. 593-612.

Bichler, M., Goeree, J., Mayer, S., and Shabalin, P. 2014. "Spectrum Auction Design: Simple Auctions for Complex Sales," *Telecommunications Policy*).

Bichler, M., Gupta, A., and Ketter, W. 2010. "Research Commentary-Designing Smart Markets," *Information Systems Research* (21:4), pp. 688-699.

Bichler, M., Shabalin, P., and Pikovsky, A. 2009. "A Computational Analysis of Linear Price Iterative Combinatorial Auction Formats," *Information Systems Research* (20:1), pp. 33-59.

Bichler, M., Shabalin, P., and Wolf, J. 2013a. "Do Core-Selecting Combinatorial Clock Auctions Always Lead to High Efficiency? An Experimental Analysis of Spectrum Auction

Designs," *Experimental Economics* (16:4), pp. 511-545.

Bichler, M., Shabalin, P., and Ziegler, G. 2013b. "Efficiency with Linear Prices? A Game-Theoretical and Computational Analysis of the Combinatorial Clock Auction," *Information Systems Research* (24:2), pp. 394-417.

Bonabeau, E. 2002. "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems," *Proceedings of the National Academy of Sciences* (99:suppl 3), pp. 7280-7287.

Brunner, C., Goeree, J.K., Holt, C.A., and Ledyard, J.O. 2010. "An Experimental Test of Flexible Combinatorial Spectrum Auction Formats," *American Economic Journal: Microeconomics* (2:1), pp. 39-57.

Carley, K.M. 1996. "Validating Computational Models," *Paper available at http://www.casos.cs.cmu.edu/publications/papers.php*.

Chambers, J.M. 1983. *Graphical Methods for Data Analysis*.

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. 2014. "**NbClust**: An *R* Package for Determining the Relevant Number of Clusters in a Data Set," *Journal of Statistical Software* (61:6). (https://doi.org/10.18637/jss.v061.i06).

Chmielewski, M.R., and Grzymala-Busse, J.W. 1996. "Global Discretization of Continuous Attributes as Preprocessing for Machine Learning," *International journal of approximate reasoning* (15:4), pp. 319-331.

Collins, J., and Gini, M. 2009. "Scheduling Tasks Using Combinatorial Auctions: The Magnet Approach," *Handbooks in Information Systems* (3:4), pp. 263-293.

Collins, J., Ketter, W., and Sadeh, N. 2010. "Pushing the Limits of Rational Agents: The Trading Agent Competition for Supply Chain Management," *AI magazine* (31:2), p. 63.

Cramton, P., Shoham, Y., and Steinberg, R. 2006. "Combinatorial Auctions,").

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), p. 319. (https://doi.org/10.2307/249008).

De Vries, S., and Vohra, R.V. 2003. "Combinatorial Auctions: A Survey," *INFORMS Journal on Computing* (15:3), pp. 284-309.

Demir, G.u., leser K, and Gini, M. 2007. "Winner Determination for Combinatorial Auctions for Tasks with Time and Precedence Constraints," *Journal of Intelligent and Fuzzy Systems* (18:3), pp. 267-280.

Dougherty, J., Kohavi, R., and Sahami, M. 1995. "Supervised and Unsupervised Discretization of Continuous Features," *Machine learning: proceedings of the twelfth international conference*, pp. 194-202.

Fan, J., and Gijbels, We. 1995. "Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation," *Journal of the Royal Statistical Society. Series B (Methodological)*), pp. 371-394.

Gilbert, N. 2008. *Agent-Based Models*. Sage.

Goodhue, D. L. 1995. "Understanding User Evaluations of Information Systems," *Management Science* (41:12), pp. 1827–1844. (https://doi.org/10.1287/mnsc.41.12.1827).

Goodhue, D. L., and Thompson, R. L. 1995. "Task-Technology Fit and Individual Performance," *MIS Quarterly* (19:2), p. 213. (https://doi.org/10.2307/249689).

He, M., Jennings, N.R., and Leung, H.-F. 2003. "On Agent-Mediated Electronic Commerce," *IEEE Transactions on knowledge and data engineering* (15:4), pp. 985-1003.

Hershberger, J., and Suri, S. 2001. "Vickrey Prices and Shortest Paths: What Is an Edge Worth?," *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pp. 252-259.

Hoffman, K., Menon, D., van den Heever, S., and Wilson, T. 2006. "Observations and near-Direct Implementations of the Ascending Proxy Auction," *Combinatorial Auctions*), pp. 415-450.

Holte, R.C. 1993. "Very Simple Classification Rules Perform Well on Most Commonly Used Datasets," *Machine learning* (11:1), pp. 63-90.

Jennings, N.R., Sycara, K., and Wooldridge, M. 1998. "A Roadmap of Agent Research and Development," *Autonomous agents and multi-agent systems* (1:1), pp. 7-38.

Jennings, N.R., and Wooldridge, M. 1998. "Applications of Intelligent Agents," in *Agent Technology*. Springer, pp. 3-28.

Jones, M.C., Marron, J.S., and Sheather, S.J. 1996. "A Brief Survey of Bandwidth Selection for Density Estimation," *Journal of the American Statistical Association* (91:433), pp. 401-407.

Kagel, J.H., Lien, Y., and Milgrom, P. 2010. "Ascending Prices and Package Bidding: A Theoretical and Experimental Analysis," *American Economic Journal: Microeconomics* (2:3), pp. 160-185.

Kaufman, L and Rousseeuw, P. J. 2009. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons.

Kleindorfer, G.B., O'Neill, L., and Ganeshan, R. 1998. "Validation in Simulation: Various Positions in the Philosophy of Science," *Management Science* (44:8), pp. 1087-1099.

Krishna, V., and Rosenthal, R.W. 1996. "Simultaneous Auctions with Synergies," *Games and economic behavior* (17:1), pp. 1-31.

Kwasnica, A.M., Ledyard, J.O., Porter, D., and DeMartini, C. 2005. "A New and Improved Design for Multiobject Iterative Auctions," *Management Science* (51:3), pp. 419-434.

Liu, H., Hussain, F., Tan, C.L., and Dash, M. 2002. "Discretization: An Enabling Technique," *Data mining and knowledge discovery* (6:4), pp. 393-423.

Macal, C.M., and North, M.J. 2005. "Tutorial on Agent-Based Modeling and Simulation," *Proceedings of the 37th conference on Winter simulation*: Winter Simulation Conference, pp. 2-15.

McAfee, R.P., and McMillan, J. 1996. "Analyzing the Airwaves Auction," *The Journal of Economic Perspectives*), pp. 159-175.

McGill, R., Tukey, J.W., and Larsen, W.A. 1978. "Variations of Box Plots," *The American Statistician* (32:1), pp. 12-16.

Miller, J.H., and Page, S.E. 2009. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton university press.

Mueller, R. 2006. "13 Tractable Cases of the Winner Determination Problem," *Combinatorial auctions*), p. 319.

Nautz, D., and Wolfstetter, E. 1997. "Bid Shading and Risk Aversion in Multi-Unit Auctions with Many Bidders," *Economics Letters* (56:2), pp. 195-200.

Naylor, T.H., and Finger, J.M. 1967. "Verification of Computer Simulation Models," *Management Science* (14:2), pp. B-92-B-101.

Park, Y.-H., and Bradlow, E.T. 2005. "An Integrated Model for Bidding Behavior in Internet Auctions: Whether, Who, When, and How Much," *Journal of Marketing Research* (42:4), pp. 470-482.

Parkes, D. 2006. "Iterative Combinatorial Auctions," *Combinatorial auctions*), pp. 41-77.

Pekeč, A., and Rothkopf, M.H. 2003. "Combinatorial Auction Design," *Management Science* (49:11), pp. 1485-1503.

Porter, D., Rassenti, S., Roopnarine, A., and Smith, V. 2003. "Combinatorial Auction Design," *Proceedings of the National Academy of Sciences* (100:19), pp. 11153-11157.

Rassenti, S.J., Smith, V.L., and Bulfin, R.L. 1982. "A Combinatorial Auction Mechanism for Airport Time Slot Allocation," *The Bell Journal of Economics*), pp. 402-417.

Ren, Y., and Kraut, R.E. 2014. "Agent-Based Modeling to Inform the Design of Multiuser

Systems," in *Ways of Knowing in HCI*. Springer, pp. 395-419.

Rousseeuw, P. J. 1987. “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis,” *Journal of Computational and Applied Mathematics* (20:C), pp. 53–65. (https://doi.org/10.1016/0377-0427(87)90125-7).

Rothkopf, M.H., Pekeč, A., and Harstad, R.M. 1998. "Computationally Manageable Combinational Auctions," *Management Science* (44:8), pp. 1131-1147.

Ruppert, D., Sheather, S.J., and Wand, M.P. 1995. "An Effective Bandwidth Selector for Local Least Squares Regression," *Journal of the American Statistical Association* (90:432), pp. 1257-1270.

Sandholm, T. 2002. "Algorithm for Optimal Winner Determination in Combinatorial Auctions," *Artificial intelligence* (135:1), pp. 1-54.

Sargent, R.G. 2005. "Verification and Validation of Simulation Models," *Proceedings of the 37th conference on Winter simulation*, pp. 130-143.

Scheffel, T., Pikovsky, A., Bichler, M., and Guler, K. 2011. "An Experimental Comparison of Linear and Nonlinear Price Combinatorial Auctions," *Information Systems Research* (22:2), pp. 346-368.

Schneider, S., Shabalin, P., and Bichler, M. 2010. "On the Robustness of Non-Linear Personalized Price Combinatorial Auctions," *European Journal of Operational Research* (206:1), pp. 248-259.

Sheffi, Y. 2004. "Combinatorial Auctions in the Procurement of Transportation Services," *Interfaces* (34:4), pp. 245-252.

Shen, Z.J.M., and Su, X. 2007. "Customer Behavior Modeling in Revenue Management and Auctions: A Review and New Research Opportunities," *Production and operations management* (16:6), pp. 713-728.

Van Horn, R.L. 1971. "Validation of Simulation Results," *Management Science* (17:5), pp. 247-258.

Wellman, M.P., Greenwald, A., and Stone, P. 2007. *Autonomous Bidding Agents: Strategies and Lessons from the Trading Agent Competition*. MIT Press.