Essays on the structural analysis of auction markets

by

Marleen Renske Marra

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at

UNIVERSITY COLLEGE LONDON (UCL)

May 2019

Declaration

I, Marleen Renske Marra, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis presents new methods, theoretical results, and empirical findings that contribute to the structural analysis of auction markets. One chapter develops a methodology to study welfare and revenue impacts of fees in auction platforms. The chapter develops and solves a structural model that describes strategic behavior of potential auction platform users. It also exploits an original dataset with 15 months of wine auctions to study these issues. Relevant model primitives are shown to be identified in the auction platform model from observed variation in reserve prices, transaction prices, and the number of bidders. Model estimates reveal significant network effects, which can be harnessed to improve both platform profitability and user welfare. Implications for competition policy are discussed as well. Another chapter focuses on nonparametric set-identification in English auctions with absentee bidding, in which the number of bidders is unknown. The chapter exploits additional identifying variation from drop-out values of absentee bidders and develops a novel nonparametric identification approach based on the stochastic spacing of order statistics. The value of the proposed method is highlighted by showing that it identifies informative bounds on policy-relevant model primitives in practice. Both chapters use original data and methods to further our understanding of auction markets and discuss directions for future research.

Impact statement

The knowledge, analysis, and insights presented in this thesis are of value both academically and for society as a whole. The game-theoretic equilibrium strategies in the two-stage auction platform game, nonparametric identification results, and computationally feasible estimation method can further the analysis of other bidding markets in future research. I am the first to empirically study endogenous entry of both bidders and sellers in an auction platform setting. My analysis serves as foundation for models suitable to address other policy concerns that are unanswered today, such as whether the consolidation of auction platforms benefits users or not, and how pricing affects users in other types of platforms.

This research comes at an opportune time: the absence of methods to quantify network effects in online platforms has been the bottleneck for antitrust policy to catch up on this increasingly salient marketplace. As such, my research pushes the frontier by providing such a method and showing that it is both feasible and of economic importance to account for entry in online platforms when evaluating welfare impacts of fees. The presented research will therefore be informative for competition authorities and regulators, and can be used to assess exact antitrust damages in court cases involving auction platforms.

As a direct benefit to society, I show that adjusting the fee structure on the BidforWine auction platform can make all parties better off. Subsidizing bidders (more) by providing them with a bid discount, and financing this by increasing seller fees, could increase platform profitability by about 30 percent while also making both winning bidders and sellers better off. I am in contact with the platform to disseminate these findings.

Further methodological contributions to the econometrics of auctions include my use of the spacing of order statistics. Both the original use of the spacing concept itself and the positive set-identification results for English auctions with absentee bidding may stimulate productive future scholarship. I show that the method delivers informative bounds on expected consumer surplus in practice, facilitating policy analysis of such auction markets with more limited observables.

Acknowledgments

I would like to express my sincere gratitude to my advisors, whose generous sharing of their time and knowledge made this thesis possible.

Andrew Chesher, your PhD lectures on the study of identification ignited my fascination for that topic and drove my research interests in that direction. It has been an absolute pleasure to learn from and interact with you, also on the few occasions when our shared interest in the arts crowded out the econometrics. I am certain that I also may thank you for your "behind the scenes" championing of my research to your colleagues, even in the years before the job market.

Lars Nesheim, thank you for your help and support in developing this thesis. It was not uncommon for our meetings to turn into brainstorm sessions, or for you to take out pen and paper in order to help me work on some issue at hand. Exemplary of your involvement is that you got to know my research so well that in some practice seminars you formulated sharper answers than I did.

Adam Rosen, you have remained a true mentor even when moving to Duke University. Thanks for your helpful feedback during our meetings at UCL, and for staying involved and helping me progress in my research afterwards. The fact that you aim to give your students the best possible start in their research careers, evidenced also by your supposedly "really great letters" and inviting me to present my work at Duke's micro-econometrics seminar, is truly exemplary.

Philip Haile, my visit to Yale University has been instrumental to the development of this thesis and our bi-weekly meetings were the highlight of that visit. You never ceased to push me to sharpen my thinking or to help me move ahead with my paper. Thank you for keeping me to the same standards as your Yale students. I also greatly appreciate that you continued to encourage me during the job market period.

Áureo de Paula, thank for the countless productive interactions. Most of our chats halted when we (you) arrived at a new weak spot in the paper, and then I would resurface only weeks later with improvements or answers. While it was frustrating at times, this process and your critical questions have significantly benefited the quality of my research. I admire your seemingly immense mental library of relevant papers and results, and I hope to develop that capacity over time as well.

I am also grateful for the support, guidance, and encouragement that I received from other faculty members at UCL. Your substantial feedback around seminars has been a real help, especially during the last stretch of the PhD.

Many thanks also to my friends at UCL and Yale for the many hilarious moments that made the PhD a lighter load. Fernanda, thinking of the day we improved our Python skills to automate submitting Hackney Council noise complaint forms, in all its misery, will make me laugh forever. I am also terribly grateful for all my dear friends in other parts of the world who have supported me along the way despite seeing less and less of me as the PhD progressed.

Finally, I would like to express appreciation to my family. To my father, for his warm continuous support and the passing of his scientific interest onto us. To my mother, for her unwavering belief in me and her inspiring zest for life. To my twin brother who beat me to becoming Dr. Marra first and had the heart to commission my “art” for his thesis cover. *Lieve Marra/Faber en Verink/Obelink families: ontzettend bedankt voor jullie begrip, steun, en interesse in mijn werk.*

Bas, man van mijn dromen, bedankt voor je onuitputtelijke liefde.

For *Renske Faber-Bakker and Gepke Marra-Visser*, in memoriam.

They would have been extremely proud and slightly relieved to see me finish this thesis.

Contents

1 Introduction **17**

2 Related literature **19**
2.1 Structural estimation of auction models 19
2.2 Competing auctions 20
2.3 Endogenous entry 21
2.4 Two-sided markets 21
2.5 Identification in English auctions 22

3 Pricing and fees in auction platforms with two-sided entry **25**
3.1 Introduction 27
3.2 Wine auctions 31
3.2.1 Data collection and description 32
3.2.2 Why listing inspection and seller selection matter 34
3.2.3 Descriptive evidence 34
3.3 A model of an auction platform for idiosyncratic goods 39
3.3.1 Model assumptions and game structure 39
3.3.2 Equilibrium strategies 42
3.4 Nonparametric identification 47
3.4.1 The distribution of bidder valuations 47
3.4.2 The distribution of seller valuations 48
3.4.3 Opportunity cost 49
3.5 Estimation method 50
3.5.1 Auction heterogeneity: homogenizing bids and reserve prices 51
3.5.2 Bidder valuations 52
3.5.3 Seller valuations 53
3.5.4 Incorporating zero reserve prices 55
3.6 Estimation results 57
3.6.1 Impact of observed wine characteristics 57
3.6.2 Estimates of idiosyncratic valuations 58
3.6.3 Entry estimates 61
3.7 Counterfactual policies 62
3.7.1 Harnessing network effects to increase platform profitability 63

3.7.2 Users are better off with alternative fee structures 66
3.7.3 Ignoring entry significantly biases welfare estimates 67
3.7.4 Larger effects in high-value auctions 69
3.8 Conclusions 71
Appendices 74
A Omitted tables and figures 74
B Testing the reserve price approximation 79
C Independent listings: additional results 80
D Further details entry equilibrium 82
E Monte Carlo simulations: a recursive algorithm 89
F Numerical approximation of the entry equilibrium 92
G Omitted proofs 94

4 Nonparametric identification in English auctions with absentee bidding 103
4.1 Introduction 105
4.2 Auction model 107
4.2.1 Absentee bidding 108
4.2.2 Equilibrium strategies 110
4.3 The identifying power of (absentee) bids and spacings 111
4.3.1 Relation of bids to values 111
4.3.2 Sample spacings and shape restrictions 113
4.4 Set-identification of bidder valuations and consumer surplus 116
4.4.1 Set-identification using sample spacings 117
4.4.2 Extension of existing method relying on two order statistics 119
4.5 Nonparametric estimation method 121
4.6 Application of methodology to wine auctions at Sotheby's 123
4.6.1 Data description 123
4.6.2 Testing independent private values 124
4.6.3 Estimation results 126
4.7 Conclusions 128
Appendices 131
H Further details absentee bidding 131
I Monte Carlo simulations 133
J Inference 137
K Omitted proofs 139

Bibliography 146

List of Figures

3.1 Estimated Cumulative Distribution Function transaction prices 37
3.2 Patterns suggesting that additional listings attract additional bidders but the mean number of bidders per listing remains constant 38
3.3 Model fit and validation 60
3.4 The number of bidders per listing is approximately Poisson distributed . . . 62
3.5 Platform revenue at counterfactual commissions, with volume constraint . . 65
3.6 Platform revenue at counterfactual seller fees, with volume constraint . . . 66
3.7 Seller and winning bidder surplus at counterfactual fees with platform objective function constraints 67
3.8 Ullage classification and interpretation 74
3.9 Listing page example: Nuits St George Les Boudouts, Domaine Leroy . . . 77
3.10 The survival probability of product listings 78
3.11 Testing equality of reserve price distribution and approximation 78
3.12 Empirical patterns suggesting that bidders do not cross-bid, thin out, and sellers do not compete 80

4.1 Estimation results 127
4.2 Estimated bounds on consumer surplus by decile of pre-auction price estimate 128
4.3 Partial outcome tree absentee bidders 131
4.4 Simulation results: Tightness of upper bound 134
4.5 Simulation results: Shape assumption & tightness lower bound 134
4.6 Simulation results: estimated bounds on consumer surplus 135
4.7 Simulation results: estimated tighter bounds on consumer surplus 135

List of Tables

3.1 Fee structure in wine auction data . . . 32
3.2 Descriptive statistics: selected auction characteristics . . . 33
3.3 Descriptive statistics: thin markets . . . 35
3.4 Descriptive statistics: non-selective bidder entry . . . 36
3.5 Estimation results: idiosyncratic valuations and entry . . . 58
3.6 Model fit and validation statistics . . . 60
3.7 Welfare impacts of commission changes in comparison to benchmark without endogenous entry . . . 70
3.8 Opposing fee structures and their differential impact on high-value lots . . . 70
3.9 First stage estimation results, main sample . . . 75
3.10 First stage estimation results, high-value sample . . . 76
3.11 Evidence stylized facts: independent listings . . . 81
3.12 Monte Carlo simulations . . . 91

4.1 Descriptive statistics of Sotheby's fine wine auctions . . . 123
4.2 P-values Kolmogorov-Smirnov tests for independence of private values . . . 125
4.3 Estimated consumer surplus by tertile of pre-auction price estimate . . . 126

Chapter 1

Introduction

Economists take an interest in auctions because they are used in practice to allocate anything from treasury bills to highway construction contracts. Also, auctions typically generate better data than other sale mechanisms. Moreover, auctions provide a unique setting to analyze individual or firm behavior with game-theoretic models as they provide relatively controlled bidding and pricing environments with clear rules and well-defined pay-offs. The structural econometric approach to understanding auction markets uses theoretical properties of the auction game to uncover what is unknown (private values) from variation in observables (bids, bidder identities). That involves a study of identification: what can we know about the relevant unobservables given the restrictions of the model? It also involves the question how to consistently estimate those model primitives with limited data? How precise are these estimates? Application of these methods to auction data then allows for the simulation of counter-factual policies to answer questions as: will the market function better if it was designed differently? New results presented in this thesis make progress on all these aspects of the structural analysis of English auctions with independent private value (IPV) bidders. Chapter 2 describes in more detail how the thesis relates to existing research.

Chapter 3 develops a new methodology to analyze welfare and revenue impacts of fees in auction platforms. It solves a two-stage game where a stage of simultaneous bidder & seller entry is followed by an IPV English auction stage. Besides solving for equilibrium strategies in this game, the chapter also discusses nonparametric identification of the latent distributions of bidder and seller valuations, and provides a computationally feasible estimation strategy. It exploits an original dataset collected over a 15 month period from auction platform BidforWine. Counterfactual fee simulations are used to show that the platform could increase its revenues by merely changing how it allocates fees between buyers and sellers. A key result is that even sellers are better off despite paying higher fees, because the entry of additional bidders drives up transaction prices on the platform. Other results indicate that the current antitrust policy benchmark, which abstracts from endogenous entry decisions, significantly underestimates welfare impacts of fee increases. These are especially valuable results as the model is parsimonious enough to apply to auction platforms with idiosyncratic goods in general. As such, the developed methodology can be used to obtain exact welfare impacts of fees in these markets - thereby overcoming a bottleneck for adequate antitrust assessment.

Chapter 4 provides a new method to set-identify the distribution of bidder valuations in IPV English auctions with absentee bidding. It first sets out the information revealed from traditional English auctions with absentee bidding, and relates bid order statistics to valuations. Then, it discusses shape restrictions on the latent valuation distributions and uses results about the spacing of order statistics to bound the distribution of valuations. As bounds can be wide in practice, a tightening approach that relies on exogenous variation in observables is advocated. The value of the method is highlighted by showing that informative bounds on expected consumer surplus are estimated in an original dataset of Sotheby's auctions. With the number of bidders or their final bids unobserved in this dataset, existing identification methods are infeasible.

Conclusions and directions for future research are provided for the two chapters separately.

Chapter 2

Related literature

This thesis contributes to various distinct strands of economic literature. Most importantly it is rooted in the literatures on identification and estimation of model primitives in auction models, and it also adds to empirical analysis of two-sided markets. This chapter provides a wider perspective on the added value of the research presented in this thesis by discussing how it relates to existing work in these fields, but does not aim to provide an exhaustive overview of all relevant research.

2.1 Structural estimation of auction models

Structural estimation of auctions uses econometric models derived from game theoretic representations of the bidding environment to interpret auction data. Due to the transparency of pay-offs and rules in real-world auctions, they provide a tight mapping between equilibrium strategies and observed outcomes (see e.g., Hendricks et al. (2003), Hendricks and Porter (1988, 2007)). The structural approach starts with Paarsch (1992, 1997), Donald and Paarsch (1993, 1996) introducing maximum likelihood, piecewise and constrained maximum likelihood estimation methods. Laffont and Vuong (1993) and Laffont et al. (1995) provide methods based on simulation for settings where the equilibrium bid function is intractable. Guerre et al. (2000) and Li et al. (2002) use two-step nonparametric procedures to estimate the valuation distribution. Haile (2001) uses instrumental variables to overcome endogeneity of the number of bidders. For an excellent introduction to the practical implementation of such methods, see Paarsch and Hong (2006).

The traditional auction setting considers a fixed set of bidders and one seller, in a set of independent auctions in which bidders have independent private information about their willingness to pay (Paarsch and Hong (2006)). This is also the setting in Chapter 4, where it is the more limited set of observables in some of these auctions that complicate identification of the latent valuation distribution. Section 2.5 sketches how my results relate to the literature on identification in English auctions. The next two sections reference papers that abstract from this baseline model by either considering that auctions may not be conducted in isolation (Section 2.2) and that the set of participants can be endogenous (Section 2.3). Literature discussed here and in Section 2.4 relate to the auction platform model with two-sided entry presented in Chapter 3.

2.2 Competing auctions

Peters and Severinov (1997) (and McAfee (1993)) develop a (dynamic) search model where sellers compete with the publicly observable reserve price. Anwar et al. (2006) show that the *cross-bidding* strategy found optimal in Peters and Severinov (2006) (when doing so is costless for bidders) explains bidding behaviour for homogeneous CPUs on eBay. The presence of competing auctions and uncertainty of rival bidder entry both agressive early bidding (entry prevention) and late bidding (learning prevention) in the structural model of Nekipelov (2007) estimated on data from pop-music CDs. Backus and Lewis (2016) provide a path-breaking framework for estimating demand in a large dynamic auction market where bidders substitute across auctions and a steady-state distribution of goods is available. Partly building on that paper, Bodoh-Creed et al. (2013, 2016) develop an empirically tractable search model with a continuum of buyers and sellers, accounting for the endogenous entry of bidders and their dynamic optimization of bids. They apply it to Kindle e-reader auctions on eBay and find that while participation costs are very low in this market, they do regulate entry and generate inefficiencies relative to holding less frequent multi-object sales.

None of the above-listed papers evaluate the impact of fees, nor do they model the entry and selection of *sellers* into the market. Albrecht et al. (2014) do study seller entry and selection in a market with a fixed number of bidders. In their model, seller entry is efficient as a business stealing effect exactly compensates the information rent from providing the additional auction. Consistent with such a model, Newberry (2015) find that in eBay auctions for Corvettes additional listings *thin out* the number of bidders per listing. Reduced form evidence from the wine auctions in my data instead reveal that additional listings attract the same number of bidders, informing the parsimonious model presented in Chapter 3 with simultaneous entry of bidders and sellers, and with costly listing inspection.

Deltas and Jeitschko (2007) also study theoretical equilibrium properties of two-sided entry in relation to the listing fee in an auction platform for the case that bidders and sellers have heterogeneous entry cost. In their model, bidder entry depends on the number of listings also through its impact on the probability that bidders find what they are looking for on the platform.[1] In Chapter 3, I show that the combination of seller selection and listing inspection cost cause the entry equilibrium to be empirically tractable and well behaved because sellers are strategic substitutes despite feedback effects on bidder entry. By contrast, Deltas and Jeitschko (2007) emphasize the instability of their entry equilibria jumping to and from different participation levels for small changes in the listing fee.

2.3 Endogenous entry

Bidder entry and its consequence for how the (single) seller should design its auction is central to a large and growing literature. Endogenous bidder entry is found to be an important and policy-relevant feature of many real-world auction markets, such as in Bajari and Hortaçsu (2003) (eBay coin auctions), Hendricks et al. (2003) (wildcat auctions of oil & gas tracts), Ye (2007) (valuable assets), and Krasnokutskaya and Seim (2011) (highway procurement). Two entry models dominate the literature; the Levin and Smith (1994) "random" entry model where bidders are uninformed about their valuation before they enter and the Samuelson (1985) selective entry models where bidders are fully informed. More flexible entry models nesting both cases are studied in Gentry and Li (2014), Roberts and Sweeting (2010, 2011, 2013), Gentry et al. (2017). Moreno and Wooders (2011) show that when bidders face heterogeneous and private entry cost, the seller revenue-maximizing reserve price and admission fee or subsidy are different from models with homogeneous entry cost or exogenous participation. Other contributions to the structural analysis of auctions with bidder entry include Li and Zheng (2009, 2012), Fang and Tang (2014), and Marmer et al. (2013).

To the best of my knowledge, I am the first to conduct an empirical analysis of auctions with endogenous entry of both buyers and sellers. Both the theoretical results presented in Chapter 3 regarding the simultaneous entry of bidders and sellers into an auction platform with idiosyncratic products and the empirical analysis of this market are novel contributions to this literature.

[1] I do not explicitly model variety although the model can be estimated for different cuts of the data when desired. Bidders are allowed to browse for free and only face entry cost after finding a listing that matches with their high-level preferences for a certain product.

2.4 Two-sided markets

Results in Chapter 3 also contribute to a growing literature providing evidence of policy-relevant interdependencies of users in various two-sided markets, including: Rysman (2004)(yellow pages), Rysman (2007) (payment cards), Gentzkow and Shapiro (2010) (media slant in newspapers), Lee (2011) (home videogames), and Bresnahan et al. (2015) (mobile apps) and Springel (2019) (electric vehicles). One difference with these studies is that in the auction marketplace both sale prices and demand are endogenous to fees set by the intermediary. My structural approach is also fundamentally different from the existing empirical literature on two-sided markets. I use the auction structure to quantify the network effect generated by bidder and seller entry into the auction platform, instead of estimating it using exogenous variation in shifters of demand from the two sides. In a sense, with this approach I provide a *microfoundation* of the entry elasticities and cross-side effects of participation that form the basis of optimal pricing in two-sided markets (e.g. the Lerner-style formula's in Rochet and Tirole (2006) and Armstrong (2006)).

Previously, Gomes (2014) and Athey and Ellison (2011) also combined the two-sided market pricing question in an auction model, but both study position auctions of advertisement space such as that offered by search engines. While Athey and Ellison (2011) focus on the Generalized English auction, Gomes (2014) studies optimal mechanism design.

2.5 Identification in English auctions

Athey and Haile (2002) show that, when adopting the standard "button auction model" of Milgrom and Weber (1982), where bidders exit observably and irreversibly as the price rises exogenously, and when bidders' private values are independent, the transaction price and number of bidders identify the latent value distribution. Haile and Tamer (2003) relax the behavioral restrictions of the button auction model and show that this information set-identifies the distribution. Song (2004) previously considered identification in IPV English auctions when the number of bidders is unobserved, specifically exploiting the fixed end time in eBay auctions. The focus of Chapter 4 is instead on the more traditional English auction setting with a flexible end time, where the auction only closes after providing bidders ample opportunity to raise the standing bid ("going, going, gone"). The method in Song (2004) does not apply directly to traditional English auctions as observables can only bound the third-highest valuation.

Generally, identification of the valuation distribution is not immediate in oral, open-outcry auctions as the drop-out value of the highest bidder is never observed. An interesting development in recent structural analysis under weaker restrictions than IPV is therefore to focus on directly identifying the policy variables that are ultimately of significance. In a symmetric private value environment where valuations are allowed to be correlated, Aradillas-López et al. (2013) set-identify the expected consumer surplus and expected seller revenue when only the transaction price and the number of bidders is observed. Coey et al. (2017) extend these results to asymmetric bidders and furthermore show that the bounds can be sharpened when bidder identities are also observed. In first-price auctions with unobserved heterogeneity, Armstrong (2013) shows that just observing the transaction price and the number of bidders is sufficient to set-identify features related to expectations of valuations such as consumer surplus. Tang (2011) bounds seller revenue in first price auctions with common values. What these papers have in common is that they investigate what economic features of interest can be (set-)identified in auctions when relaxing the auction model in important dimensions but without imposing additional parametric structure on the problem. Results presented in Chapter 4 contribute to this line of research, combining a concept from the statistics literature (*sample spacings*, see Pyke (1965, 1972)) with additional information revealed from absentee bids.

Chapter 3

Pricing and fees in auction platforms with two-sided entry

Marleen R. Marra

Abstract

Auction platforms are increasingly popular marketplaces that generate revenues from fees charged to users. The platform faces a "two-sided market" with network effects; increased seller entry raises its value to bidders, and vice versa. This means that both the platform revenue-maximizing fee structure and welfare impacts of these fees are ambiguous. I examine these issues with a new data set of wine auctions using a model with endogenous bidder and seller entry, seller selection, and costly listing inspection. I show that relevant model primitives are identified from observed variation in reserve prices, transaction prices, and the number of bidders. My estimation strategy combines methods from the auction and discrete choice literatures. Model estimates reveal significant network effects, which can be harnessed to improve both platform profitability and user surplus. Decreasing (increasing) the buyer premium (seller commission) by 15 percentage points increases platform revenues by about 30 percent. It is striking that, in the face of such fee changes, even sellers are better off as additional bidder entry drives up transaction prices. I also estimate that welfare impacts from increasing fees individually are about twice as high as when abstracting from endogenous entry and that 70-90 percent of the loss falls on sellers.

3.1 Introduction

Auction platforms provide increasingly popular marketplaces for trading goods and services, ranging from freelance jobs to vehicles to oil and gas drilling rights. Examples include: eBay.com, BidforWine.co.uk, ClassicCarAuctions.co.uk, CarsontheWeb.com, EnergyNet.com, Upwork.com, uShip.com and ComparetheManwithVan.co.uk. These auction platforms generate revenues from fees charged to buyers and sellers: eBay generated over 2 billion dollars in fee revenues in Q3 2018.[1] The platform faces a "two-sided market" with network effects given that it is more valuable to potential bidders when more sellers enter, and vice versa. This generates complications for the platform or a benevolent planner when determining how to optimally allocate fees among users. In fact, the two-sided market literature highlights that both 1) the platform revenue-maximizing fee structure, and 2) welfare impacts of those fees are theoretically ambiguous and depend on the magnitude of network effects.[2]

To study these two issues, I exploit a new data set of wine auctions and develop a structural model in which network effects arise from endogenous bidder and seller entry. A key innovation is that I leverage the transparency of payoffs in the auction game to characterize network effects in this setting.[3] This allows me to provide a tight quantitative analysis of how fee changes affect both platform profitability and user welfare. My wine auction data is representative of auction platforms for idiosyncratic goods for which bidders and sellers have private information about their willingness to pay.[4] As storage conditions and provenance of these "fine, rare, and vintage wines" are important descriptors of their quality, it is costly for bidders to inspect each listing. Empirical patterns in the data, including thin markets and independent listings, are consistent with listing inspection cost and set this environment apart from previously studied auction platforms for more homogeneous goods.[5]

While a significant literature examines implications of costly bid preparation or value discovery in auctions, it addresses markets in which a single seller can influence bidder

[1] https://investors.ebayinc.com/fast-facts/default.aspx

[2] See Rochet and Tirole (2003, 2006), Evans (2003), Wright (2004), Armstrong (2006).

[3] Previous studies that introduced a two-sided market pricing question in an auction framework are Athey and Ellison (2011) and Gomes (2014), studying position auctions.

[4] This also motivates the use of auctions rather than more convenient posted prices as the selling mechanisms on the platform. See Milgrom (1989) and Wang (1993) on auctions versus posted prices. Some auction platforms offer both auctions and posted price listings. Einav et al. (2018) find that on eBay.com, where sellers can choose between the two mechanisms, auctions are typically selected by less experienced sellers and for goods that are used or more *idiosyncratic*. This motivates my use of this term.

[5] Previous auction platform models include: Anwar et al. (2006), Peters and Severinov (1997) (see also Albrecht et al. (2012)), Nekipelov (2007), Backus and Lewis (2016) and Bodoh-Creed et al. (2013, 2016). None of these evaluate the impact of fees.

entry through optimal auction design.[6] My emphasis on selective seller entry is novel to the empirical auction literature. It generates an additional trade-off that is relevant for answering the key questions in this paper. Bidders expect lower (reservation) prices when lower-value sellers are attracted to the platform, so bidder entry depends both on the expected number and type of sellers that enter. The importance of this dynamic was first postulated in Ellison et al. (2004) but never implemented in practice. The authors hypothesize that a major reason why Yahoo! and Amazon were unsuccessful as auction platforms was their zero listing fee policy: this attracted many nonserious sellers with high reservation prices that in turn shunned bidders from their platforms. My structural model addresses this mechanism and allows me to incorporate the seller selection channel in evaluating the role of fees.

In line with the wider empirical auction literature, I exploit the relatively controlled auction environment where strategic interactions and resulting payoffs are accurately described by Bayes-Nash equilibrium properties of an incomplete information game.[7] The observed distributions of reserve prices, transaction prices and number of bidders are endogenous to the fee structure through optimal entry, bidding and reserve pricing strategies. Variation in outcomes allows for the estimation of model primitives needed to answer how fees affect user welfare in this market. As such, the wine auctions provide an opportune setting to understand the otherwise hard to quantify network effects by tracing fees through the auction platform game.

The introduction of seller selection in the auction platform model does introduce empirical challenges regarding nonparametric identification and estimation of the population distribution of seller valuations. I demonstrate that using the first order optimality condition of equilibrium reserve prices the relevant distribution of seller valuations is identified for any counterfactual fee policy that reduces expected seller surplus. Every reserve price maps to a valuation for all sellers that entered the platform, given identification of the distribution of bidder valuations from the observed second highest bid and number of bidders according to Athey and Haile (2002).[8] Only a subset of sellers currently listing

[6]See Ye (2007), Roberts and Sweeting (2010), Moreno and Wooders (2011), Krasnokutskaya and Seim (2011), Li and Zheng (2009, 2012), Fang and Tang (2014), Marmer et al. (2013), Gentry and Li (2014) and Gentry et al. (2015, 2017).

[7]See Hendricks and Porter (2007) on the close links between auction theory, empirical practice and public policy.

[8]My identification result requires the assumption that all relevant auction-level heterogeneity is observed. This is plausible given that the web scraping algorithm arguably delivered the same observables that bidders get to see when bidding on the wine. Related are Roberts (2013) and Freyberger and Larsen (2017) who use the reverse approach: variation in reserve prices traces out unobserved heterogeneity assuming that reserve prices and bids respond to common factors unobserved to the econometrician. To do so, Roberts (2013) assumes that sellers are homogeneous. Freyberger and Larsen (2017) do have het-

on the platform would enter for any counterfactual world in which platform entry is less profitable for sellers. The positive identification result does not apply for fee structures that make seller entry more profitable because a typical sample selection problem causes observables to be uninformative about valuations among sellers that did not enter.

The entry equilibrium is the unique solution to a fixed point problem in seller valuation space with a nested zero profit entry condition on the bidder side. This complicates estimation of the distribution of seller valuations because: 1) the support of the distribution of reserve prices depends on parameters and 2) the equilibrium is costly to compute for each set of candidate parameters. To address these issues I first obtain an initial estimate based on a concentrated likelihood using a consistent estimate of the entry threshold, suggested in Donald and Paarsch (1993, Footnote 4) for a similar support problem. I then solve the game once and re-estimate seller parameters based on the updated entry threshold. This algorithm is based on the Aguirregabiria and Mira (2002, 2007) nested pseudo likelihood method to solve estimation problems involving fixed point characterizations in (dynamic) games. In my case, the algorithm uses the auction structure to obtain seller parameters from a first order condition. The estimation of bidder parameters uses standard methods from the empirical auction literature, involving a first stage that controls for auction heterogeneity following Haile et al. (2003) and maximum likelihood estimation of parameters from the homogenized bidder valuation distribution as in e.g. Donald and Paarsch (1993) and Paarsch (1997).

Model estimates reveal significant network effects in this platform, which can be harnessed to improve both platform profitability and user welfare. I estimate that platform revenues can increase by up to 80 percent without reducing sale volume when implementing fee structures that subsidize buyers (more) by reducing the buyer premium while at the same time increasing the seller commission.[9] As the buyer premium is currently zero it requires providing winning bidders with a discount on the transaction price. This fully agrees with the idea that businesses in two-sided markets should subsidize the side that contributes most to profits, even if this results in negative fees.[10] Counterfactual experiments also demonstrate that all parties benefit with the adoption of some of these fee structures. For example, combining a 15 percent buyer discount with a 15 percentage point increase in seller commission and an increase in listing fee from 1.75 to 5 pounds

erogeneous sellers; the reserve price is additive in the common unobserved factor and an idiosyncratic seller-specific term and they use deconvolution to separately identify the two components.

[9] I solve for fee revenues with my static game and impose volume constraints to capture that current volume likely affects future revenues through e.g. brand familiarity or word of mouth. This approach avoids having to make stronger assumptions about the exact dynamic platform objective function.

[10] Rochet and Tirole (2003, 2006), Evans (2003), Wright (2004), Armstrong (2006).

increases platform revenues by about 30 percent. But even sellers are about 20 percent better off in this scenario despite a significant increase in seller fees. This is because the buyer discount attracts additional bidders, driving up transaction prices in the auction mechanism.

In a second set of counterfactual exercises, I focus on welfare impacts from isolated increases in buyer or seller commission. A key finding is that sellers are better off if their seller commission is increased by 5 percentage points than in the case when the buyer commission increases by the same amount. This feature of the platform setting, driven by network effects, would be missed if bidder and seller participation is considered exogenous. The magnitude of welfare impacts is also striking. For example, a 5 percentage point increase in buyer (seller) commission reduces expected surplus for winning bidders by 7 (4) percent and for sellers by 17 (15) percent. These results demonstrate that abstracting from endogenous entry and strategic interactions between platform users, as has been the norm in antitrust policy, significantly biases estimated welfare impacts of changes in the fee structure.

Wine auctions are a particularly relevant market in this context because two mayor players, auction houses Sotheby's and Christie's, have been found guilty of commission fixing in the mid-90s. Using this case for context, I estimate that it is plausible that the true antitrust injury to both parties would have been about double the estimated damages underlying the settlement of 512 million dollars (roughly 729 million dollars in 2018 prices). Especially sellers would be undercompensated: while they received only one sixth of the total settlement, about 70-90 percent of estimated damages falls on sellers regardless of which side the commission increase is charged to.

My empirical findings underscore the idea that economic principles underlying regulation in traditional markets do not necessarily apply to two-sided markets and that both sides should be evaluated in tandem. A competitive auction platform could combine high fees on one side of the market with below marginal cost prices on the other side. Both practices could be considered predatory when evaluated in isolation but they prove to be socially optimal in the two-sided market in this paper. In recent years also competition authorities and courts recognize that regulation of platform markets requires different tools and tailored solutions, but the perceived difficulty to quantify user interactions has been a bottleneck for practical application of these ideas.[11]

The rest of this paper is organized as follows. Section 3.2 provides institutional details

[11] See e.g. In re eBay Seller antitrust litigation (2008), Bomse and Westrich (2005), Tracer (2011), OECD Competition Committee (2009, 2017), Evans and Schmalensee (2013).

about online wine auctions, presents the data and empirical patterns that distinguish it from previously studied homogeneous good auctions. Section 3.3 sets out the theoretical auction model and solves for equilibrium entry, bidding, and reserve price strategies. Section 3.4 explains how to identify model primitives from available data. Details about the estimation approach are presented in 3.5 and results in 3.6. Section 3.7 presents results from counterfactual fee policies that shed light on network effects, the economic incidence of fees, and platform profitability. Concluding remarks are offered in Section 3.8.

3.2 Wine auctions

Fine wine is sold at auction in secondary markets, run by online wine platforms as well as brick-and-mortar auction houses.[12] Auction data for the empirical analysis in this paper comes from online auction platform: www.Bidforwine.co.uk (BW). It offers a marketplace for buyers and sellers to trade, akin to the eBay consumer-to-consumer format.[13] When sellers create a listing they choose the auction duration, whether or not to increase the minimum bid amount or to set a secret reserve price. They also provide wine characteristics and description, and information on delivery and insurance. When the sale is successful, they receive payment from the winning bidder, ship the wine, and receive an invoice for the amount of seller commission due. For these seller-managed lots, BW charges no buyer premium and maintains a seller commission on a sliding scale between 8.5-5.5 percent of the sale price (see Table 3.1). Upfront charges to sellers are: a 1.75 pounds listing fee, a 0.50 pounds minimum bid fee (optional, if increased), and a 0.25 pounds reserve price fee (optional, if set).

Lots are sold through an English auction mechanism with proxy bidding. Bidders submit a maximum bid and the algorithm places bids to keep the current price one increment above the second-highest bid.[14] A soft closing rule extends the end time of the auction by two minutes whenever a bid is placed in the final two minutes of the auction. There-

[12]The major platforms sold for 338 million dollars of wine in 2016, and have also been burgeoning in 2017 and 2018 (Wine Spectator (2017a,b, 2018)) The biggest players in 2016 were: Sotheby's (74 million), Zachys (66 million) and Acker, Merrall & Condit (59 million).

[13]Such seller-managed listings are the focus of this paper. They are distinct from wines consigned to the platform and sold on behalf of the seller, especially because they do not undergo quality control by the platform. BW only offers consignment services when selling a "large collection", roughly exceeding five cases, and charges higher fees for these auctions.

[14]When the highest bid is less than one increment above the second highest bid, the transaction price remains the second highest bid. This is different from the rule at eBay, where the standing price in that case would increase to the highest bid. Engelberg and Williams (2009), Hickman (2010) and Hickman et al. (2017) assess implications of this alternative bidding rule that is practically a mix between a first-price and second-price auction.

Table 3.1: Fee structure in wine auction data

	Notation	Amount / rate		Conditional on selling
Bidders:				
buyer premium	c_B	0		✓
Entry fee	e_B	£0		
Opportunity cost of time	e^o_B	estimated		
Sellers:			On part transaction price:	
Seller commission	c_S	0.085	$\leq$ £200	✓
		0.075	£200.01- £1500	✓
		0.066	£1500.01- £2500	✓
		0.055	$\geq$ £2500.01	✓
Listing fee	e_S	£1.75		
Reserve price fees	e_R	£0.75		
Opportunity cost of time	e^o_S	estimated		

Source: www.bidforwine.co.uk. Displayed fees exclude 20 percent VAT, which are included in estimation. Opportunity cost e^o_B and e^o_S are added for reference but fall outside the platform fee structure $f = \{c_B, e_B, c_S, e_S, e_R\}$. As described in the text, the reserve price fee is made up of 0.50 pounds for raising the minimum bid and 0.25 pounds for adding a secret reserve price. Different fees apply to lots consigned to and sold by the platform on behalf of sellers.

fore, there is no opportunity for a *bid sniping* strategy (bidding in the last few seconds, potentially aided by sniping software) on the BW platform.[15]

3.2.1 Data collection and description

I constructed a dataset of wine auctions by web-scraping all open auctions on BW at 30-minute intervals between January 2017 and May 2018.[16] At these intervals, I observe everything that bidders observe as well. This data collection effort resulted in a wealth of data, including: the number of bidders and bids, the current standing price, the identity of the seller, and feedback from earlier transactions. Only a quarter of listings is created by a seller with feedback, pointing to the consumer-to-consumer nature of the platform.

The repetitive recording of bids for ongoing auctions was necessary to approximate the reserve price distribution. When the seller sets a reserve price without making it public in the form of a minimum bid amount, the notifications “reserve not met” or “reserve almost met” accompany any standing price that does not exceed the reserve. I approximate the reserve price as the average between the highest standing price for which the reserve price is not met and the lowest for which it is met.[17] While only 26 percent of listings has an

[15]See Ockenfels and Roth (2006) on strategic behaviour in auctions with these two types of closing rules and Hasker and Sickles (2010) and Bajari and Hortaçsu (2004) for an overview of various explanations for bid sniping evaluated in the literature.

[16]The exact data collection times depended on when the scraping job got scheduled on the cluster, also affected by computing node failures. An example listing page is provided in Figure 3.9.

[17]If all bids would be recorded in real time, this approximation would be accurate up to half a bidding increment due to the proxy bidding system. Appendix B presents suggestive evidence that also the 30-

increased minimum bid amount, 44 percent has a (secret) reserve price, and 3 percent has both. The use of secret reserve prices in auction platforms remains a puzzle in the empirical auction literature and solving that puzzle is beyond the scope of this paper.[18] In the rest of this paper I group them together and refer to the "reserve price" as the maximum of: the minimum bid amount and the approximated secret reserve price. Of larger consequence is the choice made by a third of sellers to refrain from setting any form of reserve. This is observable to bidders by a "no reserve price" button - even before they enter the listing. The BW website encourages sellers to set no reserve price with the following argumentation: *"Bid for Wine's own statistics show that lots listed without reserve prices typically attract 50-75% more bidders and sell for up to 40% more than those with reserves."* Section 3.5.4 endogenizes the choice to set no reserve price in tandem with higher (optimal) bidder entry into no-reserve listings.

I also observe wine characteristics such as the type of wine, grape, vintage, region of origin; plus the textual description, delivery and payment information. Basic summary statistics are reported in Table 3.2. While there is a significant range in sale prices, 84 percent of all sales in the sample do not exceed the 200 pounds over which sellers pay a higher marginal seller commission. The sample includes $3,487$ auctions after excluding auctions with a "buy-now" option, that are consigned, sell spirits, or sell multiple lots at once.

3.2.2 Why listing inspection and seller selection matter

Wine sold at auction is often described as *fine, rare, and vintage wine.* A key difference with retail wines is that they are sold by individual collectors who stored the bottles either in professional warehouses or in private cellars - sometimes for decades. Sellers therefore know how much the wine is worth to them and they have their own idiosyncratic value (taste) for it. When the platform changes its fee structure, it therefore affects both the number and the type of sellers that enter. Moreover, this feeds back on how attractive the platform is for potential bidders given that more serious sellers with lower valuations set lower reserve prices.[19] This is the first paper to estimate an auction platform model with (selective) seller entry.

Listing inspection cost arise in this context because all offered wines are different. This has to do with why there is a flourishing secondary market in the first place. The

---

minute scraping interval result in a good approximation of the reserve price distribution.

[18]See e.g. Jehiel and Lamy (2015) and Hasker and Sickles (2010).

[19]Ellison et al. (2004) hypothesize that seller selection was likely a main driver for why auction sites of Amazon and Yahoo! struggled: their zero listing fees attracted non-serious sellers with high reserve prices, shunning bidders.

Table 3.2: Descriptive statistics: selected auction characteristics

	N	Mean	St. Dev.	Min	Median	Max
Transaction price	3,487	140.56	239.94	1.00	82.50	6,000.00
Is sold	3,487	0.64	0.48	0	1	1
Number bottles	3,487	3.70	4.22	1	2	72
Price per bottle if sold	2,230	74.84	124.52	0.50	35.00	2,200.00
Number of bidders	3,487	3.10	2.52	0	3	13
Seller has feedback	3,487	0.29	0.46	0	0	1
Has reserve price	3,487	0.44	0.50	0	0	1
Has increased minimum bid	3,487	0.26	0.44	0	0	1
Textual description:						
- related to storage conditions	3,487	0.17	0.38	0	0	1
- related to delivery	3,487	0.58	0.49	0	1	1
- related to *en primeur*	3,487	0.17	0.38	0	0	1
- related to expert opinion	3,487	0.51	0.50	0	1	1
Number of words in description	3,487	84.22	79.11	1	65	851

Textual description statistics are obtained using text mining with count-based evaluation. The dummy variables equal one if it contains a word that is associated with respectively "stored", "delivery", "primeur", or "parker" (referring to wine advocate Robert Parker who maintains a 50-100 point scale for fine wines); the minimum association threshold is a Pearson correlation of at least 0.3.

paramount influence of weather and harvesting conditions results in some vintages outperforming others in terms of quality.[20] Older wines can be valuable as increased scarcity of these star vintages drives up prices, given that fewer of them remain uncorked over time. Moreover, certain high-tannin wines such as red Bordeaux age well and are thought to reach their full potential only after many years. But the commodities are also perishable so that humidity and temperature control are key to deliver this potential quality. As such, assessing the wine's idiosyncratic storage conditions, provenance, *ullage* and other indicators of wine quality make it costly for bidders to bid in every auction they enter.[21]

Conceptually, also auctions of other idiosyncratic products such as second hand cars, freelance jobs or house moving trips likely involve costly listing inspection by bidders. While previous empirical studies investigate auction platforms for more homogeneous goods, this is the first to focus on goods with listing inspection cost.[22]

[20] Ashenfelter et al. (1995) and Ashenfelter (2008) predict with surprising accuracy the value of high-end Bordeaux wines using weather data from the growing and harvesting seasons.

[21] *Ullage* describes the unfilled space in a container; in wine auctions it refers to visible oxidation of the wine. For example, a "Base of Neck" fill level is better than "Top Shoulder". These apply to wines in Bordeaux-style bottles with a visible neck and shoulders; a metric classification is used for Burgundy-style bottles (see Figure 3.8).

[22] In previous literature, auction platform models are estimated using data from Kindle e-readers (in Bodoh-Creed et al. (2013, 2016)), indistinguishable CPU's (in Anwar et al. (2006)) and pop CD's (in Nekipelov (2007)).

Table 3.3: Descriptive statistics: thin markets

Number of times a product is listed									
Per market, percentiles									
10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	1	1	1	1	2	2	3	6	37
Total over 15 months, percentiles									
10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1	3	8	16	28	37	68	148	215	223

This table is based on conservative product-market specifications. In this table, products are combinations of: region of origin, wine type, and vintage decade; markets are 4 week intervals. The Independent listings section on page 36 describes other specifications used.

3.2.3 Descriptive evidence

Here, I document four empirical patterns related to the idiosyncratic nature of the goods.
1) Thin markets. The data reveals a strikingly low number of identical products per market, even when using conservative product / market definitions. All listings are active for at most 31 days, and most sellers pick the pre-set 5, 7 or 10 day duration. In this paragraph, I therefore use conservative one month periods to define a market. The BW site has filters for high level characteristics corresponding to the idea that potential bidders enter the site with at least a rough idea of the product they are looking for. As product specification, I take the combination of three high level filters: i) region of origin, ii) vintage decade and iii) wine type. For example, a red Bordeaux from the 1980s and a non-vintage Champagne are distinct products by that definition. Even with these relatively coarse product-market specifications, for 50 percent of listings this is the only one of that product offered in that market and for another 20 percent there are only two of these products available (see Table 3.3). Half of the products have been listed only 28 times during the full 15 months spanning my data, conditional on having been offered at least once.
2) Non-selective bidder entry. Whether bidders do or don't know their valuation for the listed products before they enter the platform is crucial in the way bidder entry affects outcomes.[23] Which case is likely to describe my data generating process can be tested; in a selective entry model valuations are lower in the first order stochastic dominance sense when more bidders enter the platform. Estimates presented in Figure 3.1 contest such a selective bidder entry process. It shows that estimated distributions of second-highest bids

[23] If they are fully informed before they enter, as in the Samuelson (1985) selective entry model, every additional bidder has a lower valuation so entry affects prices less than when they don't know their valuation when entering, as in Levin and Smith (1994). Roberts and Sweeting (2010) and Gentry and Li (2014) capture both scenario's as polar cases in their flexible entry models.

Table 3.4: Descriptive statistics: non-selective bidder entry

	Dependent variable: Transaction price (OLS)			
Number bidders in auction	26.650***	[1.730]	25.998***	[2.024]
Total number bidders product/market	−0.155	[0.109]	−0.282	[0.248]
Product fixed effects	No		Yes	
Observations	1,218		1,218	
Adj. R^2	0.163		0.143	

$***$: Significant at the 1% level, standard errors in square parenthesis. In this preliminary analysis, product fixed effects here are high-level observables: wine type, region of origin, decade of production; markets are 4 week intervals.

are similar for above-median and below-median bidder platform entry, evaluated separately for auctions with 2-9 bidders. [24] The same conclusion can be drawn from a reduced-form OLS regression of transaction prices on the number of bidders in the auction *and* total number of bidders on all comparable listings in the same market, also when controlling for product fixed effects in Table 3.4. Reported patterns are consistent with non-selective bidder enter and suggest that an extra bidder in an auction is associated with a transaction price that is on average 26-27 pounds higher. These documented empirical patterns are consistent with bidders needing to inspect a listing *before* learning specifics of the wine and how much they value these specifics.

3) Independent listings. Previous papers show that transaction and reserve prices in homogeneous good auctions can be affected by the number of competing listings.[25] An ordinary least squares regression analysis, detailed in Appendix C suggests that, in the wine auction data, listings are not systematically related despite ending in close proximity of each other and offering similar items. This conclusion is robust to using different

[24]Distributions are estimated from auctions without a reserve price. They are obtained with fitting nonparametric epanechnikov Kernels with optimal cross-validated bandwidths on transaction prices from auctions with below or above median bidder participation. For example, it compares transaction prices in months where non-vintage Champagne is more popular - in the sense of attracting more total bidders on this type of wine - with months where there is lower bidder interest for these wines. Not observing the pool of potential bidders precludes me from testing selection on obsevables directly as done in Roberts and Sweeting (2011, 2013).

[25]Peters and Severinov (1997) and Anwar et al. (2006) consider *cross-bidding*, motivated by the absence of listing-specific entry cost in auctions for homogeneous products. The incremental cross-bidding strategy requires bidders to always submit a bid on the auction with the lowest standing price, and only one increment above the standing price (e.g., not submit a bid once equal to their valuation). As I don't observe bidder identities, I cannot examine the incremental cross-bidding strategy directly, but a strong clue for the absence of it is that on average all bidders place only 1.7 bid (median: 1.5) so at least it cannot be a very prevalent strategy. To wit, not every bidder can be placing two bids or more in the same auction (but they could still bid once or twice in many competing auctions - if available). Another suggestion is that the majority of winning bidders that left feedback has only won an auction (and left feedback on it) once or twice (58 percent) over the entire 15 months period.

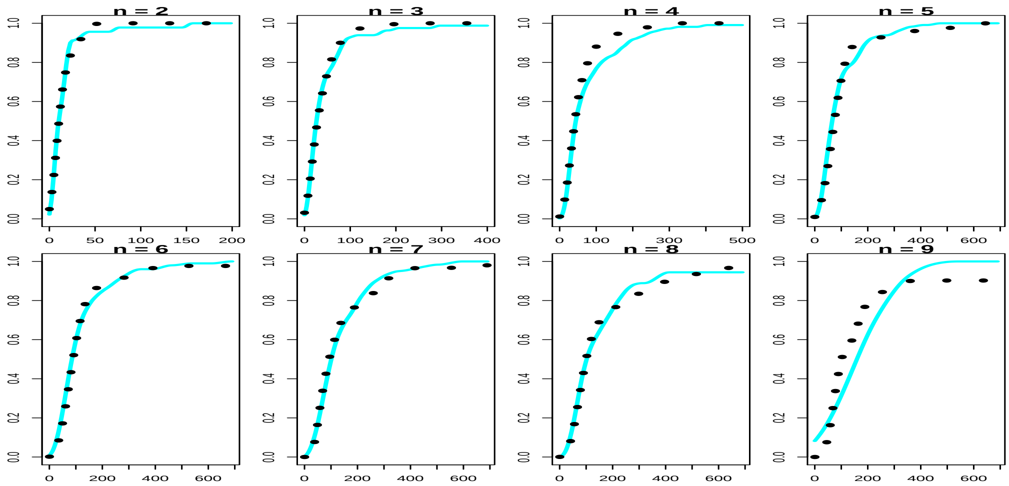


Figure 3.1: Estimated CDF transaction prices; x-axis: probability. y-axis: value

Black dotted line: estimated second-highest bid distribution for **below-median** total number bidders per product/market. **Cyan solid line**: estimated second-highest bid distribution for **above-median** total number bidders per product/market.
In these graphs, products are combinations of: region of origin, wine type, and vintage decade; markets are 4 week intervals. Figures are based on data from auctions with no reserve price in which the number of bidders is directly observed. Plots displayed by number of bidders per auction, $n = \{2, .., 9\}$. Sample sizes are too small to do this for n=10-13. The Non-selective bidder entry section on page 35 describes further details.

product and market specifications. Dependent variables analyzed are: i) the number of bidders per listing, ii) the number of bids per bidder, iii) the transaction price and iv) the reserve price. The results rely on cross-market variation in the number of listings of a certain product. The different market specifications considered are all auctions ending within a rolling window of: i) 30 days, ii) 7 days, and iii) 2 days of each other. Product specifications also vary. The coefficient on competing listings is in 68 out of the resulting 72 regressions statistically insignificant at the 10 percent level. To rule out that there are non-linear effects, the absence of a clear structural relation between the number of (competing) listings and these outcomes of interest is also confirmed with data visualizations.

The fact that reserve prices are not affected by competing listings is intuitive since most of them are kept secret. As bidders cannot select on what they cannot observe, there is no motive for sellers to compete on that margin. The absence of a *cross-bidding* strategy, as suggested by the constant number of bids per bidder, can be explained by the accumulation of listing inspection cost associated with that strategy. Overall, the fact that transaction prices do not decrease with the number of competing listings points to the absence of a "business stealing" effect and is also consistent with bidders entering and

a) Total bidders b) Bidders per listing

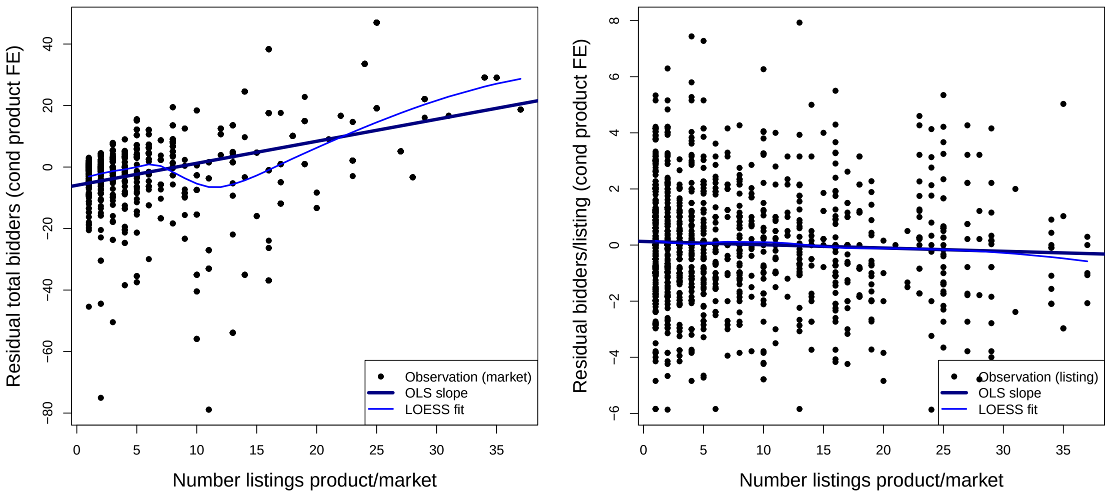


Figure 3.2: Patterns suggesting that additional listings attract additional bidders but the mean number of bidders per listing remains constant

Figures are based on data from auctions with no reserve price in which the number of bidders is directly observed. The blue solid lines represent the estimated coefficients in OLS regressions: on the left a slope of 0.7 (statistically significant at the 1 percent level) and on the right an insignificant 0. The residual total bidders in a) is obtained from a linear regression of this outcome in market m on product dummies and the residual bidders per listing in b) is obtained from a linear regression of this outcome for product p in market m on product dummies. The left-hand graph shows that, for example, markets with more listings of non-vintage champagne attract more bidders on non-vintage champagne listings while the right-hand graph suggests that bidders enter only to keep the mean number of bidders on non-vintage champagnes constant across markets. In these graphs, products are combinations of: region of origin, wine type, and vintage decade; markets are 4 week intervals.

bidding in one listing at a time.[26]

4) Network effects. Network effects describe that a product is more valuable to a group of users when it is more widely adopted by another group.[27] In our auction platform setting, network effects arise mechanically from the fact that transaction prices are endogenous to the number of bidders per listing. As bidders sort over available listings, a platform with more listings is more attractive to potential bidders c.p., and vice versa. This positive *feedback effect* is observed from the positive correlation between the number of total bidders and the availability of listings after controlling for product fixed effects (left-hand panel of Figure 3.2). The pattern also persists when controlling for a time trend.

An equilibrium prediction from a model in which (reserve) prices are unaffected by

[26]In contrast, Newberry (2015) show that in eBay auctions for Corvettes more listings result in the thinning of bidders per listing. The constant mean number of bidders per listing in my data disproves bidder thinning.

[27]See Katz and Shapiro (1985) and Rochet and Tirole (2006).

the number of listings, as shown in the next section, is that the mean number of bidders per listing is also independent of the number of listings. The right-hand panel of Figure 3.2 supports this pattern in the BW auction data. Given that the fee structure is fixed in the data, additional listings are not associated with higher cost sellers populating the platform. Network effects are such that potential bidders enter to the point of keeping the mean number of bidders per listing constant. Reported coefficients in Appendix C confirm that this result is robust to different product/market specifications (Table 3.11 Column 1).

Implications for structural model. Informed by these empirical patterns, the structural model considers a platform where bidders have a constant cost of inspecting a listing and therefore bid in one listing at a time. The game is static in correspondence with the empirical pattern of a low re-occurrence of products in subsequent months.[28] The presence of other auction platforms for wine besides BW is captured by the opportunity cost of entering and trading on BW. Hence, the (partial) equilibrium analysis is based on an implicit assumption that competing platforms keep their fee structure unchanged.[29]

3.3 A model of an auction platform for idiosyncratic goods

In this section, I model bidder and seller behavior on the platform as a static multi-stage game and study its equilibrium properties.

3.3.1 Model assumptions and game structure

Risk-neutral potential bidders and sellers consider trading on a monopoly platform with a given fee structure and furthermore have opportunity cost of doing so. Bidders have unit demands. The allocation mechanism in each listing is an English auction with flexible end time and proxy bidding. Assumptions on the matching process and model primitives that are maintained throughout are:

Assumption 1. *Bidders bid in one listing at a time and enter available listings with equal probability.*

This assumption can be justified on the basis that bidders learn about the wine's details only after they enter the product page and spend time inspecting it. Reserve prices are secret. In estimation, I implement the uniform allocation assumption conditional on

[28]Figure 3.10 in the Appendix.

[29]This is justified by BW being a small platform; the assumption would be more restrictive using eBay data unless you consider it to be a monopolist in the relevant market.

the observed "no reserve price button", allowing bidders to enter such auctions more numerously.

While the valuations of bidders and sellers may be correlated by their common appreciation of certain wine characteristics, their individual tastes are the basis of the following assumption:[30]

Assumption 2. *Conditional on the vector of observed wine attributes, variation in valuations across buyers and sellers is of a purely idiosyncratic -private values- nature. In addition, the idiosyncratic variation is independent.*

Independence is needed for identification of the distribution of idiosyncratic bidder valuations, but on the seller side it can be relaxed to unrestricted private values. The two conditional distribution functions are assumed to satisfy standard regularity conditions:

Assumption 3. *The distribution functions of idiosyncratic buyer and seller values are: i) absolutely continuous, ii) defined on a bounded support, and iii) characterized by an increasing failure rate (IFR).*

Continuity is needed for identification of the distribution of bidder valuations, but it could be omitted on the seller side. IFR is a standard restriction that guarantees uniqueness of the optimal reserve price, and is also not needed on the seller side.[31]

The valuation distributions, allocation mechanism, population sizes, and all cost (fees and opportunity cost) are common knowledge. While relatively parsimonious, this model captures the main features of an auction platform for idiosyncratic goods detailed in Section 3.2. The incomplete information structure and strategic interaction makes this suitable to study with the usual game-theoretic tools.

Timing of the game.

Entry stage	(t=1)	*Potential sellers learn their valuation and decide whether to enter and simultaneously, bidders decide whether to enter*
Auction stage	$(t = 2)$	*Sellers set a reserve price*
	$(t = 3)$	*Bidders learn their valuation and bid*

[30] The assumption is also justified by the elaborate data collection effort. In particular, it rules out the existence of wine features observable to bidders and sellers that are not excluded from their valuations and that are unobserved to the econometrician. Such features would render the private valuations of bidders and sellers affiliated in the sense of Milgrom and Weber (1982) and Aradillas-López (2016), even after controlling for observed wine attributes.

[31] With no reason to assume that their taste distributions differ, in estimation I use the same parametric restrictions on both sides (although I estimate parameters for bidders and sellers separately) and therefore I use a model with an identical set of restrictions on these distributions.

My analysis uses only one bid order statistic: the second-highest; and its relation to the second-highest valuation. English auctions generally allow for bidding strategies that complicate a tight mapping between observed bids and unobserved valuations (Haile and Tamer (2003)), and I rely on the following assumption to restrict bidder behavior:[32]

Assumption 4. *The transaction price in each auction is the greater of the second-highest bidder's willingness to pay and the reserve price.*

This assumption can be justified in different ways. It is the exact outcome of a model that imposes only the behavioral assumptions of Haile and Tamer (2003) that: 1) bidders never bid more than their valuation, and 2) never let someone else win at a price they are willing to beat, in the case of infinitesimal bidding increments and accounting for buyer's premium. With these intuitive behavioral assumptions, Assumption 4 therefore holds approximately to within one bidding increment. The assumed transaction price could also be derived from the more restrictive behavioral "button auction model" of Milgrom and Weber (1982) in the independent private values case.

Notation.
Let fee structure $f = \{c_B, c_S, e_B, e_S\}$ consist of, respectively, buyer's premium, seller commission, buyer entry cost, seller entry cost (listing fee). Opportunity cost of time equals e_B^o for potential bidders and e_S^o for potential sellers. Random vector $\mathbf{Z}$ contains auction covariates observed at the product-page. N^B, N^S, M, T respectively denote the number of: potential bidders, potential sellers, bidders (on the platform) and sellers (listings). $F_{V_0|\mathbf{Z}}$ and $F_{V|\mathbf{Z}}$ respectively denote the conditional valuation distributions for potential sellers and bidders, defined on bounded supports $V_0 \in [\underline{v_0}, \bar{v_0}]$ and $V \in [\underline{v}, \bar{v}]$. Random variables are denoted in upper case and their realizations in lower case. Furthermore, v_0 is the realized valuation of a generic seller while v_{0k} indicates the realized valuation for (potential) seller k. Similarly, v denotes the realized valuation for a generic bidder while v_i is used to denote the valuation of bidder i. Order statistics are useful as well: $X_{(n:n)}$ and $X_{(n-1:n)}$ denote respectively the highest and second-highest draw from a sample of size n from random variable X. $\mathbb{I}$ denotes the indicator function and stars denote equilibrium values. Additional notation will be introduced where necessary.

Payoffs.
The payoff for bidder i is simply his valuation v_i minus the transaction price increased

[32] Aradillas-López et al. (2013) previously adopt this assumption in an English auction model.

with buyer premium if he wins the auction. On top of that, regardless of whether he wins, by entering he foregoes entry and opportunity cost. The transaction price is the maximum of the second-highest bid and reserve price r, denoted by placeholder H here:

$$\pi_b(v_i, H) = \begin{cases} v_i - H(1+c_B) - e_B - e_B^o & : \text{for a bidder with valuation } v_i \text{ who wins} \\ -e_B - e_B^o & : \text{for a bidder with valuation } v_i \text{ who fails to win} \\ 0 & : \text{otherwise} \end{cases}$$

The payoff for a seller with valuation v_{0k} is the transaction price decreased with seller commission minus entry and opportunity cost. If he lists the good for sale but it does not sell, he only foregoes entry and opportunity cost but enjoys the value of the good:

$$\pi_s(v_{0k}, H) = \begin{cases} H(1-c_S) - e_S - e_S^o & :\text{f or a seller with valuation } v_{0k} \text{ who sells} \\ v_{0k} - e_S - e_S^o & : \text{ for a seller with valuation } v_{0k} \text{ who fails to sell} \\ v_{0k} & :\text{otherwise} \end{cases}$$

3.3.2 Equilibrium strategies

In this section, I solve for players equilibrium strategies. Considering two distinct stages of entry and auction, and given symmetry up to players' private valuations, I restrict attention to symmetric Perfect Bayesian-Nash Equilibria (PBE) in weakly undominated strategies. This equilibrium concept requires that strategies are best responses given competitors' strategies, and that beliefs are consistent with those strategies in equilibrium.

Auction stage

Conditional on entry decisions and the matching of bidders to listings, the idiosyncratic-good auction platform is made up of independent English auctions. I therefore derive standard reserve pricing (as in: Riley and Samuelson (1981)) and bidding (as in: Vickrey (1961)) strategies, up to the impact of buyer premium and seller commission.

Lemma 1. *It is a weakly undominated strategy for a bidder with valuation v to bid:*

$$b^*(v, f) \equiv \frac{v}{1+c_B} \tag{3.1}$$

Proof. This follows directly from Vickrey (1961), as bidding more than $\frac{v}{1+c_B}$ may result in negative utility and bidding less than $\frac{v}{1+c_B}$ decreases the probability of winning without affecting the transaction price in that case. □

Only the second-highest bid is relevant for this game and while other strategies are allowed to have been played by bidders with lower valuations (by Assumption 4), those strategies are not dominated by the strategy in Lemma 1 as they would lead to the same payoffs.

Lemma 2. *It is a weakly undominated strategy for a seller with valuation v_0 to set a secret reserve price that solves:*

$$r^*(v_0, f) \equiv \frac{v_0}{1 - c_S} + \frac{1 - F_{V|\mathbf{Z}}((1 + c_B)r^*(v_0, f))}{(1 + c_B)f_{V|\mathbf{Z}}((1 + c_B)r^*(v_0, f))} \tag{3.2}$$

$r^*(v_0, f)$ is increasing in c_S and decreasing in c_B.

The proof can be found in Appendix G. Note that, if $c_S = c_B = 0$, the optimal reserve price is identical to the familiar Riley and Samuelson (1981) formula for a public reserve price in auctions with a fixed number of bidders. It is easy to see why this is the case. Because $r^*(v_0, f)$ is secret, it does not affect the number of bidders in the seller's listing. This is true for any reserve price strategy of competing sellers.[33]

Section 3.5.4 introduces the choice between setting no reserve price or setting $r^*(v_0, f)$ based on additional reserve price fees, e_R, and lower expected participation in auctions with a positive reserve price.

Entry stage

With their valuations materializing in the auction stage, N^B identical potential bidders adopt a mixed strategy to enter with a probability that in equilibrium leaves their opponents indifferent between entering and staying out, as in Levin and Smith (1994). By contrast, the N^S potential sellers know their valuation v_0 so they enter selectively. Their expected surplus decreases in v_0, so they adopt the pure strategy to enter only if their valuation is below a threshold value that in equilibrium makes the marginal seller indifferent between entering and staying out given that his opponents adopt the same threshold strategy. In what follows, I denote the sellers' entry strategy by that threshold value. The following proposition summarizes key results:

[33]Sellers could be better off if they would collectively adopt (e.g., find a way to enforce) a different reserve price rule. In particular, a rule that results in a lower reserve price for any (v_0, f) increases bidder entry.

Proposition 1. The entry stage of the game results in a unique equilibrium for any fee structure. It is characterized by a bidder entry probability and seller entry threshold, $(p^*(f, v_0^*(f)), v_0^*(f))$, that jointly solve: 1) potential bidders' zero profit condition, and 2) the marginal seller's zero profit condition.

The remainder of this section derives the entry equilibrium. I first show that any candidate seller entry threshold, $\bar{v_0}$, maps to an equilibrium bidder entry probability, $p^*(f, \bar{v_0})$. Given that $p^*(f, \bar{v_0})$ is strictly decreasing in $\bar{v_0}$, sellers are strategic substitutes and the entry game reduces to a single agent discrete choice problem. I exclude a non-interesting no-trade equilibrium in which nobody enters the platform. To economize on space, a more detailed treatment is relegated to Appendix D.

Bidder entry.
Let $\pi_b(n, f, v_0)$ denote the ex-ante expected surplus for a bidder arriving in a listing with $n-1$ other bidders, fee structure f, and seller valuation v_0. The seller valuation enters π_b through optimal reserve price $r^*(v_0, f)$. In fact, this is why *seller selection* matters to bidders: $\pi_b(n, f, v_0)$ is strictly decreasing in v_0. Being unobserved to bidders, they form an expectation over V_0 using $\bar{v_0}$: $\mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}]$. They also form an expectation over the number of competing bidders in their listing, using its compound Binomial distribution, $f_N(n; p, \bar{v_0})$. From the perspective of a bidder who enters the platform, $f_N(n; p, \bar{v_0})$ combines uncertainty about: 1) the stochastic number of listings T (with realization t) given entry threshold $\bar{v_0}$, and 2) how many of $N^B - 1$ competing bidders end up in his listing when they enter the platform with probability p and sort uniformly over available listings. Combined with entry and opportunity cost, $\Pi_b(f, \bar{v_0}; p)$ denotes potential bidders' expected surplus from entering the platform:

$$\Pi_b(f, \bar{v_0}; p) = \sum_{n=0}^{N^B-1} \mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}] f_N(n; p, \bar{v_0}) - e_B - e_B^o \tag{3.3}$$

$$f_N(n; p, \bar{v_0}) = \sum_{t=0}^{N^S} \binom{N^B-1}{n} (\frac{p}{t})^n (1-\frac{p}{t})^{N^B-1-n} \binom{N^S}{t} F_{V_0|\mathbf{Z}}(\bar{v_0})^t (1 - F_{V_0|\mathbf{Z}}(\bar{v_0}))^{N^S-t} \tag{3.4}$$

Lemma 3. *Given candidate seller entry threshold $\bar{v_0}$ and fee structure f, the equilibrium bidder entry probability solves the zero profit condition:*

$$p^*(f, \bar{v_0}) \equiv \arg_{p \in (0,1)} \{\Pi_b(f, \bar{v_0}; p) = 0\} \tag{3.5}$$

Equilibrium properties are:
i) $p^*(f, \bar{v_0})$ *is unique* $\forall(f, \bar{v_0})$
ii) $p^*(f, \bar{v_0})$ *is strictly decreasing in* $(\bar{v_0}, c_B, c_S, e_B)$ *so also* $f_N(n; p^*, \bar{v_0})$ *decreases in the first-order stochastic dominance sense in* $(\bar{v_0}, c_B, c_S, e_B)$
iii) $f_N(n; p^*, \bar{v_0})$ *is invariant to changes in* N^B *or* N^S

While proofs are relegated to Appendix D and G, a short explanation suffices as intuition for these results. The zero profit condition follows from bidders being indifferent in equilibrium between staying out and entering the platform. The selection of less serious sellers (through an increase in $\bar{v_0}$) reduces bidders' expected listing-level surplus, $\mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}]$. As such, their equilibrium entry probability is lower, so that f_N places more weight on lower realizations of the number of bidders per listing. The same holds for increases in buyer premium, seller commission, and bidder entry and opportunity cost; since they all decrease expected listing-level surplus. As population sizes do *not* directly affect bidder surplus, the zero profit condition dictates that in equilibrium $f_N(n; p^*, \bar{v_0})$ remains constant. This also relates to the equilibrium prediction that is referred to in the **network effects** section on page 37: without affecting seller selection, more listings increase the number of bidders, but only to keep the mean number of bidders per listing constant.

Seller entry.
Potential sellers' expected surplus from entering the platform involves: 1) their listing-level expected surplus, and 2) an expectation over the number of bidders per listing, N, given $\bar{v_0}$ and bidders' equilibrium best-response to this threshold. Let $\Pi_s(f, v_0; p^*(f, \bar{v_0}), \bar{v_0})$ denote expected surplus for a seller with valuation v_0 when $N^S - 1$ competing sellers enter the platform if and only if their valuation is less than threshold $\bar{v_0}$:[34]

$$\Pi_s(f, v_0; p^*(f, \bar{v_0}), \bar{v_0}) = \sum_{n=0}^{N^B} \pi_s(n, f, v_0) f_N(n; p^*(f, \bar{v_0}), \bar{v_0}) - e_S - e_S^o \tag{3.6}$$

Lemma 4. *Given fee structure* f*, the equilibrium seller entry threshold solves the marginal*

[34] A slight abuse of notation is that expectation involves $f_N(n; p^*, \bar{v_0})$ (characterizing entry among $N^B - 1$ potential bidders) as defined in (3.4) instead of the distribution based on the full bidder population N^B. This avoids introduction of additional notation and the -1 will be irrelevant in the large-N^B approximation adopted for empirical tractability (page 46). In that world, the two distributions are identical by the *environmental equivalence* property of the Poisson distribution (Myerson (1998)).

seller's zero profit condition:

$$v_0^*(f) \equiv \arg_{\bar{v_0} s.t. F_{V_0|\mathbf{Z}}(\bar{v_0}) \in (0,1)} \{\Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0})) = 0\} \text{ , } \textit{with } p^*(f, \bar{v_0}) \textit{ solving } (3.5) \quad (3.7)$$

Equilibrium properties are:
i) $v_0^*(f)$ *is unique* $\forall f$
ii) $v_0^*(f)$ *is strictly decreasing in* e_B
iii) The impact of (c_B, c_S, e_S) *is ambiguous*

Proofs are provided in Appendix D and G. Key intuition is the following. Sellers have a unique best response for any competing seller entry (candidate) threshold, because $\Pi_s(f, v_0; p^*(f, \bar{v_0}), \bar{v_0})$ is strictly decreasing in v_0. Crucially, given that 1) $p^*(f, \bar{v_0})$ is strictly decreasing in $\bar{v_0}$, and 2) entry of competing sellers does not affect $\Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0}))$ in other ways, the best response function is strictly decreasing in competing sellers entry threshold. Symmetry then delivers a unique equilibrium threshold, $v_0^*(f)$, that is the fixed point in seller value space solving equation 3.7 i.e., making the marginal seller indifferent between entering and staying out.

Corollary 1. The entry equilibrium of the auction platform game is characterized by the pair of:
$(v_0^*(f), p^*(f, v_0^*(f)))$ that solves equation (3.7), which is unique for any fee structure.

Large population approximation.
The remainder of this section discusses an approximation of the entry equilibrium that is adopted for empirical tractability. A second reason is that the approximation relaxes the requirement that players know population sizes N^B and N^S, which indeed are likely to be unobserved by potential bidders and sellers.[35]

Assumption 5. *The population of potential bidders is large relative to the number of bidders on the platform:* $N^B \to \infty$ *and* $p^* \to 0$.[36]

Under this assumption, the number of bidders per listing in (3.4) is approximately Poisson distributed, and approximation error relative to the Binomial distribution is small

[35] Relatedly, given that the population of potential bidders is likely to be large relative to the actual number of bidders, the Poisson approximation of the binomial distribution is a natural one, also adopted previously in similar settings by e.g. Engelbrecht-Wiggans (2001), Bajari and Hortaçsu (2003) and Jehiel and Lamy (2015).

[36] To avoid any misinterpretation (with p^* endogenous), the population is assumed to be large and the entry probability is assumed to be small and it is not a statement about letting the population grow large or the entry probability go to 0.

(vanishing to 0 as N^B tends to infinity).[37]

Lemma 5. *With large N^B, small p^*, and T listings the number of bidders per listing has a probability mass function approximated by:*

$$f_N(k;\lambda) = \frac{\exp(-\lambda)\lambda^k}{k!},\ \lambda = \frac{N^B p^*}{T}\ ,\ \forall k \in \mathbb{Z}^+ \tag{3.8}$$

The equilibrium Poisson mean number of bidders per listing is endogenous to the fee structure and depends on seller selection: $\lambda^*(f, v_0^*(f))$. It is implicitly defined as solving potential bidders' zero profit entry condition in this slightly modified setting. Appendix D and G show this formally; here I report only the resulting equilibrium.

Corollary 2. The entry equilibrium of the auction platform game subject to the large population approximation is characterized by the pair of: $(v_0^*(f), \lambda^*(f, v_0^*(f)))$ that solves the entry problems of potential bidders and sellers, which is unique for any fee structure.

3.4 Nonparametric identification

In this section, I investigate whether model primitives are nonparametrically identified from auction observables and the set of maintained assumptions in Section 3.3.1. Model primitives are: the conditional valuation distributions $F_{V_0|\mathbf{Z}}$ and $F_{V|\mathbf{Z}}$, and opportunity costs e_S^o and e_B^o. Endogenous observables are: the number of actual bidders (A), the second-highest bid (B), and the reserve price (R). Exogenous observables are denoted by X and include: fee structure f, auction characteristics $\mathbf{Z}$, and population sizes N^B and N^S.

I adopt terminology dating back to Hurwicz (1950) and Koopmans and Reiersol (1950) (see also Chesher (2007) and Berry and Haile (2018)). The idiosyncratic-good auction platform model $\mathcal{M}$ identifies structure $[F_{V_0|\mathbf{Z}}, F_{V|\mathbf{Z}}, e_S^o, e_B^o] \equiv S_0 \in \mathcal{M}$ (admitted by the model) if and only if: $\forall S \in \mathcal{M}, S \neq S_0, F^{S_0}_{A,B,R|X} \neq F^{S}_{A,B,R|X}$, where $F^{S}_{A,B,R|X}$ indicates the distribution of outcomes (A, B, R) given other observables X that is generated by structure S.

3.4.1 The distribution of bidder valuations

Athey and Haile (2002, Theorem 1) prove identification of $F_{V|\mathbf{Z}}$ in an English auction model that places identical restrictions on this distribution up to the presence of binding

[37]In my data, it cannot be rejected at < 1 percent level that the number of bidders per listing is generated by a Poisson distribution (see figure 3.4).

reserve prices. Their proof relies on the relationship between the distribution of the second-highest draw (valuation) in a sample of known size (number of bidders) from its parent distribution and that parent distribution (see also e.g., Arnold et al. (1992)):

$$F_{V_{(n-1:n)}}(v) = n(n-1)\int_{\underline{v}}^{v} F_V(t)^{n-2}(1-F_V(t))f_V(t)dt \equiv \phi(F_V(v);n) \tag{3.9}$$

When n is known, given that $\phi(F_V(v);n)$ is strictly increasing in $F_V(v)$, F_V is identified whenever $F_{V_{(n-1:n)}}$ is. In particular, $F_V(v)$ is identified point-wise $\forall v \in [\underline{v}, \bar{v}]$ by inverting $\phi(;n)$:

$$F_V(v) = \phi^{-1}(F_{V_{(n-1:n)}}(v);n) \tag{3.10}$$

This argument also applies conditional on observed $\mathbf{Z}$, so identification of $F_{V|\mathbf{Z}}$ follows. What remains to be shown is identification of $F_{V_{(n-1:n)}}$, which is slightly different from Athey and Haile (2002) due to the presence of binding reserve prices. In auctions without a reserve price, an event that is known, order statistic $V_{(n-1:n)}$ equals the transaction price and n is observed. Hence, replacing $F_{V_{(n-1:n)}}$ with the empirical distribution F_B from auctions without a reserve price in (3.10) completes the proof.

3.4.2 The distribution of seller valuations

Given that $F_{V|\mathbf{Z}}$ is identified, in all auctions with a positive reserve price the reserve price identifies the seller's valuation in that listing. In particular, re-arranging the equilibrium reserve price strategy in Lemma 2:

$$v_0 = (1-c_S)\left(r - \frac{1-F_{V|\mathbf{Z}}(r(1+c_B))}{(1+c_B)f_{V|\mathbf{Z}}(r(1+c_B))}\right) \equiv \underline{r}, \tag{3.11}$$

where $\underline{r}$ denotes the observed scalar-valued right-hand side. Its distribution function, $F_{\underline{R}}$, trivially identifies the distribution of valuations *among sellers who enter the platform*, point-wise $\forall v \in [\underline{v_0}, v_0^*(f)]$:[38]

$$F_{\underline{R}}(v) = \frac{F_{V_0|\mathbf{Z}}(v)}{F_{V_0|\mathbf{Z}}(v_0^*(f))} \tag{3.12}$$

Dividing the number of listings, T, by the population of potential sellers, N^S, delivers the seller entry probability in the denominator of (3.12). The distribution of valuations in the

[38] $v_0^*(f)$ is the equilibrium seller entry threshold defined in (3.7). Previously, also Elyakime et al. (1994) identify seller cost using a first order optimality condition on the secret reserve price in first price auctions (in which case, the secret reserve price is equal to the seller's valuation).

population of potential sellers is identified pointwise $\forall v \in [\underline{v_0}, v_0^*(f)]$ as:

$$F_{V_0|\mathbf{Z}}(v) = F_{\underline{R}}(v)\frac{T}{N^S} \tag{3.13}$$

Without identifying variation in $v_0^*(f)$ and unless $v_0^*(f) = \bar{v_0}$, the population distribution $F_{V_0|\mathbf{Z}}(v)$ is not identified on the part of its support exceeding $v_0^*(f)$. It is worthwhile to point out that nonparametric identification of the right-truncated distribution of potential seller valuations in (3.12) is sufficient for any counterfactual that reduces $v_0^*(f)$, i.e. any scenario that reduces expected seller surplus ($\Pi_s(f, v_0; p^*(f, \bar{v_0}), \bar{v_0})$ defined in (3.6)). In any such scenario, only a subset of sellers currently trading on the platform will find it optimal to enter. As such, the distribution of valuations among sellers currently trading on the platform would be the relevant latent distribution necessary to characterize the counterfactual entry equilibrium and distributions of endogenous outcomes.

3.4.3 Opportunity cost

Opportunity cost are identified from the two zero profit conditions. In auctions with a zero reserve price the number of bidders is not truncated; observables from those auctions render the equilibrium distribution $f_N(; p^*(f, v_0^*), v_0^*(f))$ identified. Other components of expected bidder surplus $\Pi_b(f, v_0^*(f); p^*(f, v_0^*(f)))$ (defined in (3.3), here referring to its equilibrium value) are: the distribution of bidder valuations (used in the definition of $\pi_b(n+1, f, v_0)$) and the right-truncated distribution of seller valuations (to take expectations of $\pi_b(n+1, f, v_0)$ over realizations of V_0), which are both identified, and observed f. $\Pi_b(f, v_0^*(f); p^*(f, v_0^*(f)))$ is strictly decreasing in the last remaining unobservable, opportunity cost c_B^o. Hence c_B^o is identified as the value that solves the zero profit bidder entry condition in (3.5), setting $\Pi_b(f, v_0^*(f); p^*(f, v_0^*(f)) = 0$.

Similarly, the surplus for a marginal seller must by equilibrium play and the zero profit condition in (3.7) correspond to opportunity cost c_S^o. Surplus for the marginal seller, in equilibrium, $\Pi_s(f, v_0^*; p^*(f, v_0^*(f)), v_0^*(f))$ is defined in (3.6). Besides c_S^o, it depends on: the identified distribution of bidder valuations (used in the definition of $\pi_s(n, f, v_0)$), observed fees f, the identified $f_N(; p^*(f, v_0^*), v_0^*(f))$, and the value of $v_0^*(f)$. The latter is observed as the maximum implied seller valuation in (3.11). This is a valid basis for identification of $v_0^*(f)$ since identification analysis concerns a hypothetical environment with infinite data.[39] As $\Pi_s(f, v_0^*; p^*(f, v_0^*(f)), v_0^*(f))$ is strictly decreasing in the seller opportunity cost, which is the last remaining unknown, c_S^o is identified as the value that solves the marginal seller's

[39] By contrast, the finite-sample sample maximum of a noisy estimator may be far removed from the true entry threshold, as discussed in more detail in the estimation section.

zero profit entry condition in (3.12), setting $\Pi_s(f, v_0^*; p^*(f, v_0^*(f)), v_0^*(f)) = 0$.

Corollary 3. Given exogenous observables X and endogenous observables (A, B, R), the idiosyncratic-good auction platform model $\mathcal{M}$ identifies $[F_{V|\mathbf{Z}}, e_S^o, e_B^o]$ and identifies $F_{V_0|\mathbf{Z}}$ right-truncated at $v_0^*(f)$.

These positive identification results are not altered when (N^B,N^S) are unobserved and the large population assumption (Assumption 5) is added to the model. This is because: $f_N(; p^*(f, v_0^*), v_0^*(f))$ is directly identified from observables in auctions without a reserve price, the expectation over values of N in (3.3) is then over an infinite support, and the results don't rely on population sizes otherwise.

3.5 Estimation method

I estimate a parametric specification of the model, allowing me on the seller side to extrapolate beyond the support on which $F_{V_0|\mathbf{Z}}$ is identified. Parameters from the distributions of idiosyncratic bidder and seller values are estimated separately; I refer to these as bidder parameters (θ_b) and seller parameters (θ_s). Even when assuming that $F_{V|\mathbf{Z}}(; \theta_s)$ and $F_{V|\mathbf{Z}}(; \theta_b)$ are known up to finite-dimensional parameters, the fact that the entry equilibrium depends on those parameters complicates estimation. The equilibrium $v_0^*(f, \theta_s, \theta_b)$ is the solution to a fixed point problem that itself depends on a threshold-crossing problem on the bidder side, $\lambda^*(f, v_0^*(f; \theta_s, \theta_b), \theta_b)$. This equilibrium is computationally costly to compute for each set of candidate parameters, making full maximum likelihood estimation of all parameters infeasible. I adopt a multi-step estimation method that is based on:

1) controlling for auction heterogeneity $\mathbf{Z}$ (using the homogenization step in Haile et al. (2003), also used for ascending auctions in e.g., Bajari and Hortaçsu (2003) and Freyberger and Larsen (2017))

2) estimating θ_b by maximum likelihood (as in e.g. Donald and Paarsch (1996) and Paarsch (1997)), using homogenized bids

3) estimating θ_s by maximum concentrated likelihood (mentioned in Donald and Paarsch (1993, Footnote 4) to overcome a support problem in first price auctions), using homogenized reserve prices

Small sample estimation error from steps 1 and 2 affect the estimation of θ_s, especially because it involves the sample *maximum* of estimated seller values in equation 3.11. I therefore add the following steps:

4) solving for the entry equilibrium given estimated parameters

5) re-estimating seller parameters at the updated entry equilibrium

These last two steps can be iterated on until convergence, but for any number of iterations this method delivers a consistent estimate of θ_s (shown in Aguirregabiria and Mira (2002)).[40] Based on Monte Carlo simulations, the estimation includes only one iteration on steps 4-5 (see Appendix E for Monte Carlo results and Appendix F for details about the necessary numerical approximation of the entry equilibrium). The rest of this section provides estimation details. Section 3.5.4 addresses the choice for sellers to set no reserve price and estimates entry of additional bidders in these listings.

3.5.1 Auction heterogeneity: homogenizing bids and reserve prices

Considering that valuations for wines auctioned at the BW platform consist of both a common value element (due to the importance of provenance, *ullage*, the expected quality of wines from different vintages or regions) and a private "taste", and that valuations are plausibly non-negative, bidder- and potential seller valuations are taken to satisfy the following log-linear single-index structure:

$$\begin{aligned} ln(V) &= g(\mathbf{Z}) + U \\ ln(V_0) &= g(\mathbf{Z}) + U_0, \end{aligned} \tag{3.14}$$

with ln the natural logarithm and $(U, U_0, \mathbf{Z})$ mutually independent. The common $g(\mathbf{Z})$ term is interpreted as "quality". For example, it is commonly accepted that the 1961 Bordeaux vintage is better than most other vintages as a result of favourable weather conditions and that low fill levels relative to the age of the wine are bad.[41] By additivity

[40] To see why I adopt this algorithm, notice that the seller entry problem resembles a discrete choice programming problem and that the three referenced estimators based on maximum likelihood, maximum concentrated likelihood and the iterative method correspond to three solutions for solving parameters involving fixed point characterizations in the estimation of games. Respectively, those methods are based on the nested fixed point algorithm by Rust (1987), two-step methods (e.g. Pesendorfer and Schmidt-Dengler (2003), Bajari et al. (2007, 2010)) and the nested pseudo likelihood estimator of Aguirregabiria and Mira (2002, 2007). A key difference is that in my case the auction structure allows for the estimation of seller parameters from a first order condition instead of from the conditional choice (entry) probability. A second difference is that by Proposition 1, potential sellers (who know θ_s) do not need to form beliefs about the entry threshold competing sellers adopt as the entry game reduces to a single agent discrete choice problem. Hence iterating on steps (4) and (5) is considered only to potentially improve precision stemming from the fact that the true entry threshold is unknown to the econometrician, so the only "belief" that is being updated during iteration is his (hers). Due to the unique entry equilibrium that is iterated on, so the concerns expressed in Pesendorfer and Schmidt-Dengler (2010), Kasahara and Shimotsu (2012) and Egesdal et al. (2015) do not apply.

[41] It would also be feasible to adopt an alternative specification that estimates a separate $g_0(\mathbf{Z})$ for sellers. In that case, a sample of u_{0t} is obtained as the residual plus intercept from a regression of the implied seller valuation given $\hat{g(\mathbf{Z})}$ and $\hat{\theta}_b$ on $\mathbf{Z}$

of the idiosyncratic taste component, for all bidders i: $V_i = g(\mathbf{Z}) + U_i$ so that also:

$$V_{(n-1:n)} = g(\mathbf{Z}) + U_{(n-1:n)} \tag{3.15}$$

Quality is then estimated by regressing the transaction price on auction characteristics, using only data from auctions without a reserve price and with more than one bidder in which the transaction price equals the second-highest valuation.[42] Residuals from this regression (plus the intercept) are the homogenized second-highest bids used for estimation of θ_b in (3.16). On the seller side, the residualized implied seller tastes are used for the estimation of θ_s in (3.20).

3.5.2 Bidder valuations

Both U and U_0 in (3.14) are assumed to be normally distributed.[43] Following the identification argument, the mean and variance of U, $(\mu_b, k_b \in \theta_b)$, are estimated by maximum likelihood estimation in auctions with a zero reserve price. Let $\mathcal{T}$, $\mathcal{T}_{r0}$ and $\mathcal{T}_{r>0}$ denote the set of listings, listings with a zero reserve price, and listings with a positive reserve. Let $h(b_t|n_t, \mathbf{z}_t, f; \theta_b)$ denote the density of transaction prices given the number of bidders n_t, characteristics $\mathbf{z}_t$ and fees f. For all auctions with a zero reserve price it is simply the probability that the homogenized transaction price / second-highest bid b_t is the second-highest among n_t draws from $F_{V|\mathbf{Z}}$. Hence $\forall t \in \mathcal{T}_{r0}$:

$$h(b_t|n_t, \mathbf{z}_t, f; \theta_b) = n_t(n_t - 1)F_{V|\mathbf{Z}}(b_t; \theta_b)^{n_t-2}[1 - F_{V|\mathbf{Z}}(b_t; \theta_b)]f_{V|\mathbf{Z}}(b_t; \theta_b) \tag{3.16}$$

Note the tight mapping between the identification result and the estimating equation. The log likelihood of bidder parameters given data is specified as:

$$\mathcal{L}(\theta_b; \{n_t, \mathbf{z}_t, b_t, r_t\}_{t\in\mathcal{T}_{r0}}, f) = \sum_{t\in\mathcal{T}_{r0}} ln((h(b_t|n_t, \mathbf{z}_t, f; \theta_b))) \tag{3.17}$$

$$\hat{\theta}_b = \arg\max \mathcal{L}(\theta_b; \{n_t, \mathbf{z}_t, b_t, r_t\}_{t\in\mathcal{T}_{r0}}, f)$$

[42]The term refers to the homogenization step in Haile et al. (2003). This *first stage regression* is standard in the analysis of ascending auctions and used in e.g., Bajari and Hortaçsu (2003) and Freyberger and Larsen (2017).

[43]The lognormal distribution is commonly used to analyze bidding data in the empirical auction literature (adopted in a variety of settings, e.g. Paarsch (1992), Laffont et al. (1995), Haile (2001), Hong and Shum (2002)). Another common specification, the loglogistic distribution, is also considered but its heavier tails provide a slightly worse fit to nonparametric bidder values (details below).

Bidder parameters are thus estimated using data from auctions with no reserve price. I use observations from positive reserve price auctions to obtain a maximum likelihood estimate of the (unobserved) mean number of bidders in those auctions. In particular, I estimate that mean from variation in bids and number of observed bidders in positive reserve price auctions together with estimated bidder parameters from equation (3.17) and the Poisson structure. Details are provided in Section 3.5.4.

3.5.3 Seller valuations

Given estimated $\hat{\theta}_b$ and $g(\hat{\mathbf{Z}})$, a sample of implied sellers' valuations as in (3.11) is obtained, $\forall t \in \mathcal{T}_{r>0}$:[44]

$$\hat{v}_{0,t} = (1-c_S)\left(r_t - \underbrace{\frac{1-F_{V|\mathbf{Z}}(ln(\tilde{r}_t) - g(\hat{\mathbf{z}}_t); \hat{\theta}_b)}{(1+c_B) f_{V|\mathbf{Z}}(ln(\tilde{r}_t) - g(\hat{\mathbf{z}}_t); \hat{\theta}_b)}}_{\text{mark-up}} \right), \text{ and hence:} \tag{3.18}$$

$$\hat{u}_{0,t} = ln\left((1-c_S)\left(r_t - \frac{1-F_{V|\mathbf{Z}}(ln(\tilde{r}_t) - g(\hat{\mathbf{z}}_t); \hat{\theta}_b)}{(1+c_B) f_{V|\mathbf{Z}}(ln(\tilde{r}_t) - g(\hat{\mathbf{z}}_t); \hat{\theta}_b)} \right)\right) - g(\hat{\mathbf{z}}_t), \tag{3.19}$$

with $\tilde{r}_t = r_t(1+c_B)$ denoting the buyer premium-adjusted reserve price and $\hat{u}_{0,t}$ the homogenized idiosyncratic part of the implied seller value in auction t. The sample maximum of implied residual seller valuations, $\hat{\upsilon_T} = \max(\{\hat{u}_{0,t}\}_{t \in \mathcal{T}_{r>0}})$, is a consistent estimator of the seller entry threshold. Intuitively, sellers with higher residual value draws than $v_0^*(f)$ will never list so $\hat{\upsilon_T} - v_0^*(f)$ is always negative (at population values of θ_b and $g(\mathbf{Z})$) and the more sellers that do list the larger the probability that the marginal seller has a valuation equal to the threshold.[45] The same holds when observed iterations of the game tend to infinity but the number of listings in each game stays constant.

Let $h(\hat{u}_{0,t}|\mathbf{z}_t, f, v_0^*(f, \theta_b, \theta_s), \hat{\theta}_b; \theta_s)$ denote the density of $\hat{u}_{0,t}$ given bidder parameters and given the true seller equilibrium entry threshold, which follows from the relevant identification equation (3.12):

$$h(\hat{u}_{0,t}|\mathbf{z}_t, v_0^*(f, \theta_s, \theta_b), f; \theta_s) = \frac{f_{V_0|\mathbf{Z}}(\hat{u}_{0,t}; \theta_s)}{F_{V_0|\mathbf{Z}}(v_0^*(f, \theta_s, \theta_b); \theta_s)} \mathbb{I}\{\hat{u}_{0,t} \leq v_0^*(f, \theta_s, \theta_b)\} \tag{3.20}$$

A complication is that the support of the implied valuations (and reserve prices) observed

[44] $P[V \leq r|\mathbf{Z}=\mathbf{z}] = P[\exp(g(\mathbf{z}) + U) \leq r] = P[U \leq ln(r) - g(\mathbf{z})] = F_{V|\mathbf{Z}}(ln(r) - g(\mathbf{z}))$

[45] A more precise statement given that valuation distributions are continuous is that the probability that the marginal seller has a valuation within a fixed small interval around the threshold increases.

in the data depends on θ_s through its effect on $v_0^*(f, \theta_s, \theta_b)$, so that standard regularity conditions demonstrating consistency and asymptotic normality of the maximum likelihood estimate of θ_s don't apply.[46]

To address the support problem, I estimate seller parameters by maximizing a concentrated likelihood that includes the consistent estimate $\hat{v_T}$ in place of $v_0^*(f, \theta_s, \theta_b)$:[47]

$$\mathcal{L}(\theta_s; \{\hat{u}_{0,t}, \mathbf{z}_t\}_{t \in \mathcal{T}_{r>0}}, f, \hat{v_T}) = \sum_{t \in \mathcal{T}_{r>0}} ln(h(r_t | \mathbf{z}_t, \hat{v_T}, f; \theta_s)) \tag{3.21}$$
$$\hat{\theta_s}^0 = \arg\max \mathcal{L}(\theta_s; \{\hat{u}_{0,t}, \mathbf{z}_t\}_{t \in \mathcal{T}_{r>0}}, f, \hat{v_T})$$

The first order condition of the concentrated likelihood with respect to $(\mu_s, k_s \in \theta_s)$ does not depend on v_T. However, the fact that $\hat{u}_{0,t}$ depends on estimated $\hat{\theta_b}$ and $g(\hat{\mathbf{Z}})$ makes it likely that in finite samples $\hat{v_T}$ is biased. In particular, because it is the maximum of a noisily estimated sample of homogenized idiosyncratic seller valuations it likely an overestimate of the true $v_0^*(f)$. Relatedly, it introduces the possibility that the largest values of $\hat{u}_{0,t}$ incorporate the highest bias. Monte Carlo simulations show that a noisy first stage especially affects the shape parameter k_b, and in the expected direction: the initial $\hat{k_s}$ overestimates the truth as the sample of implied seller values appears more disperse. Correspondingly, the initial estimate of the seller entry threshold is also too high. Updating this threshold by solving the entry game once and then re-estimating seller parameters from a sample that excludes observations exceeding the threshold addresses that issue.[48] In particular, it involves numerical approximation of the entry equilibrium given estimated $(\hat{\theta_b}, \hat{\theta_s}^0)$ as detailed in Appendix F, resulting in equilibrium entry threshold $v_0^*(f, \hat{\theta_s}^0, \hat{\theta_b})$. Then, seller parameters are estimated by maximizing:

$$\mathcal{L}(\theta_s; \{\hat{u}_{0,t}, \mathbf{z}_t\}_{t \in \mathcal{T}_{r>0}}, f, v_0^*(f, \hat{\theta_s}^0, \hat{\theta_b})) \tag{3.22}$$

[46]This has been pointed out by Donald and Paarsch (1993) in the context of first-price auctions and addressed by Jofre-bonet and Pesendorfer (2003) for dynamic first price auctions with a Weibull specification of bidder valuations.

[47]This has been suggested e.g. in Donald and Paarsch (1993, Footnote 4) in the context of a support problem in first-price auctions.

[48]This describes the Nested Pseudo Likelihood estimator in Aguirregabiria and Mira (2002, 2007) used in discrete choice games. Roberts and Sweeting (2010) are the first to apply this algorithm to the auction literature to study auctions with selective bidder entry. Studies by Pesendorfer and Schmidt-Dengler (2010), Kasahara and Shimotsu (2012) and Egesdal et al. (2015) provide conditions under which NPL does (not) converge to the true equilibrium. A best-response stable equilibrium is a sufficient condition for the NPL algorithm to converge to it and this is certainly guaranteed (Proposition 1) by the game reducing to a single agent discrete choice problem with unique equilibrium. Aguirregabiria and Mira (2002) find that in single-agent games asymptotic efficiency is independent of the number of iterations.

This describes steps 4 and 5 in the outline at the beginning of this section. Based on results from Monte Carlo simulations, I use only one update (see Appendix E). The 0 superscript in ${\hat{\theta}_s}^0$ in (3.21) indicates that this is the initial estimate of seller parameters before solving the game and updating parameter estimates; the final estimated seller parameters are denoted by $\hat{\theta}_s$.

3.5.4 Incorporating zero reserve prices

I now address the fact that many sellers choose to set no reserve price, motivated by the additional reserve price fee c_R and (anticipated) additional entry of bidders in auctions without a reserve price.[49] The benefit of setting a reserve price is increasing in sellers' valuation so the no-reserve price choice can be described by a threshold-crossing problem.

On the seller side, let p_{r0} denote the observed share of sellers that sets no reserve price and $v^*_{0,r=0}$ the valuation of the seller that is indifferent between setting no reserve price and the optimal positive reserve price. This value depends on the fee structure (including but not limited to c_R), the equilibrium differential in bidder participation in the two types of auctions, and (θ_b, θ_s). On the bidder side, let $\lambda^*_{r>0}$ ($\lambda^*_{r=0}$) denote the equilibrium Poisson mean number of bidders in auctions with a positive reserve price (no reserve price). It is informed by $v^*_{0,r=0}$ and (θ_b, θ_s) through their impact on the expected distribution of reserve prices in auctions with positive reserves. Different bidder entry into these two types of listings is referred to as the bidder participation differential, denoted by $\hat{\beta}_b = \hat{\lambda}^*_{r=0} - \hat{\lambda}^*_{r>0}$.

Bidder participation differential

Listings with no reserve price are structurally more attractive to bidders than those with a reserve price, increasing bidder entry. Crucially, f_N in (3.8) is relaxed to the following conditional Poisson distribution:

$$f_{N|R}(k;\lambda) = \frac{\exp(-(\lambda^*_{r>0} + \beta_b \mathbb{I}\{R=0\}))(\lambda^*_{r>0} - \beta_b \mathbb{I}\{R=0\})^k}{k!}, \forall k \in \mathbb{Z}^+ \tag{3.23}$$

$\mathbb{I}\{R=0\}$ is the indicator function for the event of a zero reserve price, so that $\lambda^*_{r=0} = \lambda^*_{r>0} + \beta_b \mathbb{I}\{R=0\}$. The mean number of bidders in no reserve auctions is a consistent estimate of $\lambda^*_{r=0}$:

$$\hat{\lambda}^*_{r=0} = \frac{1}{|\mathcal{T}_{r=0}|} \sum_{t \in \mathcal{T}_{r=0}} n_t \tag{3.24}$$

[49]Note that in theory, if sellers have non-negative valuations, it is always better to set a positive reserve price (see e.g. Krishna (2009) or Riley and Samuelson (1981) or equation 3.2) *unless the number of bidders is a function of it* (Mcafee and Mcmillan (1987), Bulow and Klemperer (1996), Albrecht et al. (2012)).

A consistent estimate of $\lambda^*_{r>0}$ equals the value that maximizes the likelihood of transaction prices, b_t, and number of actual bidders, a_t, in positive reserve auctions given estimated bidder valuation parameters. In particular, the joint density of b_t, a_t *if the number of potential bidders n_t would be known*, in auctions with a positive reserve price:

$$\begin{aligned}
&h(b_t, a_t|n_t, r_t > 0, \mathbf{z}_t, f, \hat{\theta}_b) = \{F_{V|\mathbf{Z}}(\tilde{r}_t; \hat{\theta}_b)^{n_t}\}\mathbb{I}\{a_t = 0\} \qquad (3.25)\\
&\{n_t F_{V|\mathbf{Z}}(\tilde{r}_t; \hat{\theta}_b)^{n_t-1}[1 - F_{V|\mathbf{Z}}(\tilde{r}_t \hat{\theta}_b)]\}\mathbb{I}\{a_t = 1\}\\
&\{\binom{n_t}{n_t - a_t} F_{V|\mathbf{Z}}(\tilde{r}; \hat{\theta}_b)^{n_t-a_t}[1 - F_{V|\mathbf{Z}}(\tilde{r}; \hat{\theta}_b)]^{a_t}\\
&a_t(a_t - 1)F_{V|\mathbf{Z}}(b_t; \hat{\theta}_b)^{a_t-2}[1 - F_{V|\mathbf{Z}}(b_t; \hat{\theta}_b)]f_{V|\mathbf{Z}}(b_t; \hat{\theta}_b)\}\mathbb{I}\{a_t \geq 2\}
\end{aligned}$$

Note that $h(b_t, a_t|n_t, r_t > 0, \mathbf{z}_t, c_B, \hat{\theta}_b) = 0$ when $n_t = 0$. The first line covers the probability that all n_t bidders draw a valuation below the reserve price, the second line the probability that one out of n_t draw a valuation exceeding $\tilde{r}$ while the others don't (in which case $b_t = r_t$ with certainty), and the final two lines capture the probability that a_t out of n_t draw a valuation exceeding the reserve and that the second-highest out of them draws a valuation equal to b_t. Without observing n_t, a feasible specification takes the expectation over realizations of random variable $N \sim Pois(\lambda^*_{r>0})$. This is the basis of the likelihood function that $\hat{\lambda}^*_{r>0}$ maximizes:

$$g(b_t, a_t|r_t > 0, \mathbf{z}_t, f, \hat{\theta}_b; \lambda^*_{r>0}) = \sum_{n_t=a_t}^{\infty} h(b_t, a_t|k, r_t > 0, \mathbf{z}_t, f, \hat{\theta}_b) f_{N|N\geq A}(k; \lambda^*_{r>0}) \qquad (3.26)$$

$$\mathcal{L}(\lambda^*_{r>0}; \{b_t, a_t, r_t \mathbf{z}_t\}_{t\in\mathcal{T}_{r>0}}, f) = \sum_{t\in\mathcal{T}_{r>0}} ln(g(b_t, a_t|r_t > 0, \mathbf{z}_t, f, \hat{\theta}_b; \lambda^*_{r>0})) \qquad (3.27)$$

$$\hat{\lambda}^*_{r>0} = \arg\max \mathcal{L}(\lambda^*_{r>0}; \{b_t, a_t, r_t, \mathbf{z}_t\}_{t\in\mathcal{T}_{r>0}}, f)$$

No reserve price share

The share of sellers that finds it optimal to set no reserve price is observed in the data as p_{r0}. To endogenize this choice, I compute the share of sellers that should *optimally* set no reserve price given the maximum likelihood estimates of the bidder entry differential and parameter estimates, denoted by p^*_{r0}. While this share adjusts in counterfactuals, any difference between p_{r0} and p^*_{r0} for reasons outside the model is held fixed.

To calculate p^*_{r0} note that it is more beneficial for sellers with higher valuations to set a reserve price. Let $\Pi_s(f, v_0, r = 0, \hat{\lambda}^*_{r>0} + \hat{\beta}_b; \hat{\theta}_b)$ describe the parameterized expected surplus for a seller with valuation v_0 when he sets a zero reserve price and $\Pi_s(f, v_0, r > 0, \hat{\lambda}^*_{r>0}; \hat{\theta}_b)$ when he sets the optimal positive reserve price. The "screening value" $v^*_{0,r=0}$ describes the seller who is indifferent between setting a positive reserve and a zero reserve price. This is estimated by the value v_0 that solves:[50]

$$\Pi_s(f, v_0, r = 0, \hat{\lambda}^*_{r>0} + \hat{\beta}; \hat{\theta}_b) - \Pi_s(f, v_0, \hat{\lambda}^*_{r>0}; \hat{\theta}_b) + c_R = 0 \tag{3.28}$$

Details on computation of (entry) equilibrium values are provided in Appendix F.

3.6 Estimation results

3.6.1 Impact of observed wine characteristics

The homogenization (step 1) is done separately for auctions with transaction prices of at most 200 pounds, referred to as the "main sample" as they contain 80.93 percent of observations, and the remaining auctions that is referred to as the "high-value" sample.[51] Estimation is done per bottle-equivalent which delivers a better fit than the lot level. Results from this step are presented in Tables 3.9 and 3.10 in the Appendix. The estimated (sign of) coefficients for various key variables are as expected. Among other findings, results show that prices are higher for bottles sold in a case of 6 or 12, but conditional on this case effect the price is lower the more bottles are included in the lot. For bottles stored in a specialized warehouse (with optimal temperature and humidity control) and for wines in special format bottles (e.g. magnums) prices are higher. This corresponds to the idea that these wines are expected to be of higher quality.[52] The omitted *ullage* category in the Tables is "Into Neck", the best fill level, so logically all other levels deliver lower prices (some coefficients are insignificant). The tables furthermore report estimated relative values for a host of wine regions, grapes, and shipping options. These observables explain a large share of total price variation, even without controlling for the number of bidders. The adjusted R-squared is 0.530 for the main sample and 0.855 for the smaller high-value sample.

---

[50] Previous notation has abstracted from the screening value below which sellers find it optimal to set no reserve price, $v^*_{0,r=0}$. Observed implied seller values are only those $\in [v^*_{0,r=0}, v^*_0]$. Estimation of θ_s following from equation 3.21 maximizes the concentrated likelihood based on: $\frac{f_{V_0|\mathbf{Z}}(\hat{u}_{0,t};\theta_s) - p_{r0}}{v_T - p_{r0}}$. The iteration algorithm updates v_T by solving for the game keeping p_{r0} at the observed level (while updating the screening value according to updated $\hat{\theta}_s$).

[51] Seller commissions are tiered on BW (Table 3.1): for realized transaction prices exceeding 200 pounds, sellers pay one percentage point less in seller commission on the excess than for goods selling below 200 pounds. The benefit of this approach is that it generates a relevant dimension of product variety along which the welfare impacts of fee changes can be assessed, even after homogenizing bids and values.

[52] While larger bottles may also be attractive for their fun factor, their smaller surface area conditional on wine content is also associated with a lower oxidation rate.

Table 3.5: Estimation results: idiosyncratic valuations and entry

Parameters of $F_{V\|\mathbf{Z}}$, $F_{V_0\|\mathbf{Z}}$		Main	High-value	Entry equilibrium		Main	High-value
Bidders (θ_b)	μ_b	3.566	5.830	Bidders per listing	$\lambda^*_{r>0}$	4.277	4.599
		[0.031]	[0.038]			[0.009]	[0.032]
	k_b	0.908	0.553		$\lambda^*_{r=0}$	5.215	7.258
		[0.002]	[0.003]			[0.004]	[0.011]
Sellers (θ_s)	μ_s	3.702	6.211	Seller entry probability	$F_{V_0\|\mathbf{Z}}(v_0^*)$	0.859	0.816
		[0.053]	[0.054]			[0.003]	[0.006]
	k_s	1.048	0.320	Opportunity cost	$e_B^o (= e_S^o)$	6.725	20.574
		[0.019]	[0.018]			[0.154]	[0.844]

Standard errors are reported in square brackets and are obtained with 250 bootstrap repetitions.

3.6.2 Estimates of idiosyncratic valuations

Estimated parameters from the distributions of idiosyncratic (potential) bidder and seller valuations are presented in Table 3.5. Standard errors are obtained from 250 bootstrap samples (see Horowitz (2001)) and include variability from the first-stage regressions.[53] While the conditional valuation distributions for these two populations are allowed to be different, parameter estimates reveal that the two user groups are very similar in that dimension. In fact, the bootstrapped 95% confidence intervals for the mean and standard deviation of the respective distributions overlap, both in the main and high-value samples. This finding is intuitive given the consumer-to-consumer nature of the wine auction platform, also supported with the statistic that only about 30 percent of sellers has feedback from previous transactions. Of course, selective seller entry drives a wedge between their valuations and those of bidders, delivering (additional) gains from trading on the platform. In high-value samples, even conditional on observed characteristics, estimated taste distributions have a higher mean but lower dispersion.

Fit and validation.
While a log-logistic specification is a common choice to parameterize value distributions in the auction literature, suitability of this distribution has yet to be evaluated. I compare estimation results with those obtained from a different distribution. The Log-logistic distribution has a similar shape but heavier tails, which can deliver significantly different results in auction models where bidding and reserve prices depend on these tails. Table 3.6 compares model fit of estimation results obtained with both parameterizations in the main

[53]In a subset of bootstrap samples the non-linear solver finds no (or an extreme) optimizer and those cases are excluded. In my experience such cases could partly be avoided by tailoring starting values.

sample and shows that the log-logistic has a slightly better fit - especially for predicted reserve prices.

For the distribution of idiosyncratic valuations, I use the mean squared deviation between the observed second-highest bid and its predicted value (including the predicted quality level $\hat{g(\mathbf{Z})}$ from the homogenization step) as a measure of model fit. Table 3.6 provides this statistic separately for auctions with 2-10 bidders and in expectation over all number of bidders exceeding 1. Both are reported in auctions without reserve prices to focus only on the fit of the distribution of idiosyncratic bidder valuations. Results show that estimated parameters replicate the second-highest bid well, even when slicing the data in bins with 2-10 bidders. The left-hand panel in figure 3.3 supports this statement visually; in that case it compares the homogenized estimated value distribution with the nonparametric estimates for 2-8 bidders. Nonparametric estimation follows directly from the identification argument of $F_{V|\mathbf{Z}}$ in (3.9) by applying $\phi^{-1}(.;n)$ to the empirical probability that the second-highest bid in auctions with n bidders is less than v (given that $c_B = 0$ on the BW platform).

Estimation of seller parameters is evaluated by comparing the distribution of predicted values (including estimated quality) with the distribution of observed reserve prices. I report the test statistic from a two-sample Kolmogorov-Smirnov test (see equation (3.29) for details) with null hypothesis that observed reserve prices and predicted reserve prices are drawn from the same population distribution. This statistic suggests that the log-normal distribution has a better fit than the log-logistic. With a p-value of 0.101, I cannot reject the null at the five or 10 percent level. The right-hand panel in Figure 3.3 plots observed and predicted reserve prices. Values are non-homogenized so the predictions include an expectation over draws from the empirical distribution of $\hat{g(\mathbf{Z})}$. While the overall fit is good, the model seems to overestimate the share of reserve prices in the £$10 - 20, -$ bracket.

Relatedly, the implied seller value (given $\hat{\theta}_b, \hat{g(\mathbf{Z})}$) according to (3.11) is for 3.17 percent of sellers estimated to be negative. Assuming that their valuations are in fact non-negative, this could be driven by a portion of sellers setting reserve prices below the optimal levels or by small-sample estimation bias stemming from first-stage estimates $\hat{\theta}_b$ and $\hat{g(\mathbf{Z})}$ and potentially approximation error in the reserve price (see Appendix B). Note that the equilibrium mark-up is highest for sellers with the lowest valuations. It is conceivable that other sellers with low valuations set sub-optimal reserve prices. Estimation of θ_s excludes the 3.17 percent of positive reserve price auctions that violate non-negativity of ($r -$ estimated mark-up).

Table 3.6: Model fit and validation statistics

Parameters of $F_{V\|\mathbf{Z}}$, $F_{V_0\|\mathbf{Z}}$ (and $g(\mathbf{Z})$)			Entry equilibrium	
Second-highest bid	Log-logistic	Log-normal	**Probability $N = n$ (when $r = 0$)**	
Mean squared deviation with predicted			Absolute difference with Poisson	
- 2 bidders	0.584	0.578	-n=0	0.001
- 3 bidders	0.397	0.397	-n=1	0.011
- 4 bidders	0.356	0.356	-n=2	0.013
- 5 bidders	0.324	0.324	-n=3	0.002
- 6 bidders	0.250	0.250	-n=4	0.022
- 7 bidders	0.210	0.207	-n=5	0.008
- 8 bidders	0.177	0.170	-n=6	0.010
- 9 bidders	0.218	0.208	-n=7	0.002
- 10 bidders	0.214	0.202	-n=8	0.001
- Expectation over N	0.431	0.431	-n=9	0.002
			-n=10	0.001
Reserve price				
Kolmogorov-Smirnov* statistic:	0.065	0.101		

Results for auctions in the main sample. * The KS statistic refers to the Kolmogorov-Smirnov test statistic, as explained in relation to equation (3.29), with a higher KS statistic indicating a better fit. Both predicted reserve prices and predicted highest bids include draws from the estimated quality in the data, $g(\hat{\mathbf{Z}})$, and estimated valuation parameters, $(\hat{\theta_b}, \hat{\theta_s})$. The mean squared deviation statistic is the difference between predicted second-highest bid and its observed value in auctions without a reserve price. The probability that the number of bidders $N = n$ is also from the sample with no reserves.

(a) Bidder valuations (b) Reserve price

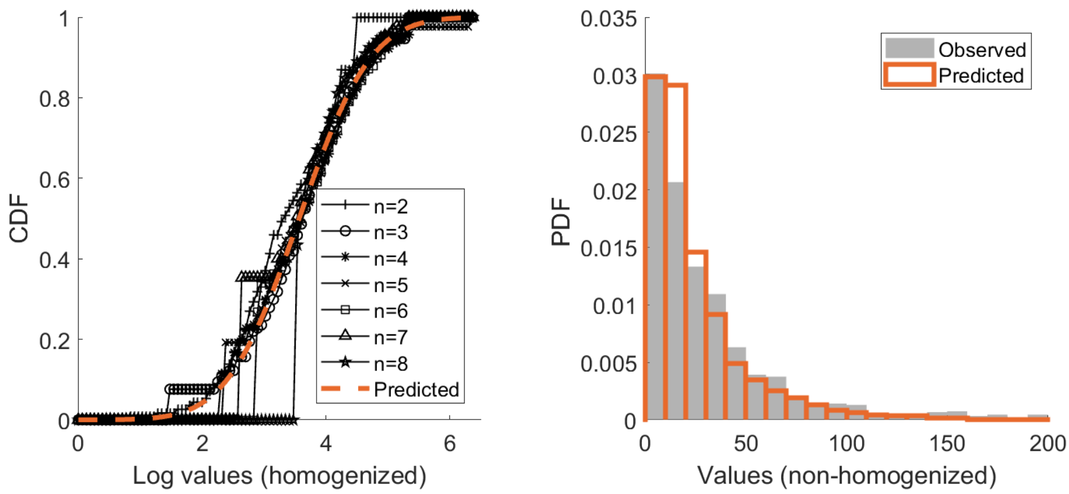


Figure 3.3: Model fit and validation. On the left, a comparison of the estimated idiosyncratic bidder valuation distribution (thick orange dashed line) in auctions with no reserve price against nonparametric estimates of this distribution in auctions with $n = 2, ..., 8$ bidders. On the right, a comparison of observed reserve prices and predicted values including estimated quality $\hat{g}(\mathbf{Z})$. As quality is estimated in the sample of auctions with no reserve price, the orange bars include out-of-sample predictions.

3.6.3 Entry estimates

The estimated Poisson parameters are $\lambda^*_{r>0} = 4.3$ and $\lambda^*_{r=0} = 5.2$, so setting no reserve price attracts on average one additional bidder into the listing (Table 3.6). It also makes intuitive sense that this participation differential is larger in the high-value sample; the probability of being the sole entrant and winning the more expensive bottle for the 1 pound opening bid contributes more to expected surplus. In high-value auctions $\lambda^*_{r>0} = 4.6$ and $\lambda^*_{r=0} = 7.3$. The estimated opportunity cost are also significantly different. They are estimated as the value that makes a potential bidder indifferent between entering and not entering, following the identification argument for e^o_B. Estimated values are 6.7 pounds for the main sample and 20.6 pounds for the high-value sample. This corresponds to the idea that listing inspection cost are significant at about 5-14 percent of the average transaction price. On the seller side, also e^o_S is identified from the zero profit condition of the marginal seller, but using its empirical counterpart is less attractive as it involves estimating seller opportunity cost from only one observation. Monte Carlo simulations (Appendix E) confirm that the value of e^o_S is a normalisation for the estimation of seller parameters, and as such I use the estimated value of e^o_B.

Fit and validation.
The right-hand side panel in Table 3.6 examines the fit of the assumed Poisson distribution with the estimated $\lambda^*_{r=0}$ and the observed Binomial distribution of the number of bidders. It shows that for all $n = 1 - 10$ the Poisson fits good. Figure 3.4 support this visually. More formally, a chi-square goodness of fit test fails to reject at the ten percent level that N is generated by a Poisson distribution, based on auctions with no reserve price in which the number of bidders is not censored (p-value 0.106).

It is of particular interest that the data does not reveal any overdispersion relative to the Poisson distribution. That would point to an entry process in which bidders enter significantly more numerously into auctions with certain characteristics - conditional on having no reserve price. Some characteristics may be slightly more popular than others, but a model with uniform sorting over listings captures the first order effects of entry behavior in the BW data.[54]

It turns out to be complicated to explain *both* the share of sellers that sets a zero reserve

[54] By contrast, estimating a model of entry in Kindle e-reader auctions on eBay, Bodoh-Creed et al. (2013) need to incorporate auction observables in their single-index conditional Poisson distribution to explain the observed pattern of a higher number of bidders in listings with certain characteristics. While it may be interesting to condition on additional characteristics in my data as well, it is not necessary to fit the data well.

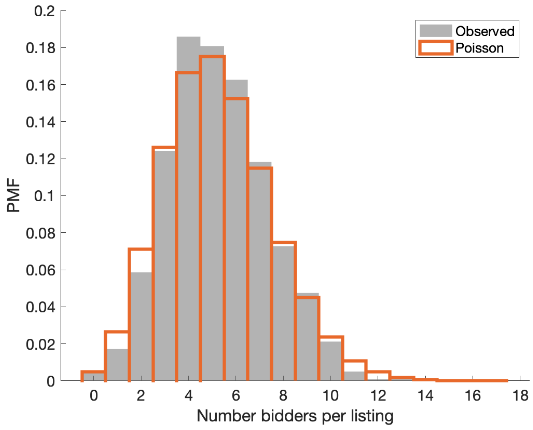


Figure 3.4: The number of bidders per listing is approximately Poisson distributed. A chi-square goodness of fit test fails to reject, at < 1 percent level, that N is generated by a Poisson distribution. The test uses data from auctions with no reserve price in which the number of bidders is not censored.

price *and* additional entry of bidders in no-reserve auctions, assuming these choices are optimal to maximize both bidder and seller expected surplus. In particular, the observed p_{r0} is equal to 0.3740 in the main sample while $p^*_{r0} = 0.6322$ calculated to be optimal given model estimates and the bidder entry differential. In the high-value sample, this bias has the same sign but a lower magnitude: $p_{r0} = 0.2459$ instead of the optimal $p^*_{r0} = 0.4246$. As the first structural model to explain endogenous entry of both bidders and sellers in an auction setting with interconnected pay-offs, also endogenizing the choice for sellers to set no reserve price is perhaps a bridge too far. Instead, in my counterfactual simulations p^*_{r0} and $\lambda^*_{r>0}$ adjust optimally subject to keeping $\hat{\beta}_s = p^*_{r0} - p_{r0}$ and the bidder participation differential $\hat{\beta}_b = \hat{\lambda}^*_{r=0} - \hat{\lambda}^*_{r>0}$ at their estimated levels.[55]

[55]One potential explanation for the inconsistency could be that the platform explicitly encourages sellers not to set a reserve price in order to attract "50-75 percent" more bidders and a "40 percent" higher transaction price. If these rules of thumb don't correspond with the actual benefit for sellers to adopt no reserve price, it may explain part of the inconsistency. I estimate that no-reserve auctions attract on average about 22 percent (main sample) to 58 percent (high-value sample) more bidders. Overall, I do consider that enodgenizing sellers' choice to set no reserve price to be important given the aim to understand the impact of fees, despite the complexity in capturing this choice perfectly in this two-sided market setting.

3.7 Counterfactual policies

This section uses model estimates to address two key indeterminacy's of two-sided markets in the context of the wine auction platform. First, how should the platform allocate fees between different platform users? Second, how do increases in fees affect users on both sides? The latter is an open question in antitrust policy that is crucial to better understand the unusual economic relationships in platform markets.[56]

3.7.1 Harnessing network effects to increase platform profitability

In this counterfactual, I simulate the game for a host of alternative fee structures. At each fee combination, I compute platform revenue, bidder and seller surplus, and the volume of sales.[57] Many changes in fees result in a trade-off between the volume of sales and platform fee revenue. Higher listing fees, for example, makes it less attractive for sellers to enter and as a result fewer listings depresses the volume of sales. Even if higher fee revenues are beneficial in a static world, if the volume of sales affects future revenues through (say) word of mouth or brand awareness, a forward-looking platform will include this statistic in their objective function.[58] I therefore estimate both the impact of alternative fee structures on static fee revenues as well as the volume of sales, and consider the problem of maximizing current volume-constrained fee revenues. This approach avoids having to impose further restrictions on the exact platform objective function.

Results show that there are significant network effects in this market that can be harnessed to improve platform profitability.[59] Contour plots in Figure 3.5 support this visually; with the levels referring to the share of current fee revenues. The top panel varies buyer and seller commission but holds flat fees at their current levels. For example, adopting any (c_B, c_S) pair that intersects on an area of the plot with level 1 (the turquoise-green colour) would result in the same platform revenue as generated by the current fee

[56]See e.g., Evans and Schmalensee (2013).

[57]Results expressed as changes with respect to baseline values and are computed in homogenized value space.

[58]This is consistent with a model of network growth with myopic users who can terminate their participation at no cost, as provided in Evans and Schmalensee (2010) to explain the emphasis of platforms such as eBay, Facebook and MySpace on network growth in their early years. The idea that sale volume is relevant is also consistent with an alternative fee structure that the platform discussed with me in general terms. I estimate that that fee structure would reduce expected static platform revenue by as much as 22 percent while increasing the sale volume by 6 percent. As such, for this to be a smart policy they would need to place considerable weight on the volume of sales.

[59]The presence of network effects is detected in the contour plot from the fact that only changing the allocation of commissions between buyers and sellers, while keeping their total levels constant, has significant effects on platform profitability. In particular, this is a requirement for the market to be two-sided by its standard definition in Rochet and Tirole (2006).

structure. The bold red line indicates the baseline volume of sales. Any fee combination to the south west of this line increases the sale volume, and any fee combination to the north east decreases volume. Note that it is not a coincidence that baseline volume and revenue (level 1) intersect at the current commission allocation, the point indicated in the Figure with a red marker.

A key take-away is that the BW platform can *increase fee revenues with up to 80 percent without reducing volume* by increasing c_S and decreasing c_B when holding current entry fees fixed (top panel in Figure 3.5). For example, increasing the seller commission by 15 percentage points and reducing the buyer premium by the same amount increases platform revenues by about 30 percent. As the buyer premium is currently zero, this would require charging a negative buyer premium, e.g. to give winning bidders a 15 percent discount on the transaction price. While negative fees may seem unintuitive, it fully agrees with the idea that businesses in two-sided markets subsidize the side that contributes most to profits even to the point of charging that side below marginal cost.[60] Commissions, which users are familiar with in wine auction platforms, can feasibly be set to negative values as they only apply to successful sales.

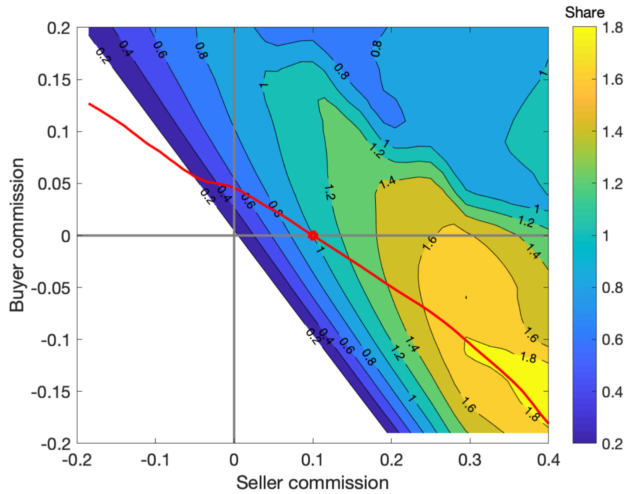
I also consider the problem with additional (self-imposed) non-negativity constraints on c_B and c_S, which is reflected in the top panel of Figure 3.5 by the two grey lines. The contour plot shows clearly that, when keeping other fees at their current levels, the platform cannot increase its revenues without reducing volume if it does not want to set a negative buyer commission. All commission combinations that do increase volume reduce platform revenue. Furthermore, even a small reduction in seller commission without additional fee changes is risky as it can lead to significantly lower platform revenue.

Figure 3.6 displays platform revenues for different combinations of seller commission and listing fee, keeping other fees at their current levels. The non-negativity constraint on the listing fee is in this case crucial to avoid sellers listing unsellable items to collect the fee. The current volume constraint is displayed in red. The plot shows that there is significant scope to increase revenues without reducing volume by increasing the listing fee. Sellers must be rather inelastic with respect to the listing fee to generate the contours observed.[61]

[60]See e.g., Rochet and Tirole (2003, 2006), Wright (2004), Armstrong (2006), Rysman (2007), Evans and Schmalensee (2013) Rysman and Wright (2002). I am not aware of any (wine) auction platform that consistently charges negative prices, although eBay does provide temporary discounts and coupons to its posted price sales.

[61]This economic reality does not seem lost on current BW management. One change to the fee structure they consider making is to increase the listing fee from 1.75 to 4 pounds. They also consider waiving this fee for sellers that set a zero reserve price. In my assessment of the discussed fee structure I find that this policy would increase the share of zero reserve price listings by 19 percent. However, I estimate that

a) Holding listing fee at current level of 1.75 pounds



b) Increasing listing fee to 5 pounds

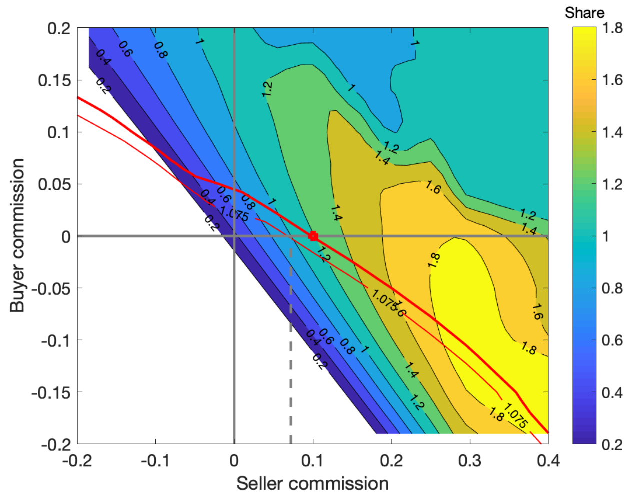


Figure 3.5: Platform revenue at counterfactual commissions, with volume constraint

Contour plots of counterfactual platform revenue as a share of current revenue, for combinations of c_B and c_S, holding entry fees (e_S, e_B) at current level (1.75,0) (top) and increasing e_S to 5 pounds (bottom). The red line indicates current volume of sales; fee combinations to the south west of this line increase volume.

Combining the above findings, the bottom panel in 3.5 displays revenue for combinations of commissions while increasing the listing fee to 5 pounds and keeping the bidder

the impact of this significant change in listing composition does by itself not affect platform revenue or volume by much. An interesting question that I leave for future work is whether and how a platform should optimally nudge sellers in their pricing decisions. This is related to the theoretical observations that when $v_0 > 0$, sellers will set too high reserve prices from the platform's perspective as sellers trade-off expected sale revenues with the value of keeping the item.

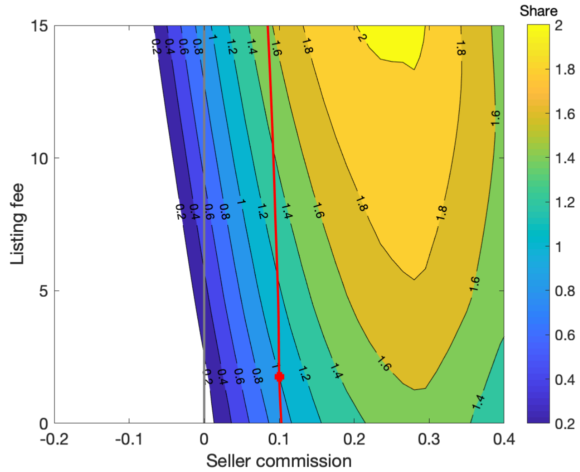


Figure 3.6: Platform revenue at counterfactual seller fees, with volume constraint

Contour plots of counterfactual platform revenue as a share of current revenue for combinations of e_S and c_S, holding (c_B, e_B) at current and constrained-optimal level (0,0). The red line indicates the current volume of sales; fee combinations to the south west of this line increase volume.

entry fee at the current level of 0. It paints a remarkable picture. Even when considering only positive commissions, platform revenues increase by roughly 20 percent without affecting volume when keeping commissions at their current levels. Alternatively, they can *increase volume by 7.5 percent without affecting revenues* by combining the listing fee increase with a reduction in seller commission of about three percentage points to 0.07.

3.7.2 Users are better off with alternative fee structures

Figure 3.7 provides contour plots for expected seller and winning bidder surplus as a share of their current levels. The dark blue and red lines tie these values in with the platform objective function; the dark blue *swirl* indicates the current level of platform revenues (any fee combination to the north east increases it), and the red curve indicates the current level of sale volume (any fee combination to the south west increases it). Estimates refer to the case that increases the listing fee to 5 pounds while keeping other fees at baseline levels. A key take-away from this picture is that for a region of combinations of c_B and c_S, implementing the negative buyer premium policy leaves both sellers and winning bidders better off. Consider the fee structure highlighted above for example, increasing the seller commission with 15 percentage points (to 25 percent) and providing winning bidders with a 15 percent discount. While it is intuitive that this fee structure increases winning bidder surplus, it also is estimated to increase seller surplus by about 20 percent. Especially the beneficial impact on sellers is striking because their fees go up significantly in this scenario. This result resonates with the idea that bidder participation is very valuable to sellers in many auction settings, as additional bidders drive up transaction prices.[62]

3.7.3 Ignoring entry significantly biases welfare estimates

In this counterfactual, I evaluate the impact of isolated changes in seller commission and buyer premium on user welfare and compare it to results obtained when using a model without entry. As this issue is relevant for antitrust policy, I use a prominent commission-fixing case involving auction giants Sotheby's and Christie's (SC) to provide context. After the conspiracy came to light, they settled with buyers and sellers for a total of 512 million dollars (roughly 729 million dollars in 2018 prices) and five sixths of this amount went to buyers. Civic case litigation makes clear that damage estimates are based on direct (alleged) overcharges and not pass through or their economic incidence and therefore abstracted from *any* interconnectedness between users or their entry decisions.[63] I estimate the welfare impacts of a five percentage point increase in buyer commission, which is most likely what the SC case settlement is based on, and compare results with those resulting from a five percentage point increase in seller commission.

The empirical results in Table 3.7 show that the five percent increase in buyer premium decreases the sale price and probability by about 4 percent, reduces the share of listings with a zero reserve price, and reduces the number of bidders per listing. Expected winning bidder surplus decreases by about 7 percent. I also estimate that sellers are significantly worse off: their expected surplus decreases by 17 percent. A finding that underlines how important it is to incorporate the interconnectedness of users in platform markets is that sellers are better off if their seller commission goes up by 5 percentage points than when the buyer commission increases by the same amount. In fact, this feature of the platform marketplace would be missed if users' endogenous entry decisions would be ignored. The columns labelled "Without entry" in Table 3.7 provide welfare impacts when bidder and seller participation is kept constant at baseline levels and only their bidding and reserve pricing strategies respond to the increases in commissions. Results from such a model suggest that sellers would instead prefer if the five percent commission increase is targeted

[62] For example, Bulow and Klemperer (1996) show that setting no reserve and attracting one additional bidder is more profitable than negotiating with fewer participants and using a reservation price. The impact of entry in the case where bidders enter selectively is theoretically ambiguous, but Roberts and Sweeting (2010) show that for in USFS timber auctions the value of increasing the pool of (potential) entrants also outweighs the value of setting a reserve price.

[63] See In re Auction Houses antitrust litigation (2001) for the full litigation text. Ashenfelter and Graddy (2005) provide a more detailed description of the case.

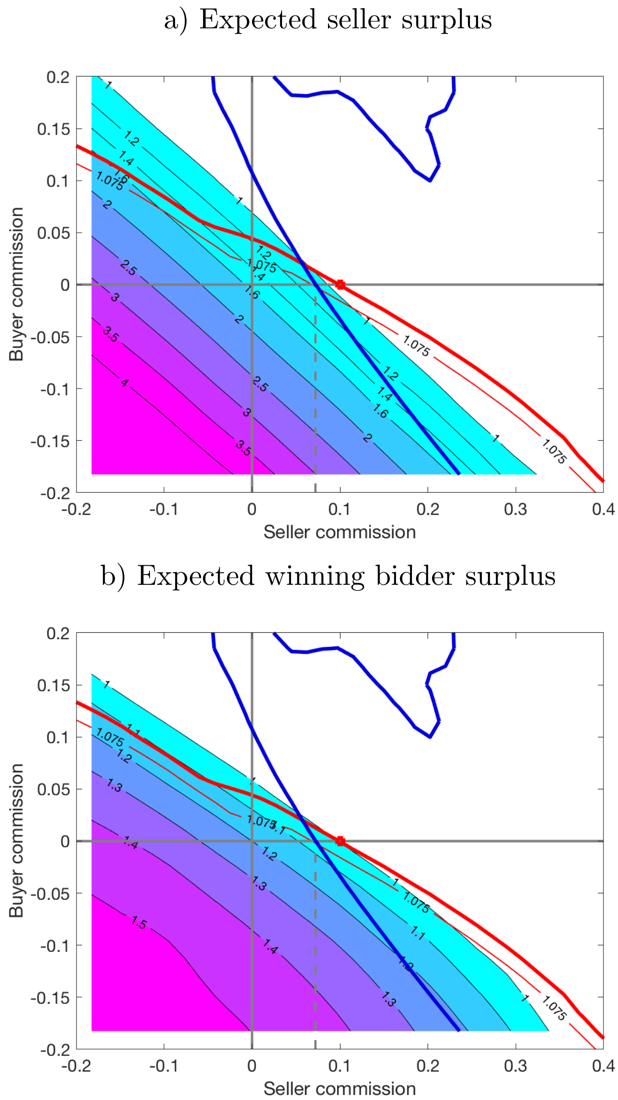


Figure 3.7: Contour plots of expected seller and winning bidder surplus showing that all parties could be better off under counterfactual fee structures.

Levels indicate the surplus as a share of current surplus, for combinations of buyer and seller commission when increasing the listing fee to 5 pounds and keeping the bidder entry fee constant at 0. The red line is the platform volume constraint: combinations to the south west increase volume. The blue *swirl* is the platform revenue constraint: combinations to the east increase revenue. An example of a fee structure that makes all parties better off is a negative 10 percent buyer premium and a 20 percent seller commission combined with the 5 pound listing fee.

to buyers.

The estimates also depart from a third model that underlies the argumentation in previous studies looking at welfare impacts of commissions in auctions in relation to the SC case. Ashenfelter and Graddy (2005) conclude: *though buyers received the bulk of the* *damages, a straightforward application of the economic theory of auctions shows that it is unlikely that successful buyers as a group were injured.* This conclusion relies on the idea that buyers reduce their bid by the amount of buyer premium and sellers accept any price so that the economic incidence of any commission or premium fully falls on sellers because they are the price-inelastic party.[64]

The simulated commission increases show that in the BW data, increasing the commission to one side of the market by 5 percent decreases expected surplus to that side more than proportionally. On top of that, the other side is affected as well. Not addressing entry in user welfare evaluations significantly underestimates expected loss in surplus. The 5 percent increase in buyer premium that likely was the basis of the SC settlement negotiations decreases total user surplus by about 12 percent, so *more than double* the premium increase.[65]

Besides the amount of total damages estimated, also awarding five-sixths of the damages to winning bidders was plausibly flawed.[66] I show that in the BW data, the economic incidence of commissions falls predominantly on sellers despite endogenous entry and also when accounting for optimal reserve prices. The loss in user surplus resulting from a 5 percent buyer premium increase is estimated to fall for 71 percent on sellers:

	c_B +5%		c_S +5%	
	Two-sided	No entry	Two-sided	No entry
Seller share of damage	0.71	0.80	0.89	0.89
Winning bidder share of damage	0.29	0.20	0.11	0.11

That the majority of the welfare loss from increasing commissions falls on sellers remains a valid conclusion also when holding bidder and seller participation fixed at baseline levels. My estimates furthermore reveal that sellers shoulder 71-89 percent of the loss in user surplus *regardless of which commission increases by 5 percentage points.*

3.7.4 Larger effects in high-value auctions

Results in the previous section are estimated with the main sample that excludes auctions exceeding 200 pounds. This section exploits the remaining data to examine whether results

[64]This argumentation also underlies evaluation of the role of commissions in (wine) auctions by Marks (2009) and McAfee (1993). I have referred to this as the "one-sided market perspective".

[65]Relating this statistic to the numbers in Table 3.7; a 5 percent buyer premium increase reduces total user surplus by 11.87 percent: $100*(((1-0.07169)*7.719+(1-0.17243)*6.757)-(7.719+6.757))/(7.719+6.757) = -11.87$.

[66]Ashenfelter and Graddy (2005) and Marks (2009) hypothesized this before based on the one-sided market perspective with inelastic sellers in which case winning bidders logically avoid a reduction in surplus by reducing their bid proportionally to any buyer premium increase.

Table 3.7: Welfare impacts of commission changes in comparison to benchmark without endogenous entry

	Baseline	Increase buyer premium 5%		Increase seller commission 5%	
		Two-sided	Without entry	Two-sided	Without entry
Buyer premium, c_B	0	0.05	0.05	0	0
Seller commission, c_S	0.10	0.102	0.10	0.15	0.15
		Percentage change w.r.t. Baseline:			
Platform revenue (x10.000)	2.715	25.774	31.734	30.944	36.064
Volume of sales (x10.000)	20.769	-10.597	-5.613	-3.882	-0.460
Expected surplus sellers (x10.000)	6.757	-17.243	-10.777	-14.665	-13.355
Expected surplus winning bidders (x10.000)	7.719	-7.169	-2.428	-3.943	-1.395
Expected surplus per seller	25.12	-15.782	-10.777	-12.387	-13.355
Expected surplus per winning bidder	28.695	-5.530	-2.428	-1.379	-1.395
Seller entry probability, $F_{V_0\|\mathbf{Z}}(v_0^*)$	0.855	-1.903	0	-2.633	0
Mean number bidders per listing, $\lambda^*_{r>0}$	4.759	-5.177	0	-1.853	0
Share listings with no reserve, p_{r0}	0.374	-8.881	0	-2.062	0
Sale probability	0.869	-3.834	-2.228	-1.13	-1.215
Sale price	91.189	-4.175	-2.286	-0.147	1.516

Table 3.8: Opposing fee structures and their differential impact on high-value lots

	Combined	High-value	Regular	Combined	High-value	Regular
	Charge seller commission			Charge buyer premium		
buyer premium, c_B	0	0	0	0.1	0.1	0.1
Seller commission, c_S	0.1	0.1	0.1	0	0	0
Platform revenue (x10.000)	5.133	2.468	2.665	4.064	1.662	2.402
Volume of sales (x10.000)	44.120	23.317	20.803	33.079	15.153	17.926
Share revenue from high-value	0.481	1	0	0.409	1	0
Share volume from high-value	0.528	1	0	0.458	1	0
Share of listings high-value	0.179	1	0	0.185	1	0
Share of bidders high-value	0.158	1	0	0.122	1	0
Expected surplus sellers (x10.000)	8.542	1.761	6.781	7.51	1.063	6.447
Expected surplus winning bidders (x10.000)	12.267	4.534	7.733	10.336	3.047	7.289
Expected surplus per seller	27.565	29.948	25.182	20.544	17.308	23.78
Expected surplus per winning bidder	11.596	17.56	5.631	11.31	16.983	5.637
Seller entry probability, $F_{V_0\|\mathbf{Z}}(v_0^*)$	0.881	0.996	0.855	0.862	0.863	0.861
Mean number bidders per listing, $\lambda^*_{r>0}$	4.559	3.745	4.748	4.023	2.543	4.455
Share listings with no reserve, p_{r0}	0.354	0.243	0.378	0.299	0.142	0.335
Sale probability	0.828	0.646	0.868	0.757	0.45	0.826
Sale price	188.21	633.052	91.072	176.159	584.581	83.657

are different in auctions that are of higher “quality”.

The first three columns in Table 3.8 report estimated platform revenue, user surplus and other relevant platform descriptors at ($c_S = 0.1, c_B = 0$) and the last three columns show these statistics for the reverse fee structure ($c_S = 0, c_B = 0.1$). The first take-away from the Table supports the importance of high value listings. When charging $c_S = 0.1$ for instance, while less than one fifth of listings and bidders is high value, those interactions generate about half of both platform revenue and volume. Columns 3 and 6 correspond to auctions with transaction prices below 200 pounds, e.g. those in the main sample central to the previous experiment. Platform revenue, total seller surplus and total surplus for winning bidders is all higher when charging sellers rather than buyers, reinforcing earlier findings to that effect. But while these differences are relatively modest, in high value auctions all parties are *significantly better off* when charging a 10 percent commission to sellers rather than buyers.

The different commission elasticities may be part of the explanation for the observed variation in fee structures in various wine auction platforms.[67] Another suggestion offered in Table 3.8 is that a platform targeting only high-value users (those offering and bidding for high-value lots) can generate similar revenues as a platform attracting lower-priced lots with only one fifth of listings.

3.8 Conclusions

In this paper I examine the welfare impacts of fees charged in auction platforms using a new dataset of wine auctions web-scraped from an online wine auction platform. Despite the platform marketplace, empirical patterns suggest an absence of dependencies between listings. This is inconsistent with theoretical predictions from previous auction platform models tailored to explain strategic behaviour in auction platforms for more homogeneous goods but can be explained by costly inspection of the idiosyncratic goods for sale. The structural auction platform model that also includes endogenous entry of bidders and sellers and seller/listing selection plausibly captures the main features of other auction platforms as well, such as those for freelance jobs and second hand cars.

Empirical findings suggest that it would be a mistake for antitrust authorities or the legal system not to consider interconnectedness of bidders and sellers on an auction platform.

[67] Many wine auction platforms take a comparable portion of transaction prices in total commissions but allocate commissions very differently to the two user groups. For example, the US-based higher-end Acker, Merrall & Condit charges 21 percent buyer- and no seller commission, the Chicago Wine Company has the reverse structure charging nothing to buyers and 28 to sellers and Winebid.com charges both sides evenly at respectively 15 and 18 percent.

Increasing fees to one side of the market affects both types of users, and network effects from entry are such that sellers are better off when the seller commission increases than when the buyer premium increases by the same amount. Results are placed in the context of a prominent commission-fixing case involving auction giants Sotheby's and Christie's. I show that relative to the benchmark rule that does not consider the economic incidence of the alleged overcharge or endogenous entry, total estimated damages are more than twice as high. Also, while sellers in the civic case received only one sixth of total damages, the economic incidence of increasing commissions by 5 percentage points falls to 71-89 percent on sellers. More conceptually, these counterfactual fee experiments highlight that it is feasible to estimate exact welfare impacts of platform fees from publicly available data, despite strategic interactions rendering these effects ex-ante ambiguous and despite the auction mechanism causing transaction prices to be endogenous to the fees.

Results also show that the platform can increase revenues from fees without reducing the volume of sales by (further) subsidizing bidders. Platform profitability can increase up to 80 percent when introducing a negative buyer premium, e.g. providing winning bidders a discount on the transaction price, paid for by a higher seller commission. When implemented right, this policy could even increase expected surplus for sellers and winning bidders on the platform. Although no auction platform currently charges negative fees, this recommendation is in line with the practice in other two-sided markets such as payment cards. I demonstrate that this fee policy could be implemented in a way that also increases expected surplus for potential bidders and sellers on the platform.

Building on the model and results in this paper, a logical next step is to consider the platform fee setting problem more explicitly. The resulting structure could be used to predict the impact of mergers between auction platforms, thereby contributing to an empirical literature studying (price-)impacts of consolidation in two-sided markets (including e.g. Chandra and Collard-Wexler (2009), Song (2013) and Fan (2013)).

Another direction for further research is to test predictions from this paper using experimental variation in fees. This would be in line with e.g. Ostrovsky and Schwarz (2016) who conduct a field experiment to show that revenues from Yahoo! ad auctions increase significantly when adopting the optimal Myerson (1981) and Riley and Samuelson (1981) reserve prices.

Acknowledgements

I am indebted to Andrew Chesher, Phil Haile, Lars Nesheim, Adam Rosen and Áureo de Paula for fruitful discussions, continued encouragement and guidance. I also thank Dan Ackerberg, Matt Backus, Dirk Bergemann, Natalie Cox, Martin Cripps, Tom Hoe, Hyejin Ku, Kevin Lang, Brad Larsen, Konrad Mierendorff, Imran Rasul, José-Antonio Espín-Sánchez, Michela Tincani, Jean Tirole, Frank Verboven, Martin Weidner and seminar participants at: Yale IO Prospectus workshop, Duke Micro-econometrics seminar, EARIE, AAWE, UCL PhD seminar and UCL structural estimation breakfast for helpful comments. I am grateful for financial support from the Economic and Social Research Council (PhD Studentship). All remaining errors are my own.

Appendices

A Omitted tables and figures

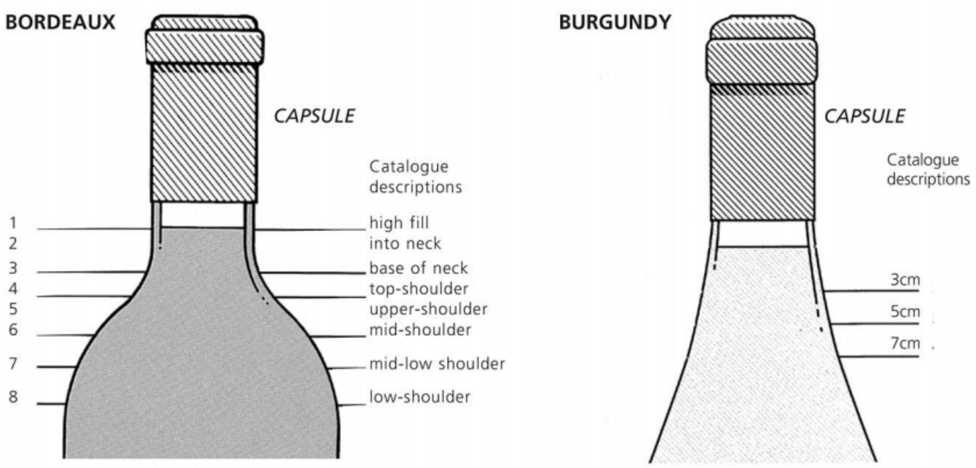


Figure 3.8: Ullage classification and interpretation

Source: Christie's (2013). Numbers refer to auction house *Christie's* interpretation of the fill levels, which are for Bordeaux-style bottles: 1) Into Neck: level of young wines. Exceptionally good in wines over 10 years old. 2) Bottom Neck: perfectly good for any age of wine. Outstandingly good for a wine of 20 years in bottle, or longer. 3) Very Top-Shoulder. 4) Top-Shoulder. Normal for any claret 15 years or older. 5) Upper-Shoulder: slight natural reduction through the easing of the cork and evaporation through the cork and capsule. Usually no problem. Acceptable for any wine over 20 years old. Exceptional for pre-1950 wines. 6) Mid-Shoulder: probably some weakening of the cork and some risk. Not abnormal for wines 30/40 years of age. 7) Mid-Low-Shoulder: some risk. 8) Low-Shoulder: risky and usually only accepted for sale if wine or label exceptionally rare or interesting. For Burgundy-style bottles where the slope of the shoulder is impractical to describe such levels, whenever appropriate [due to the age of the wine] the level is measured in centimetres. The condition and drinkability of Burgundy is less affected by ullage than Bordeaux. For example, a 5 to 7 cm. ullage in a 30 year old Burgundy can be considered normal or good for its age.

Table 3.9: First stage estimation results, main sample

Dep. var: log transaction price per bottle with 2 or more bidders	Estimate	Std. error	p-Value
Intercept	4.042	0.165	0
Case of 6 bottles	0.276	0.096	0.004
Case of 12 bottles	1.473	0.2	0
Stored in warehouse	0.506	0.211	0.016
Special format (not 75cl)	0.229	0.068	0.001
Number bottles in lot	-0.215	0.017	0
Fill level: Base of Neck (BN)	-0.215	0.068	0.002
Fill level: High Shoulder (HS)	-0.394	0.148	0.008
Fill level: Low Shoulder (LS) or worse	-0.364	0.19	0.056
Fill level: Missing	0.012	0.055	0.83
Fill level: Mid Shoulder (HS)	-0.427	0.159	0.007
Fill level: Top Shoulder (TS)	-0.401	0.136	0.003
Fill level: Very Top Shoulder (VTS)	-0.139	0.116	0.23
Duty estimate	-0.026	0.009	0.004
VAT estimate	0.008	0.006	0.173
Shiping quote	0.007	0.004	0.091
Percentage range provided shipping quotes	-0.923	0.094	0
Percentage range provided shipping quotes, squared	0.361	0.048	0
Delivers to UK	0.066	0.044	0.136
Seller accepts returns	-0.14	0.151	0.355
Bottles can be collected from seller	0.027	0.044	0.535
Buyer can *only* collect bottles from seller	-0.237	0.11	0.032
Shipping with Royal Mail	0.025	0.05	0.622
Shipping with ParcelForce	-0.155	0.046	0.001
Mentions fast shipping	0.444	0.066	0
Insurance for loss or breakage is included in shipping cost	0.098	0.041	0.016
Can pay by bank	0.336	0.089	0
Can pay by Paypall	-0.033	0.045	0.461
Can pay by cheque	-0.139	0.047	0.003
Can pay in cash	-0.088	0.113	0.436
Wine Type: Assorted	0.009	0.068	0.895
Wine Type: White	-0.247	0.079	0.002
Wine Type: Sparkling	0.176	0.105	0.094
Wine Type: Fortified	-0.161	0.125	0.198
Wine Type: Rose	-0.186	0.324	0.566
Region of origin: Tuscany	-0.375	0.106	0
Region of origin: Bordeaux	-0.199	0.069	0.004
Region of origin: Australia	-0.377	0.112	0.001
Region of origin: Burgundy	-0.154	0.095	0.106
Region of origin: Rhone	-0.061	0.094	0.518
Region of origin: Champagne	0.161	0.124	0.195
Region of origin: Provence	-0.43	0.242	0.076
Region of origin: Veneto	-0.358	0.172	0.038
Region of origin: Rioja	-0.337	0.176	0.056
Region of origin: Alsace	-0.191	0.195	0.328
Region of origin: Piedmont/Lombardy	-0.397	0.133	0.003
Region of origin: South Australia	-0.348	0.113	0.002
Region of origin: Douro	-0.349	0.195	0.074
Region of origin: Mendoza	-0.438	0.347	0.206
Region of origin: Scotland	0.228	0.26	0.381
Region of origin: Oporto	-0.217	0.197	0.272
Region of origin: Assorted	0.098	0.084	0.244
Region of origin: Bekaa Valley	0.103	0.366	0.78
Region of origin: United States	-0.103	0.197	0.6
Region of origin: California	0.002	0.159	0.992
Region of origin: Spain	-0.418	0.265	0.115
Region of origin: Cognac	0.672	0.301	0.026
Region of origin: Portugal	0.111	0.195	0.571
Region of origin: Loire	-0.759	0.268	0.005
Region of origin: Cuba	-0.115	0.346	0.74
Region of origin: Italy	-0.328	0.145	0.024
Region of origin: Oregon	-0.704	0.266	0.008
Region of origin: Ribera del Duero	-0.368	0.334	0.271
Region of origin: South Africa	-0.586	0.343	0.088
Region of origin: Islay	0.681	0.473	0.15
Region of origin: South West France	0.166	0.481	0.73
Region of origin: Other	-0.155	0.105	0.14

... Continued on next page

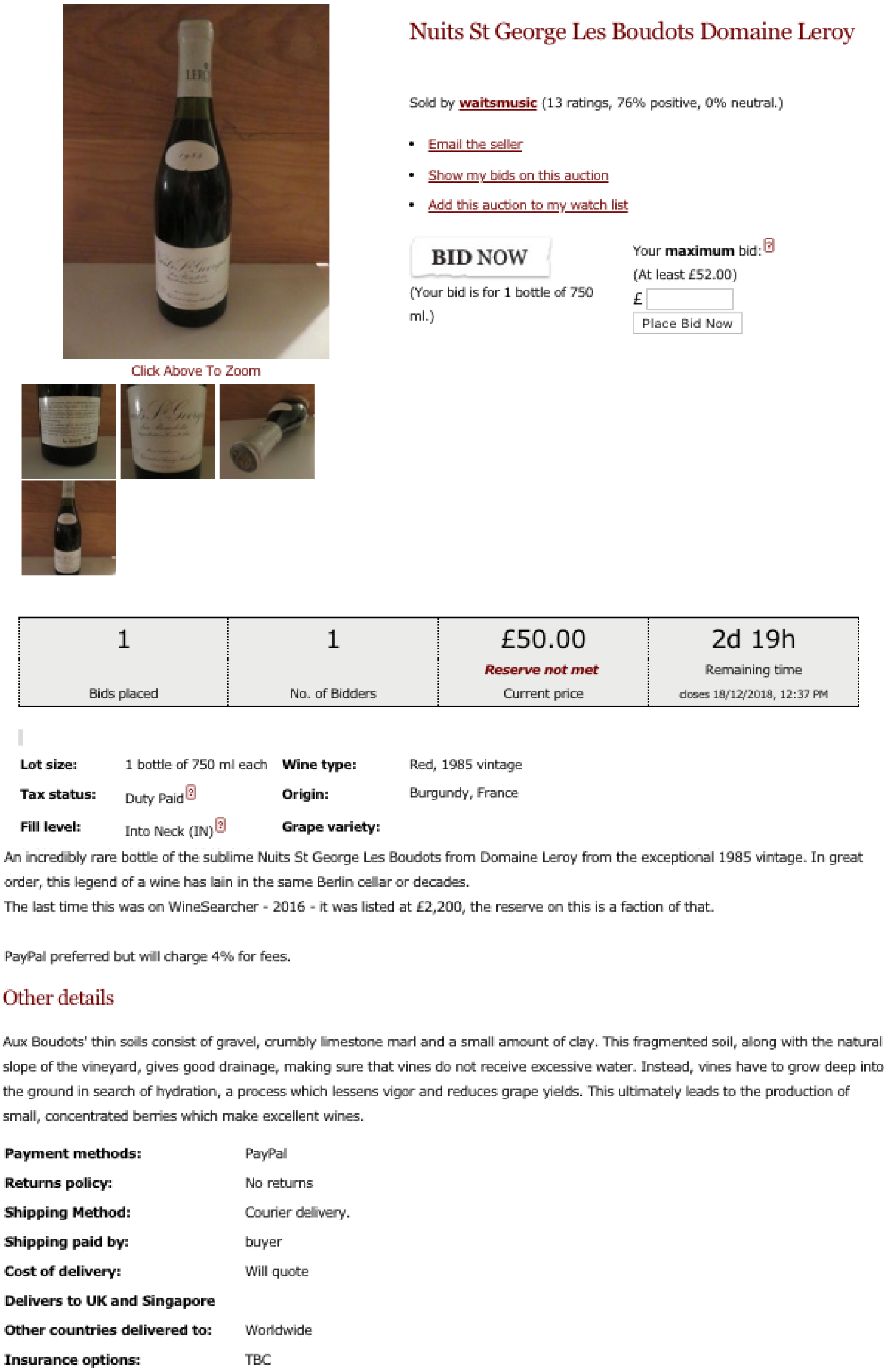
Dep. var: log transaction price per bottle with 2 or more bidders	Estimate	Std. error	p-Value
Grape: Sangiovese	0.348	0.107	0.001
Grape: Corvina	-0.027	0.238	0.91
Grape: Rhone Blend	0.034	0.123	0.783
Grape: Syrah	0.213	0.151	0.157
Grape: Bordeaux Blend	0.153	0.066	0.019
Grape: Other	-0.088	0.099	0.37
Grape: Riesling	-0.008	0.198	0.968
Grape: Nebbiolo	0.253	0.145	0.082
Grape: Cabernet Sauvignon	0.163	0.167	0.328
Grape: Chardonnay	0.207	0.117	0.076
Grape: Tempranillo	0.334	0.209	0.111
Grape: Malbec	-0.19	0.381	0.619
Grape: Pinot Noir	-0.01	0.107	0.926
Grape: Syrah/Shiraz	0.246	0.107	0.022
Grape: Port Blend	0.744	0.243	0.002
Grape: Semillon-Sauvignon Blanc Blend	0.631	0.258	0.015
Grape: Merlot	-0.337	0.193	0.08
Grape: Champagne Blend	0.673	0.228	0.003
Grape: Barbera	-0.367	0.309	0.234
Observations	2007		
Adjusted R^2	0.530		
F-statistic versus constant model (p-value)	16.8 ($8.96e^{-244}$)		

Excluded from table: popular vintage and market month fixed effects. Estimated coefficients = 141.

Table 3.10: First stage estimation results, high-value sample

Dep. var: log transaction price per bottle with 2 or more bidders	Estimate	Std. error	p-Value
Intercept	6.242	0.319	0
Case of 6 bottles	-0.881	0.105	0
Case of 12 bottles	-0.514	0.131	0
Stored in warehouse	-0.705	0.248	0.005
Special format (not 75cl)	0.393	0.145	0.007
Number bottles in lot	-0.121	0.011	0
Fill level: Missing	-0.082	0.09	0.361
Fill level: Base of Neck (BN)	0.03	0.14	0.83
Fill level: Mid Shoulder (HS)	0.1	0.229	0.661
Fill level: Top Shoulder (TS)	0.265	0.355	0.456
Fill level: Very Top Shoulder (VTS)	-0.039	0.164	0.812
Fill level: High Shoulder (HS)	0.177	0.253	0.485
Fill level: Low Shoulder (LS) or worse	-0.13	0.334	0.696
Duty estimate	0.023	0.01	0.021
VAT estimate	-0.003	0.004	0.459
Shipping quote	0.005	0.004	0.16
Percentage range provided shipping quotes	-0.192	0.321	0.549
Percentage range provided shipping quotes, squared	0.064	0.15	0.672
Delivers to UK	-0.293	0.096	0.002
Seller accepts returns	0.002	0.164	0.992
Bottles can be collected from seller	0.036	0.086	0.676
Buyer can *only* collect bottles from seller	-0.546	0.196	0.006
Shipping with Royal Mail	-0.052	0.099	0.605
Shipping with ParcelForce	-0.218	0.12	0.071
Mentions fast shipping	0.124	0.129	0.336
Insurance is included in shipping cost	-0.133	0.076	0.081
Can pay by bank	-0.146	0.211	0.491
Can pay by Paypall	-0.092	0.083	0.27
Can pay by cheque	-0.091	0.078	0.244
Can pay in cash	0.015	0.225	0.948
Observations	390		
Adjusted R^2	0.855		
F-statistic versus constant model (p-value)	18.7 (4.77^{-83})		

Excluded from table: popular vintage, wine type, region, grape and market month fixed effects. Estimated coefficients = 133.



Nuits St George Les Boudots Domaine Leroy

Sold by **waitsmusic** (13 ratings, 76% positive, 0% neutral.)

- Email the seller
- Show my bids on this auction
- Add this auction to my watch list

BID NOW

(Your bid is for 1 bottle of 750 ml.)

Your **maximum** bid:

(At least £52.00)

£

Place Bid Now

Click Above To Zoom

1	1	£50.00	2d 19h
		Reserve not met	Remaining time
Bids placed	No. of Bidders	Current price	closes 18/12/2018, 12:37 PM

Lot size:	1 bottle of 750 ml each	**Wine type:**	Red, 1985 vintage
Tax status:	Duty Paid	**Origin:**	Burgundy, France
Fill level:	Into Neck (IN)	**Grape variety:**	

An incredibly rare bottle of the sublime Nuits St George Les Boudots from Domaine Leroy from the exceptional 1985 vintage. In great order, this legend of a wine has lain in the same Berlin cellar or decades.

The last time this was on WineSearcher - 2016 - it was listed at £2,200, the reserve on this is a faction of that.

PayPal preferred but will charge 4% for fees.

Other details

Aux Boudots' thin soils consist of gravel, crumbly limestone marl and a small amount of clay. This fragmented soil, along with the natural slope of the vineyard, gives good drainage, making sure that vines do not receive excessive water. Instead, vines have to grow deep into the ground in search of hydration, a process which lessens vigor and reduces grape yields. This ultimately leads to the production of small, concentrated berries which make excellent wines.

Payment methods:	PayPal
Returns policy:	No returns
Shipping Method:	Courier delivery.
Shipping paid by:	buyer
Cost of delivery:	Will quote
Delivers to UK and Singapore	
Other countries delivered to:	Worldwide
Insurance options:	TBC

Figure 3.9: Listing page example: Nuits St George Les Boudouts, Domaine Leroy

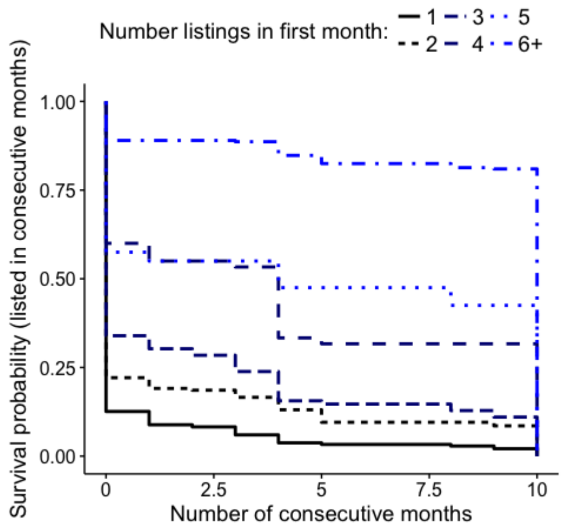


Figure 3.10: The survival probability of product listings

The survival probability is defined as the empirical probability that the same product is listed in consecutive months. For example, a survival probability of 0.2 at 5 months indicates that 20 percent of all products that were offered in the first month are also offered in months 2-6. The solid black line (bottom) pertains to the median product, of which only one is listed in a typical month, and the dash-dotted blue line (top) pertains to products in the 90th percentile of availability in a typical month. See Table 3.3 for deciles of the distribution of the number of comparable products per month. A key take-away from this plot is that about $\frac{7}{8}$th of typical listings doesn't survive one month, i.e. doesn't get listed in the next period.

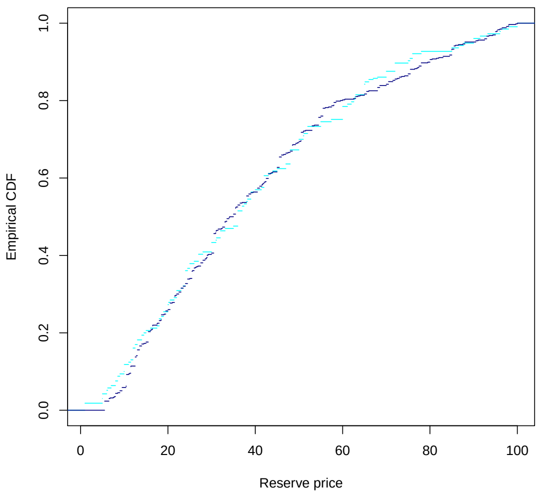


Figure 3.11: Testing equality of reserve price distribution and approximation

B Testing the reserve price approximation

I approximate the reserve price as the average between the highest standing price for which the reserve price is not met and the lowest for which it is met. If all bids would be recorded in real time, this approximation would be accurate up to half a bidding increment due to the proxy bidding system. But to relieve traffic pressure on the site I only track bids on 30-minute intervals. The reserve price approximation could be more than half a bidding increment off if the bids are not placed at regular intervals. As a compromise with constant high website traffic a separate dataset is collected that accesses open listings at 30-second intervals for the duration of two weeks, to test the reserve price approximation in the main sample.

My estimation method requires that the estimated distribution of reserve prices is consistent for its population counterpart. Equality of the distribution of approximated reserve prices in the main sample and the distribution of (approximated) reserve prices in the smaller high frequency sample is tested with a two sample nonparametric Kolmogorov-Smirnov test. To account for different listing compositions the empirical reserve price distributions are right-truncated at the 90th percentile of the high frequency reserve price sample. The null hypothesis is that the two right truncated reserve price distributions are the same. In particular, letting F_R^H and F_R^M respectively denote the empirical distribution of right truncated approximated reserve prices in the high frequency (H) and main (M) sample, the Kolmogorov-Smirnov test statistic is defined as:

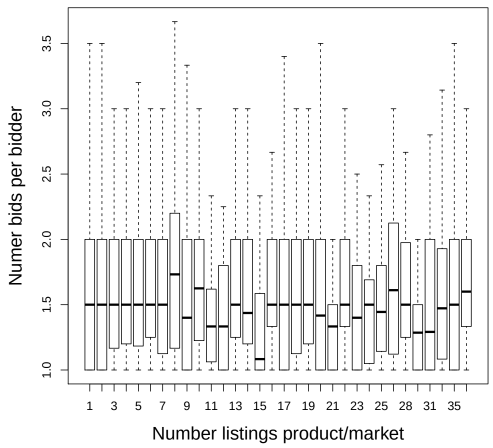
$$D_{h,m} = \sup_x |F_R^H(x) - F_R^M(x)|, \tag{3.29}$$

with $\sup_x$ the supremum function over x values and h and m respectively denoting the relevant number of observations in the high frequency and main samples, which are 330 and 596 (only for sold lots). With $D_{h,m} = 0.059$, the null cannot be rejected at the 5 pervent level ($D_{h,m} > 1.36\sqrt{(\frac{h+m}{hm}})$, the p-value = 0.4406). The two empirical distributions are plotted in Figure 3.11.

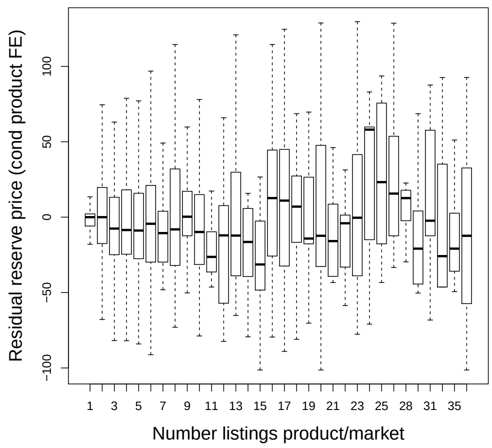
C Independent listings: additional results

This section reports descriptive statistics that point to listings being independent of each other conditional on entry and the matching of bidders to listings.

a) Number bids per bidder



b) Reserve price



c) Transaction price

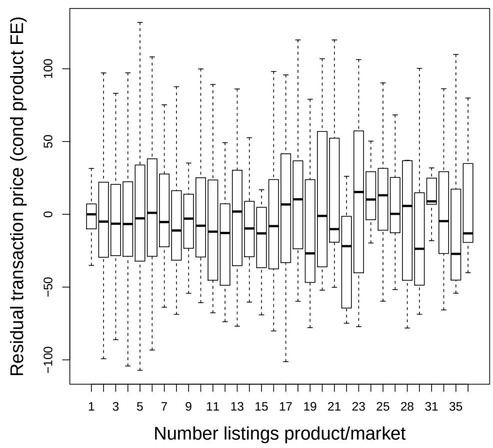


Figure 3.12: Empirical patterns suggesting that bidders do not cross-bid, thin out, and sellers do not compete

The box edges indicate the 25th to 75th percentile, the black line indicates the median, and whiskers indicate the 25th (75th) percentile minus (plus) 1.5 times the interquartile range or the sample extremes if those are less extreme. The residual reserve and transaction price in plot b) and c) are obtained from a linear regression of these outcomes Y_{pm} on product dummies α_p in: $Y_{pm} = \alpha_p + \epsilon_{pm}$. Plots display the relation between ϵ_{pm} and the number of listings of product p in market m, T_{pm}.

Table 3.11: Evidence stylized facts: independent listings

	Dependent variable:			
	Nr bidders listing	Bids per bidder	Transaction price	Reserve price
Competing listings: any wine				
- Ending within 30 days (of each other ...)	-0.001	0.00002	-0.0001	0.016
- Ending within 7 days	-0.001	-0.0002	0.008	-0.067
- Ending within 2 days	-0.003*	0.0003	0.062	-0.156
Competing listings: same wine type (e.g. red)				
- Ending within 30 days	-0.001	0.00004	0.012	0.074
- Ending within 7 days	-0.0003	-0.0002	0.084	-0.060
- Ending within 2 days	-0.003	0.001	0.167	-0.375
Competing listings: same wine region (e.g. Bordeaux)				
- Ending within 30 days	-0.004	0.001	0.155	0.265
- Ending within 7 days	-0.001	0.002	0.376	-0.336
- Ending within 2 days	-0.020**	0.006**	0.171	-0.525
Competing listings: region x wine type (eg. red Bordeaux)				
- Ending within 30 days	-0.003	0.001	0.228	-0.024
- Ending within 7 days	0.011	0.004	1.134**	-0.532
- Ending within 2 days	-0.019	0.007*	0.905	-1.994
Competing listings: region x wine type x vintage decade (e.g. red Bordeaux from 1980s)				
- Ending within 30 days	-0.012	0.001	-0.561	-0.938
- Ending within 7 days	-0.006	0.005	-0.465	-1.371
- Ending within 2 days	-0.061	0.004	-0.938	-0.669
Competing listings: subregion x wine type x vintage decade (e.g. red Margaux from 1980s)				
- Ending within 30 days	-0.009	0.001	0.433	-0.303
- Ending within 7 days	0.003	0.003	1.914**	-1.677
- Ending within 2 days	-0.034	0.007	0.775	-3.026
Product fixed effects:	Yes	Yes	Yes	Yes
Observations	1,150	2,898	2,230	2,337
Sample:	zero reserve	all	sold lots	sold lots
Adjusted R^2	0.147	0.016	0.102	0.056

Significance levels: *p<0.1; **p<0.05; ***p<0.01. All reported coefficients are from a separate regression of the dependent variable on the number of competing listings, each row of the table corresponding to a different definition, including product fixed effects. The product fixed effect relates to the product aspect of the competing listing specification (so types of wine, or regions, etc.). The adjusted R^2 is from the regression with (subregion x type x vintage decade fixed effects).

D Further details entry equilibrium

This appendix describes in more detail the entry equilibrium presented in the main text. In what follows, $\tilde{r}$ denotes the optimal reserve price increased with buyer premium, $\tilde{r} = (1 + c_B)r^*(v_0, f)$. Before knowing their valuation, the expected bidder surplus in a listing with n bidders equals:

$$\pi_b(n, f, r) \equiv \frac{1}{n}\mathbb{E}[V_{(n:n)} - \max(V_{(n-1:n)}, \tilde{r})|V_{(n:n)} \geq \tilde{r}][1 - F_{V_{(n:n)}}(\tilde{r})], \tag{3.30}$$

with the last term denoting the sale probability and the max(.) term the transaction price including buyer premium. Expected surplus for a seller with valuation v_0 in a listing with n bidders[68]:

$$\pi_s(n, f, v_0) \equiv \left(\mathbb{E}[\max(V_{n-1:n}, \tilde{r})|V_{n:n} \geq \tilde{r}](1 - c_S) - v_0\right)[1 - F_{V_{(n:n)}}(\tilde{r})] \tag{3.31}$$

The following properties are also useful for solving the entry stage; Appendix G provides proofs.

Lemma 3A. The listing-level impacts of number of bidders, commissions, and the seller's valuation are:

a) Expected listing-level bidder surplus is decreasing in n, c_B, c_S and v_0

b) Expected listing-level seller surplus is increasing in n and decreasing in c_B, c_S and v_0

Outline.

To characterize the entry equilibrium, I first consider equilibrium bidder entry as a function of a candidate seller entry threshold (Lemma 3). The equilibrium seller entry threshold is characterized as a fixed point in seller value space conditional on bidders' unique best response to it (Lemma 4). I end the section by approximating the entry equilibrium using results from Myerson (1998) on games with population uncertainty (Lemma 5).

[68]Two observations are useful to make here. The first is also made in Ginsburgh et al. (2010): in a model without entry, the allocation of (c_B, c_S) does not affect outcomes as long as a commission index $\frac{c_B+c_S}{1+c_B}$ remains constant. I mention this here because this setting is previously used to study the impact of c_B and c_S on welfare in Ashenfelter and Graddy (2003, 2005) and Marks (2009). I refer to this as a "one-sided market perspective" (on the impacts of auction fees), by the definition in Rochet and Tirole (2006) of a market in which only the level and not the allocation of fees matters. I provide a numerical example of this case in the Appendix on page 95. The platform market with entry decisions involving sunk entry cost constitutes a two-sided market (Rochet and Tirole (2006)). The second is that a seller maximizing $\pi_s(n, f, v_0)$ by his choice of reserve sets a reserve price that is too high from the platform's perspective. To see why, note that the seller trades off the expected transaction price with v_0 while the platform revenue only involves a share of the transaction price. This observation does not seem lost in practice: eBay charges a reserve price fee equal to 4 percent of the reserve price (unconditional on selling) and BW platform has flat reserve price fees.

This renders the entry game feasible to estimate while also relieving an assumption about players' common knowledge and being consistent with the data. The results in this section are summarized in the following proposition:

Proposition 1. The entry stage of the game results in a unique equilibrium for any fee structure characterized by a bidder entry probability and seller entry threshold that jointly solve: 1) potential bidders zero profit condition and 2) the marginal seller's zero profit condition. Given that any candidate seller entry threshold maps to an equilibrium bidder entry probability, which is strictly decreasing in the threshold, sellers are strategic substitutes and the entry game reduces to a single agent discrete choice problem. Assuming that the population of potential bidders is large, the equilibrium is described by the seller entry threshold and mean number of bidders per listing.

Bidder entry.

Consider first a world in which the number of listings T would be known to potential bidders. Let $\bar{v_0}$ denote a candidate seller entry threshold and $\Pi_b(T, f, \bar{v_0}; p)$ potential bidders' expected surplus from entering the platform as a function of their entry probability p, again if they knew the number of listings T:

$$\Pi_b^T(f, \bar{v_0}; p) = \sum_{n=0}^{N^B-1} \mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}] f_N^T(n; p) - e_B^o, \tag{3.32}$$

It takes the expectation of $\pi_b(n, f, v_0)$ (equation 3.30 with optimal r as in equation 3.2) over: i) possible seller values given sellers' entry threshold and ii) the number of competing bidders given their entry probability. T superscripts in $\Pi_b^T(f, \bar{v_0}; p)$ and $f_N^T(n; p)$ (equation 3.34) emphasize that they relate to the thought exercise in which T is known, while the true game's simultaneous entry requires taking the expectation over T given candidate entry threshold $\bar{v_0}$ and N^S. This is done to show more clearly that, in equilibrium, f_N^T is independent of the realization of T which implies that it must also be independent of the expectation over T. Bidding in one listing at a time, the entry problem for potential bidders is then equivalent to one in which they consider entry into a listing (also given that opportunity cost (listing inspection) e_B^o are associated with each listing). Components of equation (3.32) are then:

$$\mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}] = \int_0^{\bar{v_0}} \pi_b(n+1, f, v_0) f_{V_0|V_0 \leq \bar{v_0}}(v_0) dv_0 \tag{3.33}$$

$$f_N^T(n; p) = \binom{N^B-1}{n} (\frac{p}{T})^n (1 - \frac{p}{T})^{N^B-1-n} \tag{3.34}$$

where $f_N^T(n;p)$ denotes the Binomial probability that n out of N^B-1 competing potential bidders arrive in the same listing as the potential bidder who considers entering the platform. $\pi_b(n+1,f,v_0)$ is strictly decreasing in n due to the increasing failure rate assumption on $F_{V|\mathbf{Z}}$, which delivers a decreasing spacings property (Li (2005), discussed in more detail in the proof of Lemma 3A. Hence, the bidder entry problem is equivalent to the Levin and Smith (1994) entry equilibrium (which starts from expected bidder surplus decreasing in n). Given T, the equilibrium bidder entry probability p^{T*} solves zero profit condition:[69]

$$p^{T*}(T,f,\bar{v_0}) \equiv \arg_{p\in(0,1)} \Pi_b^T(f,\bar{v_0};p) = 0 \tag{3.35}$$

In this equilibrium the number of (competing) bidders per listing follows a Binomial distribution with mean $(N^B-1)\frac{p^{T*}}{T}$ and variance $(N^B-1)\frac{p^{T*}}{T}(1-\frac{p^{T*}}{T})$.[70]

A key property is that $\frac{p^{*T}}{T}$ is independent of T: bidders only derive positive surplus from the listing that they are matched to (e.g. T does not affect $\mathbb{E}[\pi_b(n+1,f,v_0)|V_0 \leq \bar{v_0}]$) so the zero profit condition guarantees that in equilibrium a change in T causes p^{T*} to adjust to keep f_N^T constant. The same reasoning applies when T is the stochastic outcome of the simultaneous seller entry process: the seller entry threshold only affects the equilibrium mean number of bidders per listing through $\mathbb{E}[\pi_b(n+1,f,v_0)|V_0 \leq \bar{v_0}]$ and not through its effect on the distribution of T. Lemma 3 states these points more formally and defines the bidder entry equilibrium in our simultaneous-move entry game. The following notation reflects that bidders form an expectation over T given candidate seller entry threshold $\bar{v_0}$:

$$\Pi_b(f,\bar{v_0};p) = \sum_{n=0}^{N^B-1} \mathbb{E}[\pi_b(n+1,f,v_0)|V_0 \leq \bar{v_0}] f_N(n;p,\bar{v_0}) - e_B^o \tag{3.36}$$

$$f_N(n;p,\bar{v_0}) = \sum_{t=0}^{N^S} \binom{N^B-1}{n} (\frac{p}{t})^n (1-\frac{p}{t})^{N^B-1-n} \binom{N^S}{t} F_{V_0|\mathbf{Z}}(\bar{v_0})^t (1-F_{V_0|\mathbf{Z}}(\bar{v_0}))^{N^S-t} \tag{3.37}$$

Lemma 3. Given candidate seller entry threshold $\bar{v_0}$ and fee structure f, the equilibrium

---

[69] p^{*T} is used to distinguish the entry probability from the central one pertaining to the central case where T is not known. A no-trade entry equilibrium at $p=0$ that trivially solves (3.5) always exists and excluding it (also in the remainder of this text) requires the profit-maximizing platform to set fees such that entry is profitable for players on both sides and for players not to believe that the other side enters with zero probability.

[70] For completeness this is derived in the Appendix. The variance of N would be larger when also taking the expectation over T given N^S and $\bar{v_0}$.

bidder entry probability solves the zero profit condition:

$$p^*(f, \bar{v_0}) \equiv \arg_{p\in(0,1)} \Pi_b(f, \bar{v_0}; p) = 0 \tag{3.38}$$

Equilibrium properties are:
i) $p^*(f, \bar{v_0})$ is unique $\forall (f, \bar{v_0})$
ii) $p^*(f, \bar{v_0})$ is strictly decreasing in $(\bar{v_0}, c_B, c_S, e_B)$ so also $f_N(n; p^*, \bar{v_0})$ decreases in the first-order stochastic dominance sense in $(\bar{v_0}, c_B, c_S, e_B)$
iii) $f_N(n; p^*, \bar{v_0})$ is invariant to changes in N^B or N^S

Proof. The zero profit condition follows from bidders being indifferent in equilibrium between staying out and entering the platform. The entry probability is decreasing in $\bar{v_0}$ because $\pi_b(n+1, f, v_0)$ is decreasing in v_0 (Lemma 3A) so that the expectation over v_0, $\mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}]$ is decreasing in $\bar{v_0}$. $\pi_b(n+1, f, v_0)$ is furthermore decreasing in (c_B, c_S) (Lemma 3A), so the zero profit condition dictates that higher commissions must result in a lower entry probability, and the same holds for a higher sunk entry cost e_B. Seller entry cost e_S do not affect expected bidder surplus so conditional on $\bar{v_0}$ they do not affect $p^*(f, \bar{v_0})$. Given that N^B and N^S do not affect $\mathbb{E}[\pi_b(n+1, f, v_0)|V_0 \leq \bar{v_0}]$ the zero profit condition dictates that $p^*(f, \bar{v_0})$ is such the equilibrium distribution $f_N(n; p^*, \bar{v_0})$ is invariant to N^B and N^S. □

Seller entry.
On the seller side, potential sellers' expected surplus from entering decreases in their private valuation, which makes it optimal for them to adopt a pure strategy to enter if and only if their valuation is less than threshold $\bar{v_0}$ that in equilibrium makes the marginal seller indifferent. In what follows I denote the strategy by the threshold itself. Potential sellers' expected surplus from entering the platform involves: i) their listing-level expected surplus and ii) an expectation over N given $\bar{v_0}$ and bidders' equilibrium best-response to this threshold:[71]

$$\Pi_s(f, v_0; p^*(f, \bar{v_0}), \bar{v_0}) = \sum_{n=0}^{N^B} \pi_s(n, f, v_0) f_N(n; p^*(f, \bar{v_0}), \bar{v_0}) - e_S - e_S^o \tag{3.39}$$

[71] A slight abuse of notation is that expectation involves $f_N(n; p^*, \bar{v_0})$ (explaining entry among $N^B - 1$ potential bidders) as defined in (3.37) instead of the distribution based on the full bidder population N^B. This avoids introduction of additional notation and the -1 will be irrelevant in the large-N^B approximation in which case the two distributions are identical by the *environmental equivalence* property of the Poisson distribution (Myerson (1998)).

Lemma 4. Given fee structure f, the equilibrium seller entry threshold solves the marginal sellers' zero profit condition:

$$v_0^*(f) \equiv \arg_{\bar{v_0} s.t. F_{V_0|\mathbf{Z}}(\bar{v_0}) \in (0,1)} \Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0})) = 0 \text{ , with } p^*(f, \bar{v_0}) \text{ solving (3.38)} \quad (3.40)$$

Equilibrium properties are:
i) $v_0^*(f)$ is unique $\forall f$
ii) $v_0^*(f)$ is strictly decreasing in e_B
iii) The impact of (c_B, c_S, e_S) is ambiguous

Proof. Since $\pi_s(n, f, v_0)$ is strictly decreasing in v_0 (Lemma 3), for any v_0 for which a potential seller finds it profitable to enter he would also enter with a lower value. Hence their pure strategy entry decision is characterized by a threshold value that makes the marginal seller indifferent when competing sellers adopt the same entry threshold. $\Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0}))$ denotes the expected seller surplus for the marginal seller with valuation equal to candidate threshold $\bar{v_0}$ when competing sellers follow that threshold strategy. An equilibrium exists because sellers reaction function is continuous in their own value and competing sellers threshold. Sellers have a unique best response for any competing seller entry threshold given that $\frac{\partial \Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0}))}{\partial v_0} < 0$. Given that 1) $p^*(f, \bar{v_0})$ is strictly decreasing in $\bar{v_0}$ (Lemma 4), and 2) entry of competing sellers does not affect $\Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0}))$ in other ways, the best response function is strictly decreasing in competing sellers $\bar{v_0}$: $\frac{\partial \Pi_s(f, \bar{v_0}; p^*(f, \bar{v_0}))}{\partial \bar{v_0}} < 0$. Symmetry then delivers a unique $v_0^*(f)$ that is the fixed point in seller value space that solves equation 3.40.
The equilibrium threshold must be decreasing in e_B because it reduces $p^*(f, \bar{v_0})$ for any candidate threshold (Lemma 4) and does not affect expected seller surplus otherwise. $\frac{\partial v_0^*(f)}{\partial c_B} < 0$ unless 1) the impact of a lower entry threshold on $p^*(f, \bar{v_0})$ outweighs the direct negative impact of c_B on p^* and 2) the resulting higher p^* outweighs the negative direct impact of higher c_B on $\pi_s(n, f, v_0)$. The same goes for c_S. A slightly modified argument holds for seller entry fees that do not directly impact $p^*(f, \bar{v_0})$: $\frac{\partial v_0^*(f)}{\partial e_S} < 0$ unless the positive impact of a lower entry threshold on $p^*(f, \bar{v_0})$ (increasing expected seller surplus) outweighs the direct negative impact of a higher listing fee. □

Corollary 1. The entry equilibrium of the auction platform game is characterized by the pair of: $(v_0^*(f), p^*(f, v_0^*(f)))$ that jointly solves equation (3.40), which is unique for any fee structure.

Large population approximation.
The remainder of this section discusses an approximation of the entry equilibrium that is adopted for the following reasons. First of all, it is computationally costly to implement the equilibrium literally, especially due to the expectation over realizations of T in (3.37). A second reason is that the approximation relaxes the requirement that players know population sizes N^B and N^S, which indeed are likely to be unobserved by potential bidders and sellers. Relatedly, given that the population of potential bidders is likely to be large relative to the actual number of bidders, the Poisson approximation of the binomial distribution is a natural one, also adopted previously in similar settings by e.g. Engelbrecht-Wiggans (2001), Bajari and Hortaçsu (2003) and Jehiel and Lamy (2015). Finally and not unimportant, the empirical distribution of bidders on the BW platform resembles a Poisson distribution (Figure 3.4). Formally, a chi-square goodness of fit test fails to reject at the five percent level that N is generated by a Poisson distribution, based on auctions with no reserve price in which the number of bidders is not censored.

Assumption 5. The population of potential bidders is large relative to the number of bidders on the platform: $N^B \to \infty$ and $p^* \to 0$.[72]

Under this assumption the number of bidders per listing in (3.37) is approximately Poisson distributed, and approximation error relative to the Binomial distribution is small (vanishing to 0 as N^B tends to infinity). The Poisson mean is endogenous to the fee structure. It is implicitly defined by solving a bidder zero profit entry condition in this slightly modified setting with the Poisson distributed number of bidders per listing. The result is formalized in the following Lemma.

Lemma 5.
With large N^B, small p^* and T listings the number of bidders per listing has a probability mass function approximated by:

$$f_N(k;\lambda) = \frac{\exp(-\lambda)\lambda^k}{k!}, \lambda = \frac{N^B p}{T}, \forall k \in \mathbb{Z}^+ \tag{3.41}$$

[72]To avoid any misinterpretation (with p^* endogenous), the population is assumed to be large and the entry probability is assumed to be small and it is not a statement about letting the population grow large or the entry probability go to 0.

The approximate equilibrium is the pair of $(v_0^*(f), \lambda^*(f, v_0^*(f)))$ that solves:

$$\begin{aligned} v_0^*(f) &\equiv \arg_{\bar{v_0} s.t. F_{V_0|\mathbf{Z}}(\bar{v_0}) \in (0,1)} \Pi_s(f, \bar{v_0}; \lambda^*(f, \bar{v_0})) = 0 \text{ , subject to :} \\ \lambda^*(f, \bar{v_0}) &\equiv \arg_{\lambda \in (0,\infty)} \Pi_b(f, \bar{v_0}; \lambda) = 0, \end{aligned} \tag{3.42}$$

where $\Pi_s(f, \bar{v_0}; \lambda^*(f, \bar{v_0}))$ and $\Pi_b(f, \bar{v_0}; \lambda)$ refer to players' expected surplus when the number of potential bidders follows PMF in equation (3.41). Poisson mean λ^* is endogenous to the fee structure and seller entry threshold but independent of N^S, N^B and T.

Proof. The limiting distribution of the $Binom(N^B, p)$ is Poisson with mean $N^B p$, and with uniform sorting of bidders over T listings the number of bidders per listing is Poisson with mean $\frac{N^B p}{T}$ by the *decomposition property* of the Poisson distribution as set out in Myerson (1998). Appendix G shows this formally, and also provides an alternative argument based on the same property of the Binomial distribution. Given that the number of listings T does not enter expected bidder surplus, it does not affect the equilibrium distribution of N by the same zero profit condition argument that supported that $f_N(n; p^*, \bar{v_0})$ is invariant to changes in N^B or N^S in Lemma 4. Hence all three do not impact λ^* while the fee structure and seller entry threshold do affect expected bidder surplus and therefore λ^*. □

Corollary 2.
The entry equilibrium of the auction platform game, subject to the large population approximation, is characterized by the pair of: $(v_0^*(f), \lambda^*(f, v_0^*(f)))$ that jointly solve equation (3.42), which is unique for any fee structure.

E Monte Carlo simulations: a recursive algorithm

This Section discusses the Monte Carlo simulations and provides details about numerical approximation of the entry equilibrium. Simulated auctions are structured according to the idiosyncratic-good auction platform model presented before with:

Input parameters:	**Equilibrium values:**
$g(\mathbf{Z}) = 0.5Z, Z \sim \mathcal{N}(0, 2)$	$v_0^* = 5.999, F_{V\vert\mathbf{Z}}(v_0^*) = 0.8847$
$(U_0, U) \sim \mathcal{N}(5, 0.5)$	$\lambda_{r>0}^* = 4.499, \lambda_{r=0}^* = 6.112$
$e_B^o = 10, e_S^o = 5, f = \{e_S = 5, e_B = 0, c_B = 0, c_S = 0.1\}$	
$p_{r0} = 0.10, N^S = 4000$	

Hence with the chosen valuation distributions, opportunity cost and platform fees, the seller entry threshold equals 5.599 so that in equilibrium 88.47 percent of the 4000 potential sellers enter the platform. The marginal seller, or any seller who sets a positive reserve price receives on average 4.499 bidders in his listing. Furthermore, the mean number of bidders in zero reserve auctions is calculated to be 6.112 and 10 percent of listings have no reserve price (exogenously determined here).

A recursive algorithm.
Recall the multi-step estimation method outlined on page 50. I will now provide details on steps 4 and 5:
4) solving for the entry equilibrium given estimated parameters
5) re-estimating seller parameters at the updated entry equilibrium.
When iterating on these estimation steps until convergence, they describe the iterative nested pseudo likelihood (NPL) in Aguirregabiria and Mira (2002, 2007). Roberts and Sweeting (2010) are the first to apply this algorithm to the auction literature to study auctions with selective bidder entry. Studies by Pesendorfer and Schmidt-Dengler (2010), Kasahara and Shimotsu (2012) and Egesdal et al. (2015) provide conditions under which NPL does (not) converge to the true equilibrium. A best-response stable equilibrium is a sufficient condition for the NPL algorithm to converge to it and this is certainly guaranteed (Proposition 1) by the game reducing to a single agent discrete choice problem with unique equilibrium. Aguirregabiria and Mira (2002) find that in single-agent games asymptotic efficiency is independent of the number of iterations.

Monte Carlo (MC) simulations evaluate the proposed estimation method in Steps (1)-(3) and the updating according to steps (4)-(5), either once or iterating until the maximum difference between updated and previous *seller parameters* is less than $1e^{-3}$. This choice of convergence objective is motivated by the centrality of θ_s as structural parameters in contrast to the seller entry probability or threshold. Furthermore, calculation of υ_0^* depends on opportunity cost to the marginal seller e_S^o, which, following its identification argument, is a function of both seller parameters *and* the entry threshold. If the initial estimate υ_T is an overestimate of the truth that will be reflected in low opportunity cost that reinforce a high υ_0^{*j} in the recursion. To address this, an exponentially vanishing trimming parameter τ^j is subtracted from the seller entry probability obtained in iteration j to enforce that $\upsilon_T > \upsilon_0^*$. Using $\tau^j = 0.05\exp(-j)$ the trimming parameter at $j = 1$ equals 0.05 and vanishes quickly to virtually zero in than 5 iterations.

Four sets of 500 MC experiments are conducted. The first implements the algorithm with the true value of e_S^o, the second imposes a fixed but wrong value of twice e_S^o, and the third recursively solves for e_S^{oj} given $\hat{\theta}_s^{j-1}$. Columns 1-3 of Table 3.12 show that all three ways to deal with e_S^o (the true value, a wrong value, solving recursively) deliver virtually identical estimates of θ_s. This confirms that equating seller opportunity cost to some fixed value (usually 0, for instance in Seim (2006) for a discrete choice entry game with firm heterogeneity) is truly a normalisation for the estimation of seller parameters in our setting. Another take-away is that when the first stage is estimated precisely and the resulting initial estimate υ_T is close as well, the recursive method delivers no benefits only computational cost. Intuitively, this is because seller parameters are obtained using the first order condition of optimal reserve prices that is independent of the threshold for all observations not censored by it.

Column 4 of Table 3.12 shows the benefit of updating at least once in the presence of small sample estimation bias. In this MC experiment Z is measured with noise, implemented by adding draws from $Unif(-1, 1)$ to it after simulating values and bids. The initial estimate of the seller entry probability now overestimates the truth (0.978 compared to 0.885). Especially shape parameter k_s is affected. One iteration effectively addresses this problem, and further iterations remain consistent but do not deliver benefits or improve precision.

Based on MC simulation results, estimation in the remainder of this paper is done with one update of θ_s based on solving the equilibrium and normalising $e_S^o = e_B^o$. The gray colored rows in Table 3.12 correspond to the single recursion solution.

Table 3.12: Monte Carlo simulations

	Iteration j	Truth	Given true e_S^o: Mean	SD	Med.AD	Given wrong e_S^o (double) Mean	SD	Med.AD	Estimating e_S^o Mean	SD	Med.AD	Estim. e_S^o noisy first stage Mean	SD	Med.AD
Bidder side														
μ_b		5	5	0.021	0.015	5	0.02	0.015	5	0.021	0.015	4.992	0.022	0.016
k_b		0.5	0.499	0.02	0.013	0.499	0.02	0.013	0.499	0.02	0.013	0.513	0.02	0.016
e_B^o		10	9.942	0.689	0.481	9.943	0.681	0.476	9.939	0.687	0.471	10.364	0.661	0.524
$\lambda_{r>0}^*$		4.499	4.507	0.159	0.108	4.506	0.158	0.105	4.507	0.159	0.108	4.494	0.162	0.116
$\lambda_{r=0}^*$		6.112	6.105	0.137	0.093	6.105	0.137	0.093	6.105	0.137	0.093	6.105	0.138	0.092
Seller side, include solving entry game (if $j > 0$)														
	0	5	4.998	0.014	0.01	4.998	0.014	0.01	4.998	0.014	0.01	4.987	0.009	0.013
μ_s	1	5	4.994	0.027	0.013	4.998	0.034	0.012	5	0.017	0.011	4.996	0.017	0.012
	2	5	4.994	0.027	0.013	4.998	0.032	0.011	5	0.018	0.012	5.003	0.016	0.012
	3	5	4.994	0.029	0.013	4.998	0.034	0.011	5	0.019	0.013	5.002	0.018	0.013
	J	5	4.994	0.029	0.013	4.998	0.034	0.011	5	0.019	0.013	5.002	0.018	0.013
	0	0.5	0.486	0.023	0.019	0.486	0.023	0.019	0.486	0.023	0.019	0.373	0.009	0.127
k_s	1	0.5	0.485	0.033	0.023	0.496	0.034	0.017	0.498	0.027	0.017	0.486	0.026	0.021
	2	0.5	0.481	0.035	0.02	0.493	0.032	0.016	0.493	0.026	0.017	0.477	0.023	0.026
	3	0.5	0.481	0.039	0.023	0.496	0.035	0.018	0.496	0.03	0.019	0.486	0.023	0.018
	J	0.5	0.48	0.038	0.021	0.493	0.031	0.017	0.493	0.029	0.019	0.481	0.023	0.021
e_S^o	1	5	5	0	0	10	0	5	5.174	1.935	1.244	4.104	1.879	1.547
	2	5	5	0	0	10	0	5	7.368	2.038	2.395	6.252	2.056	1.603
	3	5	5	0	0	10	0	5	8.181	2.084	3.147	6.919	2.095	2.006
	J	5	5	0	0	10	0	5	8.635	2.094	3.616	7.423	2.131	2.304
	0	0.885	0.894	0.015	0.012	0.894	0.014	0.012	0.894	0.015	0.012	0.978	0.006	0.094
$F_{V_0\|\mathbf{z}}(v_0^*)$	1	0.885	0.878	0.064	0.015	0.842	0.046	0.039	0.844	0.015	0.04	0.927	0.006	0.043
	2	0.885	0.88	0.074	0.02	0.837	0.029	0.045	0.819	0.017	0.065	0.845	0.013	0.04
	3	0.885	0.881	0.077	0.02	0.838	0.031	0.044	0.814	0.018	0.07	0.839	0.015	0.045
	J	0.885	0.881	0.077	0.021	0.837	0.028	0.045	0.81	0.021	0.075	0.832	0.016	0.053
Number iterations (J)			6.862	2.989		7.442	3.003		8.266	2.189		8.41	2.055	

Gray rows correspond to the algorithm that solves for the entry equilibrium once given initial parameter estimates, as per the adopted estimation method.

F Numerical approximation of the entry equilibrium

Computation of the surplus functions is based on Monte Carlo integration with importance sampling. For details on numerical methods see Judd (1998); I will summarize only their application. The following is implemented on homogenized values and with estimated parameters but for brevity I omit the conditioning on $\mathbf{Z}$ and $(\hat{\theta}_b, \hat{\theta}_s)$. A grid of 150 points v_0^m is drawn from F_{V_0}. For all values on that grid, and for $n = 0, 1, .., 15$ the value of $\pi_s(n, f, v_0^m)$ in (3.31) is calculated for two scenario's: when setting $r = r^*(v_0^m)$ and when setting $r = 0$. Also the value of $\pi_b(n+1, f, v_0^m)$ is calculated for $n = 0, 1, ..., 15$ and for both scenario's, which involves the expectation over values of $V_{n:n}$ and also an inner integral over realizations of $V_{n-1:n}$ conditional on $V_{n:n} = v_n$:

$$\pi_b(n, f, r) = \int_{\tilde{r}}^{\bar{v}} v_n - \max((1 + c_B)r, \int_{\underline{v}}^{v_n} v_{n-1} dF_{V_{n-1:n}|V_{n:n}=v_n}(v_{n-1})) dF_{V_{n:n}}(v_n) \tag{3.43}$$

$$F_{V_{n:n}}(v_n) = \int_{\underline{v}}^{v_n} nF_V(x)^{n-1} f_V(x) dx$$

$$F_{V_{n-1:n}|V_{n:n}=v_n}(v_{n-1}) = \int_{\underline{v}}^{v_n} \frac{(n-1)F_V(y)^{n-2} f_V(y)}{F_V(v_n)^{n-1}} dy$$

Samples from both distributions are obtained with globally adaptive quadrature (on support $[0, \bar{v}]$ with $\bar{v}$ the 1-1.0000e^{-9}th percentile of F_V) setting the inverse of the distribution equal to 250 equally spaced probabilities on $[0, 1]$, and then evaluated on 2500 points with linear interpolation, separately for each n. Then for each $r \in \{r^*(v_0^m)\}_{m=1}^{250}, 0\}$, $\pi_b(n + 1, f, r)$ is calculated by drawing 250 values from the simulated highest and second-highest valuation samples conditional on $v_n \geq \tilde{r}$ and taking averages. For the positive reserve price case, $\pi_s(n, f, v_0^m)$ and $\pi_b(n + 1, f, v_0^m)$ are linearly interpolated on finer grids with 500 points for each n. These four matrices are pre-calculated once for each set of parameter draws so essentially for each iteration (and fee combination, in the counterfactual experiment). Let $\mathbb{S}^{r=0}$ and $\mathbb{B}^{r=0}$ of dimension [16x1] denote the listing-level surplus matrices for sellers and bidders when the seller sets no reserve price and $\mathbb{S}^{r>0}$ and $\mathbb{B}^{r>0}$ the matrices of dimension [16x150] in the positive reserve price case.

The entry equilibrium solves equation (3.42), where $\Pi_b(f, \bar{v_0}, \lambda)$ includes: $\mathbb{E}[\pi_b(n_1, f, v_0)|V_0 \leq \bar{v_0}]$. In the optimal positive reserve price case it is approximated, for each n, by the elementwise product of the n+1th row of $\mathbb{B}^{r>0}$ with vector $[f_{V_0}(v_0^1), ..., f_{V_0}(v_0^m)]$ such that $f_{V_0}(v_0^m) = 0$ for $v_0^m > \bar{v_0}$, using local linear interpolation on the margin and with the resulting non-zero elements of the row summing to 1.[73] Also subtracting bidder entry cost delivers the expected bidder surplus for each number of n competing bidders, and for each $\bar{v_0}$ the equilibrium Poisson parameter $\lambda^*_{r>0}$ is approximated as the value that sets $\Pi_b(f, \bar{v_0}, \lambda) = 0$ to within a $1e^{-6}$ tolerance of either the function value or the parameter.

The calculation of $\lambda^*_{r>0}$ is nested in the seller entry equilibrium. Every candidate $\bar{v_0}$ maps to a $\lambda^*_{r>0}$. For the two columns in $\mathbb{S}^{r>0}$ corresponding to the values nearest to $\bar{v_0}$ the expected surplus is calculated by placing weight on row k corresponding to $f_N(k-1; \lambda^*_{r>0}(\bar{v_0}))$, linear interpolation to get expected surplus at $\bar{v_0}$ for each $n = k-1$, and summing over all k. The equilibrium v_0^* is approximated within a $1e^{-6}$ tolerance level.

[73] Section 3.5.4 explains that, unless all sellers find it optimal to set a zero reserve price, the marginal seller will set a positive reserve price and there is also a "screening value" of V_0 below which sellers will set no reserve price. Given that in the data only about one third of sellers sets a reserve, the entry problem is about the marginal seller who will set a positive reserve price and expects corresponding bidder entry. To economize on notation I disregard the screening value here but it is incorporated in the computation of $\lambda^*_{r>0}$ and hence in v_0^*.

G Omitted proofs

Optimal reserve price.

Proof Lemma 2. For brevity I omit conditioning on characteristics $\mathbf{Z}$, and define hat and check notation as: $\hat{x} = x(1+c_B)$ and $\check{x} = \frac{x}{1+c_B}$. Let R denote expected revenue for a seller with valuation v_0 when setting reserve price r in an auction with n bidders:

$$\begin{aligned} R = v_0 F_V(\hat{r})^n + (1-c_S) r n F_V(\hat{r})^{n-1}[1-F_V(\hat{r})] + \\ (1-c_S)\int_{\hat{r}}^{\bar{v}} \check{x} n(n-1) F_V(x)^{n-2}[1-F_V(x)] f_V(x) dx \end{aligned}$$

The three terms in the above equation for R cover three cases: i) no sale takes place, ii) a sale takes place but the second-highest bid is less than the reserve price and iii) the sale takes place and the second-highest bid exceeds the reserve. Maximizing R with respect to r:

$$\begin{aligned} \frac{\partial R}{\partial r} &= v_0 n F_V(\hat{r})^{n-1} f_V(\hat{r})(1+c_B) + (1-c_S) n F_V(\hat{r})^{n-1}[1-F_V(\hat{r})] \\ &+ (1-c_S) r n(n-1) F_V(\hat{r})^{n-2} f_V(\hat{r})(1+c_B)[1-F_V(\hat{r})] \\ &- (1-c_S) r n F_V(\hat{r})^{n-1} f_V(\hat{r})(1+c_B) \\ &- (1-c_S) r n(n-1) F_V(\hat{r})^{n-2}[1-F_V(\hat{r})] f_V(\hat{r})(1+c_B) \\ & r\{(1-c_S) n F_V(\hat{r})^{n-1} f_V(\hat{r})(1+c_B)\} = v_0 n F_V(\hat{r})^{n-1} f_V(\hat{r})(1+c_B) \\ &+ (1-c_S) n F_V(\hat{r})^{n-1}[1-F_V(\hat{r})] \end{aligned}$$

Re-arranging, this delivers the optimal reserve price $r^*(v_0, f)$ which solves:

$$r^*(v_0, f) \equiv r = \frac{v_0}{1-c_S} + \frac{1-F_V(r(1+c_B))}{(1+c_B) f_V(r(1+c_B))}$$

The optimal reserve price has the familiar Riley and Samuelson (1981) reserve price form. A unique $r^*(v_0, f)$ solves this equation $\forall (v_0, f)$ given that F_V is assumed to satisfy the increasing failure rate (IFR) property. It has the following properties:

$$\begin{aligned} i)\frac{\partial r^*(v_0, f)}{\partial v_0} > 0 \\ ii)\frac{\partial r^*(v_0, f)}{\partial c_B} < 0 \\ iii)\frac{\partial r^*(v_0, f)}{\partial c_S} > 0 \end{aligned}$$

Property ii also relies on the IFR assumption and requires (r, c_B) to be small enough s.t. $F_V(r(1+c_B))) < 1$.[74] $\square$

Non-neutrality of fees.

The fee structure is non-neutral if not just their total amount but also their allocation between different sides of the platform affect platform profitability (Rochet and Tirole (2006)). Non-neutrality is a prerequisite for the central questions in this paper to be of relevance. The presence of transaction cost, in the case of the auction platform stemming from costly listing inspection on the bidder side, entry cost, listing fees and opportunity cost for sellers, is known to generate non-neutrality (Rochet and Tirole (2006)). When taking entry as given, the fee structure is neutral in the auction stage when the non-linear commission index $\frac{c_B+c_S}{1+c_B}$ remains constant. I illustrate this point with a simple numerical example with optimal reserve prices.

Example 1. Bidder valuations are distributed according to $U[0,1]$, $n = 2$. Sellers have a valuation of v_0 and set the optimal reserve price: $r = \frac{v_0}{2(1-c_S)} + \frac{1}{2(1+c_B)}$ (solving the optimal reserve price formula with uniformly distributed valuations). The sale probability equals $1 - F(r(1+c_B))^2 = 1 - [\frac{v_0(1+c_B)}{2(1-c_S)} + \frac{1}{2}]^2$. Let X denote: $\max(\mathbb{E}[V_{n-1:n}], \frac{r}{1+c_B})$. Because bidders scale their bid down by c_B and sellers pay a commission on the transaction price the expected seller profit equals: $\pi_s = (X\frac{1-c_S}{1+c_B} - v_0)\{1 - [\frac{v_0(1+c_B)}{2(1-c_S)} + \frac{1}{2}]^2\}$. It is independent of the allocation of commissions to buyers and sellers as long as the fraction $\frac{1-c_S}{1+c_B}$ remains constant. For example, their expected surplus is the same when respectively i) $c_S = 0.1, c_B = 0$ and ii) $c_S = 0.01, c_B = 0.1$, both cases result in $\frac{1-c_S}{1+c_B} = 0.9$:

$$i)(0.9X - v_0)\{1 - [\frac{v_0}{2*0.9} - \frac{1}{2}]^2\}$$
$$ii)(\frac{0.99}{1.1}X - v_0)\{1 - [\frac{v_0 * 1.1}{2*0.99} - \frac{1}{2}]^2\}$$

In particular, the optimal reserve price adjusts to keep the sale probability constant and therefore also the expected transaction price will be the same. However, neutrality does not refer to the seller's profit but to that of the platform who designs the fee structure. Expected platform revenue is a slightly modified function accounting for the fact they retain the share $c_S + c_B$ of a transaction. The below shows that the platform fee structure is neutral when holding constant the fraction $\frac{1-c_S}{1+c_B}$ (coming from the seller maximization problem, or $\frac{c_B+c_S}{1+c_B}$ as in Ginsburgh et al. (2010)). Expected platform revenue is the same

[74] My calculations differ slightly from Ginsburgh et al. (2010) who I believe omit that a successful sale requires the highest value to exceed the reserve price *increased by buyer premium*, $r(1+c_B)$.

for fee structures i and ii:

$$i)0.1X\{1-[\frac{v_0}{2*0.9}-\frac{1}{2}]^2\}$$
$$ii)\frac{0.11}{1.1}X\{1-[\frac{v_0*1.1}{2*0.99}-\frac{1}{2}]^2\}$$

On take-away from this example is that in the auction stage sellers should be slightly better off with a say 10 percent seller commission and no buyer premium than with a 10 percent buyer premium and no seller commission. They are equally well off with a 10 percent buyer premium and a 1 percent seller commission.

Proof Lemma 3. **Expected bidder surplus increases in number bidders.**

$$i)\frac{\partial \pi_b(n,f,v_0)}{\partial n}\leq 0$$

It holds strictly in the uninteresting case when v_0 s.t. $F_V(r^*(v_0,f)(1+c_B))=1$. Otherwise, expected bidder surplus decreases in the number of bidders in a listing because F_V satisfies the increasing failure rate (IFR) property. Li (2005) prove that IFR implies decreasing spacings so that $\mathbb{E}[V_{(n+1:n+1)}-V_{(n:n+1)}]-\mathbb{E}[V_{(n:n)}-V_{(n-1:n)}]\leq 0$. This holds without a reserve price or fees and since both are independent of n this proves the statement.

$$ii)\frac{\partial \pi_b(n,f,v_0)}{\partial c_B}\leq 0$$

While expected bidder surplus conditional on a sale is independent of c_B, the probability of selling is decreasing in c_B. The optimal reserve price decreases in c_B, but not enough to keep the sale probability constant given that sellers trade this off against the sale price. Formally, denote $r=r^*(v_0,f)$ and using hat notation $\hat{r}=r^*(v_0,f)(1+c_B)$:

The sale probability and expected bidder surplus decreases in c_B and c_S.

$$\frac{\partial(1-F_{V_{(n:n)}}(\hat{r}))}{\partial c_B}=-f_{V_{(n:n)}}(\hat{r})(1+c_B)$$

which is negative because $f_{V_{(n:n)}}$ is a PDF and therefore $\in[0,1]$.The same result follows from the derivative of the sale probability with respect to c_S.As c_S does not affect bidder surplus in other ways it also follows that:

$$iii)\frac{\partial \pi_b(n,f,v_0)}{\partial c_S}\leq 0$$

Expected bidder surplus decreases in the seller valuation.

$$iv)\frac{\partial \pi_b(n, f, v_0)}{\partial v_0} \leq 0$$

This intuitively follows from the reserve price strictly increasing in v_0 and expected bidder surplus decreasing in the reserve price. The latter is necessary for sellers to have a unique optimal r^* (less than $\bar{v_0}$), so it relies on the IFR assumption. Formally, let $\hat{r} = r^*(v_0, f)(1+c_B)$ and $F_{n-1}(x|v_n) = P[V_{n-1:n} \leq x|V_{n:n} = v_n]$ with similar notation for conditional densities, and $F_n = F_{V_{n:n}}$. For $n > 0$:

$$\begin{aligned}
\pi_b(n, f, v_0) &= \mathbb{E}[V_{n:n} - \max(V_{n-1:n}|V_{n:n} \geq \tilde{r}, \tilde{r})|V_{n:n} \geq r][1 - F_n(\tilde{r})] \\
\frac{\partial \pi_b(n, f, v_0)}{\partial v_0} &= \int_{\tilde{r}}^{\bar{v}} [v_n - \underbrace{F_{n-1}(\tilde{r}|v_n)\tilde{r} - \int_{\tilde{r}}^{v_n} x f_{n-1}(x|v_n)]}_{\equiv H(v_n, \tilde{r})} f_n(v_n) dv_n \\
&= -\tilde{r} f_n(\tilde{r})(1 + c_B) - H(\tilde{r}, \tilde{r}) f_n(\tilde{r})(1 + c_B) + \int_{\tilde{r}}^{\infty} \frac{\partial H(v_n, \tilde{r}) f_n(v_n)}{\partial r} dv_n \\
&= -\hat{r} f_n(\tilde{r})(1 + c_B) + F_{n-1}(\hat{r}|v_n = \hat{r})\tilde{r} f_n(\hat{r})(1 + c_B) + \\
&\int_{\hat{r}}^{\bar{v}} f_n(v_n)[-(1 + c_B)F_{n-1}(\hat{r}|v_n) - f_{n-1}(\hat{r}|v_n)(1 + c_B)\hat{r} \\
&+ f_{n-1}(\hat{r}|v_n)\hat{r}(1 + c_B)] dv_n \\
&= \int_{\hat{r}}^{\bar{v}} -(1 + c_B)F_{n-1}(\hat{r}|v_n) f_n(v_n) dv_n
\end{aligned}$$

The third line follows from applying the Leibnitz rule. The last line follows from the fact that in the fourth line $F_{n-1}(\tilde{r}|v_n = \tilde{r}) = 1$ so that the two remaining terms on the first line cancel out and also the last two terms in square brackets on the fifth line cancelling. Given that f_n and $F_{n-1}(|v_n)$ are both always ≥ 0, $\frac{\partial \pi_b(n,f,v_0)}{\partial r} \leq 0$. The derivative equals 0 if: i) $\hat{r} \geq \bar{v}$, or ii) $n = 0$ which results in $F_{n-1}(.|v_n) = 0$ and otherwise the derivative is strictly negative.

Expected seller surplus increases in the number of bidders.

$$v)\frac{\partial \pi_s(n, f, v_0)}{\partial n} \geq 0$$

$\mathbb{E}[V_{(n-1:n)}]$ increases in n without restrictions on F_V (unlike for the bidder side in i) and the sale probability increases in n as $r^*(v_0, f)$ is independent of n and $F_{V_{(n:n)}}$ is stochastically increasing in n. No other aspects of $\partial \pi_s(n, f, v_0)$ relate to n so this delivers the result.

Expected seller surplus decreases in commissions.

$$vi)\frac{\partial \pi_s(n,f,v_0)}{\partial c_B} \leq 0$$

$$vii)\frac{\partial \pi_s(n,f,v_0)}{\partial c_S} \leq 0$$

In addition to the sale probability decreasing in c_B and c_S (in parts *iii* and *iv* above), the share of the transaction price received by the seller in the event of a sale when the reserve price is set at 0, $\frac{\mathbb{E}[V_{(n-1:n)}]}{1+c_B}(1-c_S)$ is decreasing in both c_B and c_S. The reserve price is furthermore decreasing in c_B so that completes the proof for *vi*. The reserve price is increasing in c_S but *vii* holds because if the reserve price is binding the increased c_S in $r^*(v_0, f)$ exactly cancels out with the higher share of the transaction price to pay to the platform. Commissions have no impact (*vi* and *vii* hold with equality) only in the uninteresting case when $F_{V_{(n:n)}(r^*(v_0,f))} = 1$.

Expected seller surplus decreases in v_0.

$$iv)\frac{\partial \pi_s(n,f,v_0)}{\partial v_0} \leq 0$$

Intuitively, a higher seller valuation reduces gains from trade. The decreased sale probability is not offset by a benefit from a higher v_0. Formally, relating the result to *iv* above: $H(v_n, r)$ equals expected payment to the seller if the highest valuation equals v_n and the reserve is set at r. It follows directly that $\frac{\partial \pi_s(n,f,v_0)}{\partial v_0} < 0$ because the derivative adds to the derived effect of v_0 on $\pi_b(n, f, v_0)$ the negative terms: i) $-rf_n(r)(1 + c_B)$ and ii) the contribution from the loss of higher value when the seller sells the item:

$$\pi_s(n,f,v_0) = \int_r^{\bar{v_0}} H(v_n,r)f_n(v_n)dv_n - v_0(1-F_n(r))$$

$$\frac{\partial \pi_b(n,f,v_0)}{\partial v_0} = \frac{\partial v_0}{\partial r}[\int_r^{\bar{v}} -(1+c_B)F_{n-1}(r|v_n)f_n(v_n)dv_n + (v_0-r)f_n(r)(1+c_B)]$$

This is negative due to $\frac{\partial r}{\partial v_0} > 0$ and $v_0 < r$. □

The equilibrium bidder entry probability is decreasing in the seller entry threshold.

Proof Lemma 4. It remains to be shown formally that $\frac{\partial \mathbb{E}[\pi_b(n,f,v_0)|V_0 \leq \bar{v_0}]}{\partial \bar{v_0}} < 0$, the rest of the proof is provided in the main text.

$$\begin{aligned}
\frac{\partial \mathbb{E}[\pi_b(n,f,v_0)|V_0 \leq \bar{v_0}]}{\partial \bar{v_0}} &= \pi_b(c,\bar{v_0},n) f_{V_0|V_0\leq \bar{v_0}}(\bar{v_0}) + \int_0^{\bar{v_0}} \frac{\partial \pi_b(n,f,v_0) f_{V_0|V_0\leq\bar{v_0}}(v_0)(\bar{v_0})}{\partial \bar{v_0}} dv_0 \\
&= \frac{\pi_b(c,\bar{v_0},n) f_{V_0}(\bar{v_0})}{F_{V_0}(\bar{v_0})} - \int_0^{\bar{v_0}} \frac{\pi_b(n,f,v_0) f_{V_0}(v_0) f_{V_0}(\bar{v_0})}{(F_{V_0}(\bar{v_0}))^2} dv_0 \\
&= \frac{f_{V_0}(\bar{v_0})}{F_{V_0}(\bar{v_0})}\left[\pi_b(c,\bar{v_0},n) - \int_0^{\bar{v_0}} \frac{\pi_b(n,f,v_0) f_{V_0}(v_0)}{F_{V_0}(\bar{v_0})} dv_0\right] \\
&= \frac{f_{V_0}(\bar{v_0})}{F_{V_0}(\bar{v_0})}\left[\int_0^{\bar{v_0}} (\pi_b(c,\bar{v_0},n) - \pi_b(n,f,v_0)) \frac{f_{V_0}(v_0)}{F_{V_0}(\bar{v_0})} dv_0\right]
\end{aligned}$$

The last line follows from: $\int_0^{\bar{v_0}} \frac{f_{V_0}(v_0)}{F_{V_0}(\bar{v_0})} dv_0 = 1$ and $\pi_b(c,\bar{v_0},n) \perp v_0$. Finally given that $\frac{\partial \pi_b(n,f,v_0)}{\partial v_0} \leq 0$ so $\forall v_0 < \bar{v_0}$ the derivative is negative. A qualifier is that the no trade equilibrium is excluded from consideration. □

Poisson decomposition property for number of bidders per listing

Proof Lemma 6. The proof concerns the statement that when N^B potential bidders enter a platform with T listings with probability p, the distribution of the number of bidders per listing is approximately Poisson with mean $\frac{N^B p}{T}$. Let M denote the total number of bidders on the platform, distributed Binomial($N^B p, N^B p(1-p)$). The limiting distribution of M when the population of potential bidders $N^B \to \infty$ and associated $p \to 0$ s.t. $N^B p$ remains constant is Poisson($\lambda = N^B p$). Bidders on the platform get uniformly allocated over T listings, entering each listing with probability $q = \frac{1}{T}$. Due to the stochastic number of bidders on the platform, the probability that m bidders get allocated in listing t and n enter into other listings also includes the probability that $m+n$ bidders enter the platform.

$$f_{N_t,N_{-t}}(m,n) = \frac{\exp(-\lambda)\lambda^{(m+n)}}{(m+n)!}\frac{(m+n)!}{m!n!}(q)^m(1-q)^{(n)} \tag{3.44}$$

This joint distribution can be manipulated to reach the conclusion. The $(m+n)!$ cancels out:

$$f_{N_t,N_{-t}}(m,n) = \frac{\exp(-\lambda)\lambda^{(m+n)}}{m!n!}(q)^m(1-q)^{(n)} \tag{3.45}$$

Using that $x^{(a+b)} = x^a x^b$, and rewriting the multiplications:

$$f_{N_t,N_{-t}}(m,n) = \exp(-\lambda)\frac{\lambda^n(1-q)^n}{n!}\frac{\lambda^m p^m}{m!} = \exp(-\lambda)\frac{(\lambda(1-q))^n}{n!}\frac{(\lambda q)^m}{m!} \tag{3.46}$$

Taking a convex combination of λ:

$$f_{N_t,N_{-t}}(m,n) = \frac{\exp(-\lambda q)(\lambda q)^m}{m!} \frac{\exp(-\lambda(1-q))(\lambda(1-q))^n}{n!} \tag{3.47}$$

The marginal probability of having m bidders in listing t takes the expectation over possible values of bidders allocated to other listings, N_{-t}:

$$f_{N_t}(m) = \sum_{n=0}^{\infty} \frac{\exp(-\lambda q)(\lambda q)^m}{m!} \frac{\exp(-\lambda(1-q))(\lambda(1-q))^n}{n!} = \frac{\exp(-\lambda q)(\lambda q)^m}{m!} \tag{3.48}$$

This shows that if the number of bidders on the platform follows a Poisson distribution with mean $\lambda = N^B p$, and bidders enter T listings with equal probability $q = \frac{1}{T}$, the number of bidders in each listing follows a Poisson distribution with mean $\frac{N^B p}{T}$. This is referred to as the *decomposition property* of the Poisson distribution*, e.g. in Myerson (1998). Note that this result does not require the number of bidders to be independent of the number of listings. Hence the decomposition property also applies to the auction platform where the expected number of bidders is a function of the (expected) number of listings. The t subscript is dropped from f_{N_t} as the distribution is identical for all listings $t = \{1, .., T\}$.

***Binomial decomposition property**

This result can also be derived from a decomposition property of the Binomial distribution. If N^B potential bidders enter the platform with probability p and get allocated over T listings with equal probability $\frac{1}{T}$, the below shows that the number of bidders per listing follows a Binomial distribution, $N \sim Binom(N^B \frac{p}{T}, N^B \frac{p}{T}(1-\frac{p}{T}))$ by the law of iterated expectations and iterated variance:

$$P[N=n] = \underbrace{\sum_{m=0}^{N^B} P[N=n|m]P[M=m]}_{\mathbb{E}_M[P[N=n|m]]} = \sum_{m=0}^{N^B} \binom{N^B}{m} p^m (1-p)^{N^B-m} \binom{m}{n} \frac{1}{T}^n (1-\frac{1}{T})^{m-n}$$

$$\mathbb{E}[N] = \mathbb{E}_M[\mathbb{E}[N|m]] = \mathbb{E}_M[\frac{m}{T}] = \frac{\mathbb{E}[M]}{T} = \frac{N^B p}{T}$$

The law of iterated variance states that $Var(N) = \mathbb{E}_M[Var(N|M=m)] + Var(\mathbb{E}_M[N|M=$

$m]$):

$$\begin{aligned}
\mathbb{E}_M[Var(N|M=m)] &= \mathbb{E}_M[m\frac{1}{T}(1-\frac{1}{T})] = (N^B)p\frac{1}{T}(1-\frac{1}{T}) = N^B p\frac{1}{T} - (N^B)p(\frac{1}{T})^2 \\
Var(\mathbb{E}_M[N|M=m]) &= Var(\frac{M}{T}) \\
&= (\frac{1}{T})^2 Var(M) = (\frac{1}{T})^2 N^B p(1-p) = -(\frac{1}{T})^2 N^B p^2 + (\frac{1}{T})^2 N^B p \\
Var(N) &= \mathbb{E}_M[Var(N|M=m) + Var(\mathbb{E}_M[N|M=m]) = N^B\frac{p}{T}(1-\frac{p}{T})
\end{aligned}$$

The last line follows from the last terms on the first and second line cancelling out and rearranging. The large population assumption combined with success probability of *entering in listing* t (for any $t \in \{1,..,T\}$) equal to $\frac{p}{T}$ also renders f_N Poisson with mean $\frac{N^B p}{T}$. $\square$

Chapter 4

Nonparametric identification in English auctions with absentee bidding

Marleen R. Marra

Abstract

The practice of absentee bidding in English auctions obscures the number of bidders required for nonparametric identification of the distribution of valuations, which is key to structural analysis of auction data. In this paper I exploit additional identifying variation in observables from such auctions and I develop a novel nonparametric identification approach based on the *spacings* of order statistics. In combination with a shape restriction this delivers bounds on both the latent value distribution and expected consumer surplus. The value of the proposed method is highlighted by showing that it identifies informative bounds on policy-relevant model primitives in a sample of traditional English auctions collected at Sotheby's that does not contain the number of bidders and their final bids.

"It is at least worth considering whether by a study of the distribution of intervals the statistician can give the epidemiologist any help" (epidemiologist Major Greenwood in a paper presented to the Royal Statistical Society in 1946, quote taken from Pyke (1965) in his paper on sample spacings).

4.1 Introduction

Attend a traditional English auction and you will find that it is less transparent than you remember from your game theory classes. In the Milgrom and Weber (1982) "button auction" model, the price rises exogenously and bidders observably let go of their button at their maximum willingness to pay. The auction stops when the bidder with the second-highest value drops out. Data generated from this stylized auction environment includes a vector of highest bids and the number of bidders. In practice, auction observables are more limited. Most importantly, the number of bidders is often unknown because of the common practice of absentee bidding. Bidders report their maximum willingness to pay to the auctioneer who then bids on their behalf during the live auction. This provides a challenge for structural analysis of bid data because the number of bidders is required for nonparametric identification of the valuation distribution, as established by Athey and Haile (2002) for auctions with independent private value (IPV) bidders. The proof relies on the fact that the distribution of any order statistic from an i.i.d. sample *of known size* from a continuous distribution is a unique function of its parent distribution (e.g. David and Nagaraja (2004)). Hence also in Haile and Tamer (2003), who relax the behavioural restrictions of the button auction model, the number of bidders is needed to set-identify valuations.

The main contribution of this paper is to provide a new approach to set-identify the distribution of valuations without knowing the number of bidders. I leverage observed absentee bids that reveal additional information about bidders' willingness to pay to pin down a lower bound on the third highest valuation. *Sample spacings*, a concept from the order statistics literature (Pyke (1965, 1972)) is then used to derive an upper bound on the highest valuation. I discuss various shape restrictions that can be used to obtain this bound, of which log-concavity is the weakest restriction already implied by the Increasing Failure Rate assumed in many auction studies (Bagnoli and Bergstrom (2005)). Together with observed bids this delivers bounds on the distribution of the highest valuation sufficient to bound both expected consumer surplus and seller revenue. The implied sample extremes also bound the latent distribution of valuations of primary interest in structural work. I then analyze the identifying power of additional restrictions to tighten these bounds, including exogenous variation in the number of bids. As these bounds can be wide, I discuss what features of the data generating structure are most conducive to this approach delivering bounds that are informative in practice.

Generally, identification of the valuation distribution is not immediate in oral, open-outcry auctions as the drop-out value of the highest bidder is never observed. An interesting development in recent structural analysis under weaker restrictions than IPV is therefore to focus on directly identifying the policy variables that are ultimately of significance. In a symmetric private value environment where valuations are allowed to be correlated, Aradillas-López et al. (2013) set-identify the expected consumer surplus and expected seller revenue when only the transaction price and the number of bidders is observed. Coey et al. (2017) extend these results to asymmetric bidders and furthermore show that the bounds can be sharpened when bidder identities are also observed. In first-price auctions with unobserved heterogeneity, Armstrong (2013) shows that just observing the transaction price and the number of bidders is sufficient to set-identify features related to expectations of valuations such as consumer surplus. Tang (2011) bounds seller revenue in first price auctions with common values. What these papers have in common is that they investigate what economic features of interest can be (set-)identified in auctions when relaxing the auction model in important dimensions but without imposing additional parametric structure on the problem. My results contribute to this line of research.

Also Song (2004) provides a method to nonpametrically identify the distribution of valuations in certain IPV English auctions when the number of bidders is unknown. A key insight is that observing a pair of adjacent order statistics is sufficient to point identify the distribution. As usual, the second-highest valuation is equal to the transaction price (when abstracting from bidding increments). To pin down the third-highest valuation, Song (2004) exploits the fixed end time rule used in eBay auctions. I instead focus on the more traditional English auction setting with a flexible end time, where the auction only closes after providing bidders ample opportunity to raise the standing bid ("going, going, gone").[1] The method in Song (2004) does not apply directly to traditional English auctions as observables can only bound the third-highest valuation.

To emphasize the limited information directly observable in English auctions with absentee bidding, highest bids cannot be discerned from the vector of bids when identities are not known and when live bidders may not get the chance to place a bid at their

[1]Besides being the norm in offline auctions for e.g. cattle, real estate, art, or machinery, the flexible end time rule is also used in online auctions outside of eBay. Examples in no particular order are: www.cattleusa.com, www.bonhams.com, www.taylerandfletcher.co.uk. www.sothebys.com, www.chiswickauctions.co.uk, www.liveauctioneers.com, www.i-bidder.com, www.easyliveauction.com, www.bontragerauction.com, www.auctionsnewengland.com, www.christies.com

maximum valuation. This limited information content diverges from what is assumed known in previous studies, including Paarsch (1997) using all bidders' drop-out values and the number of bidders, Haile and Tamer (2003) and Chesher and Rosen (2015, 2017) using a vector of highest bids and the number of bidders, Song (2004) using a vector of bids that includes the second and third-highest drop-out values, and Aradillas-López et al. (2013) using the second highest drop-out value and number of bidders. These papers are all ground breaking in their econometric use of bid data from English auctions, but it is problematic to apply their identification strategies to the limited bid data central to this paper. For the analysis developed here, the econometrician needs to observe only: 1) a vector of bids and 2) which bids are submitted by absentee bidders.

The rest of the paper is organized as follows. Section 4.2 introduces the theoretical auction model. Section 4.3 explains the informational content of bids from absentee bidders and the use of sample spacings to identify bounds on valuation order statistics. Section 4.4 presents identification results. Section 4.5 discusses nonparametric Kernel estimation of the relevant distribution functions. In Section 4.6 the method is applied to a new dataset of fine wine auctions in which the number of bidders and their highest bids is unknown. The Appendix contains all proofs, results from Monte Carlo simulations and a section on inference.

4.2 Auction model

The mechanism is an English auction with a flexible closing rule, fixed bidding increments and a secret reserve price. Bidders have the option to place a sealed bid ahead of the auction. Ex-ante symmetric, risk neutral bidders have unit demands and face negligible entry and bidding cost. Key assumptions on model primitives are:

Assumption: (IPV). *Bidders independently draw a valuation from a common distribution. This distribution is assumed to satisfy the following regularity conditions: i) it is absolutely continuous, ii) it is defined on the bounded support* $[\underline{v}, \bar{v}]$*, and iii) it satisfies the Increasing Failure Rate.*

The main part of this paper relies on the conditional independence assumption that is the main tenet in the structural analysis of English auctions (Paarsch and Hong (2006)). Where possible it is discussed how the results apply to auctions in which bidders have unrestricted private values.

Assumption: (Symmetry). *Bidders are symmetric. In particular, their preference for* *live or absentee bidding is independent of their valuation.*

This assumption rules out that absentee bidders differ systematically in their valuation for the auctioned objects. In practice, even in brick and mortar auction houses live bidding does usually not require a bidder to be present in the auction room as he can participate online or on the phone. As it is not necessarily more or less costly to submit an absentee bid in writing before the auction, the model assumes a random preference for absentee bidding.

I follow the convention to denote random variables in upper case and their realizations in lower case. Auction observables are denoted by random vector $\mathbf{Z}$. The marginal distribution of valuation V_i conditional on $\mathbf{Z}$ is denoted by $F_{V_i|\mathbf{Z}}$ for bidder i. $F_{\mathbf{V}|\mathbf{Z}}$ is shorthand notation for the joint distribution of valuations in auctions with n bidders. Their respective densities are: $f_{V_i|\mathbf{Z}}$ and $f_{\mathbf{V}|\mathbf{Z}}$. Symmetry makes that $F_{\mathbf{V}}$ is a joint distribution of exchangeable random variables and consequently $F_{V_i|\mathbf{Z}}$ reduces to $F_{V|\mathbf{Z}}$ $\forall i$. Independence implies that $F_{\mathbf{V}|\mathbf{Z}} = F^n_{V|\mathbf{Z}}$ which is the latent distribution targeted in structural models with IPV bidders.

The following assumption restricts variation of the unobserved number of bidders n.

Assumption: (Exogenous participation)**.** *If n varies across auctions with identical covariates, $F_{V|\mathbf{Z}}(.|n) = F_{V|\mathbf{Z}}(.|n')$ $\forall n \neq n'$, $\forall \mathbf{z} \in \mathbf{Z}$.*

The assumption formalizes what is already implied by the combination of IPV and negligible entry cost: the number of potential bidders does not vary due to auctioned items being more or less valuable. This is important to stress because without observing the number of bidders, identification of the valuation distribution requires that this distribution is either independent of the number of bidders or that the number of bidders is fixed across auctions. Exogenous participation is an assumption that has been previously imposed to aide identification or testing in various auction settings (e.g. Athey and Haile (2002), Haile and Tamer (2003), and Aradillas-López et al. (2013)). Song (2004) shows that in order to also identify the distribution of the number of bidders this number must be assumed fixed across auctions, which is more restrictive.

4.2.1 Absentee bidding

While the absentee bidding component of English auctions is hardly mentioned in the literature there are some notable exceptions: Ginsburgh (1998), Rothkopf et al. (1990) and Thiel and Petry (1995). Lucking-reiley (2000) finds that this practice has been used since at least 1878 for stamp auctions. Auction mechanisms with absentee bidding are designed in a way that does not disadvantage absentee bidders. To see why, if the opening bid would be the highest submitted absentee bid all their rent would be depleted instantly and it would be more attractive to participate in the live bidding. If only one absentee bid is submitted for an item, the opening bid must start below that value and the exact opening bid rule varies by institution. Some start at the reserve price, some start at a predetermined share of the absentee bid. The following explanations by different auction houses provide additional intuition about absentee bidding in practice. The first one is taken from the Sotheby's wine auction catalog provided for the auctions in my data:

Absentee bids are to be executed as cheaply as permitted by other bids or reserves and in an amount up to but not exceeding the specified amounts. Bids will be rounded down to the nearest amount consistent with the bidding increment. In the case of identical (rounded) bids, the earliest submitted form will take precedence.[2]

A commission/absentee bid is a figure left with the auctioneers by the interested party on the premise that we will bid on their behalf in their absence. We will only bid up to wherever the bidding stops in the room for any given item.[3]

You may hear an auctioneer say that he's got a "commission bid" or that he has a bid "on the book". What that means is that someone who is unable to attend the auction has completed a commission bidding form and entered a maximum price for which he will pay for a particular lot.[4]

The empirical observations that: i) my data contains various auctions with multiple absentee bids before live bidding starts and ii) after initial absentee bids the live bidding alternates with absentee bids motivate the following stylized representation of the data:

Bidding sequence. *With one absentee bid, the opening bid equals the minimum of the reserve price and that bid. With multiple absentee bids, the opening bid equals the minimum of the absentee bids. Subsequent bids follow fixed bidding increments until all but one absentee bidder drops out. Afterwards, live bids are alternated against the last remaining absentee bidder unless he has dropped out.*

Most importantly this guarantees that absentee bidders do not have a disadvantage compared to live bidders.[5] Accuracy of the opening bid rule cannot be verified in the data.

[2] Source: Sotheby's (2014)

[3] Source: http://www.cheffins.co.uk/en/auctions/fine-art/buying-and-selling/buying/how-to-bid

[4] Source: http://www.wilsonsauctions.com/ways-to-bid

[5] The assumed opening rule depends on the empirical application. If instead no more than two absentee bids precede live bidding, it is more plausible that the opening bid equals the second-highest absentee bid when there are more than two absentee bids.

While it is helpful to specify this sequence for analyzing bids, results do not depend on it: another rule that does not disadvantage absentee bidders and is consistent with the data is also allowed.

To see how this restriction compares to structural analysis of bid data in other IPV English auction studies, consider the two benchmark models. Paarsch (1997) assume that data is generated by the button auction model in which case all bids are assumed to be a different bidder's maximum willingness to pay. But even when relaxing the behavioral restrictions of the button auction model and using the incomplete model of Haile and Tamer (2003) for structural analysis of bid data, it is required to observe a vector of highest bids. This is explicitly stated in Chesher and Rosen (2015, 2017) who extend the work of the Haile and Tamer (2003) by providing sharp bounds on the valuation distribution. A vector of highest bids may indeed be readily observable in some English auctions. For example, eBay provides bidder identities alongside with submitted bids, but in the traditional English auctions that are the focus of this paper this information is often unavailable. Instead of needing to know highest bids and the number of bidders, the following information is necessary for the approach in this paper:

Informational requirement: *The econometrician observes: i) a vector of bids, ii) which ones are absentee bids.*

4.2.2 Equilibrium strategies

The only restriction on bidder behavior is that it is rational and that bidders are attentive. In particular, the two intuitive behavioral assumptions that define bidding strategies in the incomplete model of Haile and Tamer (2003) are imposed. While these assumptions are satisfied in all symmetric separating equilibria of the button auction model of Milgrom and Weber (1982), they also allow for alternative behavior including not bidding at all or bidding less than one's valuation.

Assumption: (Haile & Tamer bidding)**.** *Potential bidders never bid more than they are willing to pay and never let an opponent win an auction at a price they are willing to beat.*

Recall that absentee and live bidders are symmetric up to their preference for absentee bidding and their valuation draw. However, there is a crucial difference in their potential bidding strategies. Live bidders can squat, jump, not bid, bid once or many times incrementally (see Hasker and Sickles (2010) and the references therein). For absentee bidders their strategy only consists of the height of their maximum bid submitted to the auctioneer. Even under the weak behavioral assumption this bid will be truthful.

Lemma 6. *The unique symmetric Bayes Nash equilibrium in weakly undominated strategies is for absentee bidders to bid their valuation.*

The proof of Lemma 6 is relegated to the Appendix, but the result is intuitive as for absentee bidders the environment is strategically equivalent to an IPV second-price sealed bid auction. It is well-known that truthful revelation is a weakly undominated strategy in that setting (Vickrey (1961)).[6]

Note that Lemma 6 does *not* correspond to the so-called 'sealed-bid abstraction' often employed in eBay auction studies (e.g. Bajari and Hortaçsu (2003), Song (2004)).[7] Zeithammer and Adams (2010) argue that some eBay bidders tend to 'proxy bid' below their valuation, in order to potentially increase their bid at a later time, invalidating this abstraction (see also the commentary by Hortaçsu and Nielsen (2010)). Absentee bidders however cannot raise their bids at a later stage.

4.3 The identifying power of (absentee) bids and spacings

In this section, I leverage additional information revealed by absentee bids to bound an additional order statistic of the valuation distribution. The second step is to use results from the statistics literature on the stochastic spacing of order statistics depending on the shape of the underlying valuation distribution.

4.3.1 Relation of bids to values

Lemma 1 established that the maximum amount that absentee bidders report to the auctioneer is equal to their valuation. The sequence of observed bids also include intermediate bids because the auctioneer is to obtain the lot for the absentee bidder at the lowest possible price. Following the modelled bidding sequence, it is possible to determine which bids correspond to absentee bidders drop-out values rather than intermediate bids. For example, the lowest bid is one such value if there are multiple absentee bids placed before live bidding starts. The highest absentee bid if it is not the winning bid also corresponds

[6] Restriction of the strategy space to weakly undominated strategies is required for uniqueness of the equilibrium (Bikhchandani et al. (2002)) Imposing a weaker notion of "posterior implementability" as in the button auction model instead allows for losing bidders to not regret bidding less than their valuation.

[7] Bajari and Hortaçsu (2003) split the auction up into two parts; the first part is modeled as a regular English auction but the second part as a second-price sealed bid auction where bidders bid truthfully at the last-minute, so that the revelation of their private information to the other common value bidders is inconsequential. Song (2004) introduces a last 'personal monitoring opportunity' where private value bidders will bid their valuation in eBay auctions, driven by second-price sealed-bid motives.

to a drop-out value. Example 1 shows this in more detail and additional examples are provided in the appendix.

Example 2. A_i and L_j respectively denote the ith lowest overall bid, which came from an absentee bidder, and the jth lowest overall bid in the bid vector, which came from a live bidder. So the first and lowest absentee bid is A_1 if there is an absentee bidder, and the other bids follow according to the defined bidding sequence.

Auction 1: A_1, A_2, A_3, L_4, L_5

Auction 2: A_1, L_2, A_3, L_4, L_5

Auction 3: A_1, L_2, A_3

In Auction 1 there must be more than one absentee bidder as there are multiple absentee bids placed before live bidding starts, any other explanation would violate that the auctioneer acts in the best interest of the absentee bidder. Hence, A_1 is one drop-out value. A second one is A_3. Also note that L_4 and L_5 must be placed by different (live) bidders because nobody would outbid himself.

In Auction 2, live bidding starts immediately after the first bid, indicating that there is just one absentee bidder. Here, one drop-out value is A_3, and two other ones are L_4 and L_5 (again because nobody outbids himself).

In Auction 3, A_3 is the lowest drop-out value of an absentee bidder and L_2 is another drop-out value.

Practically, due to the fixed bidding increments (and rounding of absentee bids) all these drop-out values are lower bounds on the valuations of the corresponding bidders. For absentee bidders, upper bounds can be established that add two bidding increments to their identified drop-out values. To see why, consider the bid A_3 in Auction 2. If the absentee bidder for whom the auctioneer submits this bid had a valuation that was more than two bidding increments above A_3, it would show up as a higher absentee bid, replacing L_5. By contrast, due the flexibility of their bidding strategies, live bidders could have a valuation larger than two bidding increments above their drop-out value.

Let b denote the number of submitted bids in an auction and Δ the minimum bidding increment. Using order statistic notation, $B_{b:b}$ is the highest submitted bid, $B_{b-1:b}$ the second-highest and B_x the highest observed drop-out value less than $B_{b-1:b}$. As explained above, B_x is a different bid order statistic in different circumstances.[8] It is not new that

[8]In Example 1, in Auction 1 and 2 $B_x = B_{b-2:b}$ and in Auction 3 no B_x can be identified. Given the assumption of exogenous variation in the number of bidders this sample selection problem is not an issue.

the second-highest valuation is bounded between $B_{b-1:b}$ and $B_{b:b} + \Delta$ (Haile and Tamer (2003)) and it is clear that the highest valuation must exceed $B_{b:b}$. These observations explain (4.1)-(4.3) below.

I use additional information revealed by the bidding sequence including absentee bids to bound the third-highest valuation. Equation (4.4) holds by the definition of B_x. This statistic is identified in auctions where a drop-out value is observed that is less than the two highest observed bids.[9] Furthermore, nobody outbids themselves so the highest two bids are submitted by different bidders. As B_x must be submitted by another bidder by definition, this bid is a lower bound on the third-highest valuation as in (4.4).

This provides the following (incomplete) list of order statistics:[10]

$$V_{n:n} \geq B_{b:b} \equiv V^L_{n:n} \tag{4.1}$$
$$V_{n-1:n} \geq B_{b-1:b} \equiv V^L_{n-1:n} \tag{4.2}$$
$$V_{n-1:n} \leq B_{b:b} + \Delta \equiv V^U_{n-1:n} \tag{4.3}$$
$$V_{n-2:n} \geq B_x \equiv V^L_{n-2:n} \tag{4.4}$$

The observed right-hand side variables place bounds on valuation order statistics of interest. As per the above, subscrips L and U respectively denote the lower and upper bound on the order statistic pinned down by observable bids. I will now show how sample spacings can be used to bound the distribution and expected value of the highest valuation given distributions of observed ordered bids. Bounds on highest and second-highest valuations can be used directly to identify expected consumer surplus as Aradillas-López et al. (2013) do for unrestricted private value English auctions (crucially, relaxing IPV). In Section 4.4 I also investigate the usefulness of (bounds on) sample extremes in combination with additional restrictions to tighten these bounds to set-identify the distribution of valuations directly.

[9] It could be submitted by an absentee bidder or another live bidder if it can be distinguished from the highest two bids (by being submitted on the phone versus in the room, for example).

[10] Also bounding the fourth-highest valuation (a worst case bound is the opening bid but further information is available in the bidding sequence of some auctions) could be useful. In particular, it may help to evaluate what is the strongest shape assumption (introduced shortly) not falsified by the data. I have not yet implemented this in the current version of this paper. The qualification that the list of order statistics is incomplete specifically refers to the point that $\forall i \in \{0, ..., n-1\}$ $V_{n-i:n} \leq B_{b:b} + \Delta$, but these other order statistics are not useful here given that n is unobserved.

4.3.2 Sample spacings and shape restrictions

Due to the censoring problem in English auctions, the fact that the auction stops when the bidder with the second-highest valuation drops out (after adding one increment to his drop-out value), it is a standard issue that the highest valuation is never observed. I discuss three nonparametric shape restrictions of increasing power that impose enough structure to identify informative upper bounds on $V_{n:n}$. It is then trivial to translate these into lower bound on the distribution of the highest valuation. While restrictive, the three cases are general enough to be valid in a variety of empirical settings: i) log-concavity, ii) concavity, and iii) symmetry of $f_{V|\mathbf{Z}}$. Proofs of lemma's presented in this section are relegated to the appendix.

Definitions for spacings and stochastic ordering. Following Pyke (1965, 1972) spacings D_i between two adjacent order statistics are defined as: $D_i = V_{i:n} - V_{i-1:n}$, $\forall i = 2, .., n$. Normalized spacings are defined as: $\tilde{D}_i = (n-i+1)(V_{i:n} - V_{i-1:n})$, $\forall i = 2, .., n$. Both D_i and $\tilde{D}_i$ are random variables with CDF's F_{D_i} and $F_{\tilde{D}_i}$ $\forall i = 2, ..., n$. Random variable X is said to be stochastically smaller than Y if $F_X(a) \geq F_Y(a)$ $\forall a$, which also implies that $\mathbb{E}[X] \leq \mathbb{E}[Y]$. When the sample spacings $\mathbf{D} = \{D_2, D_3, .., D_n\}$ are said to be stochastically decreasing if: $F_{D_1}(a) \geq F_{D_2}(a) \geq ... \geq F_{D_n}(a)$ $\forall a$.

The first shape restriction uses the result that normalized spacings of order statistics based on an i.i.d. sample drawn from a log-concave density function are stochastically decreasing (Pyke (1965, 1972)). For reference, normalized spacings from an exponential distribution are independent and identically distributed. Also, the right tail of a log-concave density function is at most as convex as the right tail of the exponential density function.

Lemma 7. *[Log-Concave density] Let $f_{V|\mathbf{Z}}$ be (weakly) log-concave. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}$ $\forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \geq P[V^U_{n-1:n} + 2(V^U_{n-1:n} - V^L_{n-2:n}) \leq v|\mathbf{Z} = z] \equiv F^{L(lc)}_{V_{n:n}|\mathbf{Z}}(v) \tag{4.5}$$

In the case that $f_{V|\mathbf{Z}}$ is weakly log-concave (exponential) and $\Delta = 0$ $F^{L(lc)}_{V_{n:n}|\mathbf{Z}}(v)$ is only an upper bound if the bidder with the third-highest valuation had no opportunity (or reason) to drop-out at his valuation. Otherwise, (4.5) *holds with equality in this case.*

The superscript "L(lc)" in Equation 4.5 indicates that this is a lower bound for valuation distribution functions satisfying a log-concave density. Log-concavity is a commonly used restriction in the auction literature. Distribution functions that satisfy the restriction (at least for certain parameter values) include the: extreme-value, gamma, beta, Weibull and power distributions. Some distribution functions that are used in practice, including the Pareto and log-normal, do not have log-concave density functions for any parameter values (Bagnoli and Bergstrom (2005)).[11]

As log-concavity is least restrictive and often imposed in empirical auction studies $F^{L(lc)}_{V_{n:n}|\mathbf{Z}}(v)$ is the basis of results in the rest of this paper. However, to provide some additional insight into the identifying power of shape restrictions I also examine the usefulness of concavity and symmetry. At this point I'd like to emphasize that these restrictions do not exhaust the options. It is likely that other, perhaps more appropriate for some data generating processes, shape restrictions can be combined with results from the statistical literature on sample spacings to aid the structural analysis of data from auction markets. Furthermore, more can be done when having access to a vector of highest bids or to at least three order statistics. The log-concavity assumption that is commonly imposed (or implied by it: the Increasing Failure Rate assumption that is necessary for the Riley and Samuelson (1981) optimal reserve price to be unique) can be tested using spacings. It's increasing normalized spacings implication can be directly tested with such data, as more formally shown in An (1995).

Regarding concavity: spacings of order statistics based on an i.i.d. sample drawn from a concave density function are stochastically decreasing (Pyke (1965, 1972)) For reference, the right-tail of a concave density function is at most as convex as the uniform density function (on the same support). This provides a tighter lower bound on the distribution of the highest valuation as formalized in the following Lemma.

Lemma 8. *[Concave density] Let $f_{V|\mathbf{Z}}$ be (weakly) concave. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}\ \forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \geq P[V^U_{n-1:n} + V^U_{n-1:n} - V^L_{n-2:n} \leq v|\mathbf{Z} = z] \equiv F^{L(c)}_{V_{n:n}|\mathbf{Z}}(v) \tag{4.6}$$

The superscript "L(c)" in Equation 4.6 indicates that this is a lower bound for valuation distribution functions satisfying a concave density. As concavity is (much) stronger than log-concavity, this bound is (much) tighter than the bound derived under the assumption of log-concavity in Equation (4.5).

[11] Also see Bagnoli and Bergstrom (2005) for references to the use of log-concavity or restrictions implied by it in economic analysis of games including auctions and for an elaborate overview of results for log-concave densities and distributions.

Regardless of whether the density is convex or (log-)concave, and uni-modal or not, when it is symmetric the following result holds:

Lemma 9. *[Symmetric density] Let $f_{V|\mathbf{Z}}$ be symmetric. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}\ \forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \geq P[V^U_{n-1:n} + V^U_{n-1:n} - V^L_{1:n} \leq v|\mathbf{Z} = z] \equiv F^{L(s)}_{V_{n:n}|\mathbf{Z}}(v) \quad (4.7)$$

The superscript "L(s)" in Equation 4.7 indicates that this is a Lower bound for valuation distribution functions satisfying a symmetric density.

All results in this section assume that the econometrician observes a large set of independent auctions. Exogenous variation in the unobserved number of bidders (if there is variation) guarantees that it is inconsequential to select auctions with at least three bidders. Three bidders is a necessary condition for B_x to be known.

4.4 Set-identification of bidder valuations and consumer surplus

This section presents new nonparametric identification results using the derived bounds on the highest valuation, without knowing the number of bidders. I also discuss previous findings to highlight distinguishing features of different identification methods. In this section I omit auction observables $\mathbf{Z}$ for brevity.

Generally, in auctions with a binding reserve price, observables are uninformative for this distribution on the truncated part of the support (Athey and Haile (2002)). Positive nonparametric identification results however extend to the distribution that is truncated from below at the reserve price, which is typically the structural feature of interest when that reserve is known.[12] Due to the secret reserve price in my data (and typical in traditional English auctions), the distribution of interest in this paper is instead truncated at the opening bid. Since absentee bids are called out at the start of the auction, this has a similar censoring effect on potential live bidders as starting the bidding at a known reserve price. Without conditioning on $B_{1:b}$ the latent distribution of interest is a mixture that takes the expectation over opening bids.[13]

[12] Another approach is to impose parametric structure (e.g. Donald and Paarsch (1996), Paarsch (1997)), as advocated by Paarsch and Hong (2006), as the only way to calculate the optimal reserve price without imposing that it must be at least equal to the actual reserve.

[13] The mixture distribution equals: $\forall v \in [\underline{v}, \bar{v}]$: $\int \frac{F_{V|\mathbf{Z}}(v) - F_{V|\mathbf{Z}}(b_{1:b})}{1 - F_{V|\mathbf{Z}}(b_{1:b})} dF_{B_{1:b}|\mathbf{Z}}(b_{1:b}|z)$, with $dF_{B_{1:b}|\mathbf{Z}}$ the density of opening bids and $b_{1:b}$ a specific realization of random variable $B_{1:b}$. Haile and Tamer (2003,

I now briefly discuss nonparametric identification in IPV English auctions as established by Athey and Haile (2002) when the number of bidders is known. Symmetry makes that identifying any marginal CDF implies identifying the marginal CDF for all bidders, but the results generally go through for asymmetric bidders when also the identities of the bidders are known. A key insight is that the distribution of $V_{i:n}$ is a strictly increasing, known, function of its parent distribution. Hence any order statistic from an i.i.d. sample of known size from a continuous distribution uniquely pins down its parent distribution. In particular, the distribution of the $(n-i+1)$th order statistic of valuations:

$$F_{V_{i:n}}(v) = \frac{n!}{(i-1)!(n-i)!} \int_{\underline{v}}^{v} F_V(t)^{i-1}[1-F_V(t)]^{n-i} f_V(t)dt \tag{4.8}$$

Defining this monotone transformation as $\psi\left(F_V(v); n, i\right)$, $F_{\mathbf{V}}(v)$ can be obtained as: $\left[\psi^{-1}\left(F_{V_{i:n}}(v); n, i\right)\right]^n$, $\forall v \in [\underline{v}, \bar{v}]$ and where ψ depends crucially on i and n. For instance, in a button auction rationalization of the auction the highest bid (as equivalent to the second-highest valuation) and the number of bidders point-identifies the distribution of valuations. In the incomplete model of Haile and Tamer (2003), bounds on the parent distribution are identified using the known transformation in Equation (4.8) and stochastic dominance relationships of the distributions of certain bid order statistics bounding valuations.

Other objectives in structural auction analysis are pinning down expected consumer surplus (CS) and expected seller revenue (SR). The ex-ante consumer surplus in auction t only depends on the valuation of the highest bidder and the final price in each auction t: $CS^t = V^t_{(n:n)} - p^t$, with p^t being the observed total price (including commissions):

$$CS^t = \int_{B^t_{b:b}}^{\infty} x f_{V_{n:n}}(x)dx - p^t \tag{4.9}$$

Bounds on the highest valuation obtained in part with the method outlined in the previous section deliver bounds on $f_{V_{n:n}}(x)$, denoted again with U and L subscripts. Bounds on $f_{V_{n:n}}(x)$ trivially bound CS^t by stochastic dominance:

$$\int_{B^t_{b:b}}^{\infty} x f^U_{V_{n:n}}(x)dx - p^t \leq CS^t \leq \int_{B^t_{b:b}}^{\infty} x f^L_{V_{n:n}}(x)dx - p^t \tag{4.10}$$

footnote 17) present a discrete version of this mixture distribution when $F_{V|\mathbf{Z}}$ is left-truncated by a stochastic reserve price when not conditioning on that reserve price. The proposed identification strategy is only useful for counterfactuals that do not affect distribution of opening bids. As in English auctions typically the right-tail of the distribution is of interest for counterfactual policies, this is likely an inconsequential restriction. In particular, if in doubt the econometrician can also specify the latent distribution of interest to be left-truncated at another exogenous cut-off value.

Bounds on marginal distributions $F_{V_{n-1:n}}$ and $F_{V_{n:n}}$ can also be used to bound expected CS and SR as shown by Aradillas-López et al. (2013).[14]

4.4.1 Set-identification using sample spacings

The following results regard the latent distribution of valuations using sample extremes.

Lemma 10. *[Sample extremes under IPV] In symmetric IPV English auctions, observing $V^L_{1:n}$ and $V^U_{n:n}$ identifies F_V within bounds:*

$$F_{V^L_{1:n}}(v) \geq F_V(v) \geq F_{V^U_{n:n}}(v) \tag{4.11}$$

Intuitively, the valuation of any bidder must be at least as much as the valuation of the bidder with the lowest valuation but at most the valuation of the bidder with the highest valuation. Clearly, that statement does not require the independence assumption and is valid for unrestricted private values as well. In that case sample extremes bound the joint distribution of valuations. The distribution of the lowest valuation bounds the joint distribution of valuations for all bidders and the distribution of the highest valuation is equivalent to the joint distribution.

Lemma 11. *[Sample extremes under unrestricted private values] In symmetric (unrestricted) private value English auctions, observing $V^L_{1:n}$ and $V^U_{n:n}$ identifies $F_{\mathbf{V}}$ within bounds:*

$$F_{V^L_{1:n}}(v) \geq F_{\mathbf{V}}(\mathbf{v}) \geq F^U_{V_{n:n}}(v), \tag{4.12}$$

where $F_{\mathbf{V}}(\mathbf{v}) = F_{V_1,\ldots,V_n}(v,\ldots,v)$.

As a result, the proposed identification approach for latent valuations is attractive even beyond the auction environment motivating this paper where the number of bidders is unobserved but where bidders have conditionally independent private values.

Partly relaxing both the information requirements and the structural restrictions to identify bounds on the valuation distribution and the winning bidders' surplus cannot come without a cost.[15] Indeed, it is not surprising that using only the bid sample *extremes*

[14] Practically, since reserve prices are unobserved, this would multiply the surplus and revenue functions with the sale probability that is pinned down from the data, separately for each $\mathbf{Z} = \mathbf{z}$.

[15] The information set is 'partly' weakened because although the number of bidders is not required, it is assumed that all bids including which ones are from absentee bidders is known. The structural restrictions are also 'partly' weakened because an assumption on the bidding sequence (and for the lower bound on the shape of the distribution) is added. Yet as both types of added requirements are modeled to reflect actual traditional English auctions, for such environments it can be concluded that this paper imposes a less restricting structure.

rather than the infimum and supremum over a number of pinned-down bid order statistics (with IPV and a known number of bidders) in Haile and Tamer (2003) results in wider bounds on the valuation distribution; especially if the lower bound results from a fairly conservative shape restriction.

This statement by itself goes to show the importance of the number of bidders. Without that statistic it is difficult to make sense of auction data. The identification approach proposed in this paper should be considered in tandem with a strategy to sharpen the bounds. I consider identifying variation from the observed number of bids. By the exogenous participation assumption, the distribution of valuations doesn't vary due to the number of bidders. Within this context it makes sense that also variation in the number of *bids* arises due to factors exogenous to bidder valuations. Conditional on bidder valuations, the English auction model allows utter flexibility of bidding strategies in terms of the number (and timing) of bids placed by live bidders. They can actively participate from early on by placing multiple bids below their valuation, or wait with bidding initially, bid once a high amount or incrementally.[16] Furthermore, the lowest maximum bid submitted by absentee bidders is stochastic so just variation in their value draws generates different number of bids without the latent valuation distribution itself being affected.

Let the number of bids for auctions with covariates Z be denoted as random variable $B_{\mathbf{Z}}$ with realization b:

Assumption: (Conditionally Independent Number of Bids (CINB))**.** $F_{V|\mathbf{Z}} \perp B_{\mathbf{Z}}$*, so* $\forall (b \neq b^{'}) \in B_{\mathbf{Z}}$*:* $F_{V|\mathbf{Z}}(v|b) = F_{V|\mathbf{Z}}(v|b^{'})$*,* $\forall v \in [\underline{v}, \bar{v}]$ *and* $\forall \mathbf{z} \in \mathbf{Z}$*.*

$F_{V|\mathbf{Z}} \perp B_{\mathbf{Z}}$ is the exclusion restriction necessary for the tighter bounds to be valid.

Imposing CINB, those bounds (indicated with a prime) are obtained as:$\forall v \in [\underline{v}, \bar{v}]$, $\forall \mathbf{z} \in \mathbf{Z}$:

$$\begin{aligned} F^{L'}_{V|\mathbf{Z}}(v) &\equiv \max_{b \in B_{\mathbf{Z}}} F_{V^U_{n:n}|\mathbf{Z}}(v|b) \\ F^{U'}_{V|\mathbf{Z}}(v) &\equiv \min_{b \in B_{\mathbf{Z}}} F_{V^L_{1:n}|\mathbf{Z}}(v|b) \end{aligned} \tag{4.13}$$

Alternatively, if the econometrician has access to other observables that are excluded from bidder valuations the same strategy can be used to get the pointwise minimum and maximum over realizations of that random variable. This is more generally proposed as an identification strategy in Athey and Haile (2007) and also used to tighten bounds in e.g. Aradillas-López et al. (2013) who exploit exogenous variation in the number of bidders. For

[16] Cho et al. (2014) for instance provide empirical evidence for an interesting range of exercised bidding strategies including those mentioned here, although for a different bidding environment.

empirical settings where it is more plausible that valuations are stochastically increasing or decreasing in an observed variate, the same strategy applies to tighten only one of the bounds.

4.4.2 Extension of existing method relying on two order statistics

Song (2004) takes another approach to identification in IPV English auctions when the number of bidders is unknown, using the fixed closing rule of eBay auctions. The model pins down the second- and third-highest valuation and identification uses the order statistic result that the distribution of $V_{i:n}$ conditional on $V_{j:n}$ (for $j < i$) equals the distribution of $V_{i:j-1}$ from a sample truncated from below at the realization of $V_{j:n}$ (as more generally shown in David and Nagaraja (2004)). Using this additional third-highest order statistic of valuations cleverly avoids needing to know the number of bidders since the monotonic transformation depends on the distance $(j-1)$ between the order statistics. Song uses this idea to identify F_V by taking limits: $\lim_{v_j \to \bar{v}} F_{V_{i:n}|V_{j:n}}(v_i|v_j) = F_{V_{(i:j-1)}}(v_i)$ for $(j < i)$.

The limit assumption supports the identification strategy of Song (2004) (and based on her results Adams (2007) and il Kim and Lee (2014)) and is valid if the relevant third-highest order statistic is indeed sufficiently spread out. In that case, F_V is identified off auctions with few bidders and a low third-highest valuation.

I explore this result when instead all order statistics are only known within bounds, as in traditional English auctions that do not benefit (in this one sense) from a fixed closing rule.

Lemma 12. *[Based on Song (2004)] In symmetric IPV English auctions, knowing two ordered valuations $V_{i:n}$ and $V_{j:n}$, with $i = n-k$ and $j = n-l$ $(0 \leq k < l \leq n-1)$ nonparametrically identifies $F_{V|V \geq v_j}(.)$.*[17]

A complete auction model with IPV that pins down any pair of order statistics of the distribution of valuations thus does not require the number of bidders to identify the marginal distribution of valuations that is truncated at the realization of the conditioning order statistic.[18] It is then natural, in the spirit of Haile and Tamer (2003), to use a stochastic dominance argument based on monotonicity of the ψ transformation (see Equation (4.8)) to support that an *incomplete* English auction model that set-identifies

[17] Equivalently, for procurement auctions, knowing $V_{i:n}$ and $V_{j:n}$, with $i = k$ and $j = l$ $(1 \leq k < l \leq n)$ nonparametrically identifies $F_{V|V \leq v_j}(.)$

[18] I say 'pair' here to emphasize that without knowing the sample size it is important that these ordered values are either the ith and jth *highest* values or the ith and jth *lowest* values. Otherwise their distance is unknown as any number of bidders could have a valuation in between.

at least one of these order statistics delivers bounds on the truncated distribution.[19]

Lemma 13. *[Based on Song (2004) and Haile and Tamer (2003)] In symmetric IPV English auctions, knowing bounds on at least one of two ordered valuations $V_{i:n}$ and $V_{j:n}$, with $i = n - k$ and $j = n - l$ $(0 \leq k < l \leq n-1)$ nonparametrically identifies bounds on $F_{V|V \geq v_j}$.*[20]

Usefulness of this truncated distribution and informativeness of the bounds depends on the auction model and the empirical application. One can then adopt the identification at infinity strategy from Song (2004) by taking the truncation point to the lower bound of its support to deliver bounds on the untrucated distribution.

4.5 Nonparametric estimation method

Bounds on the valuation distribution are asymptotically revealed by relevant *bid* order statistics, which provide observable bounds on ordered valuations. To emphasize this, estimated bounds on F_V are denoted in terms of the distribution of ordered bids. $G_{B_{i:b}}$ denotes the distribution of the $(b - i + 1)$th highest out of b bids. Smooth nonparametric estimates are obtained using Kernels, such as for the upper bound on F_V that is estimated from the distribution of lowest bids:

$$\hat{F}_V^U(v) = \hat{G}_{B_{1:b}}(v) = \frac{1}{T} \sum_{t=1}^{T} K\left(\frac{B_{1:b}^t - v}{h_T}\right), \tag{4.14}$$

$\forall v \in [\underline{v}, \bar{v}]$. $K(u)$ is the integral of the familiar Parzen-Rosenblatt kernel density estimator: $K(u) = \int_{-\infty}^{u} k(t)dt$ with $k(.)$ the Epanechnikov Kernel function. Under the conditions that $G_{B_{1:b}}(v)$ is continuous (on its support) and $h_T \to 0$ as $T \to \infty$, Nadaraya (1964) establishes uniform convergence of this estimator.[21] Cross-validated least-squares bandwidths are selected optimally to minimize integrated mean squared error. As for the empirical

[19]While Haile and Tamer (2003) left the question of sharpness of their bounds undetermined, Chesher and Rosen (2015) prove that together with the marginal distributions of final bids its joint distribution exhausts the available information, and together they provide sharp bounds on the distribution of interest. In this context also the joint CDF of the lowest $V_{i:n}$, ..., $V_{k:n}$ conditional on $V_{j:n}$ (for $i < k < j$) is identical to the joint CDF of $V_{i:j-1}$, ..., $V_{k:j-1}$ from a sample truncated from above at $v_{j:n}$. Hence I expect that bounds on the truncated distribution based on bounds on the joint distribution of two or more ordered valuations conditional on $V_{j:n}$ leads to a weakly sharpening of the bounds obtained from their marginal (conditional) distributions. However, this is beyond the scope of the paper.

[20]Equivalently, for procurement auctions, knowing bounds on at least one of two ordered valuations $V_{i:n}$ and $V_{j:n}$, with $i = k$ and $j = l$ $(1 \leq k < l \leq n)$ nonparametrically identifies bounds on $F_{V|V \leq v_j}$.

[21]Unlike with kernel density estimation, $Th_T \to \infty$ is not a necessary condition (see Chacón and Rodríguez-Casal (2010) for a discussion)

application in this paper, this estimation strategy can easily be applied to estimate bounds conditional on auction observables. Data-driven optimal bandwidth selection extends to conditional CDF estimation using the method in Li et al. (2013), which can incorporate multi-dimensional, mixed categorical/discrete and continuous conditioning variables and automatically 'smooths out' irrelevant covariates.[22]

For the lower bound on F_V I use the strategy set out in Section 4.3 and use the weakest shape restriction that f_V satisfies log-concavity. As shown, a lower bound on the distribution of $V_{n:n}$ that also bounds the distribution of F_V from below is defined as the distribution of the following bid statistics:

$$B_{b:b} + \Delta + 2(B_{b:b} + \Delta - B_x), \tag{4.15}$$

which I denote by B_X in the estimator below. Estimation of this bound is done for all auctions in which a lower bound on the third-highest order statistic of valuations is discernable, which requires at a minimum that the number of bids exceeds three. Let $\mathbb{I}\{B_x^t \neq \emptyset\}$ indicate this event and $T_x \equiv \sum_{t=1}^{T} \mathbb{I}\{B_x^t \neq \emptyset\}$ the number of auction that satisfy the requirement. As for the upper bound, I use the Kernel estimator with cross-validated bandwidth:

$$\hat{F}_V^L(v) = \hat{F}_{V_{n:n}}^L(v) = \hat{G}_{B_X}(v) = \frac{1}{T_x} \sum_{t=1}^{T_x} K\left(\frac{B_X^t - v}{h_{T_x}}\right), \tag{4.16}$$

$\forall v \in [\underline{v}, \bar{v}]$. For the tighter bounds in Equation (4.13), the above is repeated separately for each number of bids placed after which point-wise minima and maxima are taken. The baseline simulations have a sample of $T = 1000$. Conditioning on realizations of b exacerbates the problem of small sample bias. Number of bid-specific samples are only included as long as they have at least 50 auctions. Since this excludes some auctions, the tighter bounds are calculated as the minima and maxima over the estimated b-dependent bounds *and* the unconditional bounds to utilize the data best.

I also estimate the bounds on the consumer surplus CS^t presented in Equation (4.10). Bounds on $f_{V_{n:n}}$ are estimated using Kernel density estimates with optimal cross-validated bandwidths. For the density used to calculate respectively the lower and upper bound on

[22] The "np" package in R is well equipped for nonparametric Kernel estimation and bandwidth selection and it is used for estimation of all ordered bid distributions in this paper. See Hayfield and Racine (2008) for documentation.

CS^t:

$$\hat{f_{V_{n:n}}}^U(x) = \frac{1}{Th_T} \sum_{t=1}^{T} k\left(\frac{B^t_{b:b} - x}{h_T}\right) \tag{4.17}$$

$$\hat{f_{V_{n:n}}}^L(x) = \frac{1}{Th_{T_x}} \sum_{t=1}^{T_x} k\left(\frac{B^t_X - x}{h_{T_x}}\right) \tag{4.18}$$

$\forall x \in [\underline{v}, \bar{v}]$. Here, $k(.)$ is the Epanechnikov Kernel function. Uniform consistency requires: $h_T \to 0$ and $Th_T \to \infty$ as $T \to \infty$ and for the lower bound that $h_{T_x} \to 0$ and $T_x h_{T_x} \to \infty$ as $T_x \to \infty$ (Nadaraya (1965), see also Härdle and Linton (1994) for details of the non-parametric estimation methods applied here). A lower (upper) bound on CS^t is calculated by evaluating this density for auction t on the grid from $B^t_{b:b}$ (B^t_X) to $\bar{v}$ and subtracting p_t that is observed. Essentially, the value of the integral is approximated by summing the gridpoints multiplied by their estimated probabilities that are based on the relevant bid order statistics. For the domain of integration $\bar{v}$ is set to the maximum $B_{b:b^t}$ (across all auctions) plus one standard deviation. The approximated densities are adjusted to sum to one.

The Appendix presents results from Monte Carlo simulations to provide some intuition behind the applicability of the bounds and the circumstances that lead them to be tighter and more informative in practice.

4.6 Application of methodology to wine auctions at Sotheby's

In this section the proposed identification and estimation strategies are applied to a unique dataset of fine wine auctions. As the number of bidders in this data is unknown, the method results in informative bounds on otherwise unobtainable model primitives of interest.

4.6.1 Data description

The unique dataset covers the 884 lots from the "Finest and Rarest Wines & Vintage Port" auction, held on November 19th 2014 at Sotheby's London. I collected the data by simply registering as an online bidder, recording the complete auction, and translating the video material into a dataset of bids.[23] The dataset contains all bids, which bids are

[23] The paper took a different turn after I collected the data. I went to the auction room to count the number of bidders that I needed for structural analysis of the bid data, only to find that there were just two others in that room although there was plenty of activity on the phone, online, and in the auctioneer's

Table 4.1: Descriptive statistics of Sotheby's fine wine auctions

Statistic	N	Mean	St. Dev.	Min	Median	Max
Opening bid (£)	716	1,145.68	1,797.87	50	670	30,000
Highest bid (£)	716	1,375.67	2,161.06	90	750	32,000
Increment (£)	716	77.85	149.78	10	30	2,000
Number of bids (b)	716	4.30	3.01	2	3	25
Is lot sold? (1=YES)	716	0.92	0.28	0	1	1
Number bottles per lot	716	8.28	4.79	1	7	36
Sotheby's low estimate (£)	716	1,220.64	1,787.55	80	680	26,000
Sotheby's high estimate (% higher than low estimate)	716	27.42	6.62	9.09	27.27	62.50
Vintage	672	1,998.91	10.32	1,929	2,001	2,012

The unit of observation is a single lot. The vintage is missing for Non-Vintage champagnes and for mixed lots of various vintages.

from absentee bidders, and further auction descriptives. A key descriptive of interest is the value bracket estimated by the Sotheby's wine department and that is provided for all auctions in the auction catalogue.

Of the 884 lots, 766 auctions have at least two bids and in 716 of those an absentee drop-out value can be identified. No other auctions are excluded from the sample. Table 4.1 contains descriptive statistics of the data used for estimation, with as unit of observation a single lot. For some lots the vintage is missing as the auction is either a mixed lot with multiple vintages or (a) non-vintage champagne(s).

Conforming to standard practice in the structural estimation of auctions, these 716 auctions are treated as independent of each other. This abstracts from a range of interesting issues. Some groups of lots could be described as sequential (multi-unit) auctions as they make up a 'parcel' of identical or almost identical lots. Lots in these parcels contain exactly the same wine (and vintage) but may still differ in quality[24], bottle format, or number of bottles. The winning bidder in any of the lots in the parcel has the liberty to also purchase subsequent lots for the same price. In 55 lots of the wine auction sample a bidder had this option, and in 43 of these cases it is exercised.[25]

"book" with absentee bids. Unfortunately, the wine department declined my request for access to statistics about the number of bidders.

[24]For instance, having lower ullage levels, damaged labels or otherwise unfavorable package conditions are related to potential oxidation.

[25]Related to these parcels, the empirical literature has noticed what is termed a "decreasing price anomaly" in wine (and art) auctions: identical lots that are auctioned later generally fetching a lower price. Various explanations have been provided for this (in (Ashenfelter (1989), McAfee and Vincent (1993), Beggs and Graddy (1997), Ginsburgh (1998), Ashenfelter and Graddy (2003, 2006)): lots are systematically ordered in order of decreasing quality, risk-averse bidders add a risk premium for earlier lots to cover the risk that a winning bidder exercises his option to buy subsequent lots, or absentee bidders bid sub optimally. In this analysis I abstract from any dependencies between auctions.

The auctions have a secret reserve price, which is commonplace in auctions of wine and art (Ashenfelter (1989)), as well as in online auctions (Bajari and Hortaçsu (2003)). It is known that the reserve price will be at most the value of Sotheby's low estimate for the lot (Sotheby's (2014)). Ashenfelter (1989) suggests that a secret reserve may function to weaken bidding rings.[26] As a result of not meeting the reserve, 8% of the lots in the sample remain unsold. Since this is only announced after all bids are placed, these lots are kept in the sample to estimate bounds on the joint distribution of bidders' valuations.

4.6.2 Testing independent private values

The presence of absentee bidding requires bidders to be willing to announce their bids before others do, which is a strong indication that they are not afraid of a 'winners curse' and that the assumption of private values is justified. Correlated values are allowed, but it is reasonable to expect that correlation in valuations is completely captured by auction observables, as the specialists at Sotheby's Wine Department are unlikely to be outperformed by potential buyers in establishing the current value of the wines.

I test empirically whether the IPV is a reasonable assumption or whether additional correlation cannot be ruled out. In the IPV case the conditional distribution of valuations implied by the distribution of the $(i:n)$th bid order statistic and implied by the $(i':n)$th bid order statistic should be identical $i \neq i'$ (Athey and Haile (2002)). This is also verifiable without knowing the number of bidders if three adjacent order statistics of valuations would be point-identified so that the identification strategy of Song (2004) based on conditional densities could be applied twice and two implied conditional distributions can be compared, as also exploited in il Kim and Lee (2014).[27]

Here, the conditional distributions obtained from the lowest and second-highest bid are compared, ignoring bidding increments. More specifically, given that my model is not point-identifying, an equilibrium selection assumption is needed. For the purpose of this test I rely on the assumption that the second-highest bid equals the second-highest valuation and the lowest bid is equal to the lowest valuation. Under that assumption, estimated distributions should be indistinguishable if there is no common value component. Table

[26] Eklof and Lunander (2003) finds that for English auctions of real estate seller revenue is higher with publicly announced reserve prices than with secret ones. This last finding may however be related to real estate auctions decreasing in value when auctioned multiple times. In particular, Adams (2007) (with ascending eBay auctions of Corvettes) finds that secret reserve prices are generally set for goods that lend themselves for being up for auction multiple times.

[27] In an auction setup comparable to the wine auction, they claim that the second third and fourth highest observed bids are truthful. This requires assuming that bidders are restricted to bid once and only once, in which case it is a weakly dominant strategy to bid one's valuation. I am not willing to impose that here.

Table 4.2: P-values Kolmogorov-Smirnov tests for IPV

	2	4	6	8	10	12	14	16	18	20
None (unconditional)	0.47	0	0	0	0	0	0	0	0	0
Low estimate	1	0.99	0.81	0.70	0.70	0.70	0.58	0.58	0.58	0.58
+ Nr of bottles	1	0.99	0.70	0.58	0.47	0.47	0.47	0.47	0.47	0.47
+ Vintage	1	0.99	0.81	0.58	0.58	0.58	0.58	0.58	0.58	0.58

P-values test equality unconditional distributions (first row) and conditional distributions with different covariates. When conditioning on covariates, p-values are reported from tests of the equality of the distributions at the median of covariates. Test are repeated assuming a different number of bidders in the data, results are reported in different columns.

Table 4.3: Estimated $\mathbb{E}[CS]$ by tertile of Z (averages, £ per lot)

	1st	2nd	3rd	overall
N	244	235	237	716
Highest Bid	310	794	3050	1376
Low bound E[CS]	181	179	2018	1248
95%CI Low bound	[178 , 192]	[160 , 185]	[2039 , 2116]	[872 , 1415]
Up bound E[CS] (shape)	170	330	3446	1637
95% CI Up bound (shape)	[170 , 179]	[318 , 347]	[3434 , 3560]	[1248 , 2001]

Wine auctions consumer surplus statistics using estimated valuation distributions conditioning on tertiles of Z = Sotheby's low value estimate. "N" is the number of auctions in each Z-bracket and the last column "Overall" simply takes the average of the estimated statistics in all 716 auctions. For the other columns, the mean statistic in all N auctions in that Z-bracket are reported. The estimates are in (£) per lot, which contains on average about 8 bottles. The confidence intervals are calculated from 500 bootstrap samples.

4.2 reports p-values from different nonparametric Kolmogorov-Smirnov tests to formally evaluate the IPV assumption (in fact, jointly with the equilibrium assumption). p-values indicate the probability that the two estimated distributions are as different as they are while being sampled from the same parent distribution. The number of bidders n is held fixed at different values in the arbitrary range $n \in \{2, 4, 6, 8, 10, 12, 14, 16, 18, 20\}$. Results shows that with the exception of $n = 2$, without conditioning on covariates IPV is rejected with certainty. This is not surprising given that the provided value bracket contains significant information from an expert assessment of the wine quality. Results show that it is indeed much more likely that valuations are independent *conditional* on Sotheby's low estimate. Adding as additional covariates the number of bottles in the lot and the vintage of the wine does not improve the p-values.

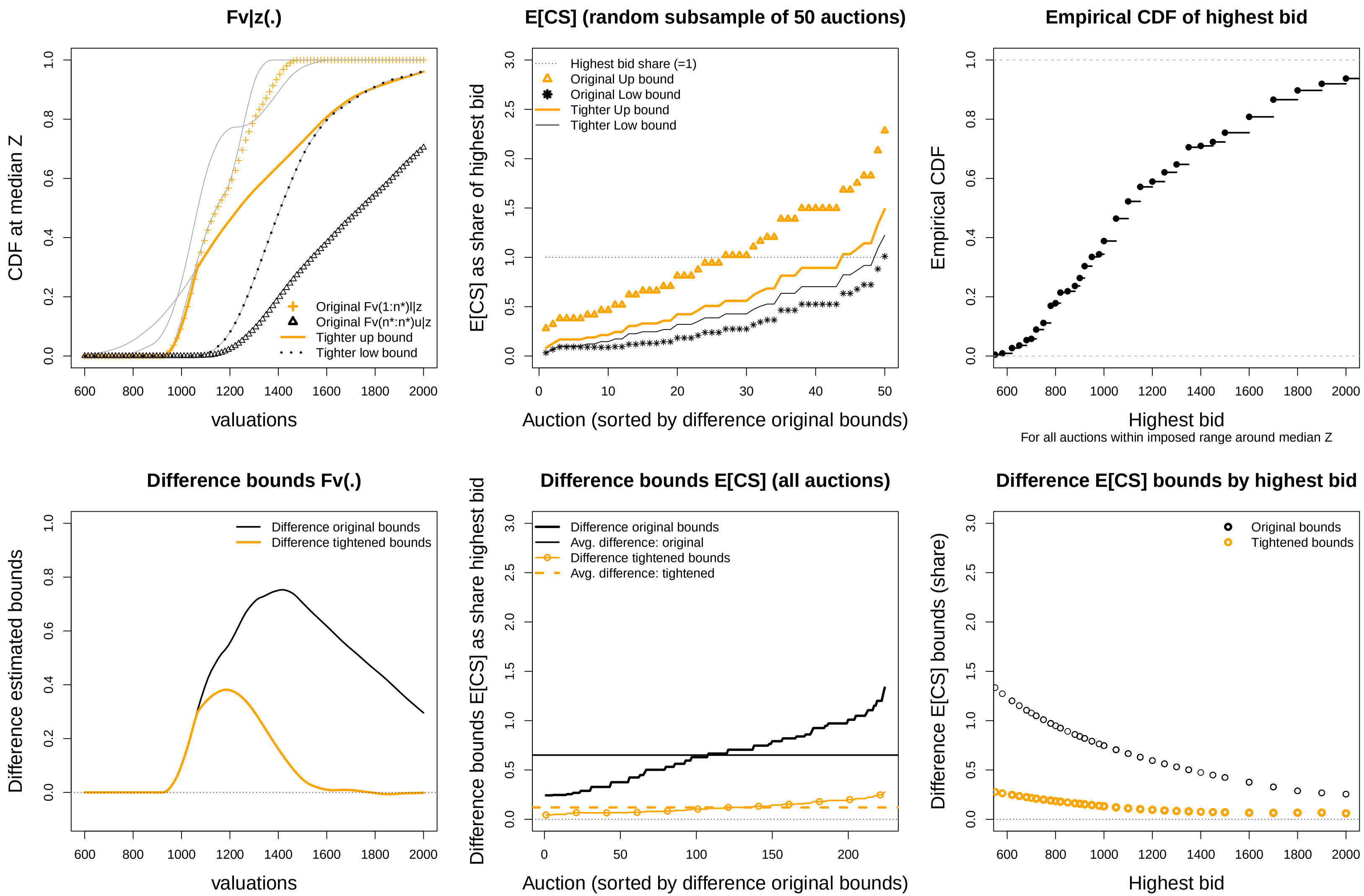


Figure 4.1: Estimation results: comparing original bounds with tighter ones based on CINB

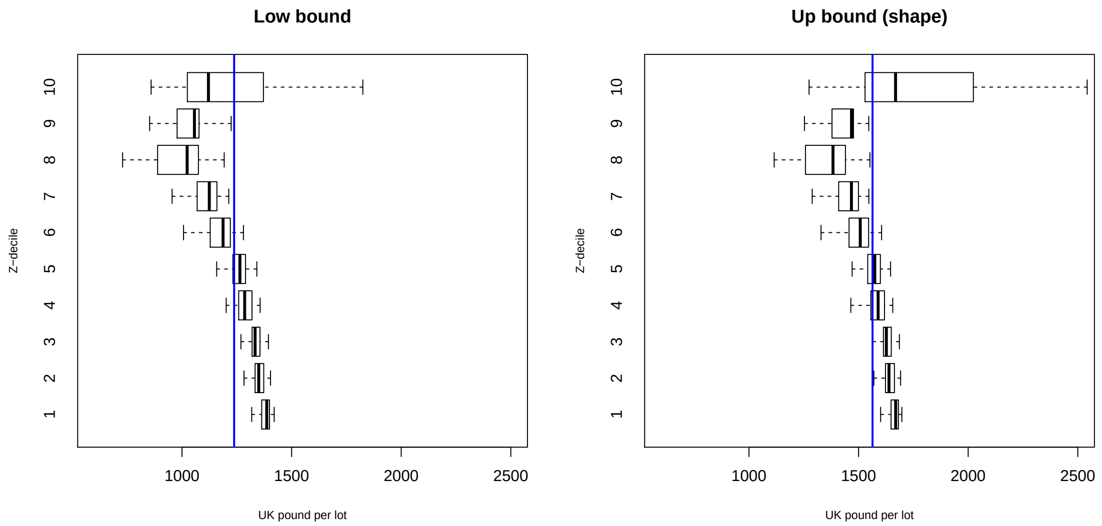


Figure 4.2: Boxplots: Estimated bounds on E[CS] wine auctions by Z-decile

Z includes Sotheby's pre-auction low estimate of the value. The vertical lines indicate the sample medians.

4.6.3 Estimation results

Conditioning on $Z =$ Sotheby's low value estimate only, sample bounds on the joint valuation distribution are estimated for the median of $Z = 1221$ pounds. As Z is continuous, conditioning on it means selecting a subset of auctions with values not too far away from its median. There is a trade-off between small sample bias and bias from estimating bounds on the joint distribution from different realizations of Z. To include enough auctions to also tighten the bounds conditional on the total number of bids, auctions are selected if the absolute difference between the selected Z and the sample realization is at most 600 pounds, which arbitrarily selects 233 out of 716 auctions for the median Z. Due to the even smaller sample size, for the tightened bounds the minimum number of auctions with a given number of total bids is set to 25 (instead of 50 in MC simulations in the Appendix). Results are displayed in Figure 4.1, also for the estimated bounds on $\mathbb{E}[CS]$. It shows that despite the relaxed information requirements data can still be informative for structural features of interest. The bounds obtained under CINB, the exclusion restriction that assumes that conditional on other observables the number of bids is independent of valuations, are significantly better than the bounds obtained using only log-concavity and sample spacings.

More details for the estimated consumer surplus are provided Table 4.3. Results highlight how informative the estimated bounds are, especially in relation to the average highest bid. Auctions are tabulated for tertiles of Z and related auctions are selected in a bandwidth of 500 pounds around those values. The (point estimate of the) lower bound on expected consumer surplus is equal to 91 percent of the average highest bid and the upper bound is not far removed from it at 119 percent of the average highest bid. Figure 4.2 shows that especially for auctions with the highest pre-auction value estimate the variability in expected surplus is the largest. Another pattern is that for mid-value auctions estimated surplus tends to be lower than for auctions with the lowest pre-auction value estimates.

Bootstrapped confidence intervals are also reported in Table 4.3, pointing to considerable small sample variation. They are obtained as follows. For both the complete dataset and the subsets by tertile of Z, 500 bootstrap samples are drawn with replacement. In all of these, the bounds on the expected surplus are estimated as before and their lowest 2.5th and highest 97.5th quantiles are reported as CI. Obtaining an analytical expression of the limiting distribution of the expected consumer surplus based on bounds that involve kernel density estimation and constructing an interval that satisfies a 95% coverage probability is beyond the scope of this draft. In the Appendix I discuss preliminary results and related literature on inference of bounds.

4.7 Conclusions

This paper explicitly models absentee bidding that is common in traditional English auctions, and characterizes how order statistics of bids relate to order statistics of valuations. It is shown that these bid order statistics identify bounds on the letent distribution of valuations without needing to know the number of bidders that is often obscured in English auctions. The identification strategy is novel in that it relies on *sample extremes* to bound the joint distribution of valuations, which not only is a solution to not knowing the number of bidders but also applies when conditional valuations are not independent.

Censoring of the highest valuation is a central problem in structural econometrics of ascending auctions, since the auction stops when the bidder with the second-highest valuation drops out. To identify a lower bound on the distribution, an identifying shape assumption is introduced. I propose three different shape restrictions of which log-concavity is the weakest and often already imposed in the auction literature. I also discuss a concavity and a symmetry restriction and how additional bid order statistics can be used to test the shape restriction imposed. Thinking about the implications of the spacings of order statistics may be a useful direction for future research involving structural analysis of auction markets. It is suggested to combine the method of spacings and sample extremes with identifying variation in the data to sharpen the bounds, such as observable variation in the number of bids or auction covariates that are excluded from valuations. I show that the identified bounds on the joint valuation distribution deliver informative bounds on the expected consumer surplus. Performance of the bounds is assessed in Monte Carlo simulations, and I propose a nonparametric estimation method based on Kernels.

To show the value of the proposed identification strategy I apply it to a unique dataset of fine wine auctions in which the number of bidders is unknown. I examine expected consumer surplus. Results include: i) auctions with the highest pre-auction value estimates have the largest variation in expected consumer surplus, ii) consumer surplus is generally lower in auctions with mid-range pre-auction value estimates than for lower pre-auction value estimates, and iii) on average bidders have an estimated consumer surplus of about 90-119 percent of the highest bid.

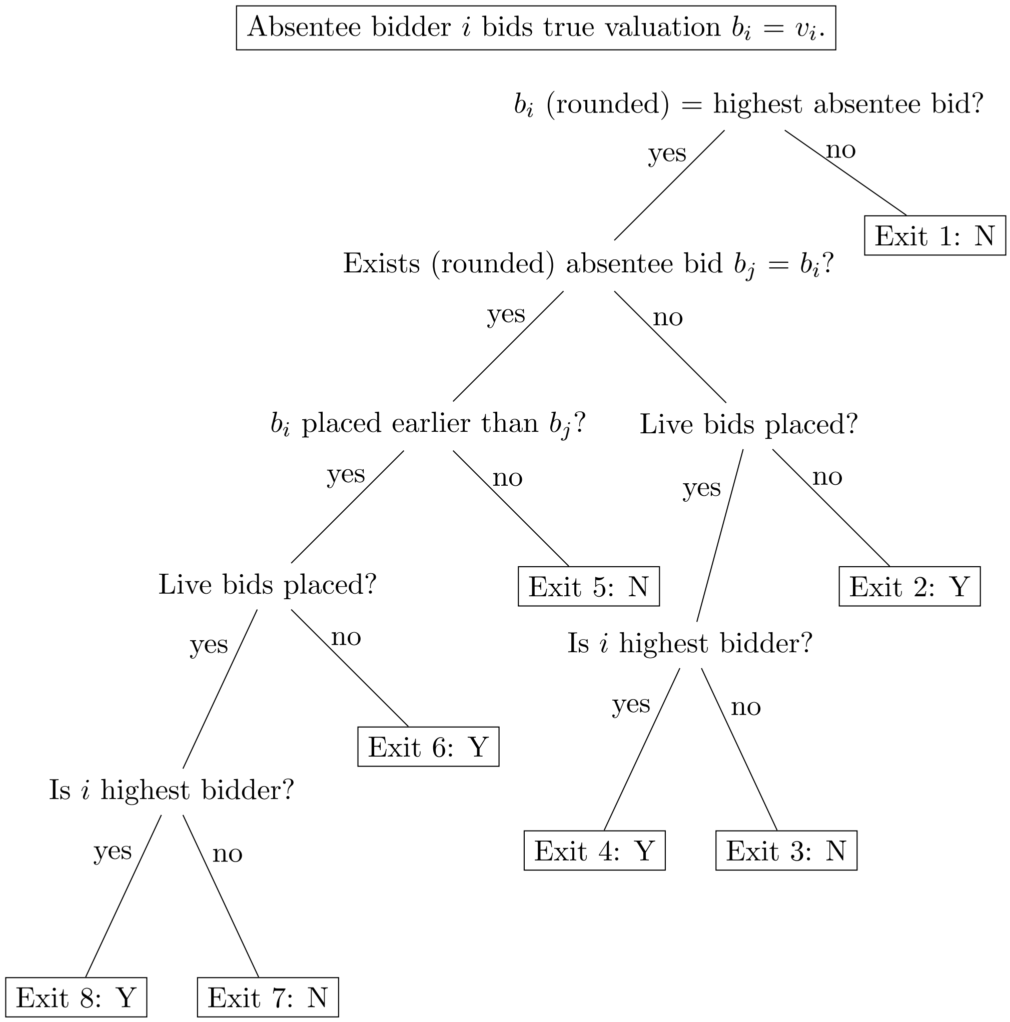
Acknowledgements

I am indebted to Adam Rosen and Andrew Chesher who provided invaluable guidance and support. I also want to thank Phil Haile, Ali Hortaçsu, Suehyun Kwon, Lars Nesheim, Áureo de Paula and conference participants at PET and EEA/ES for useful comments. Thanks to Sebastian Fahey for going over some practicalities of auctions at Sotheby's with me and to Alan Crawford for introducing us. I am also grateful for financial support from the Economic and Social Research Council (PhD Studentship). All remaining errors are my own.

Appendices

H Further details absentee bidding

Figure 4.3: Partial outcome tree absentee bidders



At exit points, the game finishes for the absentee bidder: exiting with an "N" means that he failed to win the lot and exiting with an "Y" means that he placed the highest bid. Clearly, this bidder doesn't regret not bidding less than his valuation at Exits 1, 3 and 7 because he is overbid. Exits 2 and 4 reflect a second-price auction where he doesn't regret not bidding less than his valuation, as that would mean potentially losing the auction without paying less. At Exit 5 he simply placed his bid too late. Exits 6 and 8 reflect a first-price auction where the absentee bidder doesn't regret not bidding less because he would for sure have lost the auction to the other absentee bidder with the identical bid.

Obtaining valuations from bids

Identification in English auctions relies on pinning down valuations from bid data, for auctions consistent with the model. Motivated by online, phone, and absentee bidding obscuring the number of bidders and their identities to everyone but the auction platform, the following is assumed observable:

Information Requirement. The econometrician observes: 1) the vector of bids, and 2) which bids are from absentee bidders. Not needing to observe highest bids or the number of bidders makes obtaining suitable auction data much easier and less time-consuming.

Lemma 6 established that the maximum amount that absentee bidders report to the auctioneer equals their valuation. The sequence of observed bids could however also include lower values because the auctioneer is to obtain the lot for the absentee bidder at the lowest possible price. Knowing which bids come from absentee bidders and following the modeled bidding sequence, it is possible to determine the lowest bid that is certainly truthful: it is the opening bid when there are two or more absentee bids and otherwise the value where the absentee bidder drops out if that is not the highest bid. Example 3 provides an example of how this lowest truthful commissioned bid can be identified from bidding data.

Example 3.
A_i and L_j are respectively the ith lowest overall bid, which came from an absentee bidder, and the jth lowest overall bid in the bid vector, which came from a live bidder. So the first and lowest (commissioned) bid is A_1 if there is an absentee bidder, and the other bids follow according to the assumed bidding sequence.
Auction 1: A_1, A_2, A_3, L_4, L_5
Auction 2: A_1, L_2, A_3, L_4, L_5
In Auction 1 there must be more than one absentee bidder as there are multiple commissioned bids placed before live bidding starts: the opening bid is here the lowest truthful commissioned bid. In Auction 2, live bidding starts immediately after the first bid, indicating that there is just one absentee bidder. Here, the lowest truthful commissioned bid is the value where the absentee bidder drops out: A_3. But in auctions with fixed bidding increments one has to account for "footing": where someone's highest bid is below his valuation due to the bidding sequence and increments. The absentee bidder in Auction 2 could for example have had a valuation within two increments above A_3.

I Monte Carlo simulations

This section provides some intuition about how the identified bounds vary with auction characteristics in sets of auctions simulated conform the theoretical model.

Potential bidders draw independent valuations from a common distribution, and has a fixed number of absentee and live bidders. The bidding strategy of potential absentee bidders is to submit a bid equal to their valuation if this is at least 80% of the population mean value.[28] Simulations follow the assumed bidding sequence and the algorithm records which bids are from absentee bidders. In particular, with multiple absentee bids, the bidding sequence at the lowest absentee value and places bids according to fixed increments until at most one absentee bidder has a value exceeding that bid. Bids then alternate between the last remaining absentee bidder, if there is one, and live bidders. If there are multiple live bidders with a valuation exceeding the current bid (who did not placed the current bid), the algorithm randomizes between those bidders.[29] For auctions that received at most one absentee bid, the opening bid is 80% of the mean population value (to account for a secret reserve), while the rest of the bidding process is identical.

The resulting bid vector resembles what can be directly observed from attending or recording traditional English auctions. To emphasize its limited information, highest bids cannot be discerned from the vector of bids when identities are not known and live bidders may end up not bidding when the bidding sequence passes their valuation before they get the chance to bid. The information content diverges from what is assumed known in previous studies, including Paarsch (1997) using all bidders' drop-out values and the number of bidders, Haile and Tamer (2003) and Chesher and Rosen (2015, 2017) using a vector of highest bids and the number of bidders, Song (2004) using a vector of bids that includes (for some auctions, depending on the timing of bids) the second and third-highest drop-out values, and Aradillas-López et al. (2013) using the second highest drop-out value and number of bidders. These papers are all ground breaking in their econometric use of bid data from English auctions. I list them here only to highlight how my setting departs from theirs in the assumed information available for the analysis.

Distribution functions, number of bidders, number of auctions, and bidding increments vary in the simulations to highlight how these features of the data generating process affect tightness of the identified bounds in practice. There are no covariates in the simulations.

[28]This arbitrary 80% rule can be interpreted as potential bidders accounting for the secret reserve price, which is present in the application as well as in many other traditional English auctions. This provision is solely made for the simulations to generate some variation in the number of absentee bids placed without varying the number of bidders and it can be relaxed without consequences.

[29]I stress this to make the link with real life English auctions apparent, but since bidder identities are not used in the analysis this randomization is actually redundant here.

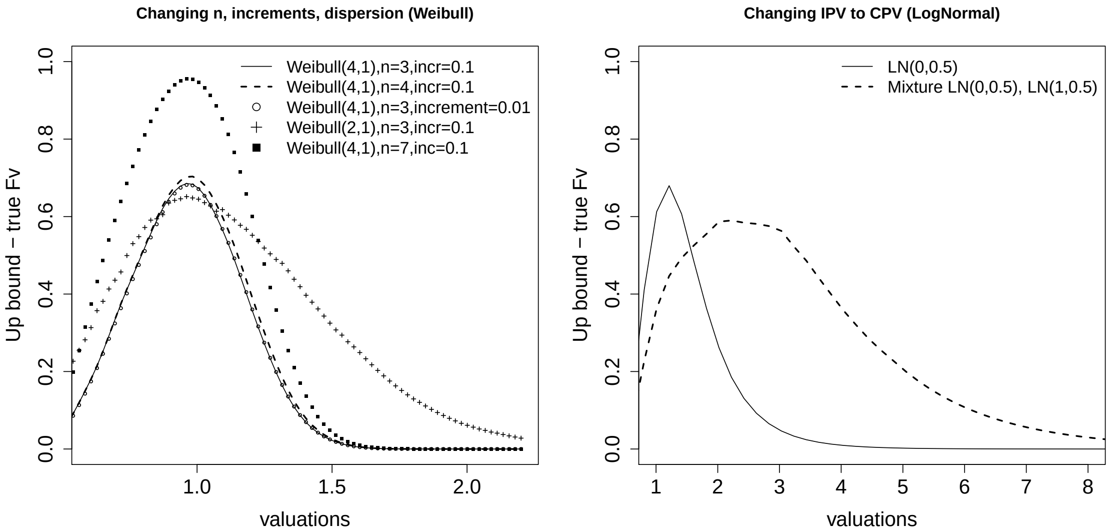


Figure 4.4: Simulation results: Tightness of upper bound

Difference between the estimated upper bound and F_V^*. The left-hand panel has IPV and shows the effect of having a higher dispersion, more potential bidders, smaller bidding increments, and estimation on a subsample with smaller n. The right-hand panel shows validity of the bounds for correlated private values, simulated as being mixture of a LogNormal(0,0.5) and Lognormal(1,0.5) with equal probability and also compared to the tightness of the bound when valuations are LogNormal(0,0.5) for all auctions.

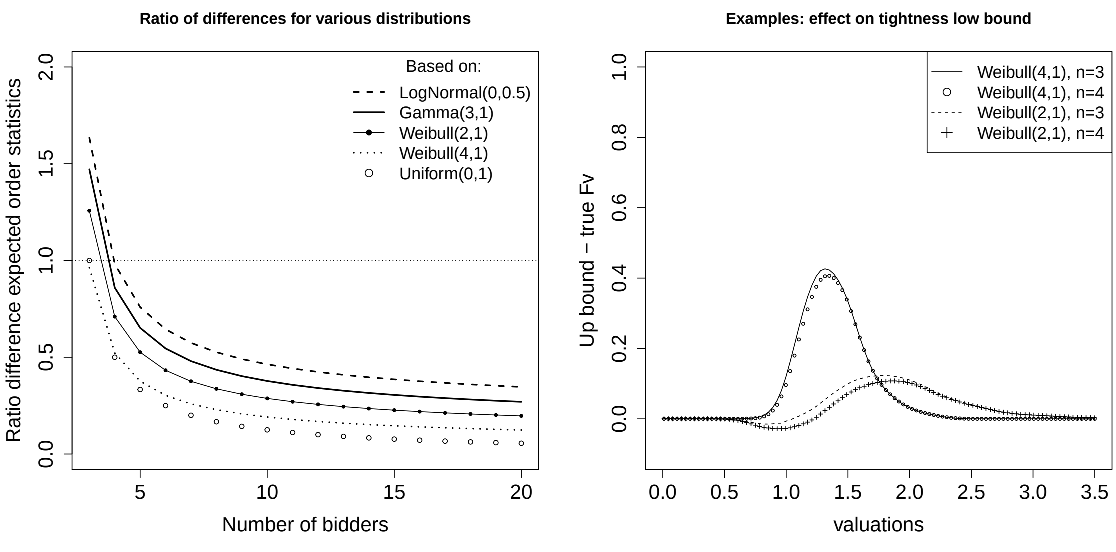


Figure 4.5: Simulation results: Shape assumption & tightness lower bound

Left-hand figure: Examples of distribution functions and the ratio of the expected spacing between the second highest and highest order statistic to the expected spacing between the lowest and second-highest order statistic ($\mathbb{E}[D(n-1,n)]/\mathbb{E}[D(1,n-1)]$) as the number of bidders increases, assuming IPV. Right-hand figure: examples of how this translates into tightness of the lower bound on F_V, with IPV.

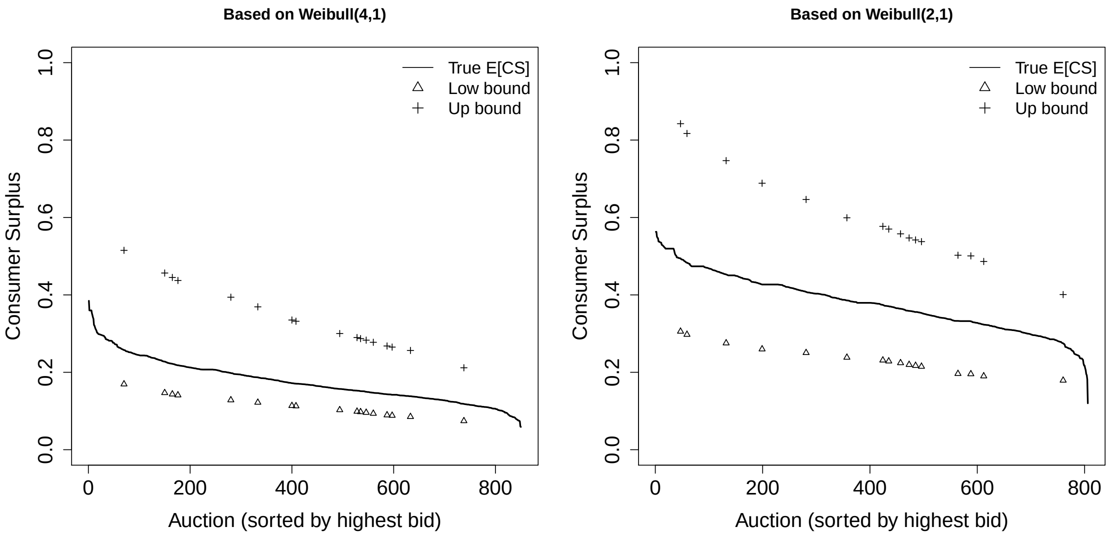


Figure 4.6: Simulation results: Bounds on E[CS]

Estimated upper and lower bounds on $\mathbb{E}[CS]$ based on two sets of simulated auctions. In the left-hand panel valuations are Weibull(4,1) and in the right-hand panel Weibull(2,1) with in both cases 2 potential live and 2 potential absentee bidders. In both cases IPV is assumed, and the displayed bounds are for random subsamples of 20 auctions.

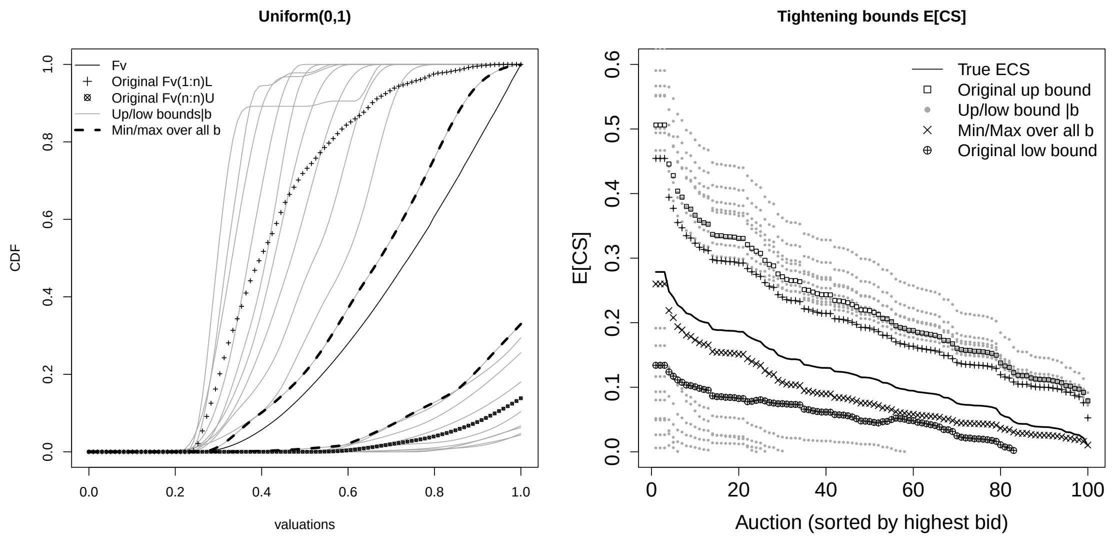


Figure 4.7: Simulation results: Tigher bounds on $\mathbb{E}[CS]$

Valuations are independently distributed U(0,1) and auctions are designed to have 4 potential live and 4 potential absentee bidders. For the the b-dependent bounds, the sample size (number of auctions) is restricted to be at least 50 as described in the text. All reported values are calculated on a 10% random subsample of all 1.000 auctions.

Figure 4.4 shows the distance between estimated upper bounds and F_V to address sensitivity to certain features of the data in empirical settings. Settings that result in tighter bounds include: i) a *relatively small* n, ii) *more correlated valuations*, iii) the valuation distribution having *a relatively low dispersion*, iv) *smaller bidding increments*, and v) *larger samples*. The left-hand panel shows results for auctions where valuations are independent draws from a Weibull(4,1) distribution with 2 potential live bidders and 2 potential absentee bidders. A higher dispersion makes the bound less tight, as exemplified by the results when valuations are Weibull(2,1) distributed. Changing the increment from 10% to 1% of the standing bid has a minimal effect, but the bound is visibly wider when estimated on the subsample with a higher n. The right-hand panel shows the validity of the upper bound for auctions with correlated private values. This is simulated as valuations being drawn from a mixture of a LogNormal(0,0.5) and LogNormal(1,0.5) with equal probability.

Figure 4.5 provides intuition for the shape assumption. The left-hand panel displays the ratio of the expected distance between the highest and second-highest valuation to the expected distance between the second-highest and lowest valuation, assuming independence. On the x-axis are the number of bidders. The shape assumption assumes that the distribution and number of bidders is such that this ratio is at most 1, which is at the horizontal dashed line. For symmetric distributions, with the number of bidders greater than 3, this is *always* satisfied. For other distributions this may be violated at low n, but as the graph shows, the ratio falls down sharply when n increases. The right-hand panel shows how this translates into the lower bound on F_V: the closer the ratio is to one without exceeding it, the sharper this bound is. For Weibull(2,1) with $n = 3$ the ratio is larger than one, which also makes that the resulting difference between the bound and F_V in the right-hand panel slightly dips below 0. This emphasizes that when the shape assumption is not satisfied, the lower bound on the joint distribution may not be valid.

Figure 4.6 shows how the bounds translate into bounds on $\mathbb{E}[CS]$. Valuations are either Weibull(4,1) in the left-hand panel or, with higher dispersion, Weibull(2,1) in the right-hand panel. Bounds are calculated for a random subsample of 20 auctions for speed. Although the distance between the truth and the bounds depends on the functional form of the valuation distribution, the upper bound that is a function of the relatively weak upper bound on the highest valuations, is generally wider than the low bound on $\mathbb{E}[CS]$.

Figure 4.7 shows that the assumption of b-independence can tighten the bounds on F_V significantly. In this example, 8 potential bidders are simulated to have independent valuations distributed as Uniform(0,1). The grey lines represent the estimated bounds conditional on the total bids b. As a tighter upper bound, the thick dashed line takes the pointwise minimum over all b-dependent estimated upper bounds and as a tighter lower bound it takes the pointwise maximum over all b-dependent estimated lower bounds, as described in the previous section. Using these tightened bounds on F_V the right-hand panel shows how this affects tightness of the bounds on $\mathbb{E}[CS]$. It supports that using an auction model with absentee bidding, allowing for correlated private values and without observing the number of bidders, informative bounds on the expected consumer surplus can be obtained with the proposed method.

J Inference

A growing literature studies inference for partially identified objects and models with moment conditions. Two main complicating issues with bound estimators are that *i*) closed form characterizations of their asymptotic distributions are often hard to obtain, and *ii*) substantial finite sample bias may cause bounds to cross. Standard bootstrapping or subsampling methods, applied as a solution to *i*), have been found to be generally inconsistent in models defined by moment inequalities but suitable alternative implementation has been proposed (by e.g. Chernozhukov et al. (2007), Andrews and Han (2009), Bugni (2010), Canay (2010)). Rosen (2008) constructs confidence intervals based on an inverted test statistic with an analytical asymptotic distribution, which does not require inferential methods but can be asymptotically conservative.

Problem *ii*) as noted by Manski and Pepper (2000, 2009) could cause estimated bounds to be much tighter than population bounds. Haile and Tamer (2003) address this issue by introducing a smoothing parameter that vanishes as the sample grows to make sure the estimated upper bound of the valuation distribution is always above the lower bound. As they remark, this problem is due to their bounds being infima and suprema of estimated objects, which is the case in many moment condition models. Chernozhukov et al. (2009) suggest an asymptotically valid inference method particularly for such models with intersection bounds. This problem concerns our tightened bounds, too, and a valid inference approach still needs to be implemented.

Inference on the population distribution is less complicated for the original (untightened) bounds in this paper as they do not involve extremum opterators and an analytical characterization of their limiting distributions can be obtained. The method proposed in Imbens and Manski (2004) is suitable for such settings and results in a confidence interval that asymptotically covers the true population parameter (instead of covering both bounds with a fixed probability as in Horowitz and Manski (2000)).

To simplify the exposition, let $F^U(v)$ and $F^L(v)$ represent the population upper and lower bounds on the object of interest that is denoted by $F(v)$ as derived in Section 4.4 and let $\hat{F}^U(v)$ and $\hat{F}^L(v)$ be the finite sample kernel estimates of these bounds. Let T denote the relevant sample sizes and h the selected bandwidth. Asymptotic properties of the estimated bounds are (for $B = \{U, L\}$):

$$(i)\sqrt{T}(\hat{F}^B(v) - F^B(v)) \xrightarrow{d} \mathcal{N}(0, \Sigma_B(v))$$
$$(ii) \text{ with probability one } \sup_{\underline{v}<v<\bar{v}} |\hat{F}^B(v) - F^B(v)| \to 0 \text{ as } T \to \infty \text{ and } h \to 0$$

Which respectively represent $\sqrt{T}$ consistency of the estimated bounds (e.g. Li and Racine (2007)) and uniform convergence with probability one (Nadaraya (1964)).[30]

Besides uniform convergence of the estimated bounds, the existence and uniform convergence of $\hat{\Sigma}_U, \hat{\Sigma}_L$, and boundedness of their population counterparts, Imbens and Manski (2004) implictly assume that $\hat{\Delta}$ is locally superefficient as noted by Stoye (2009). He provides a weaker assumption, for which a sufficient condition is that with probability one the estimated upper bound is at most the estimated lower bound. This is generally true when the bounds are ordered by construction, as they are in this paper due to their one-to-one mapping to bid order statistics. In other words, Δ is bounded from below by 0. So the Imbens and Manski (2004) confidence intervals are valid here, with the remark that they are based on the weaker notion of superefficiency as in Stoye (2009).

Having established these asymptotic properties of the estimated bounds, the method obtains a CI that covers population $F(v)$ with at least probability $(1-\alpha)$. Let $\hat{\Delta}(v) = \hat{F}^U(v) - \hat{F}^L(v)$, $\hat{\sigma}^n_B(v)$ a consistent estimate of the sample standard deviation for bound $B = \{U, L\}$ and Φ the standard normal CDF.

$$CI_\alpha^{F(v)} = \left[\hat{F}^L(v) - C_\alpha(v)\frac{\hat{\sigma}^n_L(v)}{\sqrt{T}}; \hat{F}^U(v) + C_\alpha(v)\frac{\hat{\sigma}^n_U(v)}{\sqrt{T}}\right] \tag{4.19}$$

Where $C_\alpha(v)$ is the solution to:

$$\Phi(C_\alpha(v) + \frac{\sqrt{T}\hat{\Delta}(v)}{\max(\hat{\sigma}^L(v), \hat{\sigma}^U(v))}) - \Phi(-C_\alpha(v)) = (1-\alpha) \tag{4.20}$$

Defining the identified set $\Lambda(v)$ as $\Lambda(v) = F'(v) \in [0,1] : \{F^L(v) \leq F'(v) \leq F^U(v)\}$,

[30] Note that unlike with kernel estimation of a PDF, no bias term appears in $(i), (ii)$ and convergence is faster. The limited sample variance equals: $\Sigma^n_B(v)) = \frac{1}{T}F^B(v)(1-F^B(v)) - \theta f(v)\frac{h}{T} + o(\frac{h}{T})$ for $B = \{U, L\}$ and $\theta = 2\int_\infty^\infty v'K(v')k(v')dv'$ (Li and Racine (2007)). Nadaraya (1964) established that the limiting variance of the kernel CDF estimate equals the population variance, e.g. $\Sigma_B(v) = \frac{1}{T}F^B(v)(1-F^B(v))$. I explicitly denote all relevant statistics as a function of (v) to emphasize that these are calculated at any point $v \in [\underline{v}, \bar{v}]$ and that they are not invariant over this range.

Imbens and Manski (2004) prove that $CI_\alpha^{F(v)}$ has a coverage rate that converges uniformly in $(F'(v), \Delta)$: $\lim_{T\to\infty} \inf_{F'(v)\in\Lambda,\Delta\in[0,1]} Pr(F(v) \in CI^{F(v)_\alpha} = (1-\alpha)$, which means that the probability that any $F'(v) \in CI_\alpha^{F(v)}$ is *at least* $(1-\alpha)$, which of course also holds for $F'(v) = F(v)$. The identified set adds a natural restriction on any candidate $F'(v)$ to lie within $[0,1]$ for it to be a proper CDF. In practice, the CI is restricted to lie within these bounds as well, but this doesn't affect the coverage probability because any $F'(v) \in \Lambda$ is certainly $\in [0,1]$ for any T - and thus also in the limit.

The intuition behind $CI_\alpha^{F(v)}$ in Imbens and Manski (2004) is that $F(v)$ can be arbitrarily close to either of the bounds but it is unknown to which one. If it would be close to the upper bound the asymptotic probability that $F(v) < F^L(v)$ can be ignored so that the entire α probability of mischaracterizing the population CDF can be allocated to the upper bound, and the same goes in the opposite case for the lower bound. This means that one-sided CIs need to be constructed around both bounds. Consequentially, this method also applies to the population distribution if only one bound would be identified. For example, if one wouldn't want to impose the shape assumption to obtain the lower bound, and only $F^U(v)$ based on the lowest bid would be identified. Denoting the population CDF the is only bounded from above by $\tilde{F}(.)$ and also incorporating that $\tilde{F}(.) \in [0,1]$ as discussed above, the CI is obtained as:

$$CI_\alpha^{\tilde{F}(v)} = \left[0; \min(1, \hat{F}^U(v) + C_\alpha(v)\frac{\hat{\sigma}_U^n(v)}{\sqrt{T}})\right] \tag{4.21}$$

Where $C_\alpha(v)$ is the solution to:

$$\Phi(C_\alpha(v) + \frac{\sqrt{T}\hat{F}^U(v)}{\hat{\sigma}^U(v)} - \Phi(-C_\alpha(v)) = (1-\alpha) \tag{4.22}$$

Of course, a similar CI can be constructed when only an informative lower bound would be identified. For the tighter bounds, inference still needs to be addressed. In the application I use the bootstrap method to calculate confidence intervals for estimated consumer surplus directly. See Horowitz (2001) for details on this approach.

K Omitted proofs

Truthful revelation by absentee bidders

The auction constitutes a strategic game with incomplete information, described completely by the set of bidders $\mathcal{N}$ and $\forall i \in \mathcal{N}$: a set of types $V_i \in [\underline{v}, \bar{v}]$, strategy profiles S_i with $s_i : V_i \to S_i$ and payoff functions $u_i(s_i, s_{-i}, v_i)$. The auction mechanism and private values structure are common knowledge, but other bidders' valuations are unknown. The below proves that the unique symmetric Bayes Nash equilibrium in weakly undominated strategies is for absentee bidders to bid their valuation, relying on the modeled bidding sequence.

Lemma 1. *The unique symmetric Bayes Nash equilibrium in weakly undominated strategies is for absentee bidders to bid their valuation.*

Proof Lemma 1. Let $u_i^a = u(v_i, p(b_i^a, b_{-i}, r, \Delta))$ be the utility of absentee bidder i as a function of his valuation v_i and price p that he needs to pay upon winning, which is a function of his maximum bid b_i^a, the vector of other players' maximum bids b_{-i}, secret reserve price r, and bidding increment Δ. Since quasi-linear preferences come for free with risk-neutral players who are ex-ante symmetric and not budget constrained, $u_i^a = (v_i - p(b_i^a, b_{-i}, r, \Delta))\mathbb{I}\{b_i^a = B_{(b:b)}\}$, where vector B records all bids including commissioned bids, b is the number of total bids placed, and $\mathbb{I}\{b_i^a = B_{(b:b)}\} = 1$ if i's commissioned bid is the highest bid placed in the auction and 0 otherwise.

Vector C records all commissioned bids, and c denotes the number of commissioned bids. Conform the auction rules, two identical absentee bids are both recorded but one placed earlier is ordered as a higher bid. Hence, if $C_{(c:c)} = C_{(c-1:c)}$ the highest two commissioned bids have the same value but $C_{(c:c)}$ was placed earlier. Also, $b_i^a = B_{(b:b)}$ implies that $b_i^a = C_{(c:c)}$, but the reverse is not true if there are live bids higher than b_i^a.

Conditional on winning, the absentee bidder pays:

$$p(b_i^a, b_{-i}, r, \Delta)|(b_i^a = B_{(b:b)}) \begin{cases} \max\{B_{(b-1:b)} + \Delta, r\} & \text{if } C_{(c:c)} \neq C_{(c-1:c)}, b_i^a \geq r \\ B_{(b:b)} & \text{if } C_{(c:c)} = C_{(c-1:c)}, b_i^a \geq r \\ \text{no sale} & \text{if } b_i^a < r \end{cases}$$

From this it can be deducted that absentee bidder i has no incentives to bid less than his valuation. Doing so would reduce both the probability that his bid is the highest overall bid and that it beats the reserve price, potentially reducing his utility. This loss isn't offset by having to pay a lower price upon winning. In the first case there is no equally high commissioned bid, so the hammer price is $\max\{B_{(b-1:b)} + \Delta, r\}$ which is independent of b_i^a under the requirement that the secret reserve cannot be raised (illegally). If there is an identical commissioned bid placed later, the hammer price is $b_i^a = B_{(b:b)}$. Yet, also in this case the absentee bidder has no incentive to lower his bid because wouldn't win the auction then. Finally, bidding $b_i^a > v_i$ is not optimal either as it could result in negative utility.

Hence it is an undominated strategy for player i to bid truthfully, and this holds for all absentee bidders. Another Nash equilibrium where only the absentee bidder with the highest valuation bids truthfully and the others don't bid (or bid 0) is ruled out as it is weakly dominated by the symmetric equilibrium where everybody bids truthfully (but it would be allowed with a weaker notion of posterior implementability). There are no other equilibria. $\square$

Spacings

Lemma 2. *[Log-Concave density] Let $f_{V|\mathbf{Z}}$ be (weakly) log-concave. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}$ $\forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \leq P[V^U_{n-1:n} + 2(V^U_{n-1:n} - V^L_{n-2:n}) \leq v|\mathbf{Z} = z] \equiv F^{L(lc)}_{V_{n:n}|\mathbf{Z}}(v) \tag{4.23}$$

In the case that $f_{V|\mathbf{Z}}$ is weakly log-concave (exponential) and $\Delta = 0$ $F^{L(lc)}_{V_{n:n}|\mathbf{Z}}(v)$ is only an upper bound if the bidder with the third-highest valuation had no opportunity (or reason) to drop-out at his valuation. Otherwise, (4.23) *holds with equality in this case.*

Proof Lemma 2. Normalized spacings from an exponential distribution are independent and identically distributed, while normalized spacings based on a log-concave density function are stochastically decreasing (e.g. Pyke (1965, 1972)). By the definition of D_n, $F_{V_{(n:n)}}(x) = P[V_{(n:n)} \leq x] = P[(V_{(n-1:n)} + D_n) \leq x]$ $\forall x$. By the definition of normalized spacings $\tilde{D}_n = D_n$ and $\tilde{D}_{n-1} = 2D_{n-1}$, so $P[(V_{(n-1:n)}+D_n) \leq x] = P[(V_{(n-1:n)}+\tilde{D}_n) \leq x]$. Due to $\tilde{D}_n$ being weakly stochastically smaller than $\tilde{D}_{n-1}$: $P[(V_{(n-1:n)} + \tilde{D}_n) \leq x] \geq P[(V_{(n-1:n)} + \tilde{D}_{n-1}) \leq x]$, $\forall x$ (holding with equality when valuations are drawn from the exponential distribution). This implies that $F_{V_{(n:n)}}(.) \geq F_{V_{(n-1:n)}+\tilde{D}_{n-1}}(.)$.

When $n = 3$, observing $V_{(n-1:n)}$ and $V_{(1:n)}$ identifies the normalized spacing $\tilde{D}_2$. Since by log-concavity $F_{\tilde{D}_3}(x) \geq F_{\tilde{D}_2}(x)$ $\forall x$, this establishes the lower bound as: $F_{V_{(3:3)}} \geq F_{V_{(2:3)}+\tilde{D}_2}(.)$.

When $n > 3$ the difference between $V_{(n-1:n)}$ and $V_{(1:n)}$ consists of D_{n-1} and at least one other spacing. This distance is per definition stochastically larger so the lower bound derived from it is weaker: $\forall n \geq 3$
$F_{V_{(n:n)}}(.) \geq F_{V_{(n-1:n)}+2(V_{(n-1:n)}-V_{(n-2:n)})}(.) \geq F_{V_{(n-1:n)}+2(V_{(n-1:n)}-V_{(1:n)})}(.)$, $\forall n \geq 3$. As n

doesn't need to be observed to obtain the right-hand side of this equation, this completes the proof. $\square$

Lemma 3. *[Concave density] Let $f_{V|\mathbf{Z}}$ be (weakly) concave. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}$ $\forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \leq P[V^U_{n-1:n} + V^U_{n-1:n} - V^L_{n-2:n} \leq v|\mathbf{Z} = z] \equiv F^{L(c)}_{V_{n:n}|\mathbf{Z}}(v)$$

Proof Lemma 3. Uniform spacings (spacings of an ordered sample of independent draws from a uniform distribution) are interchangeable random variables (Pyke (1965)): $F_{D_i}(.) = F_{D_j}(.)$ $\forall i, j \in \{2, .., n\}$. But spacings based on a concave f_V are stochastically decreasing since this holds for log-concave density functions (see Lemma 8) and since any concave function that is non-negative on its domain is also log-concave. This means that $\forall n \geq 3$, $\forall x$: $P[V_{(n:n)} \leq x] = P[(V_{(n-1:n)} + D_n) \leq x] \geq P[(V_{(n-1:n)} + D_{n-1}) \leq x]$; holding with equality when valuations are drawn from the uniform distribution. Hence: $F_{V_{(n:n)}}(.) \geq F_{V_{(n-1:n)}+D_{n-1}}(.)$. The spacing between the second-highest and lowest valuation is per definition weakly larger than D_{n-1} (strictly when $n > 3$) so this delivers a valid lower bound. $\square$

Lemma 4. *[Symmetric density] Let $f_{V|\mathbf{Z}}$ be symmetric. Observing $V^U_{n-1:n}$ and $V^L_{n-2:n}$ in auctions with $\mathbf{Z} = \mathbf{z}$ $\forall \mathbf{z}$, identifies a lower bound on the distribution of $V_{n:n}$, $\forall v \in [\underline{v}, \bar{v}]$:*

$$F_{V_{n:n}|\mathbf{Z}}(v) = P[V_{n:n} \leq v|\mathbf{Z} = \mathbf{z}] \leq P[V^U_{n-1:n} + V^U_{n-1:n} - V^L_{1:n} \leq v|\mathbf{Z} = z] \equiv F^{L(s)}_{V_{n:n}|\mathbf{Z}}(v)$$

Proof Lemma 4. With $n = 3$, $V_{(n-1:n)}$ coincides with median (and mean) so that $F_{D_2}(.) = F_{D_3}(.)$ as the lowest and the highest valuation are of equal distance to the median under the symmetry assumption. This implies that $F_{V_{(3:3)}}(.) = F_{V_{(2:3)}+D_2}(.)$. With $n > 3$, the median will be stochastically smaller than $V_{(n-1:n)}$ and the difference between the second-highest and lowest valuation will be stochastically larger than the difference between the highest and second-highest valuation. Hence, $F_{V_{(n:n)}}(.) \geq F_{V_{(n-1:n)}+(V_{(n-1:n)}-V_{(1:n)})}(.)$. $\square$

Bounds on the latent distribution of valuations

Lemma 5. *[Sample extremes] In symmetric English auctions, knowing (a lower bound on)* $V_{(1:n)}$ *nonparametrically identifies an upper bound on* $F_{\mathbf{V}}(.)$*. Knowing* $V_{(n:n)}$ *nonparametrically identifies* $F_{\mathbf{V}}(.)$ *and an upper bound on* $V_{(n:n)}$ *identifies a lower bound on* $F_{\mathbf{V}}(.)$*.*

Proof Lemma 5. Let $\mathbf{V}=\{V_1, V_2, ..., V_n\}$ be a sample of size n unrestricted (independent, positively correlated, negatively correlated) draws from F_V. Since $V_{(1:n)} \equiv \min(V_1, ..., V_n)$ it follows that $F_{V_{(1:n)}}(v) \geq F_{\mathbf{V}}(\mathbf{v})$, regardless of the sample dependence: $F_{V_{(1:n)}}(v) = P[\min(V_1, ..., V_n) \leq v] \geq P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_n \leq v]$. This holds with equality if $n = 1$ or valuations are perfectly correlated. Let $V^L_{(1:n)}$ be a lower bound on the 1st order statistic: $V^L_{(1:n)} \leq V_{(1:n)}$ implies $F_{V^L_{(1:n)}}(v) \geq F_{V_{(1:n)}}(v)$ so that observing a lower bound on $V_{(1:n)}$ will be informative as a weaker upper bound on $F_{\mathbf{V}}(\mathbf{v})$.

This provides an upper bound on $F_{\mathbf{V}}(\mathbf{v}) = F_{V_1,V_2,...,V_{n-1},V_n}(v, v, ..., v, v)$, which is only a diagonal slice of the full joint distribution. Now, consider a value w off the diagonal: $F_{V_1,V_2,...,V_{n-1},V_n}(v, v, ..., v, w)$. For $w > v$: $P[\min(V_1, ..., V_n) \leq v] \geq P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_{n-1} \leq v \cap V_n \leq v] \geq P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_{n-1} \leq v \cap V_n \leq w]$, meaning that $F_{V_{(1:n)}}(v)$ is still a valid upper bound for this joint probability. More generally, it holds that: $F_{V_{(1:n)}}(v) \geq F_{V_1,V_2,...,V_{n-1},V_n}(a, b, ..., y, z)$ $\forall a, b, ..., y, z \geq v, \forall v \in [\underline{v}, \bar{v}]$.

Due to the definition of the sample maximum, $V_{(n:n)} \equiv \max(V_1, ..., V_n)$, it follows that: $F_{V_{(n:n)}}(v) = P[\max(V_1, ..., V_n) \leq v] = P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_n \leq v] = F_{\mathbf{V}}(\mathbf{v})$. Let $V^U_{(n:n)}$ be an upper bound on the nth order statistic: $V^U_{(n:n)} \geq V_{(n:n)}$ implies $F_{V^U_{(n:n)}}(v) \leq F_{V_{(n:n)}}(v)$ so that observing an upper bound on the nth order statistic will be informative as a lower bound on $F_{\mathbf{V}}(\mathbf{v})$, regardless of sample dependence.

This provides a lower bound on $F_{\mathbf{V}}(\mathbf{v}) = F_{V_1,V_2,...,V_{n-1},V_n}(v, v, ..., v, v)$. Now, consider a value u off the diagonal: $F_{V_1,V_2,...,V_{n-1},V_n}(v, v, ..., v, u)$. For $u < v$:
$P[\max(V_1, ..., V_n) \leq v] = P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_{n-1} \leq v \cap V_n \leq v] \leq P[V_1 \leq v \cap V_2 \leq v \cap ... \cap V_{n-1} \leq v \cap V_n \leq u]$, meaning that $F_{V_{(n:n)}}(v)$ is a valid lower bound for this joint probability. More generally, it holds that: $F_{V_{(n:n)}}(v) \leq F_{V_1,V_2,...,V_{n-1},V_n}(a, b, ..., y, z)$ $\forall a, b, ..., y, z \leq v, \forall v \in [\underline{v}, \bar{v}]$. $\square$

Lemma 7. *[Based on Song (2004)] In symmetric IPV English auctions, knowing two ordered valuations and their distance from above* $V_{i:n}$ *and* $V_{j:n}$ *($j < i$) nonparametrically identifies* $F_{V|V \geq v_j}(v_i)$*. Equivalently, knowing two ordered valuations and their distance from below* $V_{i:n}$ *and* $V_{j:n}$ *($j > i$) nonparametrically identifies* $F_{V|V \leq v_j}(v_i)$*.*

Proof Lemma 7 (using two ordered valuations from above, the other case is omitted.) Knowing the ith and jth highest values of X, e.g. $X_{n-i+1:n}$ and $X_{n-j+1:n}$, with $i > j$, the density of $X_{n-j+1:n}$ conditional on $X_{n-i+1:n}$ equals the density of $X_{i-j:i-1}$ from an i.i.d. sample truncated from below at $X_{n-i+1:n} = x_i$:

$$f_{X(n-j+1:n)|X(n-i+1:n)}(x_j|x_i) = \frac{f_{X(n-j+1:n),X(n-i+1:n)}(x_j, x_i)}{f_{X(n-i+1:n)}(x_i)} = \tag{4.24}$$

$$\frac{\frac{n!}{(n-1)!(i-j-1)!(j-1)!}F_X(x_i)^{n-i}[F_X(x_j) - F_X(x_i)]^{i-j-1}[1 - F_X(x_j)]^{j-1}f_X(x_i)f_X(x_j)}{\frac{n!}{(n-i)!(i-1)!}F_X(x_i)^{n-1}[1 - F_X(x_i)]^{i-1}f_X(x_i)} \tag{4.25}$$

$$= \frac{(i-1)!}{(i-j-1)!(j-1)!}\left\{[1 - F_X(x_i)]\left(\frac{[F_X(x_i) - F_X(x_j)]}{[1 - F_X(x_i)]}\right)\right\}^{i-j-1} \tag{4.26}$$

$$\left\{[1 - F_X(x_i)]\left(\frac{[1 - F_X(x_j)]}{[1 - F_X(x_i)]}\right)\right\}^{j-1}\frac{[1 - F_X(x_i)]\frac{f_X(x_i)}{[1-F_X(x_i)]}}{[1 - F_X(x_i)]^{i-1}}$$

$$= \frac{(i-1)!}{(i-j-1)!(j-1)!}[F_{X|X\geq x_i}(x_j)]^{i-j-1}[1 - F_{X|X\geq x_i}(x_j)]^{j-1}f_{X|X\geq x_i}(x_j), \tag{4.27}$$

$\forall x_j > x_i$ (and 0 otherwise). Equation (4.27) follows from noticing that $[F_X(.)]/[1-F_X(x_i)]$ and $f_X(.)/[1-F_X(x_i)]$ equal respectively the CDF and PDF of X truncated from below at x_i, and that the additional $[1 - F_X(x_i)]$'s cancel out. The CDF's are related accordingly, $\forall x \in (x_i, \bar{x}]$, $\forall x_i \in [\underline{x}, \bar{x})$, $\forall i > j$:

$$F_{X(n-j+1:n)|X(n-i+1:n)}(x|x_i) = F_{X(i-j:i-1)|X\geq x_i}(x) \tag{4.28}$$

$$= \frac{(i-1)!}{(i-j-1)!(j-1)!}\int_{x_i}^{x} F_{X|X\geq x_i}(t)^{i-j-1}[1 - F_{X|X\geq x_i}(t)]^{j-1}f_{X|X\geq x_i}(t)dt \tag{4.29}$$

Again, the distribution of the $(i-j)$th order statistic from a sample of $(i-1)$ i.i.d. draws from $F_{X|X\geq x_i}(x)$ is a known and strictly monotonic function of its parent distribution, which can renders $F_{X|X\geq x_i}(x)$ identified. □

Lemma 8. *[Based on Song (2004) and Haile and Tamer (2003)] In symmetric IPV English auctions, knowing bounds on at least one of two ordered valuations $V_{i:n}$ and $V_{j:n}$, with $i = n - k$ and $j = n - l$ $(0 \leq k < l \leq n-1)$ nonparametrically identifies bounds on $F_{V|V\geq v_j}$.*[31]

[31] Equivalently, for procurement auctions, knowing bounds on at least one of two ordered valuations $V_{i:n}$ and $V_{j:n}$, with $i = k$ and $j = l$ $(1 \leq k < l \leq n)$ nonparametrically identifies bounds on $F_{V|V\leq v_j}$.

Proof Lemma 8. Let $X^l_{i:n}$ and $X^u_{i:n}$ be known lower and upper bounds on the unknown ith order statistic, and let $X_{j:n}$ be a known order statistic, with $i < j$. From $X^l_{i:n} \leq X_{i:n} \leq X^u_{i:n}$ it follows that: $F_{X^l_{i:n}}(x) \geq F_{X_{i:n}}(x) \geq F_{X^u_{i:n}}(x)$, $\forall x \in [\underline{x}, \bar{x}]$. These stochastic dominance relations remain when conditioning on the same value of $X_{j:n}$, so that $\forall x \in [\underline{x}, x_j]$, $\forall x_j \in (\underline{x}, \bar{x}]$:

$$F_{X^l_{i:n}|X_{j:n}}(x|x_j) \geq F_{X_{i:n}|X_{j:n}}(x|x_j) \geq F_{X^u_{i:n}|X_{j:n}}(x|x_j) \tag{4.30}$$

By Lemma 2, these conditional distributions identify the distributions truncated at the value of $X_{j:n}$. Since $\varphi\,(.; j, i, x_j)$ is a strictly increasing function that is identical for the conditional CDF's above (having the same j, i, x_j) it follows that on the same support:

$$F_{X^l|X \leq x_j}(x) \geq F_{X|X \leq x_j}(x) \geq F_{X^u|X \leq x_j}(x) \tag{4.31}$$

where $F_{X^l|X \leq x_j}(x)$ and $F_{X^u|X \leq x_j}(x)$ respectively refer to the CDF's identified from the lower and upper bound on $X_{i:n}$. This identifies bounds on $F_{X|X \leq x_j}(x)$.

Suppose instead that $X_{j:n}$ is unknown, but that it is known within bounds such that: $X^l_{j:n} \leq X_{j:n} \leq X^u_{j:n}$ while $X_{i:n}$ is known. By definition, $X_{(j:n)}{}^l$ must be greater than $X_{(i:n)}$ (otherwise $X^l_{j:n}$ is not a sharp lower bound when $X_{(i:n)}$ is known). Logically, $P[X_{i:n} \leq x]$ conditional on $X_{j:n} \leq x_j$ is at least as much as this probability conditional on the lower bound on $X_{j:n}$ being less than x_j, and the reverse is true for the upper bound on $X_{j:n}$. Hence, by this simple stochastic dominance argument, $\forall\ x \in [\underline{x}, x_j)$, $\forall\ x_j \in (\underline{x}, \bar{x}]$,:

$$F_{X_{i:n}|X^u_{j:n}}(x|x_j) \geq F_{X_{i:n}|X_{j:n}}(x|x_j) \geq F_{X_{i:n}|X^l_{j:n}}(x|x_j), \tag{4.32}$$

which requires that $F_{X^u_{j:n}}(x_j) > 0$. Relating to the parent distribution by the same monotonic transformation $\varphi\,(.; j, i, x_j)$ it follows that on the same support:

$$F_{X|X^u \leq x_j}(x) \geq F_{X|X \leq x_j}(x) \geq F_{X|X^l \leq x_j}(x) \tag{4.33}$$

where $F_{X|X^u \leq x_j}(x)$ and $F_{X|X^l \leq x_j}(x)$ refer to the CDF's conditional on respectively the upper and lower bound on $X_{j:n}$.

Finally, when both order statistics are only known within bounds, combining above results it follows that an upper bound on $F_{X|X \leq x_j}(x)$ is identified from $F_{X^l|X^u \leq x_j}(x)$ and a lower bound is identified from $F_{X^u|X^l \leq x_j}(x)$. $\square$

Bibliography

Adams, Christopher P. Estimating demand from eBay prices. *International Journal of Industrial Organization*, 25(6):1213–1232, 2007.

Aguirregabiria, Victor and Pedro Mira. Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica*, 70(4):1519–1543, 2002.

Aguirregabiria, Victor and Pedro Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53, 2007.

Albrecht, James, Pieter A. Gautier, and Susan Vroman. A note on Peters and Severinov, "Competition among sellers who offer auctions instead of prices". *Journal of Economic Theory*, 147(1):389–392, 2012.

Albrecht, James, Pieter A. Gautier, and Susan Vroman. Efficient entry in competing auctions. *The American Economic Review*, 104(10):3288–3296, 2014.

An, Mark Yuying. Log-concave probability distributions: theory and statistical testing. *Duke University Dept of Economics Working Paper No. 95-03*, 1995.

Andrews, Donald W. K. and Sukjin Han. Invalidity of the bootstrap and the m out of n bootstrap for confidence interval endpoints defined by moment inequalities. *The Econometrics Journal*, 12(SUPPL. 1):172–199, 2009.

Anwar, Sajid, Robert McMillan, and Mingli Zheng. Bidding behavior in competing auctions: Evidence from eBay. *European Economic Review*, 50(2):307–322, 2006.

Aradillas-López, Andrés. Nonparametric tests for conditional affiliation in auctions and other models. *Unpublished manuscript*, 2016. URL `http://www.personal.psu.edu/aza12/cond_affiliation_test.pdf`.

Aradillas-López, Andrés, Amit Gandhi, and Daniel Quint. Identification and inference in ascending auctions With correlated private values. *Econometrica*, 81(2):489–534, 2013.

Armstrong, Mark. Competition in two-sided markets. *The RAND Journal of Economics*, 37(3):668–691, 2006.

Armstrong, Timothy B. Bounds in auctions with unobserved heterogeneity. *Quantitative Economics*, 4(3):377–415, 2013.

Arnold, B., N. Balakrishnan, and Haikady N. Nagaraja. *A first course in order statistics*. Wiley & Sons, New York, 1992.

Ashenfelter, Orley. How auctions works for wine and art. *The Journal of Economic Perspectives*, 3(3):23–36, 1989.

Ashenfelter, Orley. Predicting the quality and prices of Bordeaux wine. *The Economic Journal*, 118(6):F174–F185, 2008.

Ashenfelter, Orley and Kathryn Graddy. Auctions and the price of art. *Journal of Economic Literature*, 41(3):763–786, 2003.

Ashenfelter, Orley and Kathryn Graddy. Anatomy of the rise and fall of a price-fixing conspiracy: Auctions at Sotheby's and Christie's. *Journal of Competition Law and Economics*, 1(1):3–20, 2005.

Ashenfelter, Orley and Kathryn Graddy. Art auctions. *Handbook of the Economics of Art and Culture*, 1:909–945, 2006.

Ashenfelter, Orley, David Ashmore, and Robert Lalonde. Bordeaux wine vintage quality and the weather. *Chance*, 8(4):7–14, 1995.

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

Athey, Susan and Glenn Ellison. Position auctions with consumer search. *Quarterly Journal of Economics*, 126(3):1213–1270, 2011.

Athey, Susan and Philip A. Haile. Nonparametric approaches to auctions. *Handbook of Econometrics*, 6A:3847–3965, 2007.

Backus, Matthew and Gregory Lewis. Dynamic demand estimation in auction markets. *NBER Working Paper No. 22375*, 2016.

Bagnoli, Mark and Ted Bergstrom. Log-concave probability and its applications. *Economic Theory*, 26:445–469, 2005.

Bajari, Patrick and Ali Hortaçsu. The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions. *The RAND Journal of Economics*, 34(2): 329–355, 2003.

Bajari, Patrick and Ali Hortaçsu. Economic insights from internet auctions. *Journal of Economic Literature*, 42(2):457–486, 2004.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

Bajari, Patrick, Han Hong, John Krainer, and Denis Nekipelov. Estimating static models of strategic interactions. *Journal of Business and Economic Statistics*, 28(4):469–482, 2010.

Beggs, Alan and Kathryn Graddy. Declining values and the afternoon effect: Evidence from art auctions. *The Rand Journal of Economics*, 28(3):544–565, 1997.

Berry, Steven T. and Philip A. Haile. Identification of nonparametric simultaneous equations models with residual index structure. *Econometrica*, 86(1):289–315, 2018.

Bikhchandani, Sushil, Philip A. Haile, and John G. Riley. Symmetric separating equilibria in English auctions. *Games and Economic Behavior*, 38(1):19–27, 2002.

Bodoh-Creed, Aaron, Jörn Boehnke, and Brent R. Hickman. Optimal design of two-sided market platforms: An empirical case study of eBay. *NET Institute Working Paper No. 13-22*, 2013.

Bodoh-Creed, Aaron, Jörn Boehnke, and Brent R. Hickman. How efficient are decentralized auction platforms? *Becker Friedman Institute for Research in Economics Working Paper No. 2016-2*, 2016.

Bomse, Stephen V. and Scott A. Westrich. Both sides now: Buyer damage claims in antitrust auctions involving two-sided markets. *Columbia Business Law Review*, 2005 (3):643–666, 2005.

Bresnahan, Timothy, Joe Orsini, and Pai-Ling Yin. Demand heterogeneity, inframarginal multihoming, and platform market stability: Mobile apps. *Proceedings of the 9th IDEI-TSE-IAST Conference on The Economics of Intellectual Property, Software and the Internet*, 2015.

Bugni, Federico A. Bootstrap inference in partially identified models defined by moment inequalities: Coverage of the identified set. *Econometrica*, 78(2):735–753, 2010.

Bulow, Jeremy and Paul Klemperer. Auctions versus negotiations. *The American Economic Review*, 86(1):180–194, 1996.

Canay, Ivan A. EL inference for partially identified models: Large deviations optimality and bootstrap validity. *Journal of Econometrics*, 156(2):408–425, 2010.

Chacón, José E. and Alberto Rodríguez-Casal. A note on the universal consistency of the kernel distribution function estimator. *Statistics and Probability Letters*, 80(17): 1414–1419, 2010.

Chandra, Ambarish and Allan Collard-Wexler. Mergers in two-sided markets: An application to the Canadian newspaper industry. *Journal of Economics and Management Strategy*, 18(4):1045–1070, 2009.

Chernozhukov, Victor, Han Hong, and Elie Tamer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243–1284, 2007.

Chernozhukov, Victor, Sokbae Lee, and Adam M. Rosen. Intersection bounds: Estimation and inference. *Econometrica*, 81(2):667–737, 2009.

Chesher, Andrew. Instrumental values. *Journal of Econometrics*, 139(1):15–34, 2007.

Chesher, Andrew and Adam M. Rosen. Identification of the distribution of valuations in an incomplete model of English auctions. *Cemmap Working Paper CWP30/15*, 2015.

Chesher, Andrew and Adam M. Rosen. Generalized instrumental variable models. *Econometrica*, 85(3):959–989, 2017.

Cho, Sungjin, Harry J. Paarsch, and John Rust. An empirical analysis of informationally restricted dynamic auctions of used cars. *Unpublished manuscript*, 2014. URL `https://www.semanticscholar.org/paper/An-Empirical-Analysis-of-Informationally-Restricted-Cho/1b6db454185b1d6a40a881ffa773b1f6933bc9c3`.

Christie's. Level/ullage descriptions and interpretations. 2013. URL `https://www.christies.com/Wine/Ullages{_}2013.pdf`. Accessed: 2018-08-04.

Coey, Dominic, Bradley J. Larsen, Kane Sweeney, and Caio Waisman. Ascending auctions with bidder asymmetries. *Quantitative Economics*, 8(1):181–200, 2017.

David, Herbert A. and Haikady N. Nagaraja. Order statistics. *Encyclopedia of Statistical Sciences*, 2004.

Deltas, George and Thomas D. Jeitschko. Auction hosting site pricing and market equilibrium with endogenous bidder and seller participation. *International Journal of Industrial Organization*, 25(6):1190–1212, 2007.

Donald, Stephen G. and Harry J. Paarsch. Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions. *International Economic Review*, 34(1):121–148, 1993.

Donald, Stephen G. and Harry J. Paarsch. Identification, estimation, and testing in parametric empirical models of auctions within the independent private values paradigm. *Econometric Theory*, 12(03):517–567, 1996.

Egesdal, Michael, Zhenyu Lai, and Che-lin Su. Estimating dynamic discrete-choice games of incomplete information. *Quantitative Economics*, 6:567–597, 2015.

Einav, Liran, Jonathan Levin, Chiara Farronato, and Neel Sundaresan. Auctions versus posted prices in online markets. *Journal of Political Economy*, 126(1):178–215, 2018.

Eklof, Matias and Anders Lunander. Open outcry auctions with secret reserve prices: An empirical application to executive auctions of tenant owner's apartments in Sweden. *Journal of Econometrics*, 114(2):243–260, 2003.

Ellison, Glenn, Drew Fudenberg, and Markus Mobius. Competing auctions. *Journal of the European Economic Association*, 2(1):30–66, 2004.

Elyakime, Bernard, Jean Jacques Laffont, Patrice Loisel, and Quang Vuong. First-price sealed bid auctions with secret reservation prices. *Annales d'Économie et de Statistique*, (34):115–141, 1994.

Engelberg, Joseph and Jared Williams. eBay ' s proxy bidding: A license to shill. *Journal of Economic Behavior and Organization*, 72(1):509–526, 2009.

Engelbrecht-Wiggans, Richard. The effect of entry and information costs on oral versus sealed-bid auctions. *Economic Letters*, 70(2):195–202, 2001.

Evans, David S. Some Empirical Aspects of Multi-sided Platform Industries. *Review of Network Economics*, 2(3):191–209, 2003.

Evans, David S. and Richard Schmalensee. Failure to launch: Critical mass in platform businesses. *Review of Network Economics*, 9(4):1–26, 2010.

Evans, David S. and Richard Schmalensee. The antitrust analysis of multi-sided platform businesses. *NBER Working Paper No. 18783*, 2013.

Fan, Ying. Ownership consolidation and product characteristics: A study of the US daily newspaper market. *The American Economic Review*, 103(5):1598–1628, 2013.

Fang, Hanming and Xun Tang. Inference of bidders' risk attitudes in ascending auctions with endogenous entry. *Journal of Econometrics*, 180(2):198–216, 2014.

Freyberger, Joachim and Bradley J. Larsen. Identification in ascending auctions, with an application to digital rights management. *NBER Working Paper No. 23569*, 2017.

Gentry, Matthew and Tong Li. Identification in auctions with selective entry. *Econometrica*, 82(1):315–344, 2014.

Gentry, Matthew, Tong Li, and Jingfeng Lu. Identification and estimation in first-price auctions with risk-averse bidders and selective entry. *Cemmap Working Paper CWP16/15*, 2015.

Gentry, Matthew, Tong Li, and Jingfeng Lu. Auctions with selective entry. *Games and Economic Behavior*, 105:104–111, 2017.

Gentzkow, Matthew and Jesse M. Shapiro. What drives media slant? evidence from U.S. daily newspapers. *Econometrica*, 78(1):35–71, 2010.

Ginsburgh, Victor. Absentee bidders and the declining price anomaly in wine auctions. *Journal of Political Economy*, 106(6):1302–1319, 1998.

Ginsburgh, Victor, Patrick Legros, and Nicolas Sahuguet. On the incidence of commissions in auction markets. *International Journal of Industrial Organization*, 28(6):639–644, 2010.

Gomes, Renato. Optimal auction design in two-sided markets. *The RAND Journal of Economics*, 45(2):248–272, 2014.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

Haile, Philip A. Auctions with resale markets: An application to U.S. forest service timber sales. *The American Economic Review*, 91(3):399–427, 2001.

Haile, Philip A. and Elie Tamer. Inference with an incomplete model of English auctions. *Journal of Political Economy*, 111(1):1–51, 2003.

Haile, Philip A., Han Hong, and Matthew Shum. Nonparametric tests for common values at first-price sealed-bid auctions. *NBER Working Paper No. 10105*, 2003.

Härdle, Wolfgang and Oliver Linton. Applied nonparametric methods. *Handbook of Econometrics*, 4:2295–2339, 1994.

Hasker, Kevin and Robin Sickles. eBay in the economic literature: Analysis of an auction marketplace. *Review of Industrial Organization*, 37(1):3–42, 2010.

Hayfield, Tristen and Jeffrey S. Racine. Nonparametric ecoometrics: The np package. *Journal of Statistical Software*, 27(5):1–51, 2008.

Hendricks, Kenneth and Robert H. Porter. An empirical study of an auction with asymmetric information. *The American Economic Review*, 78(5):865–883, 1988.

Hendricks, Kenneth and Robert H. Porter. An empirical perspective on auctions. *Handbook of Industrial Organization*, 3, 2007.

Hendricks, Kenneth, Joris Pinkse, and Robert H. Porter. Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *The Review of Economic Studies*, 70(1):115–145, 2003.

Hickman, Brent R. On the pricing rule in electronic auctions. *International Journal of Industrial Organization*, 28(5):423–433, 2010.

Hickman, Brent R., Timothy P. Hubbard, and Harry J. Paarsch. Identification and estimation of a bidding model for electronic auctions. *Quantitative Economics*, 8(2):505–551, 2017.

Hong, Han and Matthew Shum. Increasing competition and the winner's curse: Evidence from procurement. *The Review of Economic Studies*, 69(4):871–898, 2002.

Horowitz, Joel L. The bootstrap. *Handbook of Econometrics*, 5:3159–3228, 2001.

Horowitz, Joel L. and Charles F. Manski. Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, 95(449):77–84, 2000.

Hortaçsu, Ali and Eric R. Nielsen. Do bids equal values on eBay? *Marketing Science*, 29 (6):994–997, 2010.

Hurwicz, Leonid. Generalization of the concept of identification. *Statistical Inference in Dynamic Economic Models (T. Koopmans, ed). Cowles Commission, Monograph*, 10: 245–257, 1950.

Kim, Kyooil and Joonsuk Lee. Nonparametric estimation and testing of the symmetric IPV framework with unknown number of bidders. *Unpublished manuscript*, 2014. URL https://msu.edu/~kyookim/index_files/AUCTION_IPV.pdf.

Imbens, Guido W. and Charles F. Manski. Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857, 2004.

In re Auction Houses antitrust litigation. No. 00 Civ. 0648, LAK (S.D.N.Y. Feb. 22, 2001). 2001. URL https://casetext.com/case/in-re-auction-houses-antitrust-litigation-61. Accessed: 2018-09-12.

In re eBay Seller antitrust litigation. 545 F. Supp. 2d 1027 (N.D. Cal. 2008). 2008. URL https://www.courtlistener.com/opinion/1431968/in-re-ebay-seller-antitrust-litigation/. Accessed: 2018-09-13.

Jehiel, Philippe and Laurent Lamy. On absolute auctions and secret reserve prices. *The RAND Journal of Economics*, 46(2):241–270, 2015.

Jofre-bonet, Mireia and Martin Pesendorfer. Estimation of a dynamic auction game. *Econometrica*, 71(5):1443–1489, 2003.

Judd, Kenneth L. *Numerical methods in economics.* MIT press, 1998.

Kasahara, Hiroyuki and Katsumi Shimotsu. Sequential estimation of structural models with a fixed point constraint. *Econometrica*, 80(5):2303–2319, 2012.

Katz, Michael L. and Carl Shapiro. Network externalities, competition, and compatibility. *The American Economic Review*, 75(3):424–440, 1985.

Koopmans, Tjalling C. and O. Reiersol. The identification of structural characteristics. *The Annals of Mathematical Statistics*, 21(2):165–181, 1950.

Krasnokutskaya, Elena and Katja Seim. Bid preference programs and participation in highway procurement auctions. *The American Economic Review*, 101(6):2653–2686, 2011.

Krishna, Vijay. *Auction theory.* Academic Press, 2009.

Laffont, Jean-Jacques and Quang Vuong. Structural econometric analysis of descending auctions. *European Economic Review*, 37(2-3):329–341, 1993.

Laffont, Jean-Jacques, Herve Ossard, and Quang Vuong. Econometrics of first-price auctions. *Econometrica*, 63(4):953–980, 1995.

Lee, Robin S. Home Videogame Platforms. *Oxford Handbook of the Digital Economy*, pages 83–107, 2011.

Levin, Dan and James L. Smith. Equilibrium in auctions with entry. *The American Economic Review*, 84(3):585–599, 1994.

Li, Qi and Jeffrey S. Racine. *Nonparametric econometrics: Theory and practice*. Princeton University Press, 2007.

Li, Qi, Juan Lin, and Jeffrey S. Racine. Optimal bandwidth selection for nonparametric conditional distribution and quantile functions. *Journal of Business and Economic Statistics*, 31(1):57–65, 2013.

Li, Tong and Xiaoyong Zheng. Entry and competition effects in first-price auctions: Theory and evidence from procurement auctions. *The Review of Economic Studies*, 76(4):1397–1429, 2009.

Li, Tong and Xiaoyong Zheng. Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics*, 168(1):29–46, 2012.

Li, Tong, Isabelle Perrigne, and Quang Vuong. Structural estimation of the affiliated private value auction model. *The RAND Journal of Economics*, 33(2):171–193, 2002.

Li, Xiaohu. A note on expected rent in auction theory. *Operations Research Letters*, 33 (5):531–534, 2005.

Lucking-reiley, David. Vickrey auctions in practice: From nineteenth-century philately to twenty-first-century e-commerce. *The Journal of Economic Perspectives*, 14(3):183–192, 2000.

Manski, Charles F. and John V. Pepper. Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, 68(4):997–1010, 2000.

Manski, Charles F. and John V. Pepper. More on monotone instrumental variables. *Econometrics Journal*, 12:S200–S216, 2009.

Marks, Denton. Who pays brokers ' commissions? Evidence from fine wine auctions. *Oxford Economic Papers*, 61:761–775, 2009.

Marmer, Vadim, Artyom Shneyerov, and Pai Xu. What model for entry in first-price auctions? A nonparametric approach. *Journal of Econometrics*, 176(1):46–58, 2013.

McAfee, R. Preston. Mechanism design by competing sellers. *Journal of Economic Theory*, 61(6):1281–1312, 1993.

Mcafee, R. Preston and John Mcmillan. Auctions with a stochastic number of bidders. *Journal of Economic Theory*, 19(1):1–19, 1987.

McAfee, R. Preston and Daniel Vincent. The declining price anomaly. *Journal of Economic Theory*, 60(1):191–212, 1993.

Milgrom, Paul. Auctions and bidding: A primer. *The Journal of Economic Perspectives*, 3(3):3–22, 1989.

Milgrom, Paul R and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, 1982.

Moreno, Diego and John Wooders. Auctions with heterogeneous entry costs. *The RAND Journal of Economics*, 42(2):313–336, 2011.

Myerson, Roger B. Optimal auction design. *Mathematics of Operations Research*, 6(1): 58–73, 1981.

Myerson, Roger B. Population uncertainty and Poisson games. *International Journal of Game Theory*, 27(3):375–392, 1998.

Nadaraya, Elizbar A. Some new estimates for distribution functions. *Theory of Probability & Its Applications*, 9(3):497–500, 1964.

Nadaraya, Elizbar A. On non-parametric estimates of density functions and regression curves. *Theory of Probability & Its Applications*, 10(1):186–190, 1965.

Nekipelov, Denis. Entry Deterrence and Learning Prevention on eBay. 2007. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.407.3749&rep=rep1&type=pdf`.

Newberry, Peter W. The effect of competition on eBay. *International Journal of Industrial Organization*, 40:107–118, 2015.

Ockenfels, Axel and Alvin E. Roth. Late and multiple bidding in second price Internet auctions: Theory and evidence concerning different rules for ending an auction. *Games and Economic Behavior*, 55(2):297–320, 2006.

OECD Competition Committee. Two-sided markets, daf/comp(2009)20. 2009. URL `http://www.oecd.org/daf/competition/44445730.pdf`. Accessed: 2018-09-19.

OECD Competition Committee. Market definition in multi-sided markets, daf/comp/wd(2017)33/final. 2017. URL `http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DAF/COMP/WD(2017)22/FINAL{&}docLanguage=En`. Accessed: 2018-09-19.

Ostrovsky, Michael and Michael Schwarz. Reserve prices in internet advertising auctions: A field experiment. *Unpublished manuscript*, 2016. URL `https://web.stanford.edu/~ost/papers/rp.pdf`.

Paarsch, Harry J. Empirical models of auctions and an application to British Columbian timber sales. *University of Western Ontario, Department of Economics, UWO Department of Economics Working Paper No. 9212*, 1992.

Paarsch, Harry J. Deriving an estimate of the optimal reserve price: An application to British Columbian timber sales. *Journal of Econometrics*, 78(1):333–357, 1997.

Paarsch, Harry J. and Han Hong. *An introduction to the structural econometrics of auction data*. MIT press, 2006.

Pesendorfer, Martin and Philipp Schmidt-Dengler. Identification and estimation of dynamic games. *NBER Working Paper No. 9726*, 2003.

Pesendorfer, Martin and Philipp Schmidt-Dengler. Sequential estimation of dynamic discrete games: A comment. *Econometrica*, 78(2):833–842, 2010.

Peters, Michael and Sergei Severinov. Competition among sellers who offer auctions instead of prices. *Journal of Economic Theory*, 75(1):141–179, 1997.

Peters, Michael and Sergei Severinov. Internet auctions with many traders. *Journal of Economic Theory*, 130(1):220–245, 2006.

Pyke, Ronald. Spacings. *Journal of the Royal Statistical Society. Series B (Methodological)*, 27(3):395–449, 1965.

Pyke, Ronald. Spacings revisited. *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics*, pages 417–427, 1972.

Riley, John G. and William F. Samuelson. Optimal Auctions. *Optimal auctions*, 71(3): 381–392, 1981.

Roberts, James W. Unobserved heterogeneity and reserve prices in auctions. *The RAND Journal of Economics*, 44(4):712–732, 2013.

Roberts, James W. and Andrew Sweeting. Entry and selection in auctions. *NBER Working Paper No. 16650*, 2010.

Roberts, James W. and Andrew Sweeting. Competition versus Auction Design. *Unpublished manuscript*, 2011. URL `http://econweb.umd.edu/~sweeting/SWEETING_compaucdesign.pdf`.

Roberts, James W. and Andrew Sweeting. When should sellers use auctions? *The American Economic Review*, 103(5):1830–1861, 2013.

Rochet, Jean-charles and Jean Tirole. Two-Sided Markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

Rochet, Jean-charles and Jean Tirole. Two-sided markets: A progress report. *The RAND Journal of Economics*, 37(3):645–667, 2006.

Rosen, Adam M. Confidence sets for partially identified parameters that satisfy a finite number of moment inequalities. *Journal of Econometrics*, 146(1):107–117, 2008.

Rothkopf, Michael H., Thomas J. Teisberg, and Edward P. Kahn. Why are Vickrey auctions rare? *Journal of Political Economy*, 98(1):94–109, 1990.

Rust, John. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987.

Rysman, Marc. Competition between networks: A study of the market for Yellow Pages. *The Review of Economic Studies*, 71(2):483–512, 2004.

Rysman, Marc. An empirical analysis of payment card usage. *The Journal of Industrial Economics*, 55(1):1–36, 2007.

Rysman, Marc and Julian Wright. The economics of payment cards. *Economic Review*, 13(3):1–12, 2002.

Samuelson, William F. Competitive bidding with entry cost. *Economic Letters*, 17(1-2): 53–57, 1985.

Seim, Katja. An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics*, 37(3):619–640, 2006.

Song, Minjae. Estimating platform market power in two-sided markets with an application to magazine advertising. *Simon School Working Paper No. FR 11-1*, 2013.

Song, Unjy. Nonparametric estimation of an eBay auction model with an unknown number of bidders. *Unpublished manuscript*, 2004. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.5483&rep=rep1&type=pdf`.

Sotheby's. Finest and rarest wines & vintage port, auction in London on Wednesday 19 November 2014. 2014. Accessed: 2014-11-19.

Springel, Katalin. Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives. *Unpublished manuscript*, 2019. URL `https://www.dropbox.com/s/ix788j0b6d24j6p/kspringel_ev.pdf?dl=0`.

Stoye, Jörg. More on confidence intervals for partially identified parameters. *Econometrica*, 77(4):1299–1315, 2009.

Tang, Xun. Bounds on revenue distributions in counterfactual auctions with reserve prices. *The RAND Journal of Economics*, 42(1):175–203, 2011.

Thiel, Stuart E. and Glenn H. Petry. Bidding behaviour in second-price auctions: Rare stamp sales, 1923-1937. *Applied Economics*, 27(1):11–16, 1995.

Tracer, Daniel M. Overcharge but don't overestimate: Calculating damages for antitrust injuries in two-sided markets. *Cardozo Law Review*, 33(2):101–133, 2011.

Vickrey, William. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37, 1961.

Wang, Ruqu. Auctions versus posted-price selling. *The American Economic Review*, 83 (4):838–851, 1993.

Wine Spectator. End of year auction report 2016. 2017a. URL `http://www.winespectator.com/webfeature/show/id/End-of-Year-Auction-Report-2016`. Accessed: 2017-08-19.

Wine Spectator. Half year auction report 2017. 2017b. URL `http://www.winespectator.com/webfeature/show/id/Half-Year-Auction-Report-2017`. Accessed: 2017-08-19.

Wine Spectator. Rare wine auctions set new records in the first half of 2018. 2018. Accessed: 2018-08-04.

Wright, Julian. One-sided logic in two-sided markets. *Review of Network Economics*, 3 (1):44–64, 2004.

Ye, Lixin. Indicative bidding and a theory of two-stage auctions. *Games and Economic Behavior*, 58(1):181–207, 2007.

Zeithammer, Robert and Christopher P. Adams. The sealed-bid abstraction in online auctions. *Marketing Science*, 29(6):964–987, 2010.