Bayesian Inference in Structural Second-Price Auctions

Bertil Wegmann

Stockholm University

Doctoral Dissertation
Department of Statistics
Stockholm University
S-106 91 Stockholm

Abstract

The aim of this thesis is to develop efficient and practically useful Bayesian methods for statistical inference in structural second-price auctions. The models are applied to a carefully collected coin auction dataset with bids and auction-specific characteristics from one thousand Internet auctions on eBay. Bidders are assumed to be risk-neutral and symmetric, and compete for a single object using the same game-theoretic strategy. A key contribution in the thesis is the derivation of very accurate approximations of the otherwise intractable equilibrium bid functions under different model assumptions. These easily computed and numerically stable approximations are shown to be crucial for statistical inference, where the inverse bid functions typically needs to be evaluated several million times.

In the first paper, the approximate bid is a linear function of a bidder's signal and a Gaussian common value model is estimated. We find that the publicly available book value and the condition of the auctioned object are important determinants of bidders' valuations, while eBay's detailed seller information is essentially ignored by the bidders. In the second paper, the Gaussian model in the first paper is contrasted to a Gamma model that allows intrinsically non-negative common values. The Gaussian model performs slightly better than the Gamma model on the eBay data, which we attribute to an almost normal or at least symmetrical distribution of valuations. The third paper compares the model in the first paper to a directly comparable model for private values. We find many interesting empirical regularities between the models, but no strong and consistent evidence in favor of one model over the other. In the last paper, we consider auctions with both private-value and common-value bidders. The equilibrium bid function is given as the solution to an ordinary differential equation, from which we derive an approximate inverse bid as an explicit function of a given bid. The paper proposes an elaborate model where the probability of being a common value bidder is a function of covariates at the auction level. The model is estimated by a Metropolis-within-Gibbs algorithm and the results point strongly to an active influx of both private-value and common-value bidders.

Keywords: Asymmetry, Bid function approximation, Common values, Gamma model, Gaussian model, Markov Chain Monte Carlo, Private values, Variable selection, Internet auctions.

ISBN 978-91-7447-276-9

List of Included Papers

I **Bayesian Inference in Structural Second-Price Common Value Auctions** (with Mattias Villani)
forthcoming in Journal of Business and Economic Statistics.

II **Bayesian Inference in Structural Second-Price Auctions with Gamma Distributed Common Values**

III **Bayesian Comparison of Private and Common Values in Structural Second-Price Auctions**

IV **Bayesian Inference in Structural Second-Price Auctions with both Private-Value and Common-Value Bidders**

Acknowledgments

It is with a lot of emotions I finally put an end for my thesis here. Several people has been involved, in one way or the other, for this thesis to take form. Foremost, I would like to thank my superb supervisor Professor Mattias Villani. You are truly amazing with your constant willingness to help and your ability to make difficult problems look so easy. You inspire me a lot with your great enthusiasm and vast skills. Thanks for all helpful guidelines, fruitful discussions and for allowing me to continue with my interest in auctions when I started my doctoral studies.

Special thanks to my assistant supervisor, Professor Daniel Thorburn, who introduced me to the Department. Your expertise in many interesting discussions is greatfully appreciated. To my undergraduate supervisor Niklas Karlsson in Örebro. Thank you for recommending me to PhD studies and for all academic journeys and encouraging support. At the halfway of my doctoral studies, I gratefully acknowledge valuable comments on my Licentiate Thesis from Bjarne Brendstrup.

I am very thankful to Professor Jean-Francois Richard for the opportunity to spend a period of my research at the Department of Economics, University of Pittsburgh. The interesting discussions on the analysis of auctions have been of considerable help for my future research. Thanks also to Professor Patrick Bajari for an instructive shorter stay at the Department of Economics, University of Minnesota.

Thanks also to Camilla Villani for carefully collecting the eBay data.

Without financial support this thesis had never been written. I gratefully acknowledge the support from the Department of Statistics, Stockholm University, and the support from Helge Ax:son Johnsons stiftelse for my stay in Pittsburgh.

To all my friends and colleagues, former and present at the Department of Statistics, Stockholm University. Thank you for being such a nice group of people that are always willing to help in numerous ways. In particular, I would like to thank my friend Pär for many joyful moments and Ellinor, Daniel, Karin, Tea, Feng, Nicklas, and my fellow doctoral students for nice friendship.

I also appreciate the nice friendships from friends outside the department, including Bengan, Håkan, Ricard, Figge, Johne, Martin, Andy, Chrille, Steffo, Spiken, Niklas, Mika, Stefan A, Stina, Anders, Tobbe etc. Special thanks to Stefan for excellent friendship and great interest in my research, and to Isac for excellent friendship and your very pleasant visit in the states during my time in Pittsburgh.

Finally, all the love and support from my family are and have been invaluable for me through all types of weathers. Thanks to my sisters Bella and Elli with families for encouraging support, and to my parents, Eva and Lars, who always believe in me, support me and follow my career in the best of manners.

Contents

1 Introduction

1.1 Historical background

Auctions have been used since antiquity for the sale of a variety of objects. Ancient Romans used auctions in commercial trade to liquidate property and estate goods. It is not known if the bidding process was increasing or decreasing. However, since the word "actus" in Latin means increasing, it is assumed that auctions were held in an increasing fashion. In 193 A.D. the Praetorian Guard sell off the entire Roman Empire by means of an auction.

Nowadays, auctions are widely used in many areas and account for a huge volume of economic transactions. Private firms sell products including real estates, fresh flowers, fish, houses and cars. Governments offer contracts through procurement auctions and every week sell foreign exchange, bills and bonds through auctions. During the last decade, Internet auctions have gained wide popularity where billions of dollars are turned over every day. Probably the biggest auction in the world is currently the keyword search auctions on Google.

An auction can be summarized as a bidding mechanism, described by a set of auction rules that specifies how the winner is determined and how much the bidder has to pay. There exists many different auction rules, but four basic types of auctions, where the object is awarded to the bidder with the highest bid, are particularly common and referred to as standard auctions. They are divided into open and sealed-bid auctions. The open auctions inlude the ascending-bid or English auction and the descending-bid or Dutch auction, while the first-price and second-price auctions are sealed-bid auctions. In oral auctions, like the English auction, bidders note each other's bids and can make counteroffers. In sealed-bid auctions the bidders submit only one bid simultaneously without revealing them to others.

The English auction is the oldest auction form and typically starts with low bids and increases in small predetermined portions until only one bidder is left. The Dutch auction is the counterpart to the English auction. Here the auctioneer begins at a usually high price and gradually lowers it until someone makes a sign to claim the item at the current price. In both first-price and second-price auctions the bidder with the highest bid wins, but in a second-price auction the winning bidder pays an amount equal to the next highest bid in contrast to the payment of the highest bid in first-price auctions.

Since the pioneering work of Vickrey (1961) the theory of auctions as games of incomplete information has developed extensively, especially over the last decades. There has been a number of economic theorists that made a considerable amount of work in understanding the factors influencing auction prices for different types of goods. Klemperer (1999) outlines practical, empirical and theoretical reasons why auction theory is important. First, auctions constitute a market with a huge volume of economic transactions. This is especially true with the advent of Internet auctions in recent years where high-quality datasets are readibly available. Second, auctions are simple and can be explicitly modeled with well-defined game-theoretic forms that provide a very valuable platform for testing economic theory. Finally, through much fundamental theoretical work, auction theory has been important as a tool of understanding other methods of price formation in other competetive markets.

1.2 Valuations

Wolfstetter (1996) states that auctions are essentially used for rapid sales, to reveal the information about buyers' valuations and to prevent dishonest dealing between the seller's agent and the buyer. A key feature is the asymmetries in information. If the seller would knew the bidders' valuations, he could just offer the object to the bidder with the highest valuation at a price just below what the bidder is willing to pay. The valuations of the bidders are often classified into one of two standard paradigms: the independent private-value or the pure common-value paradigm. Within the private-value model each bidder knows his valuation (the value) and knowledge of other bidders' valuation would not affect a bidder's valuation. This is a reasonable model if the object for sale is used for consumption, e.g. a piece of furniture, a painting or private collectibles.

In the pure common-value model, bidders use their own private information to estimate the unknown value of the object that is the same for all bidders. Typical examples include objects that are derived from an unknown market price at the time of bidding, e.g. for the sale of oil contracts with an unknown amount of oil. The common-value model is a special case of a general specification called one of interdependent values. In this general setting, the bidders have only partial information regarding the value, which may be different for different bidders, and would be affected of knowing the information that other bidders posess. Each bidder uses his private information to estimate the value of the object.

1.3 Bid equilibrium

A bidder's strategy can be defined as a mapping from a conceived value of the object to a final bid. The strategic mapping from a bidder's value to a bid is the so called bid function. Nash (1951) proved for a general, finite non-cooperative game that there always exists at least one equilibrium point. In the setting of auctions this means that there always exist an equilibrium strategy that maps a bidder's value to an equilibrium bid. The consequence of the bid equilibrium is that no bidder (player) in the auction (game) can succeed with a better strategy given the strategy of the other players.

In auction theory, the equilibrium bidding strategies depend on the type of auction and the nature of the buyers and seller(s) in the auction. In this thesis, we typically consider second-price auctions with a symmetric equilibrium, in which all bidders follow the same strategy. We also assume that bidders are risk-neutral, which means that each bidder seeks to maximize his expected utility by maximizing his expected profit. The profit is defined as the difference between the bidder's value of the object and the bidder's payment. Risk-neutral bidders are commonly assumed in the literature and is a special case of risk-averse bidders, where the bidders seek to maximize their expected utility functions.

In second-price auctions with private-value bidders, Vickrey (1961) showed that it is a dominant strategy for a bidder to bid his value. The strategic problems for common-value bidders in second-price auctions are much harder, since their mapping from a value to a bid depends on the other bidders' distribution of values. The bidder faces a simple trade-off. Increasing the bid increases the probability of winning, while at the same time decreases a potential profit if the bidder wins. The expected profit for a bidder is calculated as the size of the profit from winning times the probability of winning. By maximizing the expected profit with respect to the bidder's bid, the equilibrium bid function can be derived.

In an influential article, Milgrom and Weber (1982) derive the equilibrium bid function for a symmetric second-price common value auction. In general, common value models are much more technically challenging than the models of private values. This makes it, in practice, difficult to specify distributional assumptions of valuations that yield closed-form solutions of the equilibrium bid function or at least neat implicit forms. A handful closed-form solutions have been derived, but mostly for highly specialized models, see e.g. Kagel and Levin (1986), Matthews (1984), and Levin and Smith (1991).

The lack of closed form solutions has two major drawbacks. First, it is hard to see how the bid function depends on various distributional components of the model, which makes it more difficult to bring out model characteristics. Second, to evaluate the bid function one has to make use of numerical integration, which is very time demanding. This is a crucial step for econometric analysis of auction data (e.g. likelihood/Bayesian estimation) where the equilibrium bid function has to be evaluated over and over again. Bajari and Hortacsu (2003) reduce the computational complexity significantly in their model by exploiting a linearization property, but the inverse bid function in the very complicated likelihood function still needs to be evaluated by time-consuming numerical integration.

1.4 Revenue equivalence principle

At the late seventies the major contributions came in the mechanism-design field of auction theory. Roughly during the same time, independent of each other, Myerson (1981) and Riley and Samuelson (1981) generalized Vickrey's results about the equivalence in expected seller revenue for many different auctions, including the four standard auctions. As Klemperer (1999) mention, in his broad survey of the literature in auction theory, the theorem is so fundamental that any reader who is unfamiliar with the result is strongly urged to learn it.

In short, the theorem of *revenue equivalence principle* can be described as follows. Assume that risk-neutral, private-value bidders draw their values independently from the same distribution. Then, any symmetric and increasing equilibrium of any standard auction yields the same expected revenue to the seller. The counterpart of this principle to interdependent values and affiliated (correlated) signals is called the *revenue ranking (linking) principle*, where the highest expected seller revenue is obtained in English auctions, followed by the second-price and first-price auctions.

1.5 The winner's curse

The winner's curse is by far the most highlighted phenomenon in common value auctions. The concept origins from the *curse* of winning the auction when the price exceeds the unknown market value at the time of bidding. Wilson (1969) introduced the common-value model and developed the first closed-form equilibrium analysis of the winner's curse. However, it was the three Atlantic Richfield engineers, Capen, Clapp, and Campell (1971) that introduced the name of the concept. They found out that oil firms suffered from the winner's curse during the oil era in the Gulf of Mexico after the 1950s, when business had paid off less than expected. Later, Thaler (1988) views the empirical results of the winner's curse as an anomaly in his comprehensive discussion of the concept.

The winner's curse can be clarified by the following countervailing trade-offs. Bidders in common-value auctions face effects from both competition and information perspectives. More bidders introduce more competition that gives a bidder incentives to submit a higher bid (*competition effect*). However, a bidder must also account for the risk of overestimating the value of the object if he wins, since his signal is then the highest signal among bidders. This implies that a bidder should also lower his bid when facing more bidders (*overestimation effect*). In equilibrium, the overestimation effect is always larger than the competition effect and bidders correct for the winner's curse by lowering their bids as the number of bidders increases (Krishna, 2002).

1.6 Asymmetric bidders

In recent years, auctions with asymmetric bidders have been actively studied. Asymmetry is present when at least one of the main assumptions in the modeling of auctions is dropped, e.g. assuming risk-averse bidders instead of risk-neutral, relaxing the assumption of independently drawn valuations, and allowing for both private-value and common-value components. In the mechanism-design literature of auction theory, Maskin and Riley (1985) bring out many key ideas by weakening the main assumptions on the nature of bidders and focusing on only two bidders with private values. In addition, Maskin and Riley (2000) analyze asymmetric auctions by distinguishing between *weak* and *strong* private-value bidders. Apart from private values, Goeree and Offerman (2002,2003) and Jackson (2009) analyze auctions with both private-value and common-value components, and Reny and Zamir (2004) prove the existence of equilibria in general asymmetric first-price auctions with interdependent values. In second-price auctions with both private-value and common-value bidders, Tan and Xing (2011) prove the existence of a monotone pure-strategy equilibrium.

1.7 Structural econometric auction models

Reiss and Wolak (2007) give a broad introduction to the logic of structural econometric models, including models for auctions, and compare them to other types of econometric models. Over the last decades the structural estimations of auction data have become increasingly popular. Laffont and Vuong (1996) came with major contributions in this field and emphasize that auction models are particularly suited for structural estimation, where many datasets are readily available and well-defined game forms exist.

Bajari and Hortacsu (2005) mention three conditions that must apply for structural estimation of auction data. First, the bidders' goal is to maximize their expected utility. This is basicly an assumption of rational bidders. If the bidders are risk-neutral they maximize their expected profits. Second, bidders are able to compute the relationship between their bid and the probability of winning the auction. That is, they are able to compute the optimal combination of the probability of winning and the amount of the profit if they win. Third, given their beliefs, bidders are able to correctly maximize their expected utility.

These assumptions of rationally are quite strong, but there exists a number of papers that test for necessary conditions. Guerre, Perrigne, and Vuong (2000) point out that a necessary condition for rationality in private value auction models is, in principal, to the test if the bid function is increasing in values. Paarsch and Hong (2006) survey the field of structural econometrics of auction data.

1.8 Internet auctions

Recently, over the last decade, Internet auctions have gained wide popularity. Bajari and Hortacsu (2004) argue in their survey of online auctions that auctions on the Internet grow at an impressive pace and are one of the most successful forms of electronic commerce. Lucking-Reiley (2000) survey 142 online auctions and estimate eBay as the world's largest auction site by far. At eBay, millions of items are sold every day in thousands of categories from which high-quality datasets become available to buyers and sellers through completed auction listings.

To explore the determinants of bidder and seller behaviour, Bajari and Hortacsu (2003) examine a dataset of coin auctions from eBay. According to several empirical findings for auctions with a fixed end time, e.g. Wilcox (2000) and Ockenfels and Roth (2006), bids tend to arrive very late in these auctions. In the spirit of Wilson (1977), Bajari and Hortacsu (2003) show that late bidding in their independent symmetric common value model of eBay auctions is a symmetric Nash equilibrium. In this environment each bidder is assumed to place only one bid in the very last minute of the auction, so that no other bidders have time to revise their bids. As a consequence, they estimate eBay auctions as independent second-price common value auctions.

2 Summary of the papers

2.1 Paper I: Bayesian Inference in Structural Second-Price Common Value Auctions

Structural econometric modeling of auction data has become increasingly popular in recent years, especially with the advent of high-quality datasets from Internet auctions. Bajari and Hortacsu (2003) made a number of important contributions to the field by simplifying the analysis of common value auction models with a stochastic number of bidders. They proved that a symmetric Nash equilibrium exists in their study of eBay auctions, which allowed them to model eBay coin auctions as independent second-price common value auctions. The analysis of common value models is well-known for being technically challenging under realistic model assumptions.

In this paper, we refine and extend the analysis in Bajari and Hortacsu (2003). An important obstacle in their model is the need for numerical integration to solve the equilibrium bid function. This is very time-consuming since the likelihood function of bids needs to be evaluated many times for inference. A key contribution in our approach is a very accurate approximation of the equilibrium bid as a linear function of the bidder's signal, giving fast and numerically stable evaluations of the likelihood function.

We use both simulated and real data to analyze our model. The real data was carefully collected by human inspections and contains bids and auction-specific characteristics from 1050 eBay coin auctions. To estimate the data we use an efficient Bayesian framework for variable selection that brings out the posterior probability of including a given covariate in the model. We find the publicly available book value and the condition of the auctioned object as the main determinants for bidders' valuations, whereas the eBay's detailed seller information is essentially ignored by the bidders. We also show that our approximate bid function does not distort inference for a number of economic implications, such as the bidders' correction for the winner's curse by lowering their bids. Finally, we document good out-of-sample predictions of auction prices.

2.2 Paper II: Bayesian Inference in Structural Second-Price Auctions with Gamma Distributed Common Values

The valuations in auction models are intrinsically non-negative. Nevertheless, distributional assumptions are often used in the literature that allow for negative values. In our paper, we explore this issue by proposing an extension of the Gamma model in Gordy (1998) as an alternative to the Gaussian modeling of valuations in Wegmann and Villani (2011, henceforth WV). Similar to the Gaussian model in WV, a key contribution for the Gamma model is a very accurate approximation of the equilibrium bid function. The approximate bid function is non-linear, but since the approximate inverse bid (the approximate signal) is an explicit function of a given bid, we also obtain fast and numerically stable likelihood evaluations for the Gamma model.

The Bayesian framework with variable selection and the eBay coin auction dataset in WV are used to compare the performances between the Gaussian and Gamma models. The posterior results are quite similar between the models. The Gaussian model fits the data and predicts auction prices slightly better than the Gamma model. We find evidence that this is probably due to an almost normal or at least symmetrical distribution of valuations, where the density only attains small probabilities of negative values. Finally, we document for simulated datasets with different degrees of skewness that the superiority of the Gamma model for highly skewed data diminish when the value distribution becomes more symmetrical. These are findings that agree with the results for the eBay dataset.

2.3 Paper III: Bayesian Comparison of Private and Common Values in Structural Second-Price Auctions

Most of the literature in auction theory focus on either the private value or common value paradigm. Attempts to distinguish between the paradigms have been extensive during the last couple of decades, but mostly within first-price auctions. We compare the Gaussian model for common values in WV to a directly comparable model for private values in second-price auctions. The comparison is performed in numerous ways by using the Bayesian framework on the eBay coin auction data in WV.

Both models fit the data well with a slight edge for the more robust common-value model. The private-value model is better in predicting auction prices, but the more complex common value model is more robust in predicting some auctions. We find many interesting empirical regularities between the models. First, the winner's curse effect in common values explains the differences in the estimates for the expected values. Second, the optimal minimum bids in the common value model are much closer to the actual minimum bids than the optimal choice of zero or close to zero for the private value model. Third, we find no evidence of the winner's curse in the data, since the average bids do not decrease for an increasing number of bidders. Fourth, the models seem to capture the correlation between bids equally well.

Our results indicate that the value of the object probably includes both a private and a common value component, since we find evidence for both private and common values in different ways. However, auction models with a combination of private and common values have not yet seen the light in the literature. It is indeed a promising area of future research.

2.4 Paper IV: Bayesian Inference in Structural Second-Price Auctions with both Private-Value and Common-Value Bidders

Auction models with asymmetric bidders have received much attention in recent years. Tan and Xing (2011) show the existence of a montone pure-strategy equilibrium for symmetric second-price auctions with both private-value and common-value bidders. In their setting, the dominant strategy for a private-value bidder is to bid his value, while the bid equilibrium for a common-value bidder is the solution to an ordinary differential equation (ODE) that depends on the parameters in the private-value distribution. To solve the highly complicated ODE one needs to resort to numerical integration methods which are too time-consuming to be used for statistical inference.

We assume the model in Tan and Xing (2011) and derive a very accurate approximation of the equilibrium bid function. The approximate inverse bid is an explicit function of a given bid, which virtually takes no time to evaluate, giving fast and numerically stable evaluations of the likelihood function. We use Bayesian methods to evaluate the model on the eBay coin auction data in WV. Since we do not model auctions with a minimum bid, we use data from 464 auctions where the minimum bids have a negligable effect on the bidding process.

We propose a model where the probability of being a common-value bidder is a function of auction-specific covariates. An interesting feature of this modeling is the possibility to make inference, through Bayesian variable selection, on the probability of being a private-value or a common-value bidder in a given auction. We use a Metropolis-within-Gibbs algorithm to sample from the posterior in our Bayesian inference. Our main findings are that the empirical results for the common-value distribution are essentially the same as the results for a model with only common-value bidders, whereas the estimates of the parameters in the private-value distribution are more affected by the influx of common-value bidders. Finally, there is a slightly larger probability of being a common-value bidder, but this probability does not seem to depend on the covariates at the auction level.

References

[1] Bajari, P., and Hortacsu, A. (2003), The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *The Rand Journal of Economics*, **3**(**2**), 329-355.

[2] —– (2004), Economic Insights from Internet Auctions, *Journal of Economic Literature*, **42**, 457-486.

[3] —– (2005), Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data, *Journal of Political Economy*, **113**(**4**), 703-741.

[4] Cassady, R. Jr. (1967). *Auctions and Auctioneering*, Berkeley and Los Angeles: University of California Press.

[5] Capen, E. C., Clapp, R. V., and Campbell, W. M. (1971), Competitive Bidding in High-Risk Situations, *Journal of Petroleum Technology*, **23**, 641-53.

[6] Goeree, J., and Offerman, T. (2002), Efficiency in Auctions with Private and Common Values: An Experimental Study, *American Economic Review*, **92**(**3**), 625-643.

[7] —– (2003), Competitive Bidding in Auctions With Private and Common Values, *Economic Journal*, **113**, 598-614.

[8] Gordy, M. (1998), Computationally Convenient Distributional Assumptions for Common-Value Auctions, *Computational Economics*, **12**, 61-78.

[9] Guerre, E., Perrigne, I., and Vuong, Q. (2000), Optimal Nonparametric Estimation of First-Price Auctions, *Econometrica*, **68**(**3**), 525-74.

[10] Jackson, M. (2009), Non-existence of Equilibrium in Vickrey, Second-price, and English auctions, *Review of Economic Design*, **13**, 137-145.

[11] Kagel, J.H., and Levin, D. (1986), The Winner's Curse and Public Information in Common Value Auctions, *American Economic Review*, **76**(**5**), 894-920.

[12] Klemperer, P. (1999), Auction Theory: A Guide to the Literature, *Journal of economic surveys*, **13**(**3**).

[13] Krishna, V. (2002). *Auction Theory*, San Diego: Academic Press.

[14] Laffont, J.J., and Vuong, Q. (1996), Structural Analysis of Auction Data, *The American Economic Review*, **86**(**2**), 414-20.

[15] Levin, D., and Smith, J. (1991), Some Evidence on the Winner's Curse: Comment, *The American Economic Review*, **81(1)**, 370-375.

[16] Lucking-Reiley, D. (2000), Auctions on the Internet: What's Being Auctioned, and How?, *Journal of Industrial Economics*, **48(3)**, 227-252.

[17] Maskin, E., and Riley, J. (1985), Auction Theory with Private Values, *American Economic Review*, **75**, 150-55.

[18] —— (2000), Asymmetric Auctions, *Review of Economic Studies*, **67**, 413-438.

[19] Matthews, S. (1984), Information Acquisition in Discriminatory Auctions, In Boyer, M. and R.E. Kihlstrom (eds), *Bayesian Models in Economic Theory*. Vol. 5 of *Studies in Bayesian Econometrics*. Elsevier Science Publishers, Amsterdam, pp. 181-207.

[20] Milgrom, P., and Weber, R. (1982), A Theory of Auctions and Competitive Bidding, *Econometrica*, **50**, 1089-1022.

[21] Myerson, R. (1981), Optimal Auction Design, *Mathematics of Operations Research*, **6**, 58-73.

[22] Nash, J. (1951), Non-Cooperative Games, *Annals of Mathematics*, **54(2)**, 286-295.

[23] Ockenfels, A., and Roth, A. E. (2006), Late and multiple bidding in second price Internet auctions: Theory and evidence concerning different rules for ending an auction, *Games and Economic Behaviour*, **55**, 297-320.

[24] Paarsch, J. H., and Hong, H. (2006). An Introduction to the Structural Econometrics of Auction Data, Cambridge, MA: MIT Press.

[25] Reiss, P., and Wolak, F. (2007), Structural Econometric Modeling: Rationales and Examples from Industrial Organization, In *Handbook of Econometrics*, Elsevier, edition 1, volume 6, number 6a, Heckman, J.J., and Leamer, E.E. (ed.), chapter 64.

[26] Reny, P., and Zamir, S. (2004), On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-price Auctions, *Econometrica*, **72**, 1105-1125.

[27] Riley, J., and Samuelson, W. (1981), Optimal Auctions, *American Economic Review*, **71**, 381-92.

[28] Tan, X., and Xing, Y. (2011), Auctions with both Common-value and Private-value Bidders, *Economics Letters*, **111**, 99-103.

[29] Thaler, R. (1988), Anomalies: The Winner's Curse, *Journal of Economic Perspectives*, **2(1)**, 191-202.

[30] Vickrey, W. (1961), Counterspeculation, Auctions, and Competitive Sealed Tenders, *Journal of Finance*, **16**(**1**), 8-37.

[31] Wilcox, R. (2000), Experts and Amateurs: The role of Experience in Internet auctions, *Marketing letters*, **11**(**4**), 363-74.

[32] Wilson, R. (1969), Competitive Bidding with Disparate Information, *Management Science*, **15**, 446-48.

[33] —— (1977), A Bidding Model of Perfect Competition, *Review of Economic Studies*, **44**, 511-18.

[34] Wolfstetter, W. (1996), Auctions: An Introduction, *Journal of Economic Surveys*, **10**, 367-420.

BAYESIAN INFERENCE IN STRUCTURAL SECOND-PRICE COMMON VALUE AUCTIONS

BERTIL WEGMANN AND MATTIAS VILLANI

Abstract. Structural econometric auction models with an explicit game-theoretic modeling of bidding strategies have been quite a challenge from a methodological perspective, especially within the common value framework. We develop a Bayesian analysis of the hierarchical Gaussian common value model with stochastic entry introduced by Bajari and Hortacsu. A key component of our approach is an accurate and easily interpretable analytical approximation of the equilibrium bid function, resulting in a fast and numerically stable evaluation of the likelihood function. We use a Bayesian variable selection algorithm that simultaneously samples the posterior distribution of the model parameters and does inference on the choice of covariates. The methodology is applied to simulated data and to a newly collected dataset from eBay with bids and covariates from 1000 coin auctions. We demonstrate that the Bayesian algorithm is very efficient and that the approximation error in the bid function has virtually no effect on the model inference. The model fits the data well and we document good out-of-sample predictions of auction prices.

Keywords: Bid function approximation, eBay, Internet auctions, Likelihood inference, Markov chain Monte Carlo, Normal valuation, Variable selection.

1. Introduction

Strategic bidding behavior in auctions has been a widely studied phenomenon since the pioneering work of Vickrey (1961), particularly over the last few decades (see, e.g., Wolfstetter 1996; Klemperer 1999, 2004; and Milgrom 2004 for recent surveys and a general introduction). The advances in auction theory have also found their way into the econometric analysis of auction data. It seems widely accepted that an explicit modeling of bidders' strategic considerations is a necessary condition for making economic sense of the observed patterns in the bids. The availability of high-quality auction data has increased in recent years, especially with the advent of Internet auction sites, such as eBay (see, e.g., Bajari and Hortacsu 2004 for a survey). Paarsch (1992) and Elyakime, Laffont, Loisel and Vuong (1997) have provided excellent examples of structural econometric analyses of auction data. Bajari (2005) and Paarsch and Hong (2006) have surveyed the field.

Analyzing auction data through the lens of a structural game-theoretic model is not an easy task. It has been quite a challenge to derive the strategic equilibrium bid function (i.e., the map from a bidder's conceived or estimated value of the object to his optimal bid) under realistic model assumptions. When such results are available, they typically come in a form unamendable to analytical computations, and one needs to resort to time-consuming and possibly unstable numerical methods, such as numerical integration. This is an important

Wegmann: *Department of Statistics, Stockholm University, SE-106 91 Stockholm. E-mail: bertil.wegmann@stat.su.se.* Villani: *Research Division, Sveriges Riksbank, SE-103 37 Stockholm, Sweden* and *Department of Statistics, Stockholm University. E-mail:mattias.villani@riksbank.se.* The authors thank Bjarne Brendstrup for helpful comments on a previous version of the paper and Camilla Villani for carefully collecting the eBay data.

obstacle to statistical inference as a single likelihood evaluation typically requires a repeated evaluation of the inverse bid function for all bids in the dataset.

Common value auctions have been especially difficult to analyze by structural econometric models. In common value auctions, the auctioned object has the same value to every bidder, but the common value is unknown. The bidders use private information (their signal) to infer the unknown value. In an influential work, Bajari and Hortacsu (2003; henceforth BH) made a number of important advances that substantially simplify the analysis of data from common value auctions. BH proved that it is sufficient to compute the bid function in a selected auction and then extrapolate linearly to the other auctions in the dataset. This property considerably speeds up the computation of the bid function, and thereby also likelihood evaluations. BH also show that it is optimal to place a bid in the very last seconds of commonly used Internet auction formats, such as eBay's. This, in turn, implies that Internet auctions can be modeled as sealed-bid auctions without additional strategic considerations of the timing of the bids. BH also extended the results of Milgrom and Weber (1982) to the situation with a stochastic number of bidders, an inherent feature of Internet auctions.

The present paper refines and extends the analysis of BH. Our first contribution is an accurate linear approximation of the equilibrium bid function for cases with a fixed number and a stochastic number of bidders. The approximate bid function is of a particularly simple analytical form with an interesting interpretation. It can be inverted and differentiated analytically, two extremely valuable properties for fast and numerically stable evaluations of the likelihood function. The approximation can be simultaneously evaluated for all bids in all auctions in a negligible amount of computing time.

An interesting aspect of the BH model is the use of auction-specific covariates, both in the model for the common value and in the stochastic entry process. Our second contribution is the use of a highly efficient general posterior sampling algorithm that simultaneously approximates the joint posterior distribution of the model parameters and does Bayesian variable selection among the covariates of the model. This allows us to quantify the importance of the individual covariates in the different parts of the model, and to correctly account for the uncertainty in the choice of covariates in, for example, the predictive distribution of the price. Bayesian variable selection also makes it possible to use a large number of covariates in the model because it typically reduces the dimensionality of the parameter space dramatically in every step of the Metropolis-Hastings algorithm (see Section 3.2).

Finally, we apply the methodology to a newly collected dataset with bids and auction-specific information from 1000 eBay coin auctions. The dataset was collected by a careful human visual inspection of both photos of the auctioned object and the seller's text description. We provide a partial reduced form analysis as well as structural estimates, and show that the structural model fits the data well. We also document good out-of-sample predictions of auction prices on 50 auctions that took place after the auctions in the estimation sample. The variable selection shows that the publicly available book value and the condition of the auctioned object are important determinants of bidders' valuations, whereas eBay's detailed seller information, such as bidders' subjective ratings of sellers and sellers' historical selling volumes, is essentially ignored by the bidders. The seller's posted minimum bid acts as a safeguard for the seller, to avoid large losses. We show that it is typically optimal for the seller to post a minimum bid only slightly below the seller's valuation of the object, despite the fact that a high minimum bid discourages auction entry. The estimation results are shown to be robust to a variety of modifications of the basic model.

2. A model for second price common value auctions

2.1. **General setup.** Assume that the seller sets a publicly announced *minimum bid (reservation price)*, $r \geq 0$, and that risk-neutral bidders compete for a single object using the same bidding strategy (symmetric equilibrium). The value of the object, v, is unknown and the same for each bidder at the time of bidding, but a prior distribution for v is shared by the bidders. To estimate v, each bidder relies on his or her own private information of the object modeled as a private signal, x, from a distribution, $x|v$, that is the same for all bidders (symmetric bidders). Let $f_v(v)$ denote the probability density function of v, let $f_{x|v}(x|v)$ denote the conditional probability density function of $x|v$, and let $F_{x|v}(x|v)$ denote the conditional cumulative distribution function of $x|v$. Since the auction involves symmetric bidders and a symmetric equilibrium, we can focus on a single bidder without loss of generality. The bid function can be written (BH) as

$$b(x,\lambda) = \frac{\sum_{n=2}^{\infty}(n-1)\cdot p_{n-1}(\lambda)\cdot \int_v v\cdot F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv}{\sum_{n=2}^{\infty}(n-1)\cdot p_{n-1}(\lambda)\cdot \int_v F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv},\ \text{if } x \geq x^\star \tag{2.1}$$

and 0 otherwise, where $p_{n-1}(\lambda)$ is the Poisson probability of $(n-1)$ bidders in the auction with λ as the expected value in the Poisson entry process. Bidders participate with a positive bid if their signal, x, is above the cutoff signal level, $x^\star$. Given an arbitrary bidder with signal x, let y be the maximum signal of the other $(n-1)$ bidders. The cutoff signal level is then given in implicit form as (Milgrom and Weber 1982)

$$x^\star(r,\lambda) = \inf_x \left(E_n E[v|X=x, Y<x, n] \geq r\right),$$

which gives the minimum bid, r, as

$$r(x^\star,\lambda) = \sum_{n=1}^{\infty} p_n(\lambda)\cdot \frac{\int_v v\cdot F_{x|v}^{n-1}(x^\star|v)\cdot f_{x|v}(x^\star|v)\cdot f_v(v)\ dv}{\int_v F_{x|v}^{n-1}(x^\star|v)\cdot f_{x|v}(x^\star|v)\cdot f_v(v)\ dv}. \tag{2.2}$$

The minimum bid is exogenously given by the seller and $x^\star$ is then given as the solution to (2.2).

Let v_j denote the common value in auction j, and let x_{ij} denote the signal of the ith bidder in auction j. Similar to BH, we use the following hierarchical Gaussian model

$$\begin{aligned}
v_j &\sim N(\mu_j, \sigma_j^2),\ \ j = 1,...,m,\\
x_{ij}\,|v_j &\overset{iid}{\sim} N(v_j, \kappa\sigma_j^2),\ \ i = 1,...,n_j,\\
\mu_j &= z'_{\mu_j}\beta_\mu\\
\sigma_j^2 &= \exp\left(z'_{\sigma_j}\beta_\sigma\right)\\
\lambda_j &= \exp\left(z'_{\lambda_j}\beta_\lambda\right),
\end{aligned} \tag{2.3}$$

where m is the total number of auctions, n_j is the number of bidders that bid 0 or place a positive bid in auction j, and $z_j = (z'_{\mu j}, z'_{\sigma j}, z'_{\lambda j})'$ are auction-specific covariates in the regression models for $(\mu_j, \sigma_j^2, \lambda_j)$ in auction j. In addition, let $\beta = (\beta_\mu, \beta_\sigma, \beta_\lambda)'$. BH modeled σ^2 with a linear function of covariates and thus needed to restrict the elements of β_σ to ensure that σ^2 is positive, whereas β_σ is unrestricted in our setup.

BH made the assumption that bids in parallel auctions are independent and showed that last-minute bidding is a symmetric Nash equilibrium on eBay. This allows us to model eBay auctions as independent second-price auctions. The likelihood function of bids is complicated, because some bids are unobserved. First, some bidders may draw a signal, $x < x^{\star}$, in which case they do not place a bid. Second, the highest bid is usually not observed because of eBay's *proxy bidding system* (see BH and Section 4.1 for more details). The bid distribution for a single auction is of the form:

$$f_b\left(b|\beta,\kappa,r,z,v\right) = f_{x|v}\left[\phi(b)|\beta,\kappa,r,z,v\right]\phi'(b), \tag{2.4}$$

where $\phi(b)$ is the inverse bid function. Let n_s be the number of bidders who submit a positive bid in a given auction and let $\mathbf{b} = (b_2, b_3, \ldots, b_{n_s})$ be the vector of observed bids, where $b_2 > b_3 > ... > b_{n_s}$. Then, the likelihood function for that auction is given by

$$\begin{aligned} &f_{\mathbf{b}}\left(b_2, b_3, \ldots, b_{n_s}|\beta,\kappa,r,z\right) \\ &= \sum_{i=n_s}^{\bar{N}} p_i(\lambda)\cdot\int_{-\infty}^{\infty} F_{x|v}\left(x^{\star}|\beta,\kappa,v\right)^{i-n_s}\cdot\left\{1-F_{x|v}\left[\phi\left(b_2\right)|\beta,\kappa,v\right]\right\}^{I(n_s\geq 1)} \\ &\quad\times\prod_{i=2}^{n_s} f_b\left(b_i|\beta,\kappa,r,z,v\right)\cdot f_v(v|\beta)dv, \end{aligned} \tag{2.5}$$

where $I\left(n_s \geq 1\right)$ is an indicator variable for at least one observed bid in the auction and $\bar{N}$ is an upper bound for the total number of potential bidders. Following BH, $\bar{N}$ is set to 30. If $n_s = 1$, then b_2 equals the minimum bid r.

We use Bayesian methods to estimate the model; see Section 3. A single evaluation of the posterior density (likelihood function) requires numerical integration to compute $b\left(x|\beta,r,z\right)$ in (2.1), followed by additional numerical work to invert and differentiate $b\left(x|\beta,r,z\right)$. The same applies to the computation of $x^{\star}$. This costly procedure needs to be repeated for each of the auctions in the dataset. BH cleverly exploited a linearity property of the bid function that confines a large portion of the numerical work to a single auction, which is then linearly extrapolated to the other auctions. Nevertheless, the likelihood evaluation suggested by BH is not sufficiently fast to allow its routine use for inference. Instead, we make use of an analytical approximation of the bid function (see the next section) that can be simultaneously evaluated for all bids in all auctions. This leads to much faster and numerically more stable likelihood evaluations.

Paarsch (1992) used an interesting method initially suggested by Levin and Smith (1991) to compute bid functions for a class of models, including a model with Gaussian valuations. This method assumes a diffuse prior for the auction value, which makes it harder to use an interesting covariate structure as in (2.3). Moreover, contrary to our approximation approach, the method of Levin-Paarsch-Smith (1991) is not easily extended to the case with stochastic auction entry.

2.2. **Approximation of the bid function.** We derive the following linear approximation of the bid function (see Appendix A for details):

$$b(x,\lambda) \approx c + \omega\mu + (1-\omega)x, \text{ if } x \geq x^{\star} \tag{2.6}$$

and 0 otherwise, where $c = -\frac{\sqrt{\kappa}\sigma\gamma\theta(\lambda-2)}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\omega = \frac{\frac{\kappa}{2}}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\theta = 1.96$ and $\gamma = 0.1938$. The cutoff signal can be similarly approximated by

$$x^{\star}(r,\lambda) \approx \frac{r - \sum_{n=1}^{\infty} p_n(\lambda)(\tilde{c}+\tilde{\omega}\mu)}{\sum_{n=1}^{\infty} p_n(\lambda)(1-\tilde{\omega})}, \tag{2.7}$$

where $\tilde{c} = -\frac{\sqrt{\kappa}\sigma\gamma\theta(n-1)}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$, and $\tilde{\omega} = \frac{\frac{\kappa}{2}}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$.

The approximate bid function has the following properties. First, it is a linear combination of the signal x and the publicly held value μ. The weight ω increases monotonically toward 1 as κ increases, which gives

$$b(x) \longrightarrow x \; if \; \kappa \longrightarrow 0, \quad \text{and} \quad b(x) \longrightarrow \mu \; if \; \kappa \longrightarrow \infty.$$

The higher the precision in signals, the more the bidders trust their private information and vice versa. A greater variance of the common value v implies a greater risk of drawing a large signal and thus a greater risk of overestimating the true value of the object, which is why bidders should lower their bids. The approximate bid function captures this effect well; increasing the value of σ leads to lower bids.

The approximation in (2.6) can also be used to derive the unconditional distribution of the bids, $\mathbf{b} = (b_1, ..., b_n)'$, in an auction with a given number of bids, n. It is straightforward to show that

$$\mathbf{b} \sim N\left[(c+\mu)\mathbf{1}_n, (1-\omega)^2\sigma^2\left(\kappa\mathbf{I}_n + \mathbf{1}_{n\times n}\right)\right], \tag{2.8}$$

where $\mathbf{I}_n$ is the identity matrix and $\mathbf{1}_n$ and $\mathbf{1}_{n\times n}$ denote the $n \times 1$ vector and the $n \times n$ matrix with 1s, respectively. Note that here we are ignoring the truncation that comes from $x^\star$. At a first glance, it may appear that κ is merely a factor that inflates the variance of the bids and is mainly estimated from the variance of the observed bids. Equation (2.8) shows that this is an overly simplistic view. First, the unconditional variance of the bids, $(1-\omega)^2(\kappa+1)\sigma^2$, obviously increases with κ via the factor $(\kappa + 1)$, but κ also affects $\text{var}(b)$ through ω in a nonlinear fashion that depends on λ. Second, κ determines the dependence between bids as the correlation between any pair of bids is $(1+\kappa)^{-1}$.

Figure 1 compares the exact and approximate bid functions graphically. The exact bid function is computed by numerical integration as in BH. The upper left subgraph displays the bid function and its approximation for a representative auction in the eBay dataset analyzed in Section 4. The representative auction is based on the median of the covariates in the eBay data, analyzed in Section 4, and the posterior mean of the model parameters. Rounded to the nearest integer, this gives $\kappa = 5, \mu = 22, \sigma = 9, \lambda = 4$, and $\frac{r}{\mu} = 0.5$. The other subgraphs are variations of the representative auction. The approximation of the bid function is very good in all four cases. It is not easy to assess the importance of the approximation errors in Figure 1 for practical work, but experiments in Section 3.2 and 4.3 show that inferences based on the approximate bid function are very similar to those obtained from the exact bid function.

2.3. **Discussion.** The approach in the previous subsection can be used to obtain an accurate, easily computable and interpretable approximate bid function for second-price common value auctions with Gaussian-distributed signals and values. However, it is obvious that our approach cannot be used for any arbitrary auction setup and valuation structure. To state the conditions under which it can be used, let us first define a family of distributions $\mathcal{P}$ to be *closed under multiplication* if for any $p_1, p_2 \in \mathcal{P}$ we have that $k \cdot p_1 \cdot p_2 \in \mathcal{P}$, where k is a constant. The success of the approach hinges upon i) that the distribution function $F_{x|v}(x|v)$ can be well approximated by the kernel of a density function $p_{x|v}(x|v)$, ii) that $p_{x|v}(x|v)$, $f_{x|v}(x|v)$ and $f_v(v)$ all belong to a family of distributions $\mathcal{P}$ that is closed under multiplication (with respect to v), and iii) that the kernel of any member of $\mathcal{P}$ can be integrated analytically. As we have shown, all three conditions hold for the Gaussian model, and we have also verified that they also hold when $f_{x|v}(x|v)$, $p_{x|v}(x|v)$ and $f_v(v)$ are all log normal. Moreover, since the distribution function $F_{x|v}(x|v)$ appears in other major auction formats, it is clear that the approach can at least in principle also be used to approximate the bid function in other auction setups.

3. Bayesian inference and variable selection

3.1. **Prior.** Bayesian inference combines the likelihood function in (2.5) with a prior distribution on the unknown model parameters. The numerical algorithms that we use for sampling from the joint posterior distribution (see next section) can be used with any prior. In this section, we propose a particular prior that can be used with very limited input from the user. Our prior for β_μ and β_σ is motivated by the fact that the common values $v = (v_1, ..., v_n)'$ are modeled as a heteroscedastic regression

$$v = Z_\mu \beta_\mu + \varepsilon,\ \varepsilon_i \sim N(0, \sigma_i^2), \tag{3.1}$$

where $\sigma^2 = (\sigma_1^2, ..., \sigma_n^2)' = \exp(Z_\sigma \beta_\sigma)$. We can now specify a g-prior (Zellner, 1986) for β_μ, conditional on β_σ (see Villani et al. 2009 for details)

$$\beta_\mu | \beta_\sigma \sim N[0, c_\mu (\tilde{Z}_\mu' \tilde{Z}_\mu)^{-1}] = N[0, c_\mu (Z_\mu' D Z_\mu)^{-1}],$$

where $D^{1/2} = \text{Diag}[\exp(-z_{\sigma_1}' \beta_\sigma / 2), ..., \exp(-z_{\sigma_n}' \beta_\sigma / 2)]$ and $c_\mu > 0$ is a scaling factor that determines the tightness of the prior. Setting $c_\mu = n$, where n is the number of auctions in the sample, makes the information in the prior equivalent to the information in a single auction (conditional on β_σ), which is a useful benchmark. The marginal prior for β_σ is also taken to be a g-prior

$$\beta_\sigma \sim N[0, c_\sigma (Z_\sigma' Z_\sigma)^{-1}].$$

Turning to the Poisson entry model, we use the following g-prior for β_λ

$$\beta_\lambda \sim N[0, c_\lambda (Z_\lambda' Z_\lambda)^{-1}].$$

We use an inverse-gamma prior for κ, $\kappa \sim IG(\bar{\kappa}, g)$, where $\bar{\kappa}$ is the prior mean of κ and g are the degrees of freedom.

We also allow for variable selection among the covariates by introducing point masses at 0 in the prior distribution on the regression coefficients, e.g.,

$$p(\beta_\sigma) = \begin{cases} N[0, c_\sigma (Z_\sigma' Z_\sigma)^{-1}] \text{ with probability } \pi_\sigma \\ \\ 0 \text{ otherwise,} \end{cases}$$

where π_σ is referred to as the prior inclusion probability. The user thus needs to specify the eight hyperparameters c_μ, c_σ, c_λ, $\pi_\mu, \pi_\sigma, \pi_\lambda$, $\bar{\kappa}$, and g. We typically set $c_\mu = c_\sigma = c_\lambda = c$ and $\pi_\mu = \pi_\sigma = \pi_\lambda = \pi$, thereby reducing the number of prior hyperparameters to four. Our application in Section 4 and simulations in the next section show that the posterior distribution and the variable selection inference are not overly sensitive to the exact choice of these prior hyperparameters, and that a good default value for c is $c = n$.

3.2. **A Metropolis−Hastings algorithm for variable selection.** It is clear from (2.5) that the likelihood function for second-price common value auctions is highly nonstandard and thus, the posterior distribution of the model parameters cannot be analyzed using analytical methods. The most commonly used algorithm for simulating from posterior distributions is the Metropolis−Hastings (MH) algorithm, which belongs to the Markov chain Monte Carlo (MCMC) family of algorithms (see, e.g., Gelman et al. 2004 for an introduction). At a given step of the algorithm, a proposal draw, β_p, is simulated from the proposal density $f(\beta_p | \beta_c)$, where β_c is the current draw of the parameters (i.e., the most recently accepted draw). The proposal draw, β_p, is then accepted into the posterior sample with probability

$$a(\beta_c \to \beta_p) = \min \left[1, \frac{p(\beta_p | y) / f(\beta_p | \beta_c)}{p(\beta_c | y) / f(\beta_c | \beta_p)} \right],$$

where $p(\beta|y)$ denotes the posterior density. If β_p is rejected, then β_c is included in the posterior sample. This sampling scheme produces (autocorrelated) draws that converge in distribution to $p(\beta|y)$. The $f(\beta_p|\beta_c)$ can in principle be any density, but for efficiency reasons it should be a fairly good approximation of the posterior density. One possibility is the *random walk Metropolis* algorithm, where $f(\boldsymbol{\beta}_p|\boldsymbol{\beta}_c)$ is multivariate normal density with mean $\boldsymbol{\beta}_c$ and covariance matrix $-c{\cdot}\boldsymbol{H}^{-1}$, where $\boldsymbol{H}$ is the Hessian matrix evaluated at the posterior mode and c is a scaling constant. This algorithm was used by BH. The random-walk Metropolis is a robust algorithm, but it is well known to be rather inefficient. Moreover, it is not easily extended to the case with variable selection. A more efficient alternative that can also be extended to variable selection is the *independence sampler* where $f(\boldsymbol{\beta}_p|\boldsymbol{\beta}_c)$ is the multivariate-t density $t(\hat{\boldsymbol{\beta}}, -\boldsymbol{H}^{-1}, h)$, where $\hat{\boldsymbol{\beta}}$ is the posterior mode of $p(\boldsymbol{\beta}|\boldsymbol{y})$, $\boldsymbol{H}$ is once more the Hessian matrix at the mode and h is the degrees of freedom. Here, the multivariate-t density is defined in terms of its mean and covariance matrix. We typically use $h = 10$ degrees of freedom, which we have found to work well. The posterior mode and the Hessian matrix can be easily obtained using a standard Newton−Raphson algorithm with a BFGS update of the Hessian matrix (Fletcher 1987).

Now consider setting a subset of the elements in $\boldsymbol{\beta} = (\boldsymbol{\beta}_\mu', \boldsymbol{\beta}_\sigma', \boldsymbol{\beta}_\lambda')'$ to 0 (any other value is also possible). In a regression situation, this is clearly equivalent to selecting a subset of the covariates. Let $\mathcal{J} = (j_1, ..., j_r)$ be a vector of binary indicators such that $j_i = 0$ iff the ith element of $\boldsymbol{\beta}$ is 0. We can view these indicators as a set of new parameters. For simplicity, we here assume that the elements of $\mathcal{J}$ are independent a priori with $\Pr(j_i) = \pi$ for all i, so that π is the prior probability of including the ith covariate in the model, but other priors for $\mathcal{J}$ can be handled just as easily. Appendix B describes in detail how this algorithm can be generalized to sample from the joint posterior distribution of the parameters and the variable selection indicators $\mathcal{J}$, all in a single MCMC run.

We use the mean acceptance probability and the inefficiency factor (IF) to measure the performance of the Metropolis−Hastings algorithm. The inefficiency factor is defined as $1 + 2\sum_{k=1}^{K} \rho_k$, where ρ_k is the autocorrelation at the kth lag in the MCMC chain for a given parameter and K is an upper limit of the lag length such that $\rho_k \approx 0$ for all $k > K$. The inefficiency factor approximates the ratio of the numerical variance of the posterior mean from the MCMC chain to that from hypothetical iid draws. Put differently, the IF measures the number of draws needed to obtain the equivalent result of a single independent draw. Thus, the presence of IFs close to unity indicates a very efficient algorithm.

We conducted a simulation study to evaluate the performance of the variable selection procedure, where we are particularly interested in exploring the sensitivity of the posterior inclusion probabilities to changes in prior hyperparameters $c = c_\mu = c_\sigma = c_\lambda$, and g. We set the prior inclusion probability to $\pi = 0.2$ and $\bar{\kappa} = 0.25$ throughout the simulations (see Section 4 for a sensitivity analysis with respect to π). Rather than setting the parameters in the data generating model to arbitrary values, we will here generate data from a model that mimics the estimated Gaussian model for eBay data in BH. We simulated 50 full datasets, each with 407 auctions, using the posterior mean estimates in BH as parameter values. The covariates in the model were simulated independently to mimic the summary statistics in Tables 1 and 2 in BH. The eBay auction model in BH for auction j, without a secret reserve price, can be written as (see BH for a description of the covariates)

$$\mu_j = \beta_1 \mathsf{BookVal}_j + \beta_2 \mathsf{Blemish}_j \cdot \mathsf{BookVal}_j - 2.18$$

$$\sigma_j = \beta_3 \mathsf{BookVal}_j + \beta_4 \mathsf{Blemish}_j \cdot \mathsf{BookVal}_j$$

$$log\lambda_j = \beta_5 + \beta_6 \mathsf{LogBookVal}_j + \beta_7 \mathsf{Negative}_j + \beta_8 \frac{\mathsf{MinBid}_j}{\mathsf{BookVal}_j}.$$

To check that the variable selection procedure assigns small posterior inclusion probabilities to insignificant covariates in the model, we include one superfluous covariate in each of μ_j, σ_j and λ_j, with each additional covariate drawn independently from the standard normal distribution. To speed up the computations, we use the approximate bid solution in our article. We have checked that the exact and approximate bid function gave very similar results by replicating the analysis on several simulated datasets. As an example, Table 1 displays nearly identical posterior results for a randomly chosen dataset. The empirical application on eBay coin auctions in our article also provides reassuring evidence that this approximation does not distort the inferences.

In Figure 2, posterior inclusion probabilities for each of the covariates in 50 simulated datasets are shown as box plots for different values of c. Other priors than $g = 4$ and settings with unequal values of c_μ, c_σ, and c_λ gave very similar results. As we can see in Figure 2, the inclusion probabilities for the most significant variables are all close to one and differ very little across the different priors. The inclusion probabilities for parameters β_4 and β_7 are low as these coefficients are quite close to zero in the data generating model. Figure 2 also shows that insignificant variables obtain a higher posterior inclusion probability when the prior is tighter around zero (i.e., when c is small because the cost of including a covariate with a small coefficient is lower when c is small).

We find a simple way of characterizing the posterior inclusion probabilities by *Bayesian t-ratios*, which we define as

$$t_{Bayesian} = \frac{\left|\hat{\theta}\right|}{s(\hat{\theta})},$$

where $\hat{\theta}$ is the posterior mode and $s(\hat{\theta})$ the approximate/asymptotic posterior deviation from the optimization of the posterior density. In Figure 3, the inclusion probabilities increase sharply around a threshold whose value depends on the prior hyperparameters (the dashed line in Figure 3 marks out the 1.96 threshold used in classical t-tests at a 5% significance level). As c decreases (tighter prior), the threshold moves to the left and the curve flattens out. Note that the usual rule $|t| > 1.96$ is overly generous in including covariates.

Finally, most of the Metropolis-Hastings runs gave an acceptance probability in the range 0.6-0.8, only a few below 0.1, and none below 0.25 for the prior $c_\mu = c_\sigma = c_\lambda = n$.

4. Application to eBay auction data

At eBay, millions of items are listed into thousands of categories and subcategories. EBay's search engine can be used to review recently completed auctions. The listings typically contain a detailed description of the item, the quantity sold, seller characteristics, reservation price, and the sequence of placed proxy bids and their timing. As explained in Section 2, the highest bid is typically not reported. Hence, high-quality datasets with detailed auction characteristics and all but the highest bid can be collected at eBay and be used for estimating auction models.

4.1. Description of the data.

Our dataset contains bid sequences and auction characteristics from 1000 eBay auctions of U.S. proof sets that ended between November 7 and December 19, 2007 and December 27, 2007 to January 22, 2008[1]. We exclude multi-unit objects, auctions

[1]U.S. proof sets can be defined as the specially packaged set of Proof coins issued annually and sold by the U.S. Mint.

with a *Buy It Now* option[2] and Dutch auctions. We also collected data from 50 additional auctions between January, 23 to January, 29 2008 that are used in Section 5.4 to evaluate the model's out-of-sample predictions of auction prices.

The bids recorded on eBay's Bid History Page are supposed to correspond to the final bids for each bidder. A careful inspection of the bids reveals, however, that some bids are only a tiny fraction of the object's book value, and cannot realistically represent serious final bids. Therefore, we will exclude the most extreme bids in our main analysis. A bid b is excluded if $b \leq \delta \cdot \min(\mathsf{BookValue}, \mathsf{Price})$, where $\delta = 0.25$ in the benchmark estimations (this excludes 107 bids from the 3742 bids in the sample). We also present results for the case where all bids are used in the estimation ($\delta = 0$) and for $\delta = 0.5$ (431 bids removed).

Appendix C gives a detailed description of the data, and Table 2 presents summary statistics. The last three columns of the table specify the covariates used in the models for μ, σ, and λ in the next sections.

4.2. **Preliminary analysis.** To verify that our model has appropriate functional forms for the regressions in μ, σ, and λ, we would like to plot these quantities against the (continuous) covariates. Naturally, parameters μ, σ, and λ are not known, but simple estimates can be obtained as follows.

The approximate bid solution in (3.3) as a function of (μ, σ) can be written as

$$b(x) = h\sigma + \omega\mu + (1-\omega)x,$$

where

$$h = -\frac{\sqrt{\kappa}\gamma\theta(\lambda-2)}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}, \text{ and } \omega = \frac{\frac{\kappa}{2}}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}.$$

Since $x|v \overset{iid}{\sim} N\left(\mu, \kappa\sigma^2\right)$, we have

$$b(x|v) \overset{iid}{\sim} N\left[\mu + h\sigma, (1-\omega)^2\kappa\sigma^2\right].$$

The likelihood function for μ and σ^2 is then obtained by integrating out v. It is now straightforward to verify that the maximum likelihood estimates of μ and σ^2, conditional on κ, are given by

$$(\hat{\mu}, \hat{\sigma}^2) = \left(\bar{b} - \frac{hs_b}{(1-\omega)\sqrt{\kappa}}, \frac{s_b^2}{(1-\omega)^2\kappa}\right),$$

where we make the usual correction of the sample variance, defined as $s_b^2 = \frac{\sum_{i=1}^{n}(b_i-\bar{b})^2}{n-1}$.

The estimates of μ and σ^2 are conditioned on $\kappa = 3.042$, the posterior mean in the benchmark structural model estimated in the next section. Other values for κ, e.g. the estimate in BH of $\kappa = 0.25$, did not have any major effects on the results. Note that the μ and σ estimates can only be computed for auctions with at least two bidders (if the highest bid is observed, otherwise three), so that the results in this subsection should be treated with care. Figure 4 displays scatterplots of $\hat{\mu}_j$ and $\ln \hat{\sigma}_j^2$ in auction j against the suitable transformations of the covariate $\mathsf{BookValue}$, and the relationship seems to be appropriately linear. Figure 4 also plots $\ln(n_j + 1)$, where n_j is the number of bidders with non-zero bids in auction j, against the continuous covariates in the entry model. The relationships are once more very close to linear.

As a precursor to the analysis of the full structural model in Section 5.3, we perform separate reduced form regressions with the estimates $\hat{\mu}$ and $\ln \hat{\sigma}^2$ in each auction as dependent variables, respectively. Table 3 (upper and middle parts) presents the results from a full

[2]In some auctions, buyers have the option to directly purchase the item at a certain price, in which case the auction is declared to be completed.

Bayesian analysis with variable selection using the algorithm in Section 4.2. Note that the continuous covariates have been de-meaned and a constant has been added to each model. As an example, BookStd is the covariate BookValue minus its mean. This reduces the correlation among the coefficients, which is beneficial for numerical stability. The same transformations are also used in the structural model in Section 5.3.

The covariates BookStd and interactions between book values and Unopen and MajBlem clearly belong in the model for $\hat{\mu}$, and their estimated coefficients are large and of the expected sign. Similar results were also obtained from a Bayesian Tobit regression with auction prices as the dependent variable[3], see Table 4. The strongest predictors in the regression for $\ln \hat{\sigma}^2$ are LBookStd (standardized ln(BookValue)) and the interaction between ln(BookValue) and Unopen.

Finally, the last part of Table 3 gives the result for a Poisson regression with the number of bidders with a non-zero bid as the dependent variable. Here, the covariates MinBidStd (standardized MinBid/BookValue), LBookStd, and the interactions LBook*ID, LBook*Unopen, LBook*MajBlem come out as the most relevant covariates.

The Metropolis-Hastings acceptance probabilities for all reduced form regressions were very high (70-90%), the inefficiency factors are close to unity, and the convergence was excellent.

4.3. **Estimation results from the structural model.** Tables 5 and 6 report the estimation results for the approximate and exact case of the structural model in equation (2.3). With a few exceptions, the parameter estimates, for both cases of the structural model and for the reduced form regressions in the previous section, are fairly similar. The estimated coefficients are mostly of the expected sign and are of reasonable magnitudes. The only major difference is for the estimate on Book*ID, which is highly significant in the structural model of the approximate case compared to the exact case and the reduced form regressions. These discrepancies do not seem to depend on the approximation since the reduced form regressions, which are based on the approximate bid function, give a highly insignificant result. The MH acceptance probability in the model with the exact bid function was only 10 % (compared to 38 % in the approximate case), which we partly attribute to the occasional instability in the numerical evaluation of the bid function in this case.

The covariates BookStd and LBookStd play a very central role in the models for μ and σ, respectively. The large negative sign for MinBidStd is explained by the fact that a higher minimum bid implies a higher cutoff signal, which reduces the number of positive bids. Overall, the posterior inclusion probabilities are either close to 0 or 1, which gives conclusive evidence on which covariates that are important for explaining valuations and participation in eBay auctions for common value objects. It is interesting to note that eBay's detailed seller information seems to be of little use to buyers: the covariates Power, ID and LargNeg (dummy for a large proportion of negative feedbacks from buyers) almost invariably have very small posterior inclusion probabilities. We experimented with other transformations of the negative feedback score and also transformations of the overall feedback score as substitutes for Power and ID, with unchanged results.

To check for the sensitivity of the prior hyperparameters, we repeated the estimations using several different priors. In Table 7, we use the estimated model in Table 5 as a benchmark and compare posterior means given various prior settings. Almost all parameter estimates are insensitive to changes in priors, especially in the model for λ. Notable changes only appear in the model for σ, where the estimated value of the constant increases and the estimated value of the parameter for LBookStd decreases with more prior information. The tighter

[3]In an auction where the object remains unsold, a Tobit regression uses the information that the price is smaller than the auction's minimum bid.

prior distribution around zero reduces the impact of LBookStd and the estimated constant is increased to compensate for this. The last column but one in Table 7 shows that varying the prior inclusion probability, π, does not affect the posterior mean estimates, as expected.

Finally, the last column of Table 7 gives the posterior mean estimates when the benchmark model is estimated using all bids ($\delta = 0$), see Section 5.1. The main difference in the results is the larger estimates of κ and σ, a natural consequence of the wider dispersion in bids when $\delta = 0$. Estimations with $\delta = 0.5$ (results not shown) reduced κ to 2.27. It should be noted, however, that the relation between κ and the unconditional variance of the bids is complicated by the bidders' strategic behavior. One way of seeing this is to look at the unconditional variance of the bids, $\text{var}(b) = (1-\omega)^2(\kappa+1)\sigma^2$, under the approximate bid function. $\text{var}(b)$ obviously increases with κ via the factor $(\kappa+1)$, but κ also affects $\text{var}(b)$ via ω (the weight placed on the prior mean of v) in a non-linear fashion that depends on λ. Figure 5 displays $\text{Std}(b)$ as a function of κ for $\sigma^2 = 1$, and for different values of λ. It is seen from Figure 5 that $\text{Std}(b)$ is fairly insensitive to κ over a large region. There are larger differences in $\text{Std}(b)$ across λ for a given κ.

In Table 8, we also check for the sensitivity of the posterior inclusion probabilities with respect to the prior hyperparameters. Overall, the small inclusion probabilities in the benchmark model tend to increase as the prior information becomes more precise (c decreases), which is exactly the same expected result that was established by simulation in Section 4. Once more the results are not overly sensitive to c. Furthermore, increasing the prior inclusion probability from 0.2 to 0.5 does not overturn the previous results on the variable selection inference.

4.4. **Model evaluations and predictions.** We will now evaluate the in-sample fit of the model by comparing the observed data to simulated data from the estimated benchmark model. Given the observed auction-specific covariates, we simulated 100 new complete datasets for each of 100 systematically sampled posterior draws of the model parameters. This gives us $10,000$ full datasets, each with bids from 1000 auctions.

Following BH, we compare the observed and simulated data through two summary statistics: within-auction bid dispersion and cross-auction heterogeneity. The within-auction dispersion is defined as the difference between the highest observed bid and the lowest bid divided by the auctioned item's book value, and cross-auction heterogeneity is investigated by histograms of the bids divided by the corresponding book value in that auction. As we can see in Figures 6 and 7, the observed within-auction bid dispersion and the cross-auction heterogeneity are very well captured by the model. This is in contrast to BH where the within-auction bid dispersion is severely under-estimated and the cross-auction bid dispersion is over-estimated.

A more severe test of the model is to evaluate its out-of-sample predictions. To test this, we use the estimated benchmark model in Section 5.3 to predict a fresh dataset of 50 additional auctions of *U.S.* proof sets, all completed in the week directly following the estimation sample. Given the covariates from these auctions, we simulated price distributions for each auction in a similar way as above. The parameters are drawn from the posterior distribution, which was computed by only using the previously analyzed 1000 auctions. Figures 8 and 9 display the predictive distributions for the 50 out-of-sample auctions. Note that these distributions have three components: i) a probability that the item is not sold (Pr(No)), ii) a point mass at the minimum bid (Pr(Min)), which is the price when there is a single bidder in the auction, and iii) a continuous price density when there is more than one bidder. The predictive price distributions look reasonable and capture the observed prices very well in most cases. This fact together with the good fit of the bid dispersion strongly indicates that our estimated eBay auction model is quite accurate in explaining seller and bidder behavior at eBay.

4.5. **Economic implications.** The winner's curse is by far the most highlighted phenomenon in common value auctions where bidders face effects from both information and competition perspectives.[4] More bidders lead to more competition which gives a bidder incentives to submit a higher bid (*competition effect*). However, a bidder must also account for the risk of overestimating the value if he wins, since his signal is then the highest signal among all bidders. As such, a bidder should also lower his bid when facing more bidders (*overestimation effect*). In equilibrium, the overestimation effect is always larger than the competition effect and bidders correct for the winner's curse by lowering their bids as the number of bidders increases (Krishna, 2002); see Figure 10 for an illustration based on the representative auction described in Section 3. The approximate counterpart to Figure 10 is given in Figure 11. Figure 11 for the approximate case is quite similar to the exact case in Figure 10, except for high signals located above a certain signal where the approximate bid functions intersect. However, such high signals lie almost three standard deviations above the expected value, and are therefore highly improbable. To investigate this more generally, let us look at the first derivative of the approximate bid function with respect to n. This gives

$$b_n'(x) = \frac{\gamma\kappa\left[x-\mu-\frac{\sigma\theta(2+\kappa)}{\sqrt{\kappa}}\right]}{2\left[\gamma(n-2)+1+\frac{\kappa}{2}\right]^2}. \tag{4.1}$$

It can be verified that the unconditional distribution of signals is given by $X \sim N(\mu,(\kappa+1)\sigma^2)$, and, as such, $x = \mu + d\sigma\sqrt{\kappa+1}$ represents signals that deviate from the expected value μ with d standard deviations. By replacing x by $\mu + d\sigma\sqrt{\kappa+1}$, the derivative of the approximate bid function in (5.1) can be written as

$$b_n'(x) = \frac{\gamma\kappa\sigma\sqrt{\kappa+1}\left[d-\frac{\theta(2+\kappa)}{\sqrt{\kappa(\kappa+1)}}\right]}{2\left[\gamma(n-2)+1+\frac{\kappa}{2}\right]^2},$$

which is positive for any κ if $d > \frac{\theta(2+\kappa)}{\sqrt{\kappa(\kappa+1)}} > \frac{\theta(2+\kappa)}{1+\kappa} > \theta = 1.96$. According to the estimation results in Table 5, let $\kappa = 3$. Then, $b_n'(x) > 0$ for $d > 2.83$, showing that the winner's curse is present in the approximate bid function with a probability close to unity.

It is expected that more competition (an increased number of bidders) should increase sellers' revenue. The mechanism-design literature also suggests a bidder's expected profit to increase with the magnitude of his signal, see Gordy (1998). Nevertheless, counter-examples are often available. For example, Matthews (1984) shows that the seller revenue can decrease with an increasing number of bidders if the signals come from a Pareto distribution. Therefore, it is interesting to present comparative statics on the bidder's expected profit and expected seller revenue using the hierarchical Gaussian model. To simulate the expected seller revenue we utilize the same Monte Carlo techniques as in Gordy (1998).

In a second price common value auction, a bidder's expected profit for a given signal x is given by

$$\Pi(x) = \int_{-\infty}^{x} (v(x,y) - b(y))\, f_{Y|X}(y|x)\; dy. \tag{4.2}$$

This integral can be solved by using Gaussian quadrature methods. As we can see in Figure 12, the bidder's expected profit increases with signals x, and decreases with more competition as n increases. However, according to Figure 13, the bidder's expected profits do not monotonically increase with a higher precision in signals as κ decreases. By intuition, this is an expected

[4]See Thaler (1988) for a careful discussion.

result. Gordy (1998) suggests that higher precision in signals only increase $\Pi(x)$ to a certain point and will eventually decrease $\Pi(x)$ when signal precision becomes too narrow. In the limit, as $\kappa \longrightarrow 0$, signals become perfectly precise. Thus, the true unknown value of the object becomes common knowledge and the bidders face Bertrand competition, which results in zero expected profits. The precision in signals of about $\kappa = 1.5$ was estimated to give the highest expected profits for sufficiently high values of signals x. Lower and higher precision from this point results in a decline of $\Pi(x)$.

To compute the posterior distribution of the expected seller revenue, we found that 100000 auctions were good enough for convergence. Figures 14 and 15 show that the expected seller revenue increases with n and μ as expected. In addition, we also noted that a higher variance of the common value v decreases the expected seller revenue and increases a bidder's expected profit. Intuitively, the bidders' fear of overestimation increases with the uncertainty in v which makes them lower their bids. Overall, we find no evidence of pathological behavior whatsoever.

Finally, since entry into the eBay auction is stochastic, the seller faces two counter-acting effects. First, the seller needs to post a sufficiently low minimum bid to encourage a sufficient number of bidders to participate in the auction. Second, to protect himself from a very low sale price, the seller also needs to post a sufficiently high minimum bid. A risk-neutral seller will choose the minimum bid that maximizes his expected revenue, given by

$$E[Revenue] = Pr(Sale|r)E[Price|r] + (1 - Pr(Sale|r))\mu,$$

where μ is assumed to represent the seller's valuation of the object if he fails to sell it. To obtain the optimal minimum bid for a seller, we simulated 500 auctions from the representative auction setup, described in Section 2.2, for each of a 1000 posterior draws from the variable selection algorithm, giving a total of $500,000$ simulated expected revenues for the representative auction. We then repeated this procedure for different values of the $MinBid/BookVal$ ratio. Figure 16 shows that the seller's optimal minimum bid is obtained just above the book value of the auctioned object. The expected seller revenue up to this point is fairly constant, so the choice between lower minimum bids does not seem to be of any greater importance. In Figure 17, the probability intervals of the number of bidders shrink as the ratio of $MinBid/BookVal$ increases. Eventually, already at $MinBid/BookVal = 1.1$, the probability becomes nearly 1 for a zero number of bidders.

5. Conclusions

Our paper contributes to the technically challenging econometric analysis of (second price) common value auctions. Building on the seminal work of Bajari and Hortacsu (2003), we propose a Bayesian framework that can be used to analyze auction datasets on a routine basis using a prior distribution with a small number of easily specified prior hyperparameters. One of the key features of our approach is a linear approximation of the otherwise highly complicated equilibrium bid function. The approximate bid function is analytically invertible and differentiable and can therefore be used for a fast and numerically stable evaluation of the likelihood function. Our proposed posterior sampling algorithm has the ability to handle variable selection among the model's covariates. The algorithm is documented to be very efficient on real and simulated data.

We collected a high-quality dataset from coin auctions on eBay and analyzed it with both reduced form regressions and a structural model for second price common value auctions. The results pointed strongly to book values as the most important predictor of common values. The minimum bid turned out to be the main determinant of the number of bids in an auction. Interestingly, the detailed seller information provided by eBay, and eBay's feedback score system seemed to be of very little value to the buyers. These covariates did consistently get a low posterior inclusion probability in the estimations. Finally, unopened coin envelopes attracted a high bidding activity and were sold at unusually high prices.

The estimated eBay model captured the within-auction bid dispersion and the cross-auction heterogeneity very well. A more severe test was to evaluate the out-of-sample predictions. The predictive price distributions looked reasonable and captured the observed prices very well in most cases. Overall, this fact together with the good fit of the bid dispersion indicates strongly that our estimated eBay auction model is quite accurate in explaining seller and bidder behavior at eBay.

Finally, possible extensions for future research could be to take into account and model cross-auction heterogeneity. A careful inspection of the bids revealed that some bids are only a tiny fraction of the object's book value, and cannot realistically represent serious final bids. One possible explanation could stem from the presence of incremental bidding. Ockenfels and Roth (2006) argue that late bidding may also arise out of equilibrium as a best reply to incremental bidding. Another possibility could be the effect that bidders are searching the eBay marketplace for low-price auctions, see Sailer (2006) for a non-parametrical identification of bidding costs in settings where the bidder searches with a reservation bid for low-price auctions. Other possible extensions include auctions with both a private and a common value element of the object, multiunit objects, or auctions with risk-averse bidders that we do not cover in this paper.

Appendix A. The linear approximation of the bid function

We first focus on the case with a known number of bidders, n, and then generalize the results to the case with stochastic entry. The derivation of the linear approximation is divided into three steps.

Step 1. By substituting $t = \frac{x-v}{\sqrt{\kappa}\sigma}$, the bid function for a known number of bidders becomes

$$b(x) = x - \sqrt{\kappa}\sigma \frac{\int_{-\infty}^{\infty} te^{-t^2}\Phi^{n-2}(t)e^{-\frac{1}{2\sigma^2}(x-\sqrt{\kappa}\sigma t-\mu)^2}\ dt}{\int_{-\infty}^{\infty} e^{-t^2}\Phi^{n-2}(t)e^{-\frac{1}{2\sigma^2}(x-\sqrt{\kappa}\sigma t-\mu)^2}\ dt}, \tag{A.1}$$

where $\Phi(\cdot)$ is the standard normal *cdf*.

Step 2. Let $h(t|\gamma,\theta) = e^{-\gamma\cdot(t-\theta)^2}$ be the approximating function to the standard normal distribution function $\Phi(t)$ on $[-a,a]$ (see Figure 18). The approximation error outside this interval is likely to only have a minor effect on the approximation of $b(x)$, because we choose a in the tails of the signal density (see later). The approximation constants γ and θ are obtained by minimizing the maximum divergence between $h(t|\gamma,\theta)$ and $\Phi(t)$ over $[-a,a]$, that is,

$$(\hat{\gamma},\hat{\theta}) = \min_{\gamma,\theta}\left(\max_{t} |h(t|\gamma,\theta) - \Phi(t)|\right).$$

For $a = 2$, we obtain

$$(\hat{\gamma},\hat{\theta}) = (0.1938, 1.9600).$$

Figure 18 shows that the strictly increasing $\Phi(t)$ is well approximated by the Gaussian density at most values over the target domain $[-2,2]$. The figure also shows that choosing $a = 1$ gives a worse approximation in the tails, and $a = 3$ does somewhat better in the tails than $a = 2$, but loses out in the main part of the distribution. The posterior inferences in the eBay data used in Section 4 are strikingly similar for $a = 1, 2$ or 3, but the value $a = 2$ gave the best approximation of the exact bid function; results are available from the authors upon request. Thus, we use $a = 2$ in this work.

Step 3. Replacing $\Phi(t)$ by $h(t|\hat{\gamma},\hat{\theta})$, the approximated bid function becomes

$$b(x) \approx x - \sqrt{\kappa}\cdot\sigma\cdot\frac{\int_{-\infty}^{\infty} t\cdot e^{-t^2}\cdot e^{-(n-2)\hat{\gamma}(t-\hat{\theta})^2}\cdot e^{-\frac{1}{2\sigma^2}\cdot(x-\sqrt{\kappa}\sigma t-\mu)^2}\ dt}{\int_{-\infty}^{\infty} e^{-t^2}\cdot e^{-(n-2)\hat{\gamma}(t-\hat{\theta})^2}\cdot e^{-\frac{1}{2\sigma^2}\cdot(x-\sqrt{\kappa}\sigma t-\mu)^2}\ dt}.$$

By completing the squares of the exponential functions, the bid function $b(x)$ can be simplified to

$$b(x) \approx x - \sqrt{\kappa}\cdot\sigma\cdot\frac{\int_{-\infty}^{\infty} t\cdot e^{-m_1(t-m_2)^2}\ dt}{\int_{-\infty}^{\infty} e^{-m_1(t-m_2)^2}\ dt} = x - \sqrt{\kappa}\cdot\sigma\cdot\frac{\mathrm{E}(t)}{1} = x - \sqrt{\kappa}\cdot\sigma\cdot m_2, \tag{A.2}$$

where $m_1 = 1+(n-2)\hat{\gamma}+\frac{\kappa}{2}$, and $m_2 = \frac{(n-2)\hat{\gamma}\hat{\theta}+\frac{\sqrt{\kappa}(x-\mu)}{2\sigma}}{1+(n-2)\hat{\gamma}+\frac{\kappa}{2}}$. Note that the constants of the normal kernel in the numerator and the denominator cancel out. Substituting the expression for m_2 gives the approximate bid function for a known number of bidders. If $n = 2$, then $\Phi^{n-2}(t) = 1$, and the bid function in A.1 can be computed exactly, and equals the approximate solution in (A.2).

Turning to the model with stochastic entry, it is straightforward to show that the same type of approximations can be used for the minimum bid in (2.2)

$$r(x^{\star},\lambda)\approx\sum_{n=1}^{\infty}p_n(\lambda)\left(c_r+\omega_r\mu+(1-\omega_r)x^{\star}\right),$$

which can be solved for $x^{\star}$ as

$$x^{\star}(r,\lambda)\approx\frac{r-\sum_{n=1}^{\infty}(c_r+\mu\omega_r)p_n(\lambda)}{\sum_{n=1}^{\infty}(1-\omega_r)p_n(\lambda)}.$$

The same approach cannot be used to approximate the bid function in (2.1). We could do the approximation term by term in the summation, but then the bid function can no longer be inverted analytically. One way of proceeding is to note that the bid function can be expressed as

$$b(x)=\frac{\mathrm{E}_{n|\lambda}\left[(n-1)g_1(n)\right]}{\mathrm{E}_{n|\lambda}\left[(n-1)g_2(n)\right]},\tag{A.3}$$

where $g_1(n)=\int_{-\infty}^{\infty}v\cdot F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)dv$, $g_2(n)=\int_{-\infty}^{\infty}F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)dv$, and $\mathrm{E}_{n|\lambda}$ denotes the expectation with respect to the Poisson distribution of the number of bidders, n. A first-order Taylor expansion of $(n-1)g_1(n)$ and $(n-1)g_2(n)$ around $n=\lambda$ then gives

$$b(x)=\frac{\mathrm{E}_{n|\lambda}\left[(n-1)g_1(n)\right]}{\mathrm{E}_{n|\lambda}\left[(n-1)g_2(n)\right]}\approx\frac{g_1(\lambda)}{g_2(\lambda)},\tag{A.4}$$

where the ratio $g_1(\lambda)/g_2(\lambda)$ can now be approximated with the linear approximation in (2.6) with $n=\lambda$. Figure 1 and the estimation results in Section 4.3 verify that this gives quite an accurate approximation of the true bid function. This is because $g_1(n)$ and $g_2(n)$ are very similar functions and thus, the approximation errors in the numerator and denominator in (A.4) cancel out (see Tierney and Kadane 1986 for similar results in a more general setting).

Appendix B. A generalization of the MH algorithm

Starting with George and McCulloch (1993) and Smith and Kohn (1996), there has been a number of algorithms that simultaneously draw the regression coefficients from the posterior and does variable selection, all in a single run of the sampler. In particular, Nott and Leonte (2004) propose an efficient algorithm for variable selection in generalized linear models (GLM). Nott and Leonte's algorithm requires that the gradient and hessian matrix are available in closed form (which is the case for GLMs). The algorithm presented below is of similar form, but can be applied to any problem as long as the likelihood and prior can be evaluated numerically.

We present the algorithm for a general setting where β contains all r model parameters and D denotes the available data. Consider now setting a subset of the elements in β to zero (any other value is also possible). In a regression situation, this is clearly equivalent to selecting a subset of the covariates. Let $\mathcal{J}=(j_1,...,j_r)$ be a vector of binary indicators such that $j_i=0$ iff the ith element of β is zero. We can view these indicators as a set of new parameters. For simplicity, we shall here assume that the elements of $\mathcal{J}$ are independent a priori with $\Pr(j_i)=\pi$ for all i, so that π is the prior probability of including the ith covariate in the model.[5] The following algorithm samples β and $\mathcal{J}$ simultaneously using an extended

[5] A perhaps better way of viewing this is that the prior on the coefficients is given by a two-component mixture density with one of the components degenerated at zero (Smith and Kohn, 1996).

Metropolis-Hastings algorithm. The algorithm uses the following proposal density

$$f(\beta_p, \mathcal{J}_p|\beta_c, \mathcal{J}_c) = g(\beta_p|\mathcal{J}_p, \beta_c)h(\mathcal{J}_p|\beta_c, \mathcal{J}_c),$$

where β_p and $\mathcal{J}_p$ are the proposed values for β and $\mathcal{J}$, β_c and $\mathcal{J}_c$ are the current values for β and $\mathcal{J}$, h is the proposal distribution for $\mathcal{J}$, and g is the proposal density for β conditional on $\mathcal{J}_p$. The Metropolis-Hastings acceptance probability then becomes

$$a[(\beta_c, \mathcal{J}_c) \to (\beta_p, \mathcal{J}_p)] = \min\left(1, \frac{p(D|\beta_p, \mathcal{J}_p)p(\beta_p|\mathcal{J}_p)p(\mathcal{J}_p)/g(\beta_p|\mathcal{J}_p, \beta_c)h(\mathcal{J}_p|\beta_c, \mathcal{J}_c)}{p(D|\beta_c, \mathcal{J}_c)p(\beta_c|\mathcal{J}_c)p(\mathcal{J}_c)/g(\beta_c|\mathcal{J}_c, \beta_p)h(\mathcal{J}_c|\beta_p, \mathcal{J}_p)}\right),$$

where $p(D|\beta_p, \mathcal{J}_p)$ is the likelihood of the observed data conditional on β with zeros given by $\mathcal{J}_p$, $p(\beta|\mathcal{J})$ is the prior of the non-zero elements of β, and $p(\mathcal{J})$ is the prior probability of $\mathcal{J}$.

We will propose $\mathcal{J}_p$ using the two updating steps: i) randomly picking a small subset of $\mathcal{J}_p$ and then always proposing a change of the selected indicators (metropolized move), and ii) randomly picking a pair of covariates (one currently in the model and the other currently not in the model) and proposing a switch of their corresponding indicators (switch move). Our experience is that these two simple updating rules work well in practise. Note also that the acceptance probability for these updates simplifies to

$$a[(\beta_c, \mathcal{J}_c) \to (\beta_p, \mathcal{J}_p)] = \min\left(1, \frac{p(D|\beta_p, \mathcal{J}_p)p(\beta_p|\mathcal{J}_p)p(\mathcal{J}_p)/g(\beta_p|\mathcal{J}_p, \beta_c)}{p(D|\beta_c, \mathcal{J}_c)p(\beta_c|\mathcal{J}_c)p(\mathcal{J}_c)/g(\beta_c|\mathcal{J}_c, \beta_p)}\right). \tag{B.1}$$

More sophisticated ways of proposing can easily be implemented, e.g. the adaptive scheme in Nott and Kohn (2005), where the history of $\mathcal{J}$-draws is used to adaptively build up a proposal for each indicator.

The proposal density $g(\beta_p|\mathcal{J}_p, \beta_c)$ is obtained as follows. First, by including all covariates, we approximate the posterior with the $t(\hat{\beta}, -H^{-1}, h)$ density described in Section 3.2 for the case without variable selection. We then propose $\beta_p|\mathcal{J}_p$ from this multivariate t distribution conditional on the zero restrictions dictated by $\mathcal{J}_p$. Assume for notational simplicity that the elements in β have been rearranged so that $\beta = (\beta_0', \beta_p')'$ where β_0 are the p_0 zero-restricted elements of β under $\mathcal{J}_p$, and β_p are the non-zero parameters. Decompose $\hat{\beta}$ and $P = -H^{-1}$ conformably with the decomposition of β as

$$\begin{aligned} \hat{\beta} &= (\hat{\beta}_0', \hat{\beta}_p')' \\ \hat{P} &= \begin{pmatrix} P_{00} & P_{0p} \\ P_{p0} & P_{pp} \end{pmatrix}. \end{aligned}$$

Using results from the conditional distributions of subsets of multivariate-t variables (see e.g. Bauwens, Lubrano and Richard (1999, Theorem A.16)), we can now propose β_p conditional on $\beta_0 = 0$ from

$$\beta_p|(\beta_0 = 0) \sim t[\hat{\beta}_p + P_{pp}^{-1}P_{p0}\hat{\beta}_0, P_{pp}, c(\beta_0), v + p_0].$$

$$c(\beta_0) = 1 + \hat{\beta}_0'(\hat{P}_{00} - \hat{P}_{0p}\hat{P}_{pp}^{-1}\hat{P}_{p0})\hat{\beta}_0,$$

using the parametrization of the t distribution in Bauwens et al. (1999). A similar algorithm has recently been suggested by Giordani and Kohn (2008) in their adaptive sampling framework. They propose to use a mixture of multivariate normals as a proposal density rather than a multivariate t. They document good performance on simulated data.

Appendix C: Description of eBay covariates

The following list describes each covariate for our collected data of 1000 eBay auctions:

Book Value: The price of the item as reported by the large Internet coin seller *Golden Eagle Coins* at http://www.goldeneaglecoin.com.

Minor Blemish: Dummy variable, coded as 1 if the proof set had a minor damage on the box or packaging according to careful subjective assessment of the item using the seller's detailed descriptions and pictures.

Major Blemish: Dummy variable, coded as 1 if at least one coin were missing or if other major imperfections were present.

Seller's Feedback Score: Buyer's rating of the seller in terms of reliability and timeliness in delivery.

Seller's Negative Feedback Score: Buyer's negative ratings of the seller.

PowerSeller: Dummy variable, coded to be 1 if the seller is ranked among the most successful sellers in terms of product sales and customer satisfaction on eBay.

ID Seller: Dummy variable, coded to be 1 if the seller's ID is verified. In the ID verification process, members' identity is established by cross-checking their contact information in consumer and business databases, which helps both buyers and sellers to trust each other.

UnOpen: Dummy variable, coded to be 1 if the proof set is delivered sealed and unopened in its original envelope. Most common for proof sets that originate from the fifties or sixties.

LargeNeg: Dummy variable, coded to be 1 if the seller's negative feedback score is at least one per cent of the seller's overall feedback score.

References

[1] Bajari, P., and A. Hortacsu (2003), The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *The Rand Journal of Economics*, **3(2)**, 329-355.

[2] —— (2004), Economic Insights from Internet Auctions, *Journal of Economic Literature*, **42**, 457-486.

[3] —— (2005), Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data, *Journal of Political Economy*, **113(4)**, 703-741.

[4] Bauwens, L. Lubrano, M., and Richard, J-F. (1999),*Bayesian Inference in Dynamic Econometric Models*. Oxford University Press: Oxford.

[5] Elyakime, B., Laffont J.J., Loisel, O., and Vuong, Q. (1997) Auctioning and Bargaining: An Econometric Study of Timber Auctions With Secret Reservation Prices, *Journal of Business & Economic Statistics*, **15(2)**, 209-20.

[6] Fletcher, R. (1987), *Practical Methods of Optimization*, Wiley, New York.

[7] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis, 2nd edition*, Chapman & Hall, New York.

[8] George, E., and McCulloch, R. (1993), Variable selection via Gibbs sampling, *Journal of the American Statistical Association*, **88**, 881-889.

[9] Giordani, P., and Kohn, R. (2008), Adaptive MCMC, Forthcoming in Advances in Econometrics.

[10] Gordy, B. M. (1998), Computationally Convenient Distributional Assumptions for Common-Value Auctions, *Computational Economics*, **12**, 61-78.

[11] Kagel, J.H., and Levin, D. (1986), The Winner's Curse and Public Information in Common Value Auctions, *American Economic Review*, **76(5)**, 894-920.

[12] Klemperer, D. P. (1999), Auction Theory: A Guide to the Literature, *Journal of Economic Surveys*, **13(3)**, 227-286.

[13] —— (2004), *Journal of Economic Surveys*, **13(3)**, 227-286.

[14] Krishna, V. (2002). *Auction Theory*, San Diego: Academic Press.

[15] Matthews, S. (1984), Information Acquisition in Discriminatory Auctions, In Boyer, M. and R.E. Kihlstrom (eds), *Bayesian Models in Economic Theory*. Vol. 5 of *Studies in Bayesian Econometrics*. Elsevier Science Publishers, Amsterdam, pp. 181-207.

[16] Milgrom, P. (2004). *Putting Auction Theory to Work*, Cambridge, England: Cambridge University Press.

[17] Milgrom, P., and Weber, R. (1982), A Theory of Auctions and Competitive Bidding, *Econometrica*, **50**, 1089-1022.

[18] Nott, D. J., and Kohn, R. (2005), Adaptive sampling for Bayesian variable selection, *Biometrika*, **92**, 747-763.

[19] Nott, D. J., and Leonte, D. (2004), Sampling schemes for Bayesian variables selection in generalized linear models, *Journal of Computational and Graphical Statistics*, **13**, 362-382.

[20] Ockenfels, A., and Roth, A. E. (2006), Late and multiple bidding in second price Internet auctions: Theory and evidence concerning different rules for ending an auction, *Games and Economic Behavior*,

55, 297-320.

[21] Paarsch, J. H. (1992), Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions, *Journal of Econometrics*, **51**, 191-215.

[22] Paarsch, J. H., and Hong, H. (2006). An Introduction to the Structural Econometrics of Auction Data, Cambridge, MA: MIT Press.

[23] Sailer, K. (2006), Searching the eBay Marketplace, mimeo, Department of Economics, University of Munich.

[24] Smith, M., and Kohn, R. (1996), Nonparametric regression using Bayesian variable selection, *Journal of Econometrics*, **75**, 317-344.

[25] Thaler, H. R. (1988), Anomalies: The Winner's Curse, *Journal of Economic Perspectives*, **2**(**1**), 191-202.

[26] Vickrey, W. (1961), Counterspeculation, Auctions, and Competitive Sealed Tenders, *The Journal of Finance*, **16**(**1**), 8-37.

[27] Wolfstetter, E. (1996), Auctions: An Introduction, *Journal of Economic Surveys*, **10**(**4**), 367-420.

[28] Zellner, A. (1986), On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, 233–243. North-Holland/Elsevier.

TABLE 1. Comparing the inferences from the exact and approximate bid functions on simulated data.

Parameter	Mean		Std Dev		Incl prob	
	Exact	Approx	Exact	Approx	Exact	Approx
$\ln \kappa$	−1.131	−1.122	0.115	0.112	1.000	1.000
β_1	0.996	0.995	0.005	0.005	1.000	1.000
β_2	−0.134	−0.136	0.032	0.035	0.911	0.892
β_3	0.953	0.949	0.013	0.013	1.000	1.000
β_4	−0.300	−0.304	0.072	0.069	0.908	0.911
β_5	1.253	1.241	0.101	0.105	1.000	1.000
β_6	0.221	0.224	0.021	0.021	1.000	1.000
β_7	−0.008	−0.005	0.033	0.029	0.119	0.086
β_8	−2.358	−2.369	0.166	0.166	1.000	1.000

Note: $c = c_\mu = c_\sigma = c_\lambda = n$, $\bar{\kappa} = 0.25$, $g = 4$, and $\pi = 0.2$.

TABLE 2. Summary statistics of the eBay data

Variable	Mean	Std Dev	Min	Max	μ	σ	λ
Book Value($)	32.30	37.26	7.50	399.50	x	x	x
Price/Book Value, no Blemish	0.76	0.30	0.16	3.63			
Price/Book Value, Blemish	0.61	0.24	0.05	1.20			
Minor Blemish	0.09				x	x	x
Major Blemish	0.03				x	x	x
Number of Bidders	3.74	2.70	0	15.00			
Minimum Bid/Book Value	0.35	0.34	0	1.34			x
Seller's Feedback Score	2427	2883	0	12568			
PowerSeller	0.53				x	x	x
ID Seller	0.06				x	x	x
Unopen	0.10				x	x	x
%Sold	0.86						

Table 3. Posterior inference for the reduced form regressions

Coeff	Covariate	Mean	Stdev	t-ratio	Incl Prob	IF
κ		5.499	–	–	–	–
μ	Const	24.913	0.298	83.710	1.000	4.8320
	BookStd	0.698	0.008	89.448	1.000	3.017
	Book*Power	0.030	0.012	2.523	0.645	2.321
	Book*ID	−0.003	0.022	−0.156	0.041	–
	Book*Unopen	0.249	0.017	14.616	1.000	2.964
	Book*MinBlem	0.032	0.017	1.841	0.184	–
	Book*MajBlem	−0.234	0.022	−10.793	1.000	2.691
	Book*LargNeg	−0.004	0.013	−0.270	0.037	–
	σ_ε	5.902	0.170	–	–	2.338
	R^2	94.6%				
σ	Const	2.553	0.060	42.562	1.000	2.307
	LBookStd	1.961	0.075	26.299	1.000	2.175
	LBook*Power	−0.017	0.033	−0.511	0.056	–
	LBook*ID	0.089	0.081	1.090	0.076	–
	LBook*Unopen	0.263	0.044	5.927	1.000	1.790
	LBook*MinBlem	0.044	0.058	0.764	0.056	–
	LBook*MajBlem	−0.160	0.076	−2.099	0.259	–
	LBook*LargNeg	0.083	0.048	1.738	0.167	–
	σ_ε	1.279	0.036	–	–	2.3156
	R^2	56.1%				
λ	Const	1.056	0.023	46.978	1.000	1.694
	Power	−0.031	0.037	−0.859	0.010	–
	ID	−0.401	0.093	−4.299	0.997	1.304
	Unopen	0.444	0.049	9.0538	1.000	1.379
	MinBlem	−0.027	0.055	−0.486	0.005	–
	MajBlem	−0.235	0.090	−2.615	0.111	1.615
	LargeNeg	0.085	0.056	1.525	0.011	–
	LBookStd	−0.113	0.028	−3.998	0.973	1.416
	MinBidStd	−1.894	0.074	−25.675	1.000	2.797

Upper: linear regression with $\hat{\mu}$ as dependent variable.

Middle: linear regression with $\ln\hat{\sigma}^2$ as dependent variable.

Lower: Poisson regression with the number of bidders with a non-zero bid as dependent variable.

$\hat{\mu}$ and $\hat{\sigma}^2$ are the MLE conditional on the posterior mean of κ from the structural model.

TABLE 4. Posterior inference for the Tobit price regression

Covariate	Mean	Stdev	t-ratio	Incl Prob	IF
Const	27.177	0.230	95.824	1.000	1.789
BookStd	0.757	0.007	86.353	1.000	1.796
Book*Power	−0.005	0.009	−0.557	0.039	–
Book*ID	0.032	0.021	1.676	0.118	–
Book*Unopen	0.321	0.018	17.322	1.000	2.063
Book*MinBlem	0.003	0.014	0.291	0.035	–
Book*MajBlem	−0.151	0.023	−6.673	1.000	1.601
Book*LargNeg	0.006	0.013	0.398	0.032	–
σ_ε	6.664	0.158	–	–	1.938

Tobit regressions with auction prices possibly censored at the minimum bid as dependent variable.

Table 5. Posterior inference for the structural model with the approximate bid function

Parameter	Covariate	Mean	St Dev	t-ratio	Incl prob	IF
κ	–	5.499	0.772	–	–	7.889
μ	Const	28.272	0.245	-14.014	1.000	6.062
	BookStd	0.740	0.010	61.024	1.000	3.615
	Book*Pow	0.033	0.015	2.260	0.064	–
	Book*ID	0.128	0.036	2.775	0.900	4.197
	Book*Unopen	0.372	0.029	12.023	1.000	3.771
	Book*MinBlem	-0.022	0.021	-1.594	0.010	–
	Book*MajBlem	-0.252	0.030	-8.277	1.000	3.306
	Book*LargNeg	-0.003	0.018	0.635	0.004	–
$log(\sigma^2)$	Const	3.997	0.071	54.553	1.000	7.018
	LogBookStd	1.262	0.038	29.549	1.000	5.003
	LBook*Pow	0.043	0.018	2.810	0.220	–
	LBook*ID	0.108	0.040	3.173	0.481	3.946
	LBook*Unopen	0.211	0.027	6.947	1.000	4.862
	LBook*MinBlem	-0.028	0.027	-0.902	0.012	–
	LBook*MajBlem	0.036	0.040	1.026	0.007	–
	LBook*LargNeg	0.035	0.021	1.303	0.017	–
$log(\lambda)$	Const	1.193	0.021	40.055	1.000	3.634
	Pow	0.009	0.035	0.163	0.005	–
	ID	-0.177	0.110	-1.656	0.030	–
	Unopen	0.323	0.048	6.711	1.000	3.387
	MinBlem	-0.049	0.048	-1.222	0.008	–
	MajBlem	-0.151	0.085	-2.097	0.019	–
	LargNeg	0.055	0.049	0.912	0.012	–
	LogBookStd	-0.038	0.027	-1.473	0.018	–
	MinBidStd	-1.433	0.056	-22.291	1.000	3.521

Note: $c = n, \bar{\kappa} = 0.25, g = 4$, and $\pi = 0.2$. The last column displays the inefficiency factors.

Table 6. Posterior inference for the structural model with the exact bid function

Parameter	Covariate	Mean	St Dev	t-ratio	Incl prob	IF
κ	-	4.479	0.345	-	-	5.129
μ	Const	29.345	0.409	-8.558	1.000	35.355
	BookStd	0.756	0.015	50.200	1.000	10.767
	Book*Pow	0.049	0.017	2.202	0.351	-
	Book*ID	0.120	0.054	2.393	0.078	-
	Book*Unopen	0.464	0.037	12.044	1.000	5.388
	Book*MinBlem	-0.033	0.022	-2.044	0.005	-
	Book*MajBlem	-0.224	0.029	-6.723	1.000	5.028
	Book*LargNeg	-0.029	0.024	0.368	0.016	-
$log(\sigma^2)$	Const	4.474	0.050	112.648	1.000	46.660
	LogBookStd	1.536	0.034	43.820	1.000	6.607
	LBook*Pow	0.053	0.021	3.902	0.507	31.367
	LBook*ID	0.091	0.040	3.409	0.125	-
	LBook*Unopen	0.247	0.023	9.574	1.000	8.293
	LBook*MinBlem	0.005	0.019	-0.228	0.009	-
	LBook*MajBlem	-0.007	0.035	-0.283	0.009	-
	LBook*LargNeg	0.045	0.017	2.260	0.239	-
$log(\lambda)$	Const	1.274	0.024	40.934	1.000	4.631
	Pow	-0.028	0.040	-0.792	0.008	-
	ID	-0.133	0.111	-1.405	0.010	-
	Unopen	0.322	0.048	6.566	1.000	3.999
	MinBlem	-0.067	0.055	-2.067	0.006	-
	MajBlem	-0.162	0.089	-1.738	0.023	-
	LargNeg	0.064	0.047	0.110	0.010	-
	LogBookStd	-0.119	0.029	-4.371	0.986	6.691
	MinBidStd	-1.587	0.079	-21.566	1.000	7.088

Note: $c = n, \bar{\kappa} = 0.25, g = 4$, and $\pi = 0.2$. The last column displays the inefficiency factors.

Table 7. Sensitivity analysis - Posterior means

Coeff	Covariate	Bench	$c = n/16$	$c = n/64$	$c = n/256$	$\pi = .5$	$\delta = 0$
κ		5.499	5.875	5.626	**3.535**	5.519	**7.165**
μ	Const	28.272	28.165	28.138	28.330	28.175	28.469
	BookStd	0.740	0.732	**0.714**	**0.625**	0.738	0.743
	Book*Power	0.033	0.031	0.031	0.022	0.031	0.033
	Book*ID	0.128	0.124	0.106	0.075	0.131	0.134
	Book*Unopen	0.372	0.361	0.340	**0.275**	0.369	0.369
	Book*MinBlem	−0.022	−0.025	−0.021	−0.021	−0.024	−0.034
	Book*MajBlem	−0.252	−0.252	−0.252	−0.238	−0.252	−0.247
	Book*LargNeg	−0.003	−0.002	0.002	−0.003	−0.005	0.007
σ	Const	3.997	4.010	4.006	3.878	3.972	**4.197**
	LBookStd	1.262	1.214	**1.114**	**0.932**	1.263	**1.426**
	LBook*Power	0.043	0.046	0.043	0.037	0.047	0.040
	LBook*ID	0.108	0.096	0.069	**0.009**	0.111	0.111
	LBook*Unopen	0.211	0.190	0.171	**0.133**	0.201	0.175
	LBook*MinBlem	−0.028	−0.024	−0.004	**0.039**	−0.031	0.012
	LBook*MajBlem	0.036	0.039	0.048	0.067	0.027	**−0.053**
	LBook*LargNeg	0.035	0.039	0.040	0.046	0.041	0.010
λ	Const	1.193	1.187	1.180	1.174	1.192	1.203
	Power	0.009	0.001	0.006	0.007	−0.012	0.022
	ID	−0.177	−0.135	−0.075	−0.066	−0.178	−0.183
	Unopen	0.323	0.330	0.335	0.334	0.327	0.320
	MinBlem	−0.049	−0.051	−0.051	−0.060	−0.055	−0.015
	MajBlem	−0.151	−0.157	−0.149	−0.156	−0.162	−0.248
	LargeNeg	0.055	0.073	0.069	0.048	0.075	0.054
	LBookStd	−0.038	−0.029	−0.014	−0.057	−0.041	**0.022**
	MinBidShare	−1.433	−1.449	−1.478	−1.537	−1.435	−1.525

The bench column displays the posterior mean of each parameter (conditional on inclusion) for the benchmark model, the other columns are results for variations of the benchmark model.
Bold numbers indicate that the estimate lies outside the benchmark model's 95% posterior credibility interval.

TABLE 8. Sensitivity analysis - Posterior inclusion probabilities

Coeff	Covariate	Bench	$c = n/16$	$c = n/64$	$c = n/256$	$\pi = .5$	$\delta = 0$
κ		–	–	–	–	–	–
μ	Const	1.000	1.000	1.000	1.000	1.000	1.000
	BookStd	1.000	1.000	1.000	1.000	1.000	1.000
	Book*Power	0.064	**0.384**	**0.501**	**0.325**	**0.297**	0.046
	Book*ID	0.900	0.918	0.938	**0.790**	0.919	**0.763**
	Book*Unopen	1.000	1.000	1.000	1.000	1.000	1.000
	Book*MinBlem	0.010	0.045	0.070	**0.135**	0.041	0.030
	Book*MajBlem	1.000	1.000	1.000	1.000	1.000	1.000
	Book*LargNeg	0.004	0.018	0.051	0.061	0.016	0.007
σ	Const	1.000	1.000	1.000	1.000	1.000	1.000
	LBookStd	1.000	1.000	1.000	1.000	1.000	1.000
	LBook*Power	0.220	**0.671**	**0.804**	**0.913**	**0.600**	0.149
	LBook*ID	0.481	**0.620**	0.403	**0.076**	**0.759**	**0.664**
	LBook*Unopen	1.000	1.000	1.000	1.000	1.000	1.000
	LBook*MinBlem	0.012	0.038	0.044	**0.372**	0.050	0.006
	LBook*MajBlem	0.007	0.039	**0.112**	**0.625**	0.030	0.016
	LBook*LargNeg	0.017	0.113	**0.289**	**0.824**	0.098	0.006
λ	Const	1.000	1.000	1.000	1.000	1.000	1.000
	Power	0.005	0.023	0.045	0.058	0.014	0.005
	ID	0.030	0.074	0.062	0.097	0.112	0.032
	Unopen	1.000	1.000	1.000	1.000	1.000	1.000
	MinBlem	0.008	0.022	0.045	**0.117**	0.031	0.004
	MajBlem	0.019	0.067	**0.124**	**0.305**	0.064	**0.295**
	LargeNeg	0.012	0.050	0.078	0.103	0.042	0.003
	LBookStd	0.018	0.035	0.041	**0.632**	0.053	0.006
	MinBidShare	1.000	1.000	1.000	1.000	1.000	1.000

The bench column displays the Posterior inclusion probabilities for the benchmark model, the other columns are results for variations of the benchmark model. Bold numbers indicate a difference from bench by more than 0.1.

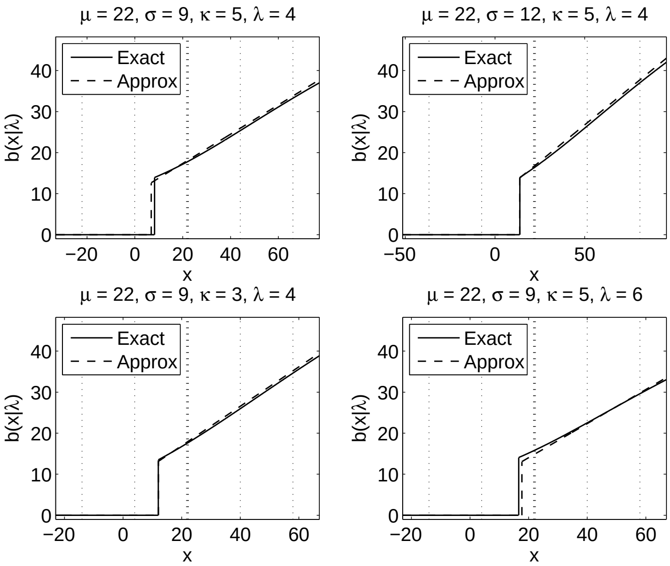

FIGURE 1. Examining the accuracy of the approximate bid function in the Gaussian model for different configurations of parameter values. The vertical lines in the figures mark out the mean (thick dotted) $\pm$ 1 and 2 standard deviations (thin dotted) in the unconditional distribution of the signals, x.

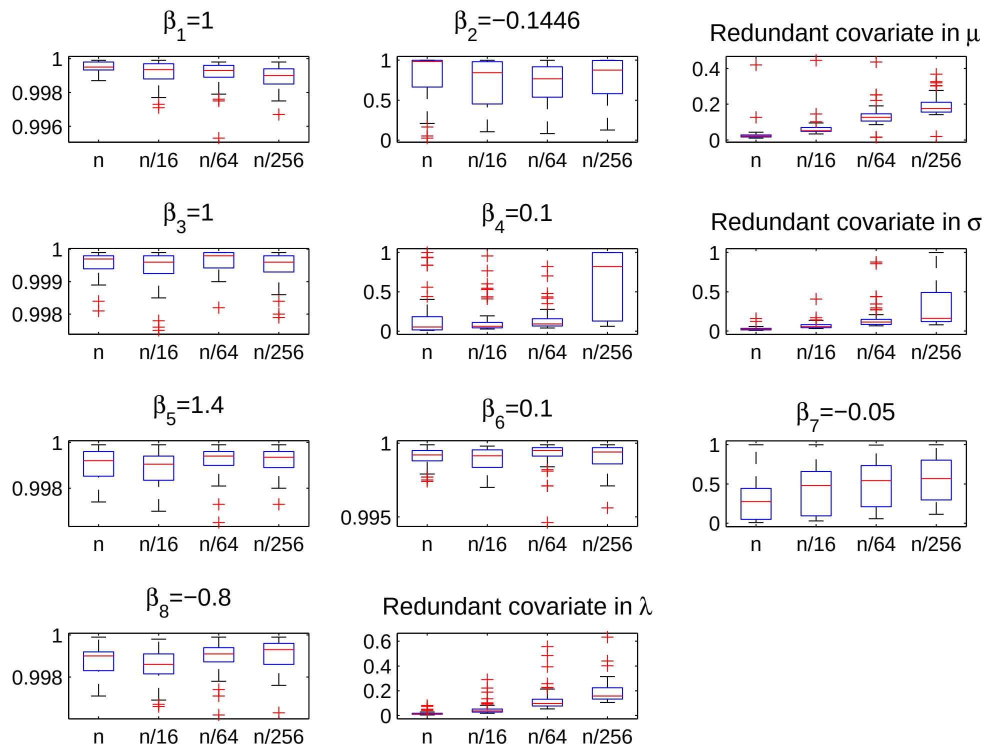

FIGURE 2. Posterior variable selection probabilities for the simulated data with different values of $c = c_\mu = c_\sigma = c_\lambda$ on the horizontal axis. $\bar{\kappa} = 0.25$, $g = 4$ and $\pi = 0.2$ were used in all simulations.

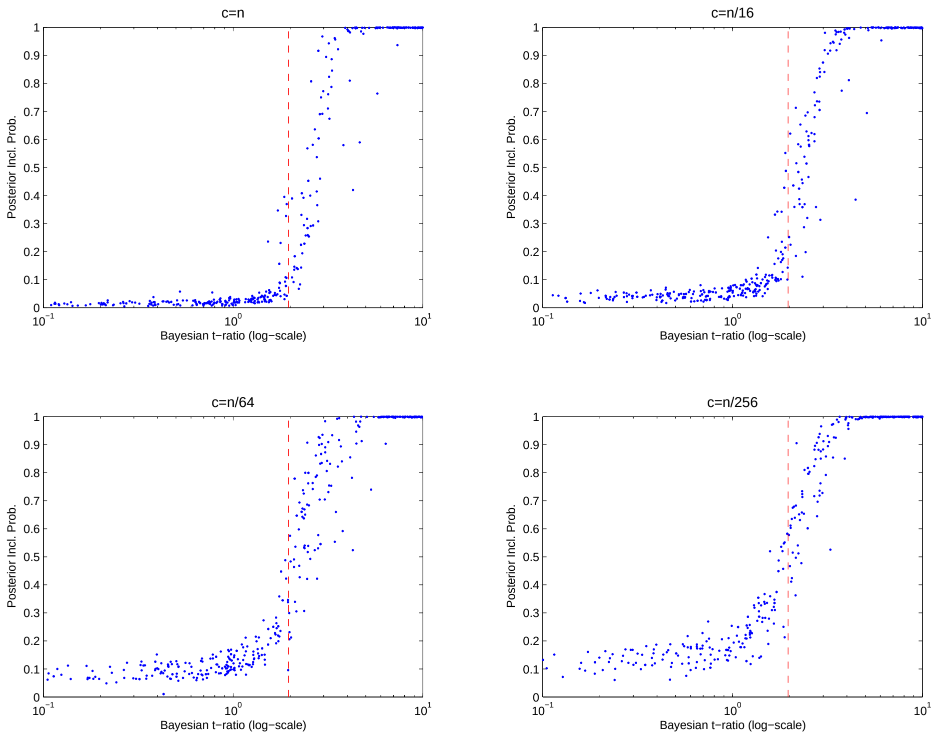

FIGURE 3. The relation between Bayesian t-ratios and the posterior inclusion probabilities for the simulated datasets. Each subplot corresponds to a different value of the prior hyperparameter $c = c_{\mu} = c_{\sigma} = c_{\lambda}$.

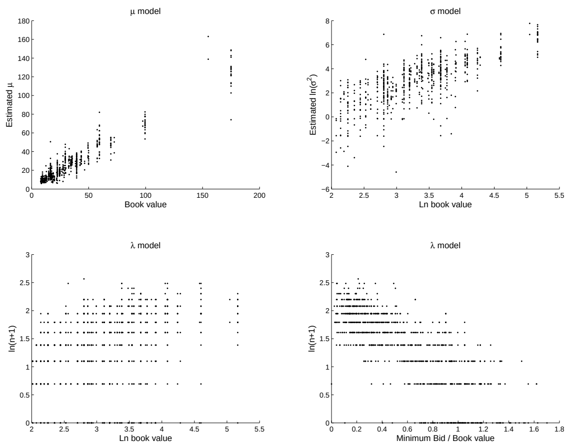

FIGURE 4. Scatterplots of $\hat{\mu}, \ln \hat{\sigma}^2$, and $\ln(n+1)$ against the suitable transformations of the covariate BookValue, and $\ln(n+1)$ against the continuous covariates in the entry model.

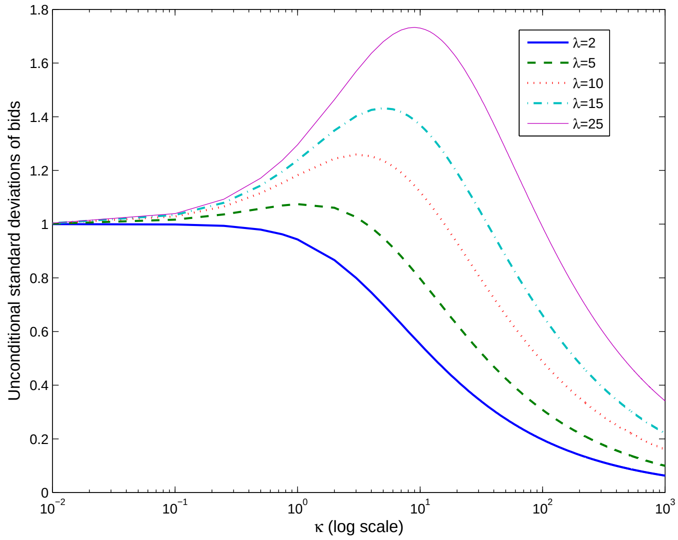

FIGURE 5. The unconditional standard deviation of bids as a function of the variance scale parameter κ.

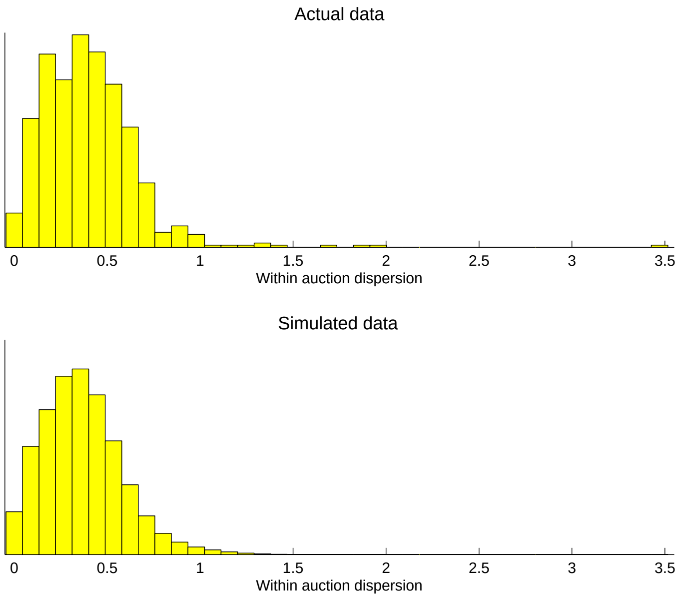

FIGURE 6. Posterior predictive check of the within-auction dispersion. The within-auction dispersion is defined as the difference between the highest observed bid and the lowest bid divided by the auctioned item's book value.

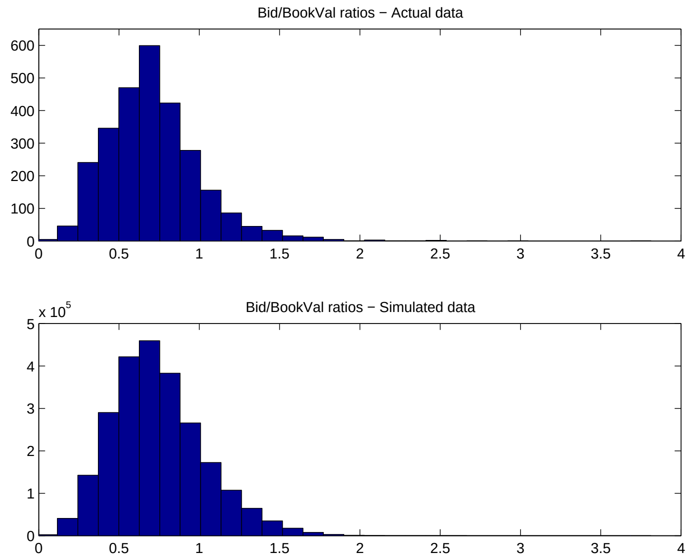

FIGURE 7. Posterior predictive check of the cross-auction heterogeneity. The figure displays histograms of the bids divided by the corresponding book value in that auction.

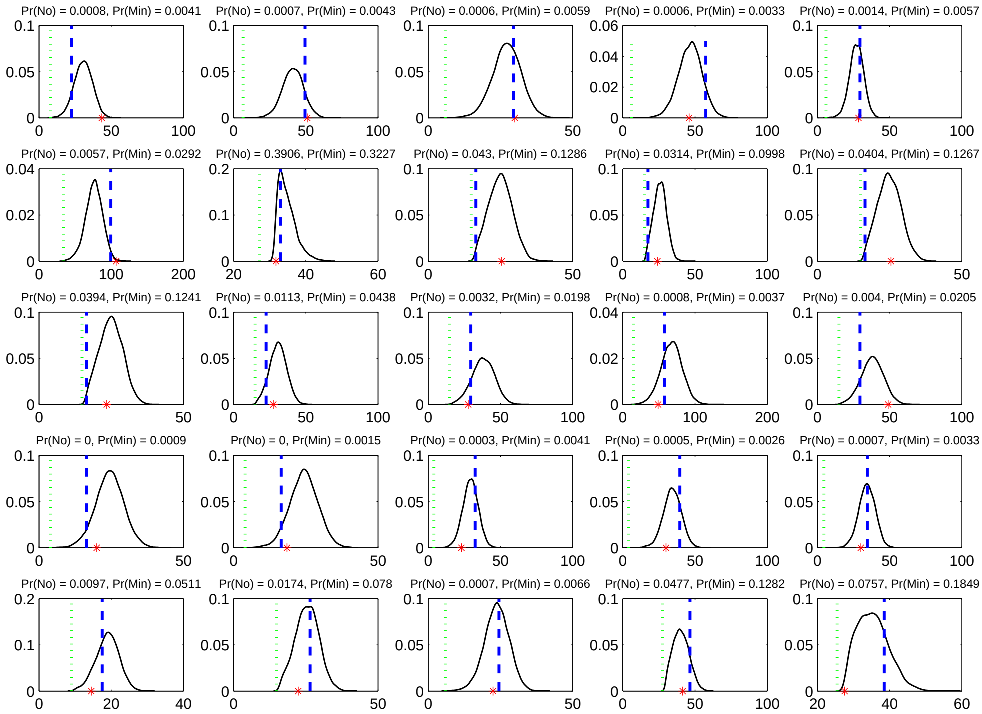

FIGURE 8. Out-of-sample predictions for the first 25 auctions in the evaluation sample. Each subplot displays the realized price marked out by a star (if the item is sold), the minimum bid (vertical dotted line), and the book value (vertical dashed line). Pr(No) and Pr(Min) are the predictive probabilities of no bids and a single bid (in which case the winner pays the seller's posted minimum bid), respectively. The solid curve is the predictive price density conditional on at least two bids.

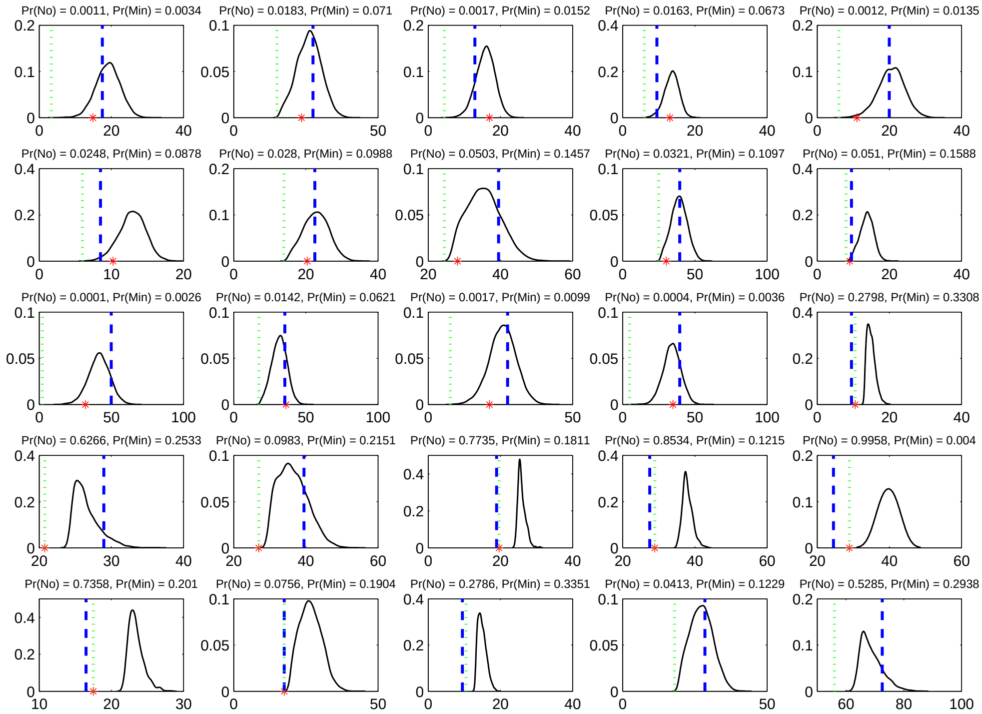

FIGURE 9. Out-of-sample predictions for the last 25 auctions in the evaluation sample. Each subplot displays the realized price marked out by a star (if the item is sold), the minimum bid (vertical dotted line), and the book value (vertical dashed line). Pr(No) and Pr(Min) are the predictive probabilities of no bids and a single bid (in which case the winner pays the seller's posted minimum bid), respectively. The solid curve is the predictive price density conditional on at least two bids.

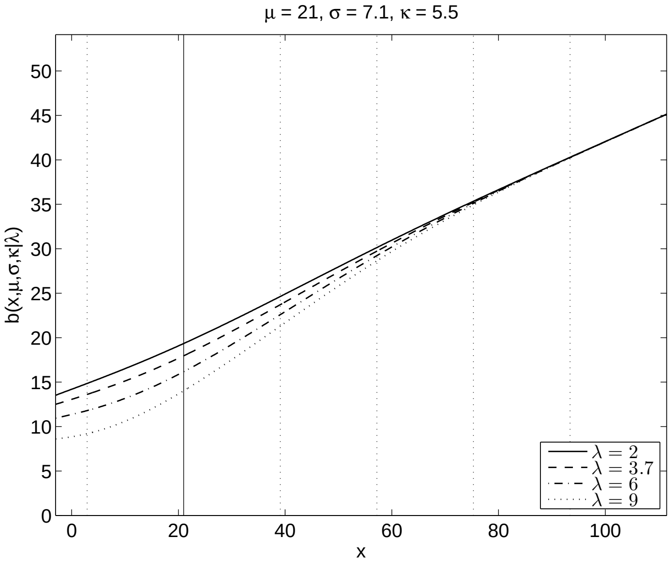

FIGURE 10. The winner's curse effect of the exact bid function with a stochastic number of bidders. Dotted vertical lines represent the positions of different number of standard deviations from the expected signal μ in the unconditional distribution of x, and the solid vertical line represent the position of μ. The representative auction, described in Section 3, is used as a benchmark.

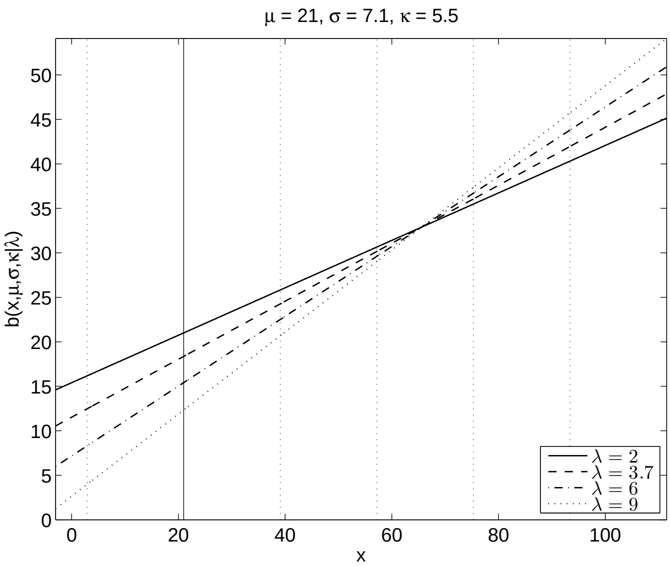

FIGURE 11. The winner's curse effect of the approximate bid function with a stochastic number of bidders. Dotted vertical lines represent the positions of different number of standard deviations from the expected signal μ in the unconditional distribution of x, and the solid vertical line represent the position of μ. The representative auction, described in Section 3, is used as a benchmark.

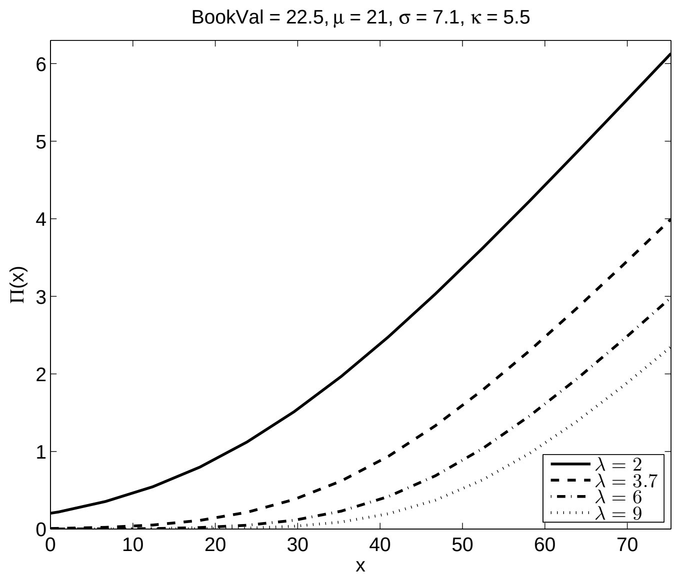

FIGURE 12. The bidder's expected profit for a different number of expected bidders λ. The representative auction, described in Section 3, is used as a benchmark.

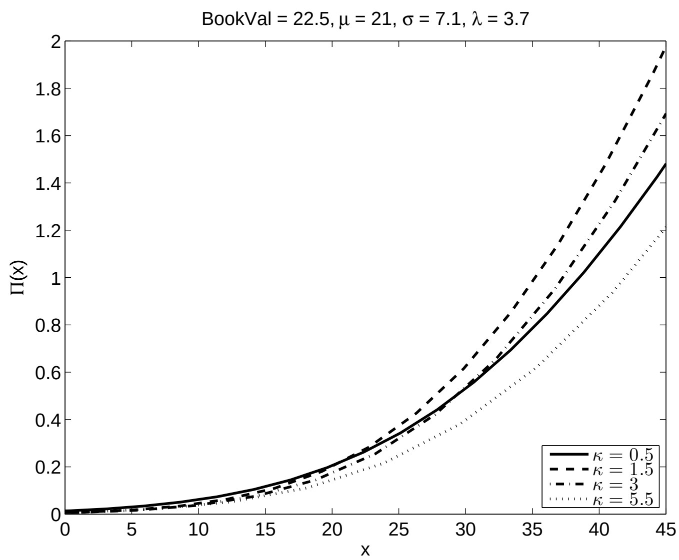

FIGURE 13. The bidder's expected profit for different values of κ, the variance scale parameter for signals. The representative auction, described in Section 3, is used as a benchmark.

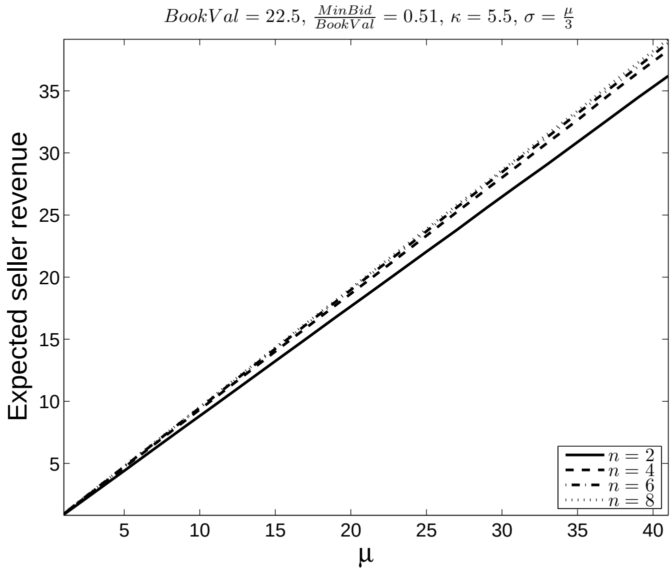

FIGURE 14. The expected seller revenue using the exact bid function. Each curve corresponds to a different number of bidders n. The representative auction, described in Section 3, is used as a benchmark.

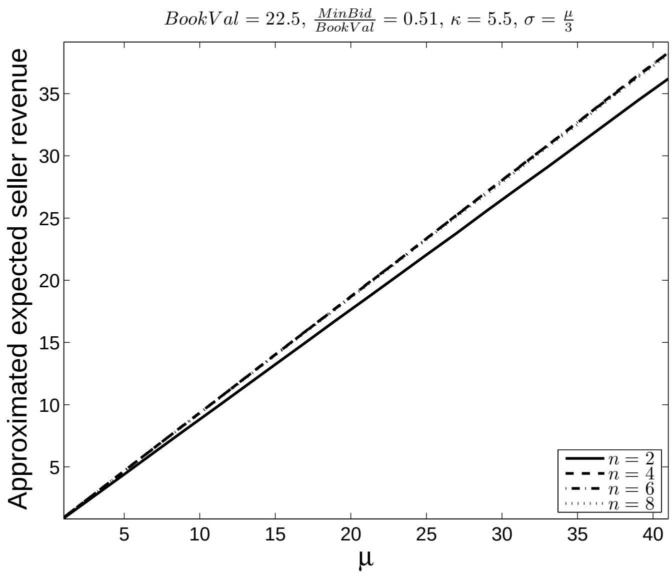

FIGURE 15. The expected seller revenue using the approximate bid function. Each curve corresponds to a different number of bidders n. The representative auction, described in Section 3, is used as a benchmark.

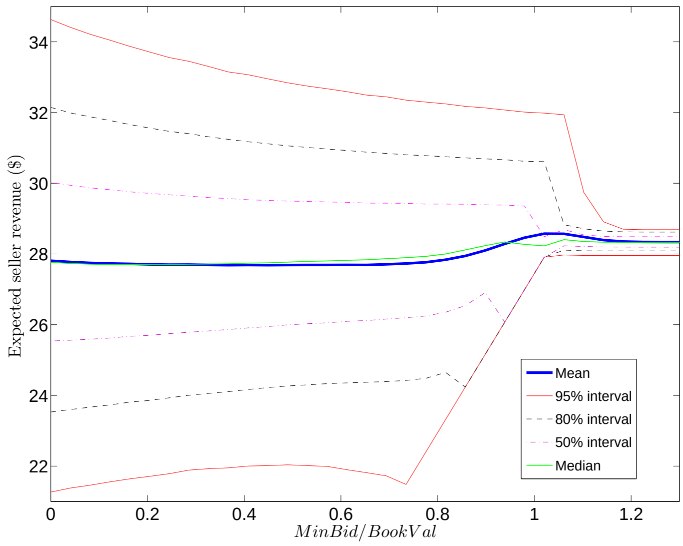

FIGURE 16. Posterior distribution of approximate expected seller revenue in the eBay data as a function of the seller's posted minimum bid divided by the book value. All other covariates are fixed to their values in the representative auction. The approximate bid function was used. It is assumed that the seller obtains a revenue equal to μ if the object is not sold.

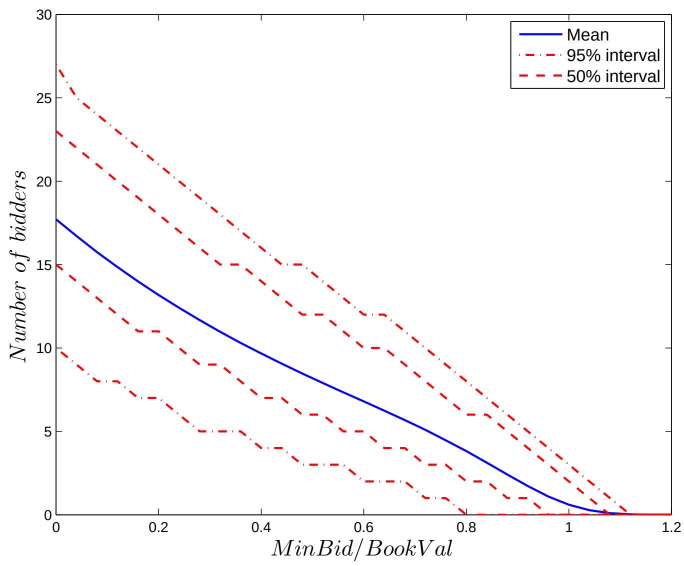

FIGURE 17. Posterior distribution of the number of bidders in the eBay data as a function of the seller's posted minimum bid divided by the book value. All other covariates are fixed to their values in the representative auction. The approximate bid function was used.

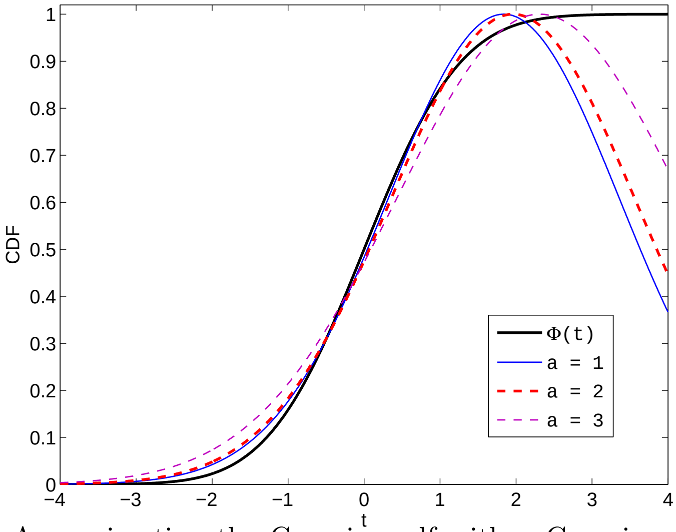

FIGURE 18. Approximating the Gaussian cdf with a Gaussian pdf over the interval [-a,a].

BAYESIAN INFERENCE IN STRUCTURAL SECOND-PRICE AUCTIONS WITH GAMMA DISTRIBUTED COMMON VALUES

BERTIL WEGMANN

ABSTRACT. Our paper explores possible limitations of the Gaussian model in Wegmann and Villani (2011) due to intrinsically non-negative values. The relative performance of the Gaussian model is compared to an extension of the Gamma model in Gordy (1998) within the symmetric second price common value model. A key feature in our approach is the derivation of an accurate approximation of the bid function for the Gamma model, which can be inverted and differentiated analytically. This is extremely valuable for fast and numerically stable evaluations of the likelihood function. The general MCMC algorithm in WV is utilized to estimate WV's eBay dataset from 1000 auctions of U.S. proof coin sets, as well as simulated datasets from the Gamma model with different degrees of skewness in the value distribution. The Gaussian model fits the data slightly better than the Gamma model for the particular eBay dataset, which can be explained by the fairly symmetrical value distribution. The superiority of the Gamma to the Gaussian model is shown to increase for higher degrees of skewness in the simulated datasets.

KEYWORDS: Bayesian variable selection, Bid function approximation, eBay, Gamma model, Gaussian model, Markov Chain Monte Carlo, Non-negative common values.

1. INTRODUCTION

Structural econometric models of auction data have become increasingly common in the literature; see e.g. Bajari (2005) and Paarsch and Hong (2006) for recent surveys. As emphasized by Laffont and Vuong (1996), auction models are particularly well suited for structural estimation since many datasets are available and well-defined game forms exist. This is especially true in the field of Internet auctions, such as eBay, where high quality datasets are readibly available.

Most of the work in econometric models of auction data analyzes either the private or the common value model. Within the private value paradigm, each bidder knows his own valuation of the object and will not be affected by knowing the valuations of the other bidders. In a common value auction the value of the object is unknown but the same for all bidders, and each bidder uses his private information (the signal) to estimate the unknown value. Good examples of structural econometric auction models are Paarsch (1992), Elyakime, Laffont, Loisel and Vuong (1997), and Bajari and Hortacsu (2003, henceforth BH).

Within each paradigm, the values of the objects are often assumed to follow a certain distribution that is commonly known to all bidders. In most empirical work, the outcomes from one or several value distributions are treated in isolation without comparing the performance between the different distributions. Our paper assumes the common value model, and shows how the relative performance between two different distributional setups differs substantially for various types of data.

BH model Internet coin auctions at eBay as independent second price common value auctions with stochastic entry, and assume symmetric bidders and a symmetric Nash equilibrium.

Department of Statistics, Stockholm University, SE-106 91 Stockholm. E-mail: bertil.wegmann@stat.su.se.

The values of the object are assumed to follow a hierarchical Gaussian model where the mean, variance, and expected number of bidders are functions of covariates. Wegmann and Villani (2011, henceforth WV) refine and extend the analysis in BH. WV derive an accurate linear approximation of the bid function that can be inverted and differentiated analytically. This is extremely valuable for fast and numerically stable evaluations of the likelihood function. Moreover, WV use a general Metropolis-Hastings algorithm for Bayesian variable selection to quantify the importance of individual covariates in the model. The model appears to fit the data well, and the out-of-sample predictive performance is good.

Since values are intrinsically non-negative, it can be argued that the assumption of normally distributed values in BH is untenable. The Gaussian model is nevertheless likely to be a very useful approximation when the coefficient of variation (CV) of the value distribution is moderate or small, so that the distribution has a small probability on negative values. The aim of this paper is to explore this issue by contrasting the Gaussian model with an extension of the Gamma model in Gordy (1998). We use the same dataset as in WV from 1000 eBay auctions of U.S. proof coin sets. Other distributions than Gamma could also be used on $(0, \infty)$, but since the Gamma model is closed under multiplication, we are able to derive a suitable and accurate approximate solution of the bid function for the Gamma case. This approximation is non-linear, but of a simple form that can be inverted and differentiated analytically.

The parameter estimates for the two models are nearly the same, in both signs and magnitudes, but the Gaussian model performs slightly better in predicting auction prices and the MCMC algorithm is more efficient in the Gaussian case. The reason why the Gaussian distribution does so well is that the inferred signals in this particular eBay dataset are fairly symmetric and essentially bounded away from zero. To explore the limitations of the Gaussian model for other possible datasets, we conducted a simulation study where the data generating process had increasingly more skewness in the signals. As expected, the Gamma model is shown to be superior to the Gaussian model in this case, and the superiority is increasing for higher degrees of skewness.

2. Two hierarchical models

Based on the theoretical considerations in BH, we model eBay auctions as independent second-price auctions with stochastic entry. To account for asymmetric or positive-valued distributions, we extend the Gamma model in Gordy (1998) and compare this model to the Gaussian model in WV.

2.1. General setup.

Assume that the seller sets a publicly announced *minimum bid (public reserve)*, $r \geq 0$, and that risk-neutral bidders compete for a single object using the same bidding strategy (symmetric equilibrium). The value of the object, v, is unknown and the same for each bidder at the time of bidding, but a prior distribution for v is shared by the bidders. To estimate v, each bidder relies upon her own private information of the object to receive a private signal x from the same distribution $x|v$ (symmetric bidders). Let $f_v(v)$ denote the probability density function of v, $f_{x|v}(x|v)$ the conditional probability density function of $x|v$, and $F_{x|v}(x|v)$ the conditional cumulative distribution function of $x|v$. Since the auction involves symmetric bidders and a symmetric equilibrium, we can focus on a single bidder without loss of generality. The bid function can be written as (see BH for an implicit solution)

$$b(x, \lambda) = \begin{cases} \dfrac{\sum_{N=2}^{\infty}(N-1)\cdot p_{N-1}(\lambda)\cdot\int_v v\cdot F_{x|v}^{N-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv}{\sum_{N=2}^{\infty}(N-1)\cdot p_{N-1}(\lambda)\cdot\int_v F_{x|v}^{N-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv}, & \text{if } x \geq x^\star \\ 0, & \text{otherwise,} \end{cases} \tag{2.1}$$

where $p_{N-1}(\lambda)$ is the Poisson probability of $(N-1)$ potential bidders in the auction with λ as the expected value of the Poisson entry process. Bidders participate with a positive bid if their signal, x, is above the cut-off signal level $x^{\star}$. Given an arbitrary bidder with signal x, let y be the maximum signal of the other $(N-1)$ bidders. The cut-off signal level is then given in implicit form as (Milgrom and Weber, 1982)

$$x^{\star}(r,\lambda) = \inf_{x} \left(E_N E[v|X=x, Y<x, N] \geq r \right),$$

which gives the minimum bid, r, as

$$r(x^{\star},\lambda) = \sum_{N=1}^{\infty} p_N(\lambda) \cdot \frac{\int_v v \cdot F_{x|v}^{N-1}(x^{\star}|v) \cdot f_{x|v}(x^{\star}|v) \cdot f_v(v)\ dv}{\int_v F_{x|v}^{N-1}(x^{\star}|v) \cdot f_{x|v}(x^{\star}|v) \cdot f_v(v)\ dv}. \tag{2.2}$$

Note that the seller's publicly announced minimum bid, r, is only written as a function of the cut-off signal for tractability.

In both the Gaussian and the Gamma models, presented below, the expected value μ and the variance σ^2 exist in the distribution of v. Moreover, $E[x|v] = v$, and $E\left[\frac{1}{x}|v\right] = \frac{1}{v}$ in the Gaussian and Gamma models, respectively. Similarly to BH, we specify regression models for $(\mu_j, \sigma_j^2, \lambda_j)$ in auction j as

$$\begin{aligned} \mu_j &= z'_{\mu j}\beta_\mu \\ \sigma_j^2 &= \exp\left(z'_{\sigma j}\beta_\sigma\right) \\ \lambda_j &= \exp\left(z'_{\lambda j}\beta_\lambda\right), \end{aligned} \tag{2.3}$$

where $z_j = (z'_{\mu j}, z'_{\sigma j}, z'_{\lambda j})'$ are auction-specific covariates in auction j.

The likelihood function of bids is complicated since some bids are unobserved. First, potential bidders with signals x below $x^{\star}$ do not place any bid. Second, the highest bid is usually not observed because of eBay's proxy bidding system (see BH for a detailed description). The bid distribution for a single auction is of the form:

$$f_b\left(b|\boldsymbol{\beta}, \boldsymbol{\eta}, r, \boldsymbol{z}, v\right) = f_{x|v}\left[\phi\left(b\right)|\boldsymbol{\beta}, \boldsymbol{\eta}, r, \boldsymbol{z}, v\right]\phi'\left(b\right), \tag{2.4}$$

where $\phi(b)$ is the inverse bid function, and $\boldsymbol{\eta}$ is a vector of additional parameters in the model. Let n be the number of participating bidders who submit a positive bid above r in an arbitrary auction, and let $\mathbf{b} = (b_2, b_3, \ldots, b_n)$ be the vector of observed bids where $b_2 > b_3 > \cdots > b_n$. Then, the likelihood function for an arbitrary auction is given by

$$\begin{aligned} &f_{\mathbf{b}}\left(b_2, b_3, \ldots, b_n|\boldsymbol{\beta}, \boldsymbol{\eta}, r, \boldsymbol{z}\right) \\ &= \sum_{i=n}^{\bar{N}} p_i(\lambda) \cdot \int_{-\infty}^{\infty} F_{x|v}\left(x^{\star}|\boldsymbol{\beta}, \boldsymbol{\eta}, v\right)^{i-n} \cdot \left\{1 - F_{x|v}\left[\phi\left(b_2\right)|\boldsymbol{\beta}, \boldsymbol{\eta}, v\right]\right\}^{I(n\geq 1)} \\ &\qquad \times \prod_{i=2}^{n} f_b\left(b_i|\boldsymbol{\beta}, \boldsymbol{\eta}, r, \boldsymbol{z}, v\right) \cdot f_v(v|\boldsymbol{\beta})dv, \end{aligned} \tag{2.5}$$

where $I(n \geq 1)$ is an indicator variable for at least one bidder in the auction, and $\bar{N}$ is an upper bound for the total number of potential bidders. For the sake of tractability, let $\bar{N} = 30$, as in BH. If $n = 1$, b_2 is replaced by the minimum bid r.

We use the same Bayesian methods as in WV to estimate the models, see Section 3.1. A single evaluation of the posterior (likelihood) requires numerical integration to compute $b(x|\beta, r, z)$ in (2.1), followed by additional numerical work to invert and differentiate $b(x|\beta, r, z)$. The same applies to the computation of $r(x^{\star}|\beta, z)$ in (2.2). This costly procedure needs to be repeated for each of the auctions in the dataset. Instead, we make use of bid approximations for both models which leads to much faster and numerically stable likelihood evaluations. We first review the linear approximation of the bid function in WV for the hierarchical Gaussian model. Then, we derive the approximate bid function for the hierarchical Gamma model below.

2.2. **Gaussian model.** Let v_j denote the common value in auction j, and let x_{ij} denote the signal of the ith bidder in auction j. The hierarchical Gaussian model can be defined as

$$\begin{aligned} v_j &\sim N(\mu_j, \sigma_j^2),\;\; j=1,...,m, \\ x_{ij}\,|v_j &\overset{iid}{\sim} N(v_j, \kappa\sigma_j^2),\;\; i=1,...,N_j, \end{aligned} \tag{2.6}$$

where m is the total number of auctions, and N_j the number of potential bidders that bid zero or place a positive bid in auction j. To get much faster and numerically stable likelihood evaluations, we use the linear approximation of the bid function in WV,

$$b(x,\lambda) \approx \begin{cases} c+\omega\mu+(1-\omega)x, \text{ if } x \geq x^\star \\ 0, \text{ otherwise,} \end{cases} \tag{2.7}$$

where $c = -\frac{\sqrt{\kappa}\sigma\gamma\theta(\lambda-2)}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\omega = \frac{\frac{\kappa}{2}}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\theta = 1.96$ and $\gamma = 0.1938$. In addition, WV show that the cut-off signal level can be similarly approximated by

$$x^\star(r,\lambda) \approx \frac{r-\sum_{n=2}^{\infty} p_n(\lambda)(\tilde{c}+\tilde{\omega}\mu)}{\sum_{n=2}^{\infty} p_n(\lambda)(1-\tilde{\omega})}, \tag{2.8}$$

where $\tilde{c} = -\frac{\sqrt{\kappa}\sigma\gamma\theta(n-1)}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$, and $\tilde{\omega} = \frac{\frac{\kappa}{2}}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$. Using the approximate bid function, the distribution of the bids in (2.4) simplifies to

$$b|v \in N[c+\mu, (1-\omega)^2\kappa\sigma^2], \tag{2.9}$$

which speeds up the likelihood evaluation even more.

2.3. **Gamma model.** A drawback of the Gaussian model in the previous section is that the common value, the signals and the bids can be negative. It can be argued that the Gaussian distribution serves as a good approximation when the mean is at least a couple of standard deviations away from zero, which is a common situation in practice. This is clearly not always the case, however, and there are situations when it would be better to use a distribution with non-negative support. Therefore, we will extend and analyze a model in Gordy (1998) based on the Gamma distribution.

The Gamma model is more conveniently written in terms of *inverse* signals, $s_{ij} = \frac{1}{x_{ij}}$,

$$\begin{aligned} v_j &\sim Gamma(\xi_j,\psi_j),\;\; j=1,...,m,\;\; E\,(v_j)=\mu_j=\frac{\xi_j}{\psi_j},\;\; Var(v_j)=\sigma_j^2=\frac{\xi_j}{\psi_j^2} \\ s_{ij}\,|v_j &\overset{iid}{\sim} Gamma(\tau,\tau v_j),\;\; i=1,...,N_j,\;\; E\,(s_{ij}\,|v_j\,)=\frac{1}{v_j},\;\; Var\,(s_{ij}\,|v_j\,)=\frac{1}{\tau v_j^2}, \end{aligned} \tag{2.10}$$

where τ is a precision parameter, m is the total number of auctions, and N_j is the number of potential bidders that bid zero or place a positive bid in auction j. The bid function for the

Gamma model is of the form (see Gordy (1998) for the case with a fixed number of bidders)

$$b(x,\lambda) = \begin{cases} \frac{\sum_{N=2}^{\infty}(N-1)\cdot p_{N-1}(\lambda)\cdot\int_0^\infty v\cdot(1-F_{s|v}(1/x|v))^{N-2}\cdot f_{s|v}^2(1/x|v)\cdot f_v(v)\ dv}{\sum_{N=2}^{\infty}(N-1)\cdot p_{N-1}(\lambda)\cdot\int_0^\infty (1-F_{s|v}(1/x|v))^{N-2}\cdot f_{s|v}^2(1/x|v)\cdot f_v(v)\ dv}, & \text{if } x \geq x^\star \\ 0, \text{ otherwise}, \end{cases} \tag{2.11}$$

and the minimum bid function for the Gamma model is given by

$$r(x^\star,\lambda) = \sum_{N=1}^{\infty} p_N(\lambda)\cdot\frac{\int_0^\infty v\cdot(1-F_{s|v}(1/x^\star|v))^{N-1}\cdot f_{s|v}(1/x^\star|v)\cdot f_v(v)\ dv}{\int_0^\infty (1-F_{s|v}(1/x^\star|v))^{N-1}\cdot f_{s|v}(1/x^\star|v)\cdot f_v(v)\ dv}. \tag{2.12}$$

Gordy (1998) obtains a finite series expansion of the bid function for the Gamma model (see $B_2(x)$ in equation 7 in Gordy's article). Gordy's solution for the bid function is elegant and fast, but there are two reasons why it is of limited use in a likelihood-based approach. First, the likelihood function depends on the *inverse* bid function which has to be solved numerically for each bid in every auction. Second, the solution is restricted to the set of positive integers for τ, which makes it hard to link τ to covariates. The equilibrium bid function for the Gamma model can be approximated in a similar way as for the Gaussian model in WV, see Appendix. The approximate bid function is given by

$$b(x) \approx \frac{[\xi+2\tau+(\lambda-2)\xi_\tau]\cdot x}{\psi x+2\tau+(\lambda-2)\psi_\tau}, \tag{2.13}$$

where

$$\xi_\tau = -0.1659 - 0.0159\cdot\tau + 0.2495\cdot\log(1+\tau),$$

and

$$\psi_\tau = 0.3095 - 0.0708\cdot\tau + 1.0241\cdot\log(1+\tau),$$

for $0.1 \leq \tau \leq 10$. If necessary, other values of τ can be handled similarly, see the Appendix for details.

The approximate bid function in Equation (2.13) has the following properties. First, the more precise is the public information (larger ψ), the less weight is placed on the bidder's private signal. To see this, replace ξ with $\psi\mu$ in Equation (2.13) to obtain

$$b(x) \approx \frac{\psi\mu+2\tau+(\lambda-2)\xi_\tau}{\psi+\frac{2\tau+(\lambda-2)\psi_\tau}{x}}, \tag{2.14}$$

which yields

$$b(x) \longrightarrow \mu\ if\ \psi \longrightarrow \infty \ \text{ and }\ b(x) \longrightarrow \frac{2\tau+(\lambda-2)\xi_\tau}{2\tau+(\lambda-2)\psi_\tau}\cdot x\ if\ \psi \longrightarrow 0.$$

The approximate bid function also satisfies

$$b(x) \longrightarrow \mu\ if\ \tau \longrightarrow 0,$$

and

$$b(x) \approx \frac{\left[2+\frac{\xi}{\tau}+(\lambda-2)\frac{\xi_\tau}{\tau}\right]\cdot x}{2+\frac{\psi x}{\tau}+(\lambda-2)\frac{\psi_\tau}{\tau}} \longrightarrow x\ if\ \tau \longrightarrow \infty.$$

The last two results are easily proved by noting that our approximation routine described in the Appendix implies that

$$\left(\frac{\xi_\tau}{\tau}, \frac{\psi_\tau}{\tau}\right) \longrightarrow (0,0)\ if\ \tau \longrightarrow \infty,$$

and

$$(\xi_\tau, \psi_\tau) \longrightarrow (0,0) \; if \; \tau \longrightarrow 0.$$

The approximate minimum bid function is given by

$$r(x^\star, \lambda) \approx \frac{[\xi + \tau + (\lambda - 1)\xi_\tau] \cdot x^\star}{\psi x^\star + \tau + (\lambda - 1)\psi_\tau}. \tag{2.15}$$

Figure 1 compares the exact and approximate bid function graphically. The exact bid function is computed by numerical integration as in BH. The upper left sub-graph displays the bid function and its approximation for a representative auction in the eBay dataset analyzed in Section 3. The representative auction is based on the median of the covariates in the eBay data, and the posterior mean of the model parameters. Rounded to the nearest integer, this gives $\tau = 3, \mu = 21, \sigma = 8, \lambda = 4$, and $\frac{r}{\mu} = 0.5$. The other sub graphs are variations from the representative auction. The approximation deteriorates somewhat with increasing λ, but is still very accurate, irrespective of τ and σ.

3. Model comparison on eBay coin auction data

The performance between the Gaussian and Gamma models is here compared using the same dataset as in WV from 1000 eBay auctions of U.S. proof coin sets; see WV for a detailed description of the data. The unknown model parameters in each model are estimated with the same Bayesian methods and priors as the benchmark model in WV, which are briefly summarized below.

3.1. Prior distribution and a Metropolis-Hastings algorithm for variable selection.

The likelihood function in (2.5) is combined with a prior distribution on the unknown model parameters to form the posterior density of each model. A generalization of the Metropolis-Hastings algorithm (a Markov Chain Monte Carlo (MCMC) algorithm) is used to simultaneously do Bayesian variable selection among the auction covariates and sample the posterior of the model parameters (WV, Section 4.2 and Appendix B). Variable selection involves adding point masses at zero in the prior distributions; see Smith and Kohn (1996).

The prior distribution for β in the Gaussian model is given by a g-prior (Zellner, 1986) for β_μ, conditional on β_σ, as (see Villani et. al. 2009 for details)

$$\beta_\mu | \beta_\sigma \sim N[0, c(Z'_\mu D Z_\mu)^{-1}],$$

where $D^{1/2} = \text{Diag}[\exp(-z'_{\sigma_1}\beta_\sigma/2), ..., \exp(-z'_{\sigma_n}\beta_\sigma/2)]$, and $c > 0$ is a scaling factor equal to the number of auctions in the data for the benchmark model. The marginal g-prior for β_σ is given by

$$\beta_\sigma \sim N[0, c(Z'_\sigma Z_\sigma)^{-1}].$$

The Gamma model in Section 2.3 cannot be written as a heteroscedastic linear regression, and a similar characterization of the prior for β_μ and β_σ is therefore not possible. We will instead assume that $\beta_\mu \sim N[0, c_\mu(Z'_\mu Z_\mu)^{-1}]$ independently from $\beta_\sigma \sim N[0, c_\sigma(Z'_\sigma Z_\sigma)^{-1}]$ in the Gamma model.

We use a g-prior for β_λ in the Poisson entry model, given by β_λ

$$\beta_\lambda \sim N[0, c_\lambda(Z'_\lambda Z_\lambda)^{-1}].$$

We use an inverse Gamma prior for κ in the Gaussian model, $\kappa \sim IG(\bar{\kappa}, g)$, where $\bar{\kappa} = 0.25$ is the prior mean of κ and $g = 4$ are the degrees of freedom. Similarly, we assume the prior $\tau \sim IG(\bar{\tau} = 0.25, h = 4)$ in the Gamma model. Finally, the prior inclusion probability of a given covariate is set to 0.2.

3.2. **Estimation results.** The models are estimated using the dataset of 1000 eBay coin auctions in WV. The dataset was carefully collected by human inspections of auction-specific covariates and bid sequences from auctions that ended between November 7 to December 19, 2007 and December 27, 2007 to January 22, 2008. The book value of the object and the minimum bid divided by the book value in each auction are defined by Book and MinBidShare. Remaining covariates are dummy variables that are coded to be 1 if a certain characteristic is present in a given auction. The largest sellers are described by Pow (PowerSellers) and ID describes a seller whose identity is verified on eBay. MinBlem (MajBlem) describes if the object has a minor (major) damage, and UnOpen pertains to an object which is sold in its original sealed envelope. A large negative feedback score for a seller is described by LargeNeg.

Similarly to WV, we perform some transformations on the covariates before the analysis to better match the functional form of the model and reduce the correlation between some of the slope coefficients and intercepts. We use the notation, $x_d = x - \bar{x}$ and $\text{L}x = ln(x)$, where x is a covariate. As an example, LBook_d is the deviation of log book value from the mean log book value.

Table 1 reports our posterior results for the Gaussian and Gamma models. With a few exceptions, the results are fairly similar between the models. The covariates Book_d and LBook_d are the main drivers in the models for μ and σ^2, respectively. The large negative coefficient on MinBidShare_d in the Poisson entry process is due to the fact that a higher minimum bid implies a higher threshold for potential bidders to participate in the auction. In general, the posterior inclusion probabilities are either close to 0 or 1, which gives good indications of which covariates that are of importance in the models. Neither of the posterior inclusion probabilities for $\text{Book} \cdot \text{ID}$, $\text{LBook} \cdot \text{Pow}$, and $\text{LBook} \cdot \text{ID}$ are close to 1 in both models. This suggests that eBay's detailed seller information is not an obvious significant source for how the expected value or the standard deviation of the object's unknown common value is affected.

The major differences between the models are in the parameter estimates of κ, and τ, and between the parameter estimates in the models for σ^2. This is probably due to the differences in the model setups. Parameter σ^2 is used in both the distributions of v and $x|v$ for the Gaussian model compared to the Gamma model with σ^2 only used in the distribution of v. In addition, v is defined as the expected value in the distribution of $x|v$ for the Gaussian model, whereas v is used in both the expected value and the variance of the inverse signal distribution for the Gamma model. It seems that the lower value of τ, compared to κ, compensates for the generally higher values of the parameter estimates in σ^2 for the Gamma model.

The posterior inclusion probabilities for most of the covariates are very close in magnitude between the models, especially for the covariates in λ. The major differences in inclusion probabilities for Book*ID, LBook*Pow, and LBook*ID are due to only minor differences of the *Bayesian t-ratio*,

$$t_{Bayesian} = \frac{\left|\hat{\theta}\right|}{s(\hat{\theta})},$$

where $\hat{\theta}$ and $s(\hat{\theta})$ are the posterior mode and the approximate/asymptotic posterior deviation, respectively, from the optimization of the posterior density. WV show that the inclusion probabilities are very sensitive to changes around a certain threshold of the Bayesian t-ratios. The inclusion probabilities increase sharply above the threshold value (slightly larger than the 1.96 threshold in classical t-tests at a 5% significance level). Hence, differences in inclusion probabilities which are not close to either 0 or 1 should not be treated as a sign of very different estimation results between the models.

The performance of the MCMC algorithm is better in the Gaussian case. The mean acceptance probability is 56% for the Gaussian model compared to 30% for the Gamma model, and the inefficiency factors (IF) are lower in the Gaussian case. In WV, the IF is defined as the number of draws needed to obtain the equivalent result of a single independent draw. IFs close to unity is therefore an indication of a very efficient algorithm. The minimum, the median, and the maximum values of IF are $3.03, 3.80$, and 11.54 for the Gaussian model, respectively, compared to $5.08, 19.30$, and 36.80 for the Gamma model.

3.3. **In-sample fit.** Similar to BH and WV, we use a posterior predictive analysis to evaluate the in-sample fit of the models, i.e. we compare the observed data to simulated data from the estimated models (Gelman et al., 2004). Given the observed auction-specific covariates, we simulated 100 new complete datasets for each of a 100 systematically sampled posterior draws of the model parameters. This gives us $10,000$ full datasets, each with bids from 1000 auctions.

Following BH, we compare the observed and simulated data through two summary statistics: *within-auction bid dispersion* and *cross-auction heterogeneity*. The within-auction dispersion is defined as the difference between the highest observed bid and the lowest bid divided by the book value of the auctioned item, and the cross-auction heterogeneity is investigated by histograms of the bids divided by the corresponding book value in that auction. As we can see in Figures 2 and 3, the observed within-auction bid dispersion and the cross-auction heterogeneity are very well captured by both the Gaussian and the Gamma models, and the differences between the two models are small. This is in contrast to BH where the within-auction bid dispersion is under-estimated and the cross-auction heterogeneity is highly over-estimated. BH suggest adding unobserved heterogeneity to the model as a way of improving the in-sample fit of the model, but our results in Figures 2 and 3 suggest that such extensions are not needed to capture the variation in the bids.

3.4. **Out-of-sample predictions.** The differences between the models are most apparent in their out-of-sample predictions. Following WV, we use our estimated Gaussian and Gamma models to predict a dataset of 48 additional auctions of *U.S.* proof sets, which were not used in the estimation process. Given the covariates from these auctions, we simulated predictive price distributions for each auction in a similar way as for the simulated datasets of the in-sample fits. The predictive distributions have three components: i) a probability that the item is not sold, ii) a point mass with probability at the minimum bid (which is the final price when there is a single bidder in the auction), and iii) a continuous price density conditional on there being at least two bidders in the auction.

The predictive price distributions for both models are displayed on top of each other for each auction in Figures 4 and 5. Auction $1-37$ contains at least two bids, $38-45$ a single bid, and $46-48$ no bids. Both models perform well in predicting auction prices with the actual price located near the center of the visible continuous price distributions. The variances of the predictive price densities for the Gaussian model appear to be smaller than for the Gamma model, where the price distributions are lower and skewed to the left.

To quantify the visual results in Figures 4 and 5, we compute the log predictive score (LPS) for each auction in the 48 out-of-sample auctions. Since not all auctions end with a realized price, we compute two versions of the LPS: LPS_c evaluates the log predictive density in the 37 auctions with at least two bids, while LPS_d is a multinomial type of evaluation that focuses on three discrete events: i) no bids, ii) one bid (where price = minimum bid) and iii) more than one bid. LPS_d is evaluated on all 48 out-of-sample auctions.

Let p_{ij} be the probability of j bids in auction i, and let

$$I_{ij} = \begin{cases} 1, \text{ if there are } j \text{ actual bids in auction } i \\ \\ 0, \text{ if there are not } j \text{ actual bids in auction } i, \end{cases}$$

where $i = 1, \ldots, m_t,\ j = 0, 1$, and m_t is the number of test auctions used to evaluate the predictions. Let $\mathbf{I} = (I_{10}, I_{11}, I_{12}, I_{20}, \ldots, I_{m_t 2})$ be the vector of observed indicator variables, and let $\mathbf{p} = (p_{10}, p_{11}, p_{12}, p_{20}, \ldots, p_{m_t 2})$ be the vector of predictive probabilities for each auction, where $I_{i2} = (1 - I_{i0} - I_{i1})$, and $p_{i2} = (1 - p_{i0} - p_{i1})$ are the indicator variable and the predictive probability for at least two bids in auction i, respectively. Then, the discrete version of the LPS is defined as

$$LPS_d = \frac{\mathbf{I} \cdot \mathbf{p}'}{m_t}. \tag{3.1}$$

Table 2 presents the LPS$_d$ for the 48 test auctions. In a majority of the auctions, the Gaussian model attains a higher score than the Gamma model, even if the mean value of LPS$_d$ is lower across all auctions for the Gaussian model. Generally, the scores from both models are much lower for the 11 auctions with a maximum of one bid, where the Gaussian model attains a substantially lower score than the Gamma model. In all, when it comes to the discrete part of the predictive distribution, this suggests a slight edge for the Gaussian model in auctions with at least two bids, and a slight edge for the Gamma model whenever there is a maximum of one bid in the auction.

The continuous version of the LPS is defined by

$$LPS_c = \frac{\sum_{i=1}^{m_t^\star} \log\left(\frac{\tilde{p}(y_i)}{p_{i,2}}\right)}{m_t^\star}, \tag{3.2}$$

where y_i is the realized price in auction i evaluated in the predictive price distribution $\tilde{p}(\cdot)$, and $m_t^\star$ is the number of test auctions with at least two bids. The Gaussian model outperforms the Gamma model here. As can be seen in Table 2, the Gaussian model performs best in a large majority of the auctions, and the score for the Gaussian model is substantially higher compared to the Gamma model. In Figures 6 and 7 the discrepancies in scores are obvious. The Gaussian model attains, in general, higher predictive probabilities at the realized prices than the Gamma model.

The reason for the relatively good performance of the Gaussian model is probably that the true value distribution is close to normal, or at least symmetric, for this particular dataset. To infer the magnitude of the skewness in the value distribution, we obtained the values of μ and σ for the Gamma model in each of the 48 test auctions. Specifically, given the covariates in an arbitrary auction and the posterior means for the Gamma model in Table 1, the typical degree of skewness, S_k, can be calculated for each auction as

$$S_k = \frac{2}{\sqrt{\xi}} = 2\frac{\sigma}{\mu} = 2c_\nu,$$

where c_ν is the coefficient of variation. Figure 8 displays a boxplot for the different degrees of skewness in the 48 test auctions. In most auctions, the degree of skewness is about $0.7 - 0.8$, which suggests that the value distribution is fairly symmetric compared to the results in the next section for the least skewed model. In that upcoming section, we simulated data with different degrees of skewness in the value distribution to investigate the relative performance between the Gaussian and Gamma models.

4. Model comparisons on skewed simulated data

We conducted a simulation study to compare the performance between the Gaussian and Gamma models using positive values and different degrees of skewness, S_k, in the data-generating process. Specifically, three Gamma models with $S_k = 2$ (the exponential distribution), $S_k = 1.5$, and $S_k = 0.5$ (the least skewed model) were used to simulate 25 full datasets of bids and auction-specific covariates in 1000 auctions for each model. Then, each dataset for each degree of skewness was estimated by both the Gaussian and Gamma models.

4.1. **In-sample fit.** Since the book value does not exist as a covariate for the simulated data, we exclude the book value from the definitions of the within-auction dispersion and the cross-auction heterogeneity. Hence, the within-auction dispersion is here defined as the difference between the highest observed bid and the lowest bid, while the cross-auction heterogeneity is displayed as histograms of the bids across all auctions.

Figures 9 to 11 display the median of the within-auction dispersion and the cross-auction heterogeneity across the 25 simulated datasets for each of the three degrees of skewness. The Gamma model fits the data very well in all cases. As the degree of skewness increases, the fit of both the within-auction dispersion and the cross-auction heterogeneity becomes worse for the Gaussian model. However, in Figure 11 for the least skewed data, the Gamma model fits the cross-auction heterogeneity (the distribution of bids) only slightly better than the Gaussian counterpart. This is probably due to the quite symmetrical distribution of bids for the actual data in the figure. In fact, whenever the distribution of bids is symmetric, the Gaussian model fits the data very well, since the bid distribution of the Gaussian model in Equation (2.9) is normal for the approximate bid function in Equation (2.7). In Figures $8-9$, the distribution of bids is highly skewed, and the Gaussian fit of the data becomes highly skewed to the right.

4.2. **Out-of-sample predictions.** As the degree of skewness increases in the data-generating process, the variance in the common value distribution becomes larger. Therefore, to compare the LPS_c in a more appropriate way for different degrees of skewness, we adjuste the LPS_c in this case to

$$LPS_c = \frac{\left(\sum_{i=1}^{m_t^\star} \log\left(\frac{\tilde{p}(y_i)}{p_{i,2}}\right)\right)/m_t^\star}{\sqrt{Var(v)}} = \frac{\psi S_k}{2}\frac{\sum_{i=1}^{m_t^\star} \log\left(\frac{\tilde{p}(y_i)}{p_{i,2}}\right)}{m_t^\star}. \tag{4.1}$$

Table 3 reports the LPS for each degree of skewness. The Gamma model outperforms the Gaussian model in the continuous part of the predictive price distribution, except for the lowest degree of skewness where the Gamma model is only slightly better than the Gaussian model. On average, the mean and standard deviation of LPS_c are higher and lower, respectively, for the Gamma model. However, for roughly half of the 25 datasets, the Gaussian model has a higher LPS_c than the Gamma model in a majority of auctions, regardless of degree of skewness. To get a closer look at possible differences, we compared the predictive price densities of the Gamma and the Gaussian model in several datasets for each degree of skewness.

In general, the densities for the Gamma model were more concentrated and skewed to the left than the Gaussian counterpart where the densities were more symmetrical and more located to the right. This resulted in substantial differences between the LPS_c of the models for realized prices to the left in the predictive distribution, but only minor differences to the right. In the left-hand part of the predictive price densities, the Gamma model was clearly the dominating model in LPS_c, while the Gaussian model only had a slightly higher LPS_c than the Gamma model in the right-hand part. As the degree of skewness increased, these differences became larger.

Turning to the discrete part of the predictive distribution, we can discern from Table 3 that the Gamma and the Gaussian model alternate in attaining the lowest scores for at least two bids and a maximum of one bid. Whenever there are at least two bids the Gamma model is, on average, outperforming the Gaussian model with substantially higher scores, in contrast to the opposite effect for a maximum of one bid where the Gamma model, on average, attains lower scores. This is consistent with the result from the eBay dataset where the Gaussian model is generally the dominating model. The mean and the precision of the scores are increasing in the degree of skewness for at least two bids, while the opposite effect pertains to the Gamma model for a maximum of one bid.

Moreover, the differences in scores decrease as the data becomes more symmetrical, and for the most symmetrical case, the models are once more equally good as for the continuous part of the predictive distribution. In fact, since the Gaussian model performs better than the Gamma model for the eBay data, and since both models seem to perform equally well for $S_k = 0.5$, there is possibly a threshold, less than $S_k = 0.5$, when the Gaussian model becomes the best fitted model for lower degrees of skewness. In all, the results from the simulated data show that the Gamma model is clearly outperforming the Gaussian model for highly skewed data, but the differences become smaller when the degree of skewness decreases.

5. Conclusions

An inherent feature of econometric auction models is that the value of the auctioned object is intrinsically non-negative. Nevertheless, there often exist value distributions in the literature that allow for negative values, even if such distributions are untenable. Our paper explores this issue by contrasting the Gaussian model in WV with an extension of the Gamma model in Gordy (1998). A key feature in our approach is an accurate, analytical approximation of the bid function for the Gamma model, which can be inverted and differentiated analytically. This is extremely valuable for a fast and numerically stable evaluation of the likelihood function.

We utilized the general MCMC algorithm for Bayesian variable selection in WV to compare the relative performance between the Gaussian and Gamma models, using WV's eBay dataset from 1000 auctions of U.S. proof coin sets. In general, the Gaussian model was slightly better than the Gamma model in fitting this data, and in predicting auction prices on the 48 out-of-sample auctions. This is probably due to an almost normal or at least symmetric value distribution, where the density is bounded away from zero with small probabilities of negative values. In fact, we obtained small values for the coefficient of variation and the degree of skewness in most auctions, which support this explanation.

To explore possible limitations of the Gaussian model for other degrees of skewness in the value distribution, we simulated 25 datasets from three Gamma models with different degrees of skewness. Then, the relative performance between the Gaussian and Gamma models was compared for each case. The Gamma model was clearly outperforming the Gaussian model for the two most skewed models, but was only slightly better for the least skewed model. This agrees with the result from the eBay dataset. As the value distribution becomes more symmetric the relative performance of the Gaussian model increases compared to the Gamma model. This suggests that it is of importance to make use of either model, or both, depending on the skewness of the data.

Appendix. Approximating the bid function in the Gamma model

The key ingredient in our approach is to approximate the survival function $(1-G_{s|v}(1/x|v))$ by a Gamma probability density function over the whole interval $(0, \infty)$. This is superior to the approximation of the bid function for the Gaussian model in WV, where a certain interval needs to be obtained for the approximation of the standard normal distribution function. By substitution, the distribution function of $s|v$ becomes

$$F_{s|v}(1/x|v) = \int_0^{1/x} \frac{(\tau v)^\tau}{\Gamma(\tau)} l^{\tau-1} e^{-\tau v \cdot l} dl = \int_0^{\tau \cdot \frac{v}{x}} \frac{1}{\Gamma(\tau)} t^{\tau-1} e^{-t}\, dt.$$

Hence, the distribution function $F_{s|v}$ depends on the parameter τ through the support $\frac{v}{x}$. Let $h\left(\frac{v}{x}|\xi_\tau, \psi_\tau\right) = \frac{\psi_\tau^{\xi_\tau+1}}{\Gamma(\xi_\tau+1)}\left(\frac{v}{x}\right)^{\xi_\tau} \cdot e^{-\psi_\tau \cdot \frac{v}{x}}$ be the approximating Gamma *p.d.f.* to $(1 - F_{s|v}(1/x|v))$. Then, given an arbitrary τ, the approximation constants ξ_τ and ψ_τ are obtained by minimizing the maximum divergence between $h\left(\frac{v}{x}|\xi_\tau, \psi_\tau\right)$ and $(1 - F_{s|v}(1/x|v))$, i.e.

$$(\hat{\xi_\tau}, \hat{\psi_\tau}) = \min_{\xi_\tau, \psi_\tau} \left(\max_{\frac{v}{x}} \left| h\left(\frac{v}{x}|\xi_\tau, \psi_\tau\right) - \left(1 - F_{s|v}(1/x|v)\right) \right| \right).$$

The approximation constants can, in principle, be calculated for any τ, but, in practice, it is more convenient to tabulate them over a grid $\mathcal{T} = (\tau_1, ..., \tau_T)$ of values for τ. Since the approximation of the bid function needs to be solved for any τ in the estimation process, we model $\{\hat{\xi_\tau}, \hat{\psi_\tau}\}_{\tau \in \mathcal{T}}$ as a multivariate regression with several functions of τ as independent variables. The fit of the regression is improved if the grid $\mathcal{T}$ is not too wide, and we will here choose a grid that covers all relevant values of τ in our datasets. The best multivariate regression model (according to adjusted R^2) is

$$\hat{\xi_\tau} = -0.1659 - 0.0159 \cdot \tau + 0.2495 \cdot \log(1+\tau),$$

and

$$\hat{\psi_\tau} = 0.3095 - 0.0708 \cdot \tau + 1.0241 \cdot \log(1+\tau)$$

for $0.1 \leq \tau \leq 10$ with R^2 equal to 99.5% and 99.2%, respectively.

Now, by replacing $(1 - F_{S|V}(1/x|v))$ with $h_v(v|\hat{\xi_\tau}, \hat{\psi_\tau})$, the approximate bid function for a known number of bidders becomes

$$\begin{aligned}
b(x) &\approx \frac{\int_0^\infty v \cdot v^{(N-2)\hat{\xi}_\tau + 2\tau + \xi - 1} \cdot e^{-\left(\frac{1}{x}(\hat{\psi}_\tau(N-2)+2\tau)+\psi\right)\cdot v}\, dv}{\int_0^\infty v^{(N-2)\hat{\xi}_\tau + 2\tau + \xi - 1} \cdot e^{-\left(\frac{1}{x}(\hat{\psi}_\tau(N-2)+2\tau)+\psi\right)\cdot v}\, dv} \\
&= \frac{\int_0^\infty Gamma(v|\xi'_\tau + 1, \psi'_\tau)\, dv}{\int_0^\infty Gamma(v|\xi'_\tau, \psi'_\tau)\, dv} = \frac{\Gamma(\xi'_\tau + 1)}{\psi_\tau'^{(\xi'_\tau+1)}} \cdot \frac{\psi_\tau'^{\xi'_\tau}}{\Gamma(\xi'_\tau)} \\
&= \frac{\xi'_\tau}{\psi'_\tau} = \frac{\left[\xi + 2\tau + (N-2)\hat{\xi}_\tau\right] \cdot x}{\psi x + 2\tau + (N-2)\hat{\psi}_\tau},
\end{aligned}$$

where $\xi'_\tau = (N-2)\hat{\xi}_\tau + 2\tau + \xi$, $\psi'_\tau = \frac{1}{x}(\hat{\psi}_\tau(N-2) + 2\tau) + \psi$ and $Gamma(v|\xi'_\tau, \psi'_\tau)$ denotes the *pdf* of the Gamma distributed variable v with parameters ξ'_τ, and ψ'_τ.

Using the same approach as for the Gaussian model in WV, Appendix A, this result can be generalized to the case with a stochastic number of bidders by substituting N for λ, giving the approximate bid function for the Gamma model in Equation (2.13).

References

[1] Bajari, P., and A. Hortacsu (2003), The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *The Rand Journal of Economics*, **3**(**2**), 329-355.

[2] —— (2005), Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data, *Journal of Political Economy*, **113**(**4**), 703-741.

[3] Elyakime, B., Laffont J.J., Loisel, O., and Vuong, Q. (1997) Auctioning and Bargaining: An Econometric Study of Timber Auctions With Secret Reservation Prices, *Journal of Business & Economic Statistics*, **15**(**2**), 209-20.

[4] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis, 2nd edition*, Chapman & Hall, New York.

[5] Gordy, B. M. (1998), Computationally Convenient Distributional Assumptions for Common-Value Auctions, *Computational Economics*, **12**, 61-78.

[6] Laffont, J.J., and Vuong, Q. (1996), Structural Analysis of Auction Data, *American Economic Review*, **86**(2), 414-20.

[7] Milgrom, P., and Weber, R. (1982), A Theory of Auctions and Competitive Bidding, *Econometrica*, **50**, 1089-1022.

[8] Paarsch, J. H. (1992), Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions, *Journal of Econometrics*, **51**, 191-215.

[9] Paarsch, J. H., and Hong, H. (2006). An Introduction to the Structural Econometrics of Auction Data, Cambridge, MA: MIT Press.

[10] Smith, M., and Kohn, R. (1996), Nonparametric regression using Bayesian variable selection, *Journal of Econometrics*, **75**, 317-344.

[11] Villani, M., Kohn, R., and Giordani, P. (2009), Regression Density Estimation using Smooth Adaptive Gaussian Mixtures, *Journal of Econometrics*, **153**, 155-173.

[12] Wegmann, B., and Villani, M. (2011), Bayesian Inference in Structural Second-Price Common Value Auctions, *forthcoming in Journal of Business and Economic Statistics.*

[13] Zellner, A. (1986), On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, 233–243. North-Holland/Elsevier.

TABLE 1. Comparing the posterior inference from the Gaussian and Gamma models on the eBay data.

Parameter	Covariate	Mean		St Dev		Incl Prob	
		Gaussian	Gamma	Gaussian	Gamma	Gaussian	Gamma
κ/τ	-	5.499	2.997	0.772	0.111	1.000	1.000
μ	Const	28.273	28.307	0.245	0.304	1.000	1.000
	Book_d	0.740	0.747	0.010	0.012	1.000	1.000
	Book·Power	0.033	0.046	0.015	0.018	0.064	0.107
	Book·ID	0.128	0.052	0.036	0.039	0.900	0.017
	Book·Sealed	0.372	0.488	0.029	0.051	1.000	1.000
	Book·MinBlem	−0.022	0.002	0.021	0.028	0.010	0.008
	Book·MajBlem	−0.252	−0.269	0.030	0.040	1.000	1.000
	Book·NegScore	−0.003	−0.020	0.018	0.025	0.004	0.009
$\log(\sigma^2)$	Const	3.997	4.314	0.071	0.038	1.000	1.000
	LBook_d	1.262	1.276	0.038	0.026	1.000	1.000
	LBook·Power	0.043	0.069	0.018	0.020	0.220	1.000
	LBook·ID	0.042	0.032	0.040	0.067	0.481	0.011
	LBook·Sealed	0.211	0.362	0.027	0.019	1.000	1.000
	LBook·MinBlem	−0.028	−0.057	0.027	0.026	0.012	0.039
	LBook·MajBlem	0.036	0.063	0.040	0.049	0.007	0.017
	LBook·NegScore	0.035	0.042	0.021	0.027	0.017	0.050
$\log(\lambda)$	Const	1.193	1.234	0.021	0.022	1.000	1.000
	Power	0.009	−0.028	0.035	0.029	0.005	0.012
	ID	−0.177	−0.197	0.110	0.078	0.030	0.048
	Sealed	0.323	0.331	0.048	0.048	1.000	1.000
	MinBlem	−0.049	−0.042	0.048	0.048	0.008	0.009
	MajBlem	−0.151	−0.115	0.085	0.097	0.019	0.015
	NegScore	0.055	0.086	0.049	0.047	0.012	0.022
	LBook_d	−0.038	−0.036	0.027	0.021	0.018	0.031
	MinBidShare_d	−1.433	−1.380	0.056	0.059	1.000	1.000

NOTE: The prior hyperparameters are set to $c = n$, $\bar{\kappa} = \bar{\tau} = 0.25$, $g = h = 4$, and $\pi = 0.2$, as for the benchmark model in WV. $x_1 \cdot x_2$ denotes the interaction of x_1 and x_2.

TABLE 2. Log predictive scores (LPS) for the Gaussian and Gamma models on the eBay data.

Test auction	Measure	Gaussian	Gamma	% Gaussian wins
$1-37$	LPS$_c$	-3.547	-3.684	27/37
$1-48$	LPS$_d$	-0.526	-0.515	25.5/47
$1-37$	LPS$_d$	-0.212	-0.238	20.5/37
$38-48$	LPS$_d$	-1.687	-1.539	5/11

NOTE: Auctions 1-37 contain at least two bids, and auctions 38-48 a maximum of one bid. LPS$_d$ for auction 43 is undefined ($-\infty$) for the Gamma model, and is therefore excluded from the calculations of LPS$_d$. The Gaussian and Gamma models are equally good in LPS$_d$ for auction 21.

TABLE 3. Log predictive scores (LPS) for the Gaussian and Gamma models on simulated data with different degrees of skewness, S_k.

Score	Statistics	$S_k = 2$		$S_k = 1.5$		$S_k = 0.5$	
		Gaussian	Gamma	Gaussian	Gamma	Gaussian	Gamma
LPS$_c$	Mean	-2.430	-2.138	-2.814	-2.361	-3.173	-3.157
	St Dev	0.254	0.186	0.833	0.166	0.109	0.095
LPS$_d$	Mean	-0.886	-0.877	-0.764	-0.755	-0.363	-0.364
	St Dev	0.069	0.081	0.070	0.084	0.092	0.093
LPS$_d$	Mean	-0.584	-0.449	-0.383	-0.301	-0.111	-0.111
$Bids \geq 2$	St Dev	0.057	0.042	0.035	0.031	0.011	0.011
LPS$_d$	Mean	-1.372	-1.567	-1.742	-1.924	-2.612	-2.630
$Bids \leq 1$	St Dev	0.120	0.126	0.103	0.127	0.244	0.264

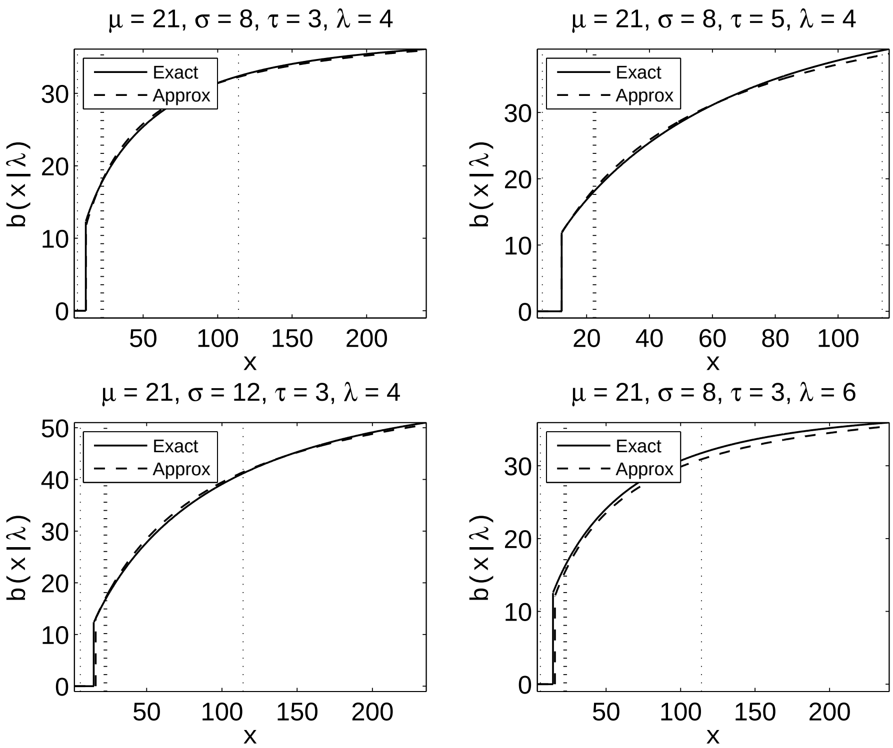

FIGURE 1. Examining the accuracy of the approximate bid function in the Gamma model for different configurations of parameter values. The vertical lines in the figures are the mean (thick dotted), and the 5% and 95% percentiles (thin dotted) in the unconditional distribution of the signals, x.

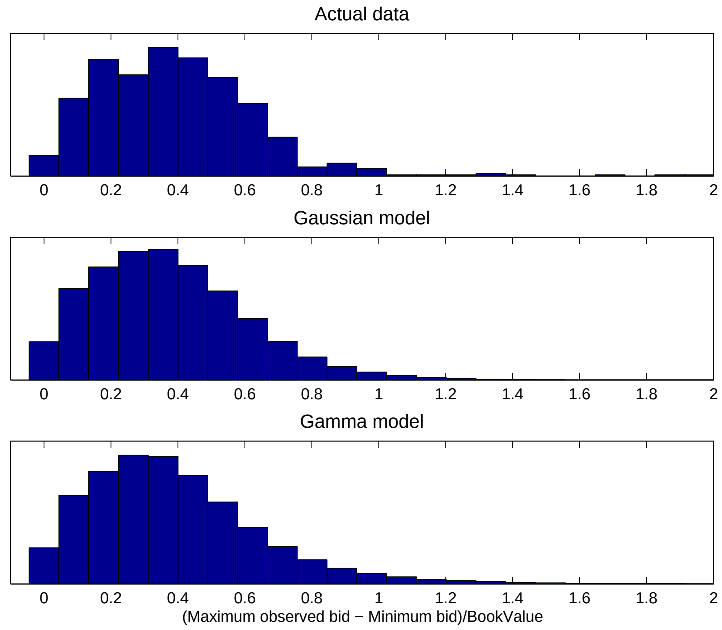

FIGURE 2. Posterior predictive analysis of the models' fit of the within-auction variation.

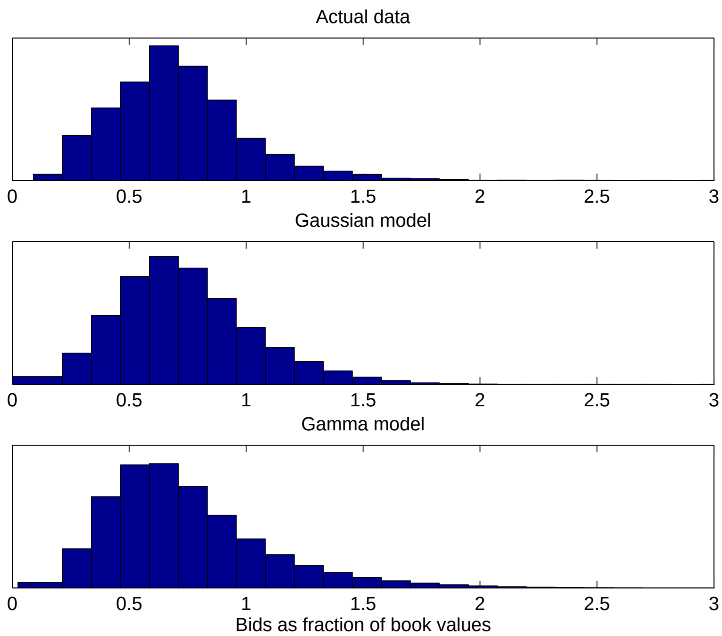

FIGURE 3. Posterior predictive analysis of the models' fit of the cross-auction heterogeneity.

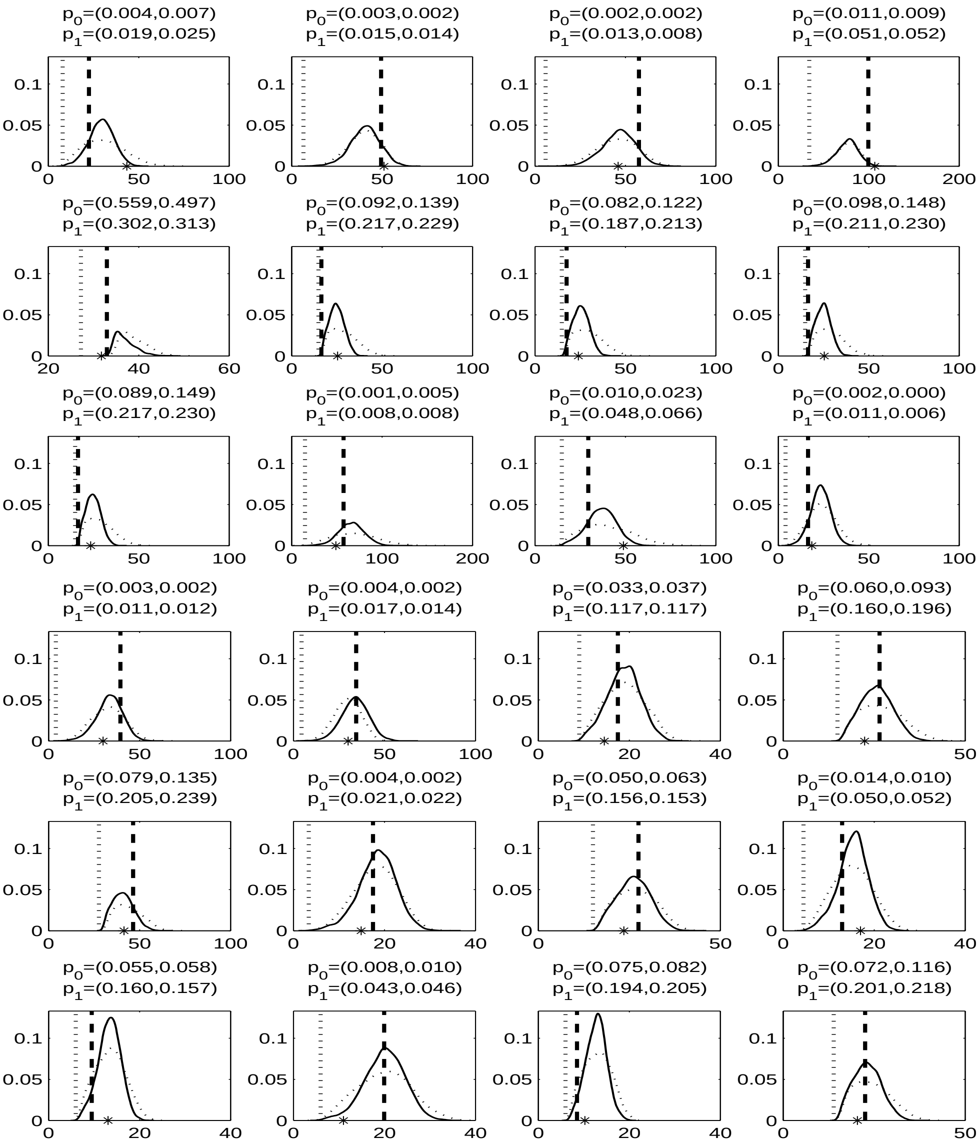

FIGURE 4. Predictive distributions for the auction price in auctions $1-24$ in the evaluation sample. The values of p_0 and p_1 in the titles are the predictive probabilities of zero and one bid (where the price is not observed) for the Gaussian (first number) and Gamma models (second number). The densities are the predictive densities of the price when the auction has at least two bidders (the integral of the density is the probability of at least two bids) for the Gaussian (solid line) and Gamma (dotted line) models. The vertical lines indicate the minimum bid (dotted) and the book value (dashed). The realized price (if observed) is displayed by the star symbol.

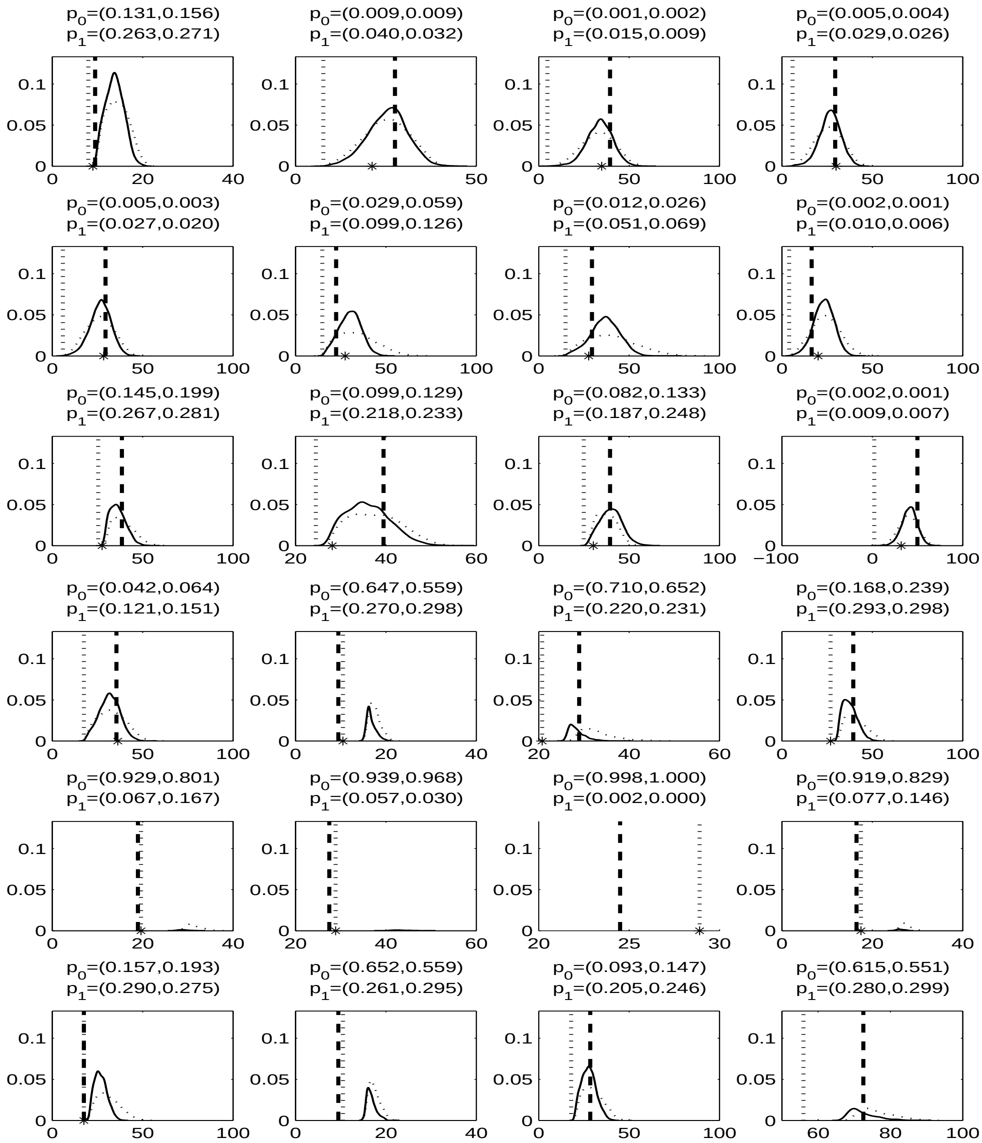

FIGURE 5. Predictive distributions for the auction price in auctions $25-48$ in the evaluation sample. The values of p_0 and p_1 in the titles are the predictive probabilities of zero and one bid (where the price is not observed) for the Gaussian (first number) and Gamma models (second number). The densities are the predictive densities of the price when the auction has at least two bidders (the integral of the density is the probability of at least two bids) for the Gaussian (solid line) and Gamma (dotted line) models. The vertical lines indicate the minimum bid (dotted) and the book value (dashed). The realized price (if observed) is displayed by the star symbol.

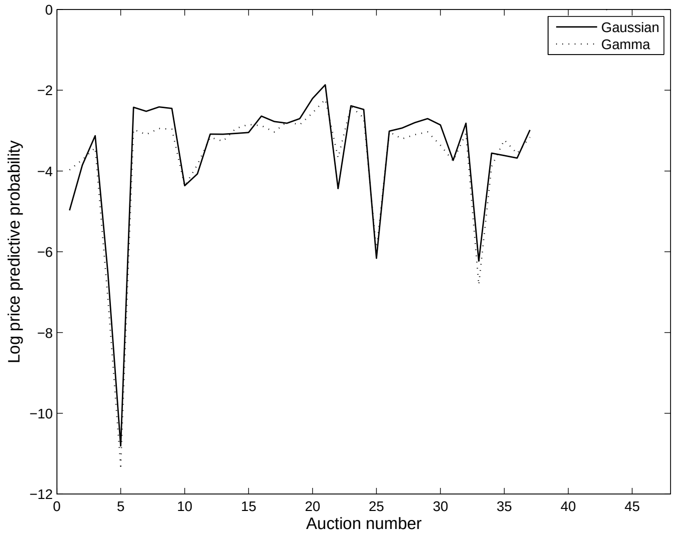

FIGURE 6. Log price predictive probabilities evaluated at the actual price in each auction, conditional on at least two bids for the Gaussian and Gamma models. Auctions 38-45 contain a single bid, and auctions 46-48 no bids.

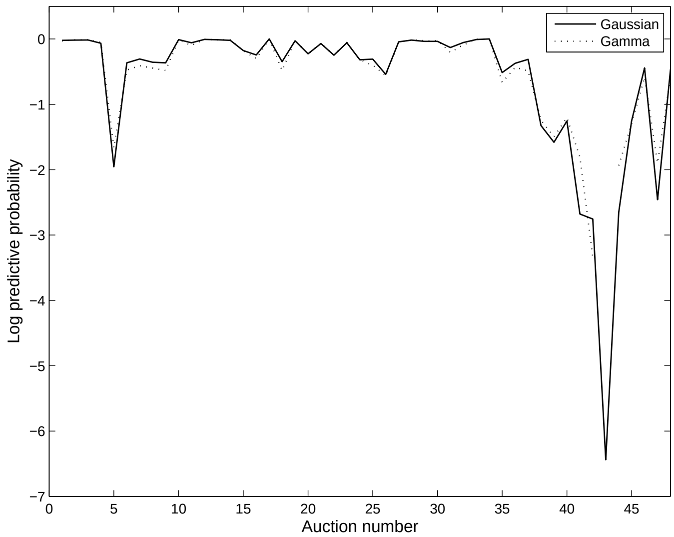

FIGURE 7. Log predictive probabilities for the actual outcome of either no bids, one bid, or at least two bids in each auction for the Gaussian and Gamma models. Auctions 1-37 contain at least two bids, auctions 38-45 a single bid, and auctions 46-48 no bids. The log predictive probability in auction 43 for the Gamma model is not displayed since the probability equals zero.

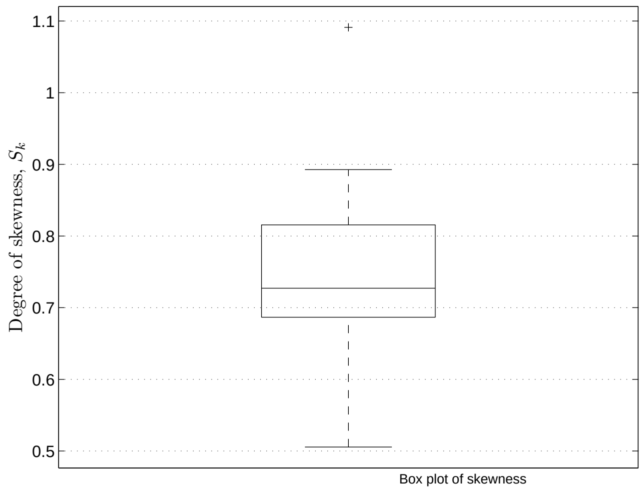

FIGURE 8. Box plot for the estimated skewness in the 48 test auctions.

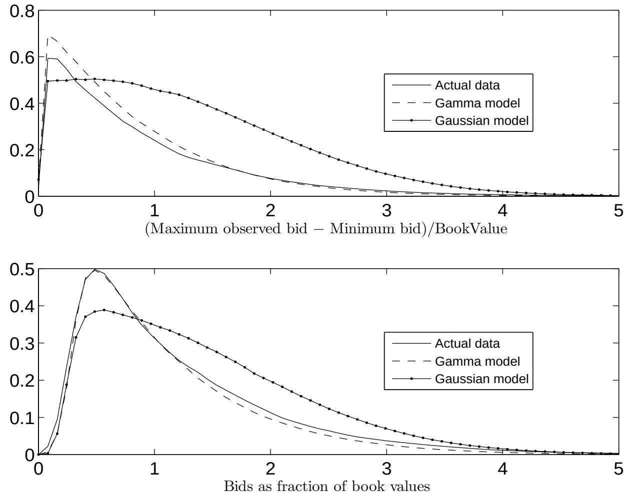

FIGURE 9. Posterior predictive analysis of the models' fit of the within-auction variation (upper) and cross-auction heterogeneity (lower). The degree of skewness for the simulated data, S_k, is equal to 2.

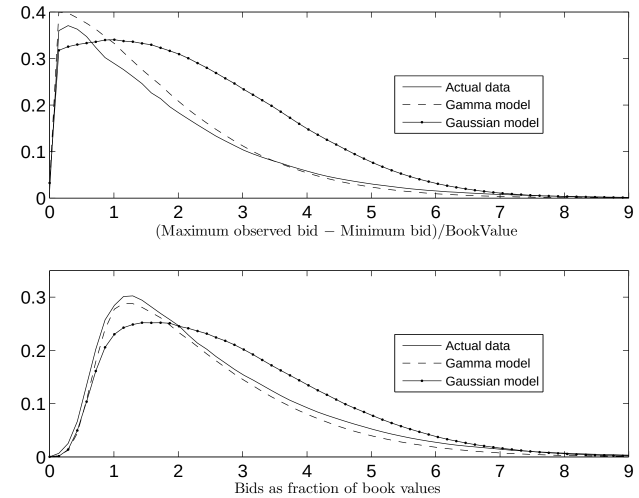

FIGURE 10. Posterior predictive analysis of the models' fit of the within-auction variation (upper) and cross-auction heterogeneity (lower). The degree of skewness for the simulated data, S_k, is equal to 1.5.

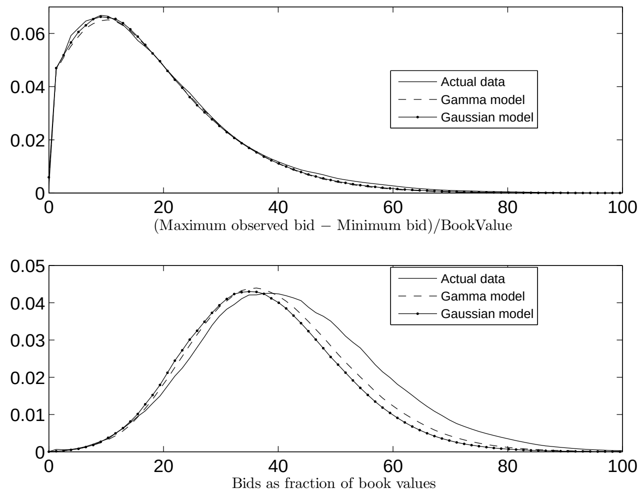

FIGURE 11. Posterior predictive analysis of the models' fit of the within-auction variation (upper) and cross-auction heterogeneity (lower). The degree of skewness for the simulated data, S_k, is equal to 0.5.

BAYESIAN COMPARISON OF PRIVATE AND COMMON VALUES IN STRUCTURAL SECOND-PRICE AUCTIONS

BERTIL WEGMANN

Abstract. We compare the performance of the Gaussian second-price common value (CV) model in Wegmann and Villani (2011) to a comparable independent private value (IPV) version of that model. The two models are contrasted on a dataset from 1050 Internet coin auctions at eBay. The models are evaluated along several dimensions, such as parameter inference, in-sample fit, and accuracy of out-of-sample predictive density forecasts. Both models fit the eBay data well with a slight edge for the more robust CV model. We do not find any evidence of a winner's curse effect in the eBay data, which speaks in favor of the IPV model. However, the optimal minimum bids in the CV model are clearly closer to the actual minimum bids in the eBay data than the optimal choice of no minimum bid in the IPV model. The IPV model predicts auction prices slightly better in most auctions, while the CV model is much better at predicting auction prices in more unusual auctions. The robustness of the CV model is also supported by a small simulation study, where the CV model performs relatively better on simulated data from the IPV model than the IPV model fitted to CV data.

Keywords: Bayesian variable selection, Common values, eBay, Gaussian model, Markov Chain Monte Carlo, Private values, The winner's curse.

1. Introduction

The behavior of bidders in an auction hinges critically upon the theoretical settings of values. Traditionally, researchers in auction theory have been working within either the independent private value (henceforth IPV) or the pure common value (henceforth CV) model, and these two models are by far the most commonly used models in empirical studies. In the IPV setting, each bidder knows his value of the object and is not influenced by the values of the other bidders. In the CV model, the value is unknown but the same for all bidders, and each bidder uses his own private information (the signal) to estimate the unknown value.

Bajari and Hortacsu (2003, henceforth BH) use some empirical regularities to argue that coin auctions at eBay possess a common value component. Armantier (2002) points out that both paradigms are plausible in most cases. Attempts to distinguish between the paradigms have been under considerable attention, especially within first-price auctions. In a seminal piece of work, Paarsch (1992) uses structural econometric modeling to discriminate between IPV and CV models in first-price auctions. As a test for common values, Paarsch (1992) examines whether winning bids decrease with the number of bidders. However, Pinkse and Tan (2005) show that bids can increase or decrease in both common and private value models within first-price auctions. In second-price auctions, this obstacle does not apply (Athey and Haile (2002)). We compare a second-price CV model to a directly comparable IPV model. Comparisons between models in second-price auctions are rare in the literature and our use of comparable CV and IPV models is novel.

Sareen (1999) uses the posterior odds ratio to decide between the CV and IPV paradigms in reversed first-price auctions where the bidder with the lowest bid wins the auction. While the

Department of Statistics, Stockholm University, SE-106 91 Stockholm. E-mail: bertil.wegmann@stat.su.se.

posterior odds ratio is commonly used in Bayesian model comparisons, it critically depends on the choice of prior, especially when the prior is less informative (Kass (1993)). To overcome this difficulty, Geweke and Keane (2007) and Villani et al. (2009) use cross-validation of the log predictive density score (LPDS) to choose among model specifications.

We evaluate the IPV and CV models along several dimensions in an attempt to empirically decide between the models in second-price auctions. In particular, we compare the in-sample fit and out-of-sample log predictive score (LPS) between the models and examine whether the winner's curse exists in the data. Following BH, we model eBay auctions as independent second-price auctions with stochastic entry. The data is taken from Wegmann and Villani (2011, henceforth WV) and consists of bids and auction specific covariates from 1050 eBay coin auctions. The mean and variance of the value distribution and the expected number of bidders are modeled as functions of covariates. To quantify the importance of the covariates, we use the generalized Metropolis-Hastings algorithm with variable selection in WV.

The CV model in BH and WV is contrasted to a comparable IPV model with a similar structure. The equilibrium bid function for common values is complicated as compared to the bidding strategy in second-price IPV auctions where the dominant strategy is to bid one's value (Vickrey, 1961). However, we assume the approximate bid function in WV which is linear in the signal. The evaluation of the likelihood function for observed bids is then fast and numerically stable for both the IPV and CV models, and can therefore be routinely used for analyzing auction data. WV show that the approximate bid function for the CV model is very accurate.

The parameter estimates are especially different for some coefficients of the covariates in the regression models for the mean and variance of the value. The deviations in the mean are partly attributed to the winner's curse effect, while the differences in the variance are mainly explained by intrinsically different valuation structures between the models.

Both models fit the eBay data very well, but the CV model performs slightly better in fitting the within-auction bid dispersion and cross-auction heterogeneity in bids. The winner's curse effect does not seem to be present in the eBay auctions, which speaks in favor of the IPV model. However, the actual minimum bids set by the seller in the eBay data are much closer to the optimal minimum bids in the CV model, compared to no minimum bid as the optimal choice in the IPV model. We propose a model and estimate the correlation between observed bids in the eBay data to be 0.5. Simulated data from the estimated IPV and CV models capture this correlation fairly well. When it comes to predictive ability, the IPV model is slightly better in most auctions. However, the CV model is much better in predicting more unusual auctions. Overall, there is a slight edge for common values in the eBay data.

We also conducted a small simulation study to examine how well the models fit simulated IPV and CV data. In general, it seems that the CV model is more robust than the IPV model, in the sense that the CV model fits IPV data much better than the IPV model fits CV data.

2. Private and Common Value Models in Second-Price Auctions

Motivated by the theoretical arguments in BH, we model eBay auctions as independent second-price auctions with stochastic entry. In this model, risk-neutral bidders compete for a single object using the same bidding strategy, the seller sets a publicly announced *minimum bid* (*reservation price*), $r \geq 0$, and the winner of the auction pays the second highest bid. We evaluate the performances between the private and common value paradigms by assuming two models for values. The hierarchical Gaussian CV model in WV is used for common values, and by straightforward modifications of this model we define a comparable Gaussian IPV model. Any other distribution can be used to model the values. However, Wegmann (2011) documents that the Gaussian CV model fits the eBay data well, and does better out-of-sample predictions than a comparable Gamma model for common values.

2.1. **The IPV model.** Within the IPV paradigm, each bidder knows the value, x, to himself and is not affected by the values of the other bidders. It is a weakly dominant strategy for a bidder to bid his value (Vickrey, 1961), which gives the bid function

$$b(x) = \begin{cases} x, & \text{if } x \geq r \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

Let x_{ij} denote the value of the ith bidder in auction j, and let λ_j be the expected number of potential bidders in auction j. The Gaussian IPV model is given by

$$\begin{aligned} x_{ij} &\overset{iid}{\sim} N(\mu_j, \sigma_j^2), \;\; j = 1, \ldots, m, , \;\; i = 1, ..., N_j \\ \mu_j &= z'_{\mu j}\beta_\mu \\ \sigma_j^2 &= \exp\left(z'_{\sigma j}\beta_\sigma\right) \\ \lambda_j &= \exp\left(z'_{\lambda j}\beta_\lambda\right), \end{aligned} \tag{2.2}$$

where m is the total number of auctions, N_j is the number of potential bidders in the jth auction, and $z_j = (z'_{\mu j}, z'_{\sigma j}, z'_{\lambda j})'$ are auction-specific covariates in the regression models for $(\mu_j, \sigma_j^2, \lambda_j)$ in auction j.

Some bids in the eBay auctions are unobserved, which complicates the likelihood function. Potential bidders with values below the minimum bid r will not place a bid, and the highest bid is typically not observed because of eBay's proxy bidding system (see BH for a detailed description). Since bidders bid their value, the distribution of potential bids, $f_b(b|\mu, \sigma^2)$, equals the distribution of values. Let $f_x(\cdot|\mu, \sigma^2)$ and $F_x(\cdot|\mu, \sigma^2)$ be the probability density function and the cumulative distribution function of x, respectively, let n be the number of participating bidders who submit a positive bid above r in an arbitrary auction, and let $\mathbf{b} = (b_2, b_3, \ldots, b_n)$ be the vector of observed bids where $b_2 > b_3 > \cdots > b_n$. Then, the likelihood function for an arbitrary auction of observed bids becomes

$$\begin{aligned} &f_{\mathbf{b}}\left(b_2, b_3, \ldots, b_n|\mu, \sigma^2, \lambda, r\right) \\ &= \sum_{i=n}^{\bar{N}} p_i(\lambda) \cdot F_x\left(r|\mu, \sigma^2\right)^{i-n} \cdot \left\{1 - F_x\left(b_2|\mu, \sigma^2\right)\right\}^{I(n\geq 1)} \cdot \prod_{j=2}^{n} f_x\left(b_j|\mu, \sigma^2\right), \end{aligned} \tag{2.3}$$

where $p_i(\lambda)$ is the Poisson probability of i potential bidders in the auction with λ as the expected value, $I(n \geq 1)$ is an indicator variable for at least one participating bidder in the auction, and $\bar{N}$ is an upper bound for the total number of potential bidders. As in BH, let $\bar{N} = 30$ for tractability. If $n = 1$, b_2 is replaced by r.

2.2. **The CV model.** Within the CV paradigm the value of the object, v, is unknown and the same for each bidder, but a prior distribution for v is shared by the bidders. To estimate v, each bidder uses his own private information of the object to receive a private signal x from the same distribution, $x|v$. Let $f_v(v)$ denote the probability density function of v, $f_{x|v}(x|v)$ the conditional probability density function of $x|v$, and $F_{x|v}(x|v)$ the conditional cumulative distribution function of $x|v$. The bid function can then be written as (see BH for an implicit solution)

$$b(x,\lambda)=\begin{cases}\frac{\sum_{n=2}^{\infty}(n-1)\cdot p_{n-1}(\lambda)\cdot\int_v v\cdot F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv}{\sum_{n=2}^{\infty}(n-1)\cdot p_{n-1}(\lambda)\cdot\int_v F_{x|v}^{n-2}(x|v)\cdot f_{x|v}^2(x|v)\cdot f_v(v)\ dv}, & \text{if } x\geq x^{\star}\\ 0, \text{otherwise.}\end{cases}\tag{2.4}$$

Bidders participate with a positive bid if their signal, x, is above the cut-off signal level, $x^{\star}$. Given an arbitrary bidder with signal x, let y be the maximum signal of the other $(n-1)$ participating bidders. The cut-off signal level is then given in implicit form as (Milgrom and Weber, 1982)

$$x^{\star}(r,\lambda)=\inf_x\left(E_nE[v|X=x,Y<x,n]\geq r\right),$$

which gives

$$r(x^{\star}|\lambda)=\sum_{n=1}^{\infty}p_n(\lambda)\cdot\frac{\int_v v\cdot F_{x|v}^{n-1}(x^{\star}|v)\cdot f_{x|v}(x^{\star}|v)\cdot f_v(v)\ dv}{\int_v F_{x|v}^{n-1}(x^{\star}|v)\cdot f_{x|v}(x^{\star}|v)\cdot f_v(v)\ dv}.\tag{2.5}$$

Note that the seller's exogenously given minimum bid, r, is only written as a function of the cut-off signal level for tractability.

Let v_j denote the common value in auction j, and let x_{ij} denote the signal of the ith bidder in auction j. Our CV model is then of the form:

$$\begin{aligned}v_j &\sim N(\mu_j,\sigma_j^2),\ \ j=1,...,m,\\ x_{ij}\,|v_j &\overset{iid}{\sim} N(v_j,\kappa\sigma_j^2),\ \ i=1,...,N_j,\\ \mu_j &= z_{\mu j}'\beta_\mu\\ \sigma_j^2 &= \exp\left(z_{\sigma j}'\beta_\sigma\right)\\ \lambda_j &= \exp\left(z_{\lambda j}'\beta_\lambda\right),\end{aligned}\tag{2.6}$$

using the same notation as for the IPV model. The censoring of missing bids also applies in the likelihood function for the common values, i.e. the highest bid is typically not observed and bidders with signals x below $x^{\star}$ do not place a bid. The bid distribution for an arbitrary auction is of the form:

$$f_b\left(b|\beta,z,v\right)=f_{x|v}\left[\phi\left(b|\beta,z\right)|v,\kappa,\sigma\right]\phi^{'}\left(b|\beta,z\right),\tag{2.7}$$

where $f_b\left(b|\beta,z,v\right)$ is the probability density function of the bids conditional on (β,z,v), and $\phi\left(b|\beta,z\right)$ is the inverse bid function given (β,z). The likelihood function for an arbitrary auction is given by

$$\begin{aligned}&f_{\mathbf{b}}\left(b_2,b_3,\ldots,b_n|\mu,\sigma^2,\lambda,\kappa,r\right)\\ &=\sum_{i=n}^{\bar{N}}p_i(\lambda)\cdot\int_{-\infty}^{\infty}F_{x|v}\left(x^{\star}|v,\kappa,\sigma\right)^{i-n}\cdot\left\{1-F_{x|v}\left[\phi\left(b_2|\beta,r,z\right)|v,\kappa,\sigma\right]\right\}^{I(n\geq1)}\\ &\quad\times\prod_{i=2}^{n}f_b\left(b_i|\beta,r,z,v\right)\cdot f_v(v|\mu,\sigma)dv.\end{aligned}\tag{2.8}$$

A single evaluation of the likelihood function in (2.8) requires numerical integration to compute $b\left(x,\lambda|\beta,z\right)$ in (2.4) and $r\left(x^{\star},\lambda|\beta,z\right)$ in (2.5), followed by additional numerical work to invert and differentiate $b\left(x,\lambda|\beta,z\right)$ and $r\left(x^{\star},\lambda|\beta,z\right)$. This is very time-consuming and needs to be repeated for each of the auctions in the eBay dataset. Instead, to get much faster and numerically stable likelihood evaluations, we use the accurate approximation in WV, given by

$$(2.9) \qquad b(x,\lambda) \approx \begin{cases} c + \omega\mu + (1-\omega)x, & \text{if } x \geq x^\star \\ \\ 0, \text{otherwise}, \end{cases}$$

where $c = -\frac{\sqrt{\kappa}\sigma\gamma\theta(\lambda-2)}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\omega = \frac{\frac{\kappa}{2}}{\gamma(\lambda-2)+1+\frac{\kappa}{2}}$, $\theta = 1.96$ and $\gamma = 0.1938$, and

$$x^\star(r,\lambda) \approx \frac{r - \sum_{n=2}^{\infty} p_n(\lambda)(\tilde{c} + \tilde{\omega}\mu)}{\sum_{n=2}^{\infty} p_n(\lambda)(1-\tilde{\omega})},$$

where $\tilde{c} = -\frac{\sqrt{\kappa}\sigma\gamma\theta(n-1)}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$, and $\tilde{\omega} = \frac{\frac{\kappa}{2}}{\gamma(n-1)+\frac{1}{2}+\frac{\kappa}{2}}$. Using this approximation of the bid function, the distribution of the bids conditional on v in (2.7) simplifies to

$$(2.10) \qquad b(x,\lambda|v) \overset{iid}{\sim} N[c+\mu, (1-\omega)^2\kappa\sigma^2],$$

which speeds up the likelihood evaluation further.

3. Comparisons between the competing paradigms on eBay coin auction data

We use the dataset in WV with 1050 eBay auctions of U.S. proof coin sets to compare the performance of the IPV and CV models. A detailed description of the data is given in WV. The unknown model parameters in each model are estimated with the same Bayesian methods and priors as the benchmark model in WV, which are briefly summarized below.

3.1. Prior distribution and a Metropolis-Hastings algorithm for variable selection.

The likelihood function in either (2.3) or (2.8) is combined with a prior distribution on the unknown model parameters to form the posterior distribution of each model. Then, a generalization of the Metropolis-Hastings algorithm (a Markov Chain Monte Carlo (MCMC) algorithm) is used to simultaneously do Bayesian variable selection among the auction covariates and sample the posterior of the model parameters (WV, Section 4.2 and Appendix B). Variable selection involves adding point masses at zero in the prior distributions; see Smith and Kohn (1996).

The prior distribution for β is given by a g-prior (Zellner, 1986) for β_μ, conditional on β_σ, as

$$\beta_\mu|\beta_\sigma \sim N[0, c(z'_\mu D z_\mu)^{-1}],$$

where $D^{1/2} = \text{Diag}[\exp(-z'_{\sigma_1}\beta_\sigma/2), ..., \exp(-z'_{\sigma_n}\beta_\sigma/2)]$, and $c > 0$ is a scaling factor equal to the number of auctions in the data for the benchmark model. Marginal g-priors for β_σ and β_λ are given by

$$\beta_\sigma \sim N[0, c(z'_\sigma z_\sigma)^{-1}],$$

and

$$\beta_\lambda \sim N[0, c(z'_\lambda z_\lambda)^{-1}].$$

We use an inverse Gamma prior for κ, $\kappa \sim IG(\bar{\kappa}, g)$in the CV model where $\bar{\kappa} = 0.25$ and $g = 4$. Finally, the prior inclusion probability of a given covariate is set to 0.2.

3.2. **Estimation results.** The models are estimated using the whole dataset of 1050 eBay coin auctions in WV. The dataset was carefully collected by human inspections of auction-specific covariates and bid sequences from auctions that ended between November 7 to December 19, 2007 and December 27, 2007 to January 29, 2008. The book value of the object and the minimum bid divided by the book value in each auction are defined by Book and $\mathsf{MinBidShare}$. The remaining covariates are dummy variables that are coded to be 1 if a certain characteristic is present in a given auction. The largest sellers are described by Pow (PowerSellers) and ID describes a seller whose identity is verified on eBay. $\mathsf{MinBlem}$ ($\mathsf{MajBlem}$) describes if the object has a minor (major) damage, and UnOpen pertains to an object which is sold in its original sealed envelope. A large negative feedback score for a seller is described by $\mathsf{LargeNeg}$.

Table 1 reports the posterior results for both the IPV and CV models. The posterior inclusion probability is the posterior probability of including a given covariate in the model. The inefficiency factor (IF) is a measure of the numerical efficiency of the MCMC algorithm, defined as the number of posterior draws needed to obtain the equivalent of a single independent draw. As can be discerned from the posterior inclusion probabilities, μ is mainly determined by the constant, the book value of the object and the covariates $\mathsf{Book}{\cdot}\mathsf{UnOpen}$, and $\mathsf{Book}{\cdot}\mathsf{MajBlem}$ in both models. The major differences in μ between the models are in the posterior mean estimates of the constant and the coefficient of Book_d. Using the posterior mean estimates, it can be verified that μ is higher in all auctions in the CV model. Moreover, when Book_d becomes larger μ increases more rapidly in the CV than in the IPV model.

This is in contrast to the differences between the models in posterior mean estimates of the regression coefficients in σ^2, where the main drivers are $\log(\mathsf{Book})_d$ and $\log(\mathsf{Book}){\cdot}\mathsf{UnOpen}$. Here, σ^2 increases more rapidly with the book value in the IPV model. Hence, it seems that the mean-variance ratio, $\frac{\mu}{\sigma^2}$, in the IPV model is more sensitive to different book values than the common value counterpart. In Figure 1 we can see that in general this is true. The mean-variance ratio decreases more rapidly in the IPV model as the book value of the object increases. The differences in mean-variance ratios can partly be explained by the different bidding strategies in the two models.

First, the larger estimate in the CV model for the coefficient of Book_d in μ is partly attributed to the presence of the winner's curse phenomenon in the CV model. When the book value of the object increases, the expected value, μ, and the variance of the values, σ^2, increase in both models. Within the CV model a higher variance brings more uncertainty and thereby a higher risk of overestimating the value of the object upon winning. To avoid this (the winner's curse phenomenon), a bidder in a CV auction needs to lower his bid when the variance increases. This is in contrast to the IPV model where the bid is equal to the value regardless of an increment in σ^2 and μ. Hence, μ needs to change more in the CV model when the book value of the object changes to accomodate the bid shading effect in the CV model.

Second, the variance of bids, $var(b)$, also depends on κ and λ in the CV model. This suggests that σ in the CV model does not need to be that sensitive to changes in the main driver $\log(\mathsf{Book})_d$ as compared to σ in the IPV model, where the variance of bids is solely determined by σ. The estimated coefficient of $\log(\mathsf{Book})_d$ is larger and thereby more sensitive in the IPV model, which is consistent with the suggestion.

The expected number of bidders, λ, is mainly determined by the covariates UnOpen and $\mathsf{MinBidShare}$. The major differences in posterior mean estimates between the models are in the constant and the coefficient of $\mathsf{MinBidShare}$, where a higher minimum bid more strongly discourages entry in the CV model.

3.3. **In-sample fit.** We evaluate the in-sample fit of the models using both graphical and analytical methods. The observed dataset is compared to simulated data from each model.

We simulated $10,000$ full datasets with bids in 1050 auctions, using the observed covariates and the parameters systematically sampled from the posterior sample in Section 3.2. Figure 2 presents the within-auction bid dispersion as the difference between the highest observed bid and the lowest bid divided by the book value of the object in each auction. In Figure 3, cross-auction heterogeneity is presented as histograms of the bids divided by the corresponding book value in each auction.

As we can see in both figures, the observed within-auction bid dispersion and cross-auction heterogeneity are very well captured by both models. However, it is not easy to judge which of the models that fits the data best. One way of accomplishing such a comparison is to compute the Hellinger distance between the actual and simulated distribution in the figures. The Hellinger distance for distributions f_1 and f_2 can be written as (Nikulin (2001))

$$H(f_1, f_2) = \left(\frac{1}{2}\int_{-\infty}^{\infty}\left(\sqrt{f_1(x)} - \sqrt{f_2(x)}\right)^2 dx\right)^{1/2},$$

where $0 \leq H(f_1, f_2) \leq 1$.

The probability density functions f_1 and f_2 are estimated by a normal kernel density estimation with automatic bandwidth selection, see Silverman (1986). Table 2 presents the Hellinger distance between the simulated distribution and the actual data distribution for each model. The distances are very small for both models, but the CV model is slightly better at fitting the within-auction bid dispersion and cross-auction heterogeneity.

3.4. **Out-of-sample predictions.** The predictive ability of the models is compared by a Bayesian version of cross-validation. We partition the eBay dataset into five equally large subsets. The first 4/5 of the data are used to estimate the model (the training set) and the observations in the excluded fifth partition (the testing set) are subsequently predicted. This is repeated five times with a new fifth as the evaluation data each time. To quantify the predictive ability, we compute the log predictive score (LPS) for each auction in the test sets. As in Wegmann (2011), the predictive distribution is partitioned into discrete and continuous components.

The discrete part of the distribution is computed by the multinomial predictive probability of no bids, one bid, and at least two bids. Following Wegmann (2011), let $p_{i,j}$ be the probability of j bids in auction i, and let

$$I_{i,j} = \begin{cases} 1, & \text{if there are } j \text{ actual bids in auction } i \\ \\ 0, & \text{if there are not } j \text{ actual bids in auction } i, \end{cases}$$

where $i = 1, \ldots, m_t,\ j = 0, 1$, and m_t is the number of test auctions used to evaluate the predictions. Let $\mathbf{I} = (I_{1,0}, I_{1,1}, I_{2,0}, \ldots, I_{m_t,1})$ be the vector of observed indicator variables, and let $\mathbf{p} = (p_{1,0}, p_{1,1}, p_{2,0}, \ldots, p_{m_t,1})$ be the predictive probabilities for each auction. Then, the discrete LPS is defined as

$$LPS_d = \frac{\sum_{i=1}^{m_t}\sum_{j=0}^{2} I_{i,j} \cdot \log p_{i,j}}{m_t}, \tag{3.1}$$

where $I_{i,2} = (1 - I_{i,0} - I_{i,1})$, and $p_{i,2} = (1 - p_{i,0} - p_{i,1})$ are the indicator variable and the predictive probability for at least two bids in auction i, respectively.

The LPS for the continuous part of the predictive price distribution is defined as

$$LPS_c = \frac{\sum_{i=1}^{m_t^\star} \log\left(\frac{\tilde{p}(y_i)}{p_{i,2}}\right)}{m_t^\star}, \tag{3.2}$$

where y_i is the realized price in auction i evaluated in the predictive price distribution $\tilde{p}(\cdot)$, and $m_t^\star$ is the number of test auctions with at least two bids. Note that the predictive price distribution is evaluated conditional on the presence of at least two bidders. Hence, the continuous part of LPS does not depend on the predictive probabilities of the number of bids in the discrete part of LPS.

Table 3 presents the LPS from each model and the proportion of auctions (% wins) with a lower LPS for each model as compared to the other model. In a large majority of the auctions the LPS are higher for the IPV model. However, in auctions where the CV model attains a lower LPS, the differences in LPS are often large. This is especially true for the discrete part of the LPS, which can also be seen from Table 3. The mean value of LPS_d for the CV model is considerably larger, even if the LPS_d are larger in most of the auctions within the IPV model. Therefore, it seems that the simpler IPV model has a slightly better predictive ability in most auctions, but at the cost of being less robust in some auctions as compared to the more complex common value counterpart.

3.5. **Existence of the winner's curse.** Upon winning in a CV auction, the risk of overestimating the value of the object increases with the number of bidders, which induces bidders to lower their bids. At the same time, the bidders need to bid more aggressively and place larger bids due to more competition. In equilibrium, the effect of overestimating the value of the object is larger than the effect of competition, and bidders react to this winner's curse effect by lowering their bids. This is in contrast to the second-price IPV auction where it is a dominant strategy to bid one's valuation, no matter how many bidders that are present. The empirical implication of the winner's curse phenomenon is that the average bid in an N-person auction should be lower than an otherwise equivalent $(N-1)$-person auction.

In the light of these differences between the models, we regressed the observed bids divided by the book value on the number of observed bids and other covariates in a similar way as in BH. Table 4 presents the estimated regressions on the eBay data. The results were obtained using the MH-algorithm for variable selection, described in Section 3.1. The bids do not seem to depend on the number of bidders. The estimates and the posterior inclusion probabilities for both NBids and $\mathsf{NBids{\cdot}MinBidShare}_d$ are very small and close to zero, which speaks in favor of the IPV model. To control for omitted factors in the estimated regression model, we also used other transformations of the number of bidders. However, the estimated results from these models were very similar to the specified model in Table 4.

3.6. **Correlation between observed bids.** WV show that the correlation between bids in the CV model is approximately $\frac{1}{\kappa+1}$ if we ignore missing bids because of eBay's proxy bidding system; see BH for a detailed description, and the truncation of bids that comes from $x^\star$. If all bids had been observed, the correlation between bids would have been a clear discriminant between the two models, since all bids in the IPV model are uncorrelated. However, missing bids introduce correlation between the observed bids in the IPV model. Consequently, if we want to estimate the actual correlation between the observed bids, we need other tools.

Let $\mathbf{b}_j = (b_{j1}, \ldots, b_{jn_j}),\ j = 1, \ldots, m,\ i = 1, \ldots, n_j$ be the vector of n_j observed bids in the jth auction. Assume,

$$\mathbf{b}_j \overset{iid}{\sim} N(\theta_j 1_{n_j}, \Sigma_j)$$

$$\theta_j = z_{\theta j}' \beta_\theta,$$

where $z_{\theta j}$ are auction-specific covariates in the regression model for θ_j, and Σ_j is an equi-variance equi-correlation covariance matrix

$$\Sigma_j = \sigma_j^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \rho \\ \rho & \rho & \cdots & 1 \end{pmatrix}.$$

The distribution of the mean bid is then given by

$$\bar{b}_j \overset{iid}{\sim} N\left(\theta_j, \sigma_j^2 \frac{1+\rho(n_j-1)}{n_j}\right),$$

and it can be shown that

$$\sum_{i=1}^{n_j} \left(b_{ij} - \bar{b}_j\right)^2 \overset{iid}{\sim} \sigma_j^2 (1-\rho)^2 \chi^2(n_j - 1).$$

The MLE of $\sigma_j^2|\rho$ is given by $\hat{\sigma}_j^2|\rho = (1-\rho)^{-2} \frac{\sum_{i=1}^{n_j}\left(b_{ij}-\bar{b}_j\right)^2}{n_j-1}$. Now, by substituting σ_j^2 in the distribution of $\bar{b}_j$ with its estimate, the concentrated log-likelihood function for an arbitrary auction with respect to β_θ and ρ is given by

$$\begin{aligned} p(\mathbf{b}_1, \ldots, \mathbf{b}_{m^\star}|\hat{\sigma}, \rho) &\overset{\beta_\theta,\rho}{\propto} m^\star \log(1-\rho) - \frac{1}{2}\sum_{j=1}^{m^\star} \log(1+\rho(n_j-1)) \\ &- \frac{(1-\rho)^2}{2} \sum_{j=1}^{m^\star} \frac{n_j(n_j-1)\left(\bar{b}_j - \theta_j\right)^2}{(1+\rho(n_j-1))\sum_{i=1}^{n_j}(b_{ij}-\bar{b}_j)^2}, \end{aligned} \tag{3.3}$$

where $m^\star$ is the number of auctions with at least two observed bids.

Using a standard optimization algorithm, we maximize the above log-likelihood with respect to β_θ and ρ, using all bids and auction-specific covariates from all auctions with at least two bidders in the eBay dataset. Since we are only interested in the estimate of ρ, we found that it was sufficient to control for the covariates in μ and $\mathsf{MinBidShare}$ in the regression model for θ. This resulted in $\hat{\rho} = 0.5023$.

We evaluate the posterior predictive distribution (see e.g. Gelman et. al. (2004)) of the correlation between observed bids in the IPV and CV models. Given the observed auction-specific covariates and the posterior modes from the estimation of each model, we simulated 100 full datasets with observed bids from each of the 1050 eBay auctions. Figure 4 presents histograms of the 100 estimated correlations between observed bids that were obtained from the simulated datasets. It seems that the estimated correlations are slightly biased within both models. While the bias is somewhat smaller in the IPV model, the variance of the estimated correlations is slightly smaller in the CV model. Therefore, the models seem to be more or less equally good at capturing the estimated correlation between bids in the eBay data.

3.7. **Optimal minimum bids.** Sellers at eBay frequently set a minimum bid. A low minimum bid attracts more bidders, but also endangers a low sale price. A risk-neutral seller will choose the minimum bid that maximizes the expected revenue,

$$\mathsf{E}[\mathsf{Revenue}] = \mathsf{Pr}(\mathsf{Sale}|r) \cdot \mathsf{E}\left(\mathsf{Price}|r\right) + [1 - \mathsf{Pr}(\mathsf{Sale}|r)] \cdot \mathsf{ResidualValue},$$

where $\mathsf{ResidualValue}$ is the value of the object to the seller if the object is not sold. Using the median of the covariates in the eBay dataset as a representative auction, we simulate data in a similar way as in Section 3.3 and compute the posterior distribution of the expected seller revenue.

The upper parts of Figure 5 show the expected seller revenue as a function of the minimum bid divided by the book value (MinBidShare) for each model. The optimal minimum bid in the IPV model is zero irrespective of the seller's residual value. This is contrary to the CV model where the optimal minimum bid is a large fraction of the object's book value for each residual value. The optimal minimum bid is as high as 95% of the object's book value when the residual value is μ, 70% when the residual value is 70% of μ, and zero when the residual value is below or equal to 0.6μ (not shown).

In the lower parts of Figure 5, the probability intervals of the number of bidders shrink as MinBidShare increases in each model. The probability intervals are very similar between the models, but in the IPV model the number of bidders is slightly larger for each value of MinBidShare. When MinBidShare is below 0.6 (0.4) in the private (common) value model, the sale probability is close to unity and the seller's expected revenue is therefore essentially the same irrespective of the residual value.

When MinBidShare decreases from 0.6 in the IPV model, the seller's expected revenue increases montonically. The absence of bid shading and almost a guarantee for sale in the IPV model will automatically lead to higher expected selling prices when the number of bidders increases. In the CV model, bidders shade their bids to account for the winner's curse when the number of bidders increases. When MinBidShare decreases from 0.4, it seems that the seller's expected revenue is about the same. Hence, the increased revenue from having more competing bidders is offset by the bidders' shading of bids.

The mean of MinBidShare is 0.58, which is close to the optimal minimum bids in Figure 5 for both 0.8μ and 0.7μ as residual values in the CV model. In Figure 6, the optimal MinBidShare is compared to the actual MinBidShare outcome in each auction for the CV model. Except for some auctions with zero as the optimal MinBidShare, we can see that it is optimal for the seller to choose MinBidShare around his residual value. On average, the optimal MinBidShare is rather close to the actual MinBidShare when the residual value is either 0.9μ or 0.8μ. These facts together with zero or close to zero as the optimal minimum bid for each auction in the IPV model clearly speak in favor of the CV model.

4. Model robustness on simulated private and common value data

If the data originate from the CV model, we would certainly expect the CV model to fit the data better than the IPV model. It is more difficult to say how much worse the IPV model would do. Motivated by this, we try to infer by simulation how robust the models are on either IPV or CV data. Although the CV model is more complex, the IPV model is not nested in the CV model. Using the covariates and posterior mode estimates from the eBay dataset, we simulated 10 datasets of bids from each model. For each dataset, the posterior distribution of the parameters is computed under both models.

4.1. **Estimation results.** Table 5 presents the mean values of the posterior means over the 10 datasets of IPV and CV data. Most of the mean estimates in the IPV and CV models differ only slightly for the different types of data. The major differences in both models are for the estimated constants and the estimated coefficients of MinBidShare in λ. In the CV model, there are also substantial differences in the posterior means for κ and the estimated constant in σ.

Interestingly, the differences in κ can be explained by the zero-expected correlation between potential bids in each auction for the IPV data (since private signals are independently distributed). The correlation between any pair of potential bids equals $\frac{1}{\kappa+1}$ in the CV model, see WV. Using the mean estimates of κ from Table 5, the estimated correlation for the CV data becomes 0.14 as compared to 0.07 for the IPV case. As expected, the correlation between potential bids is smaller and close to zero in the IPV data.

It is straightforward to show from equation (2.10) that the unconditional variance of bids, $\mathrm{Var}(b)$, is equal to $(1-\omega(\kappa,\lambda))^2(\kappa+1)\sigma^2$ in the CV model. Hence, the variance of bids in the CV model also depends on κ and λ in addition to the variance, σ^2, in the IPV model. Table 6 presents the mean of the variance of bids in the IPV and CV data for each model, using the mean of the covariates in Table 5. Within each model, $\mathrm{Var}(b)$ seems to be very similar between the IPV and CV data, while $\mathrm{Var}(b)$ seems to be very different between the models for each type of data. The differences in bid variances between the models are probably explained by the inability of the IPV model to adjust the bid variances for different values of λ.

Table 7 presents $\mathrm{Var}(b)$ for different values of λ within the CV model. Overall, the bid variances are rather similar across λ for the different types of data. Using the median covariates in the eBay data and the mean estimates for the CV model in Table 1 gives $\lambda = 3.7$ and $\mathrm{Var}(b) = 37.33$. This is very close to the mean values of $\mathrm{Var}(b)$ in Table 6 for the IPV model, which suggests that the IPV model estimates the bid variances around the median value of the expected number of bidders λ.

4.2. **In-sample fit.** The in-sample fit of the models is evaluated by comparing the distribution of the actual data to simulated data from the estimated models. Using the same settings as in Section 3.3, 10,000 full datasets with bids were simulated from each estimated dataset within each model. This gives 10 different Hellinger distances for both the IPV and the CV datasets within each model. Table 8 presents the mean of 10 Hellinger distances of the within-auction bid dispersion and cross-auction heterogeneity for the different types of data and model.

The CV model seems to be more robust than the IPV model, although the IPV model is not nested in the CV model. Regardless of which model that generates the data, the Hellinger distance of the cross-auction heterogeneity is lower for the CV model. This might be explained by the more complex CV model, where the correlation between any pair of potential bids is determined by κ. Therefore, the possible clustering between correlated bids can be more easily captured within the CV model as compared to the simpler IPV model. The IPV model cannot account for correlated bids. Even if the data come from the IPV model, the bids in some auctions can be unevenly distributed and perhaps also spuriously correlated by random. This is especially a risk in auctions with very few bids, which is often the case in eBay auctions.

Another reason for the differences in the Hellinger distances can simply be that each model needs to balance the fit of both the cross-auction heterogeneity and the within-auction bid dispersion simultaneously in the data. In Table 8, we can see that the fit of the within-auction bid dispersion is traded off for a better fit of the cross-auction heterogeneity, as we move from CV to IPV data. The opposite occurs for the CV model.

In general, it seems easier for the IPV model to fit within-auction bid dispersion and much easier for the CV model to fit cross-auction heterogeneity. The IPV model can only control the variation in bids through σ, while the CV model has a richer set of parameters to balance the fit between within and cross-auction heterogeneities in a better way.

5. Conclusions

In this paper, we evaluate the empirical performance of the IPV and CV paradigms in second-price auctions. We apply the general MCMC algorithm for Bayesian variable selection in WV. By straightforward modifications of the CV model in BH and WV, we propose a comparable IPV model. The relative performance of the models is evaluated on the eBay coin auction dataset in WV, and model robustness is examined on simulated IPV and CV data. Both models fit the eBay dataset well with a slight edge for the CV model. The IPV model is slightly better at predicting auction prices in most auctions, but the CV model does a much better job in predicting more unusual auctions.

We discern several interesting empirical regularities between the estimated models. In particular, the differences in the expected value of the object are explained by the winner's curse effect in common value auctions. According to our reduced form analysis with bids regressed on functions of the number of bidders, it seems that the winner's curse phenomenon is not apparent in the eBay data. The bids do not depend on the number of bidders, which is evidence against a common value framework. However, the optimal reservation prices in the CV model are very close to the actual reservation prices in the eBay data compared to the zero reservation price which is shown to be optimal in the IPV model.

We propose a model to estimate the correlation between observed bids. In the eBay data, the estimated correlation is roughly 0.5. This is captured fairly well by both models for one hundred replicated eBay datasets within each model. The CV model is much better at fitting CV data than the IPV model, while the IPV model is only slightly better at fitting IPV data. Hence, the CV model seems to be more robust to different types of data.

To conclude, we find evidence for both private and common values in different ways. The value of the object probably includes both a private and a common value component. Unfortunately, the game-theoretic models with a combination of private and common values are not yet ready for an empirical analysis of auction data. Certainly, creating such models is a well warranted area of future research. Meanwhile, we need to choose between the private or the common value paradigm, or use Bayesian model averaging between IPV and CV models. Model averaging can result in a better predictive ability of auction prices, but at the cost of a more complex interpretation of the valuation structure.

References

[1] Armantier, O. (2002), Deciding Between the Common and Private Value Paradigms: an Application to Experimental Data, *International Economic Review*, **43(3)**, 783-803.

[2] Athey, S., and Haile, P. (2002), Identification of Standard Auction Models, *Econometrica*, **70(6)**, 2107-2140.

[3] Bajari, P., and Hortacsu, A. (2003), The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *The Rand Journal of Economics*, **3(2)**, 329-355.

[4] Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004), *Bayesian Data Analysis*. Chapman & Hall/CRC.

[5] Geweke, J., and Keane, M. (2007), Smoothly mixing regressions, *Journal of Econometrics*, **138(1)**, 252-291.

[6] Kass, R.E. (1993), Bayes factors in practice, *The Statistician*, **42(5)**, 551-560.

[7] Krishna, V. (2002), *Auction Theory*. San Diego: Academic Press.

[8] Milgrom, P., and Weber, R. (1982), A Theory of Auctions and Competitive Bidding, *Econometrica*, **50(5)**, 1089-1022.

[9] Nikulin, M.S. (2001), Hellinger distance. In *Hazewinkel, M., Encyclopaedia of Mathematics*. Springer.

[10] Paarsch, J. H. (1992), Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions, *Journal of Econometrics*, **51**$(\mathbf{1-2})$, 191-215.

[11] Pinkse, J., and Tan, G. (2005), The Affiliation Effect in First-Price Auctions, *Econometrica*, **73(1)**, 263-277.

[12] Sareen, S. (1999), Posterior Odds Comparison of a Symmetric Low-Price, Sealed-Bid Auction within the Common-Value and the Independent-Private-Values Paradigms, *Journal of Applied Econometrics*, **14(6)**, 651-676.

[13] Silverman, B. W. (1986), Density estimation for statistics and data analysis. In *Monographs on Statistics and Applied Probability*. London: Chapman and Hall.

[14] Smith, M., and Kohn, R. (1996), Nonparametric regression using Bayesian variable selection, *Journal of Econometrics*, **75(2)**, 317-343.

[15] Vickrey, W. (1961), Counterspeculation, Auctions, and Competitive Sealed Tenders, *The Journal of Finance*, **16(1)**, 8-37.

[16] Villani, M., Kohn, R., and Giordani, P. (2009), Regression Density Estimation using Smooth Adaptive Gaussian Mixtures, *Journal of Econometrics*, **153(2)**, 155-173.

[17] Wegmann, B. (2011), Bayesian Inference in Structural Second-Price Auctions with Gamma Distributed Common Values, manuscript, Department of Statistics, Stockholm University.

[18] Wegmann, B., and Villani, M. (2011), Bayesian Inference in Structural Second-Price Common Value Auctions, *forthcoming in Journal of Business and Economic Statistics.*

[19] Zellner, A. (1986), On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, 233-243. North-Holland/Elsevier.

Table 1. Comparing the estimation results between the IPV and CV models on the eBay dataset.

Par	Covariate	Mean		Std		Incl prob		IF	
		IPV	CV	IPV	CV	IPV	CV	IPV	CV
κ	–	–	5.551	–	0.725	–	1.000	–	5.833
μ	Const	22.781	28.310	0.225	0.243	1.000	1.000	8.707	8.360
	Book_d	0.566	0.741	0.008	0.011	1.000	1.000	3.129	5.372
	Book·Pow	0.030	0.033	0.010	0.014	0.404	0.073	1.810	–
	Book·ID	−0.004	0.123	0.031	0.034	0.005	0.917	–	4.928
	Book·Unopen	0.259	0.375	0.020	0.027	1.000	1.000	3.680	4.203
	Book·MinBlem	−0.006	−0.015	0.016	0.017	0.007	0.005	–	–
	Book·MajBlem	−0.240	−0.248	0.024	0.029	1.000	1.000	2.710	4.350
	Book·LargNeg	−0.007	−0.005	0.014	0.017	0.004	0.006	–	–
$\log(\sigma^2)$	Const	3.871	3.995	0.035	0.068	1.000	1.000	5.297	8.079
	$\log(\text{Book})_d$	1.406	1.276	0.035	0.038	1.000	1.000	2.595	6.652
	log(Book)·Pow	0.037	0.042	0.017	0.018	0.121	0.228	–	–
	log(Book)·ID	0.111	0.103	0.037	0.038	0.577	0.460	2.281	5.535
	log(Book)·Unopen	0.239	0.193	0.023	0.026	1.000	1.000	3.317	6.478
	log(Book)·MinBlem	−0.014	−0.001	0.026	0.024	0.012	0.005	–	–
	log(Book)·MajBlem	0.008	0.036	0.035	0.038	0.010	0.010	–	–
	log(Book)·LargNeg	0.021	0.037	0.024	0.021	0.010	0.033	–	–
$\log(\lambda)$	Const	1.287	1.202	0.022	0.020	1.000	1.000	2.265	3.859
	Pow	−0.033	0.003	0.033	0.034	0.008	0.004	–	–
	ID	−0.173	−0.142	0.085	0.084	0.035	0.022	–	–
	Unopen	0.326	0.360	0.046	0.045	1.000	1.000	2.439	4.563
	MinBlem	−0.049	−0.022	0.062	0.044	0.008	0.004	–	–
	MajBlem	−0.159	−0.165	0.089	0.095	0.024	0.021	–	–
	LargNeg	0.101	0.074	0.056	0.041	0.032	0.018	–	–
	$\log(\text{Book})_d$	0.024	−0.038	0.025	0.028	0.006	0.016	–	–
	MinBidShare_d	−1.141	−1.417	0.060	0.053	1.000	1.000	2.566	4.468

Note: $x_1 \cdot x_2$ denotes the interaction of x_1 and x_2, and $x_d = x - \bar{x}$.

Table 2. The Hellinger distance of the within-auction bid dispersion and the cross-auction heterogeneity for each model on the eBay coin auction data.

Model	Within	Cross-auction
IPV	0.0774	0.0894
CV	0.0763	0.0632

TABLE 3. Log predictive scores (LPS) on the eBay coin auction data for the IPV and CV models.

Model	LPS_c		LPS_d	
	Mean	% wins	Mean	% wins
Pri	−3.640	75.7	−0.926	73.3
Com	−3.781	24.3	−0.575	26.7

TABLE 4. Observed bids divided by their book values are regressed on the number of bidders (NBids) and other covariates.

Covariate	Mean	Std	Incl prob	IF
NBids	0.000	0.001	0.067	–
NBids·MinBidShare$_d$	−0.005	0.008	0.339	1.456
MinBidShare$_d$	0.657	0.047	1.000	36.946
Const	0.856	0.009	1.000	14.460
Pow	0.001	0.005	0.055	–
ID	0.000	0.002	0.004	–
UnOpen	0.162	0.013	1.000	2.887
MinBlem	0.000	0.002	0.009	–
MajBlem	−0.193	0.025	1.000	2.326
LargNeg	0.000	0.001	0.006	–
σ_ϵ	0.245	0.003	–	2.034

Note: $x_1 \cdot x_2$ denotes the interaction of x_1 and x_2, and $x_d = x - \bar{x}$.

TABLE 5. Comparing the mean values of the estimated posterior means from 10 datasets of IPV and CV data.

Par	Covariate	IPV model		CV model	
		IPV Data	CV Data	IPV Data	CV Data
κ	–	–	–	12.705	6.061
μ	Const	23.569	23.493	28.523	28.352
	Book_d	0.565	0.565	0.749	0.743
	Book·Pow	0.021	0.004	0.008	0.002
	Book·ID	0.000	0.011	0.085	0.073
	Book·Unopen	0.286	0.267	0.377	0.373
	Book·MinBlem	0.000	0.000	−0.001	0.000
	Book·MajBlem	−0.194	−0.189	−0.202	−0.238
	Book·LargNeg	0.000	−0.010	0.000	−0.001
$\log(\sigma^2)$	Const	3.646	3.761	4.314	3.935
	$\log(\text{Book})_d$	1.568	1.536	1.386	1.322
	log(Book)·Pow	0.000	−0.001	0.002	0.000
	log(Book)·ID	0.043	0.016	0.020	0.004
	log(Book)·Unopen	0.225	0.237	0.165	0.163
	log(Book)·MinBlem	−0.002	0.001	0.001	0.000
	log(Book)·MajBlem	−0.009	−0.075	−0.003	−0.020
	log(Book)·LargNeg	−0.001	−0.001	0.001	0.000
$\log(\lambda)$	Const	1.073	0.926	1.047	0.905
	Pow	0.001	0.007	0.005	0.001
	ID	−0.056	−0.003	−0.002	0.001
	Unopen	0.359	0.456	0.334	0.431
	MinBlem	−0.003	0.003	0.000	0.001
	MajBlem	−0.038	−0.183	−0.033	−0.195
	LargNeg	0.005	0.000	0.001	0.000
	$\log(\text{Book})_d$	−0.023	−0.011	0.000	−0.007
	MinBidShare_d	−1.642	−2.047	−1.607	−2.091

Note 1: The posterior means are calculated from all of the draws in the posterior sampling. This is contrary to the results in Table 1, where the posterior means were obtained given that the covariate was included in the model.

Note 2: $x_1 \cdot x_2$ denotes the interaction of x_1 and x_2, and $x_d = x - \bar{x}$.

TABLE 6. The mean estimates of the variance of bids, $var(b)$, in the IPV and CV data using the mean of the covariates for each model in Table 5.

Model	IPV Data	CV Data
IPV	38.3	43.0
CV	24.6	25.3

Table 7. The unconditional variance of bids, $var(b)$, in the CV model is presented for IPV and CV data and the eBay dataset. The posterior mean estimates from Table 1 and 5, and the mean of the covariates in σ are used in the calculations.

	$\lambda = 2$	$\lambda = 3$	$\lambda = 3.7$	$\lambda = 4$	$\lambda = 5$	$\lambda = 7$	$\lambda = 9$
IPV Data	18.95	25.63	30.68	32.91	40.68	57.33	74.97
CV Data	22.24	28.85	33.59	35.63	42.46	56.02	69.11
eBay Data	24.97	32.19	37.33	39.53	46.87	61.28	75.02

Table 8. The mean of the 10 Hellinger distances for the within-auction bid dispersion and the cross-auction heterogeneity of each model on the simulated IPV and CV datasets.

	Hellinger Within		Hellinger Cross-auction	
Model	PriData	ComData	PriData	Comdata
IPV	0.0711	0.0679	0.0814	0.1356
CV	0.1037	0.0615	0.0571	0.0666

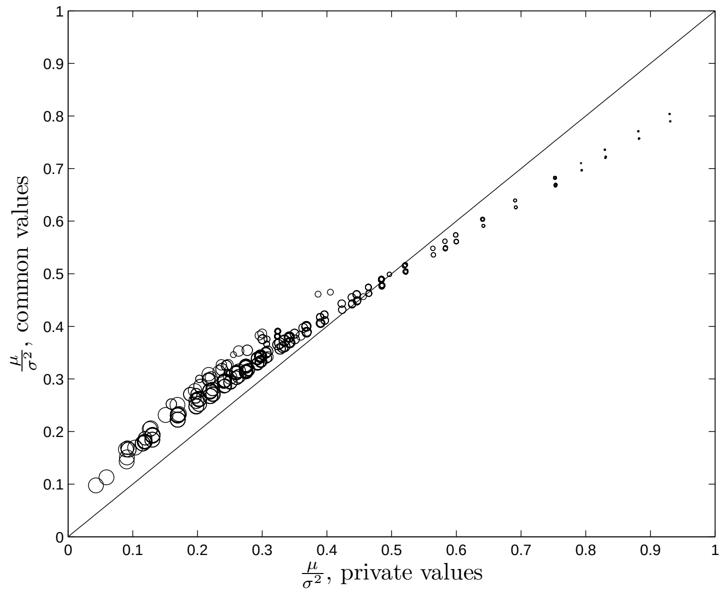

FIGURE 1. The mean-variance ratio, $\frac{\mu}{\sigma^2}$, based on the auction-specific covariates and the posterior mean estimates from the IPV and CV models. The size of the bubbles is proportional to the book value of the auctioned object.

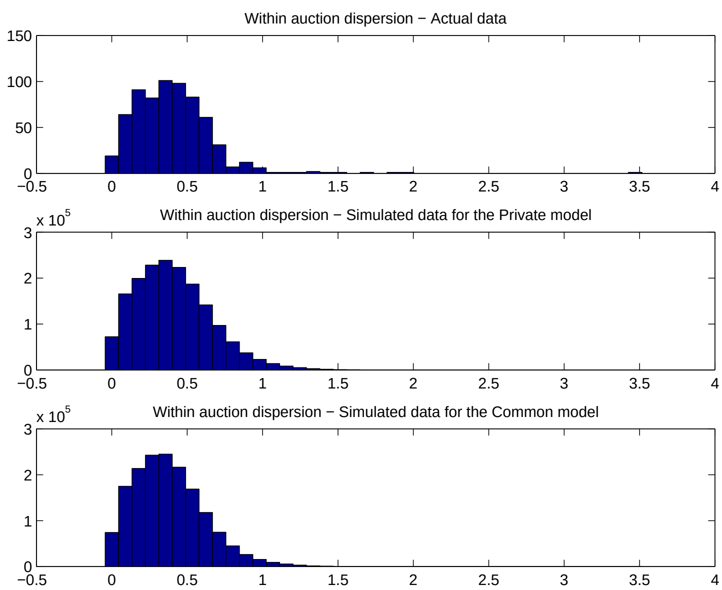

FIGURE 2. Posterior predictive comparison of the within-auction dispersion for the IPV and CV models. The within-auction dispersion is defined as the difference between the highest observed bid and the lowest bid divided by the auctioned item's book value.

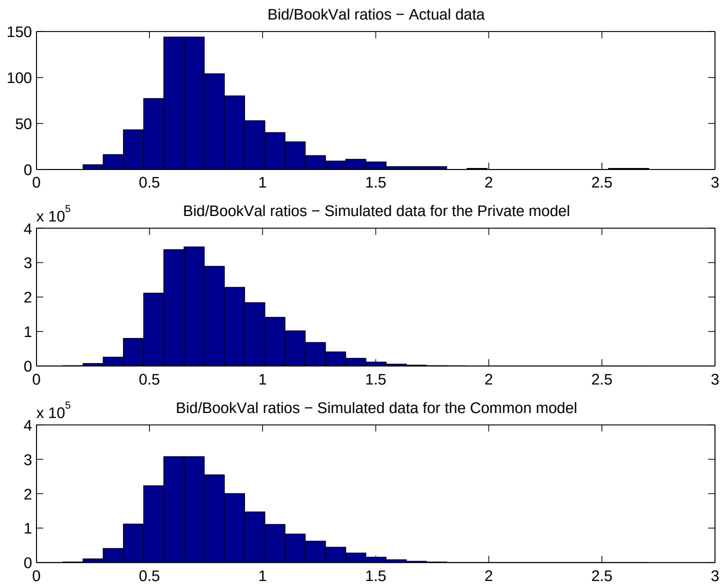

FIGURE 3. Posterior predictive comparison of the cross-auction heterogeneity for the IPV and CV models. Cross-auction heterogeneity is defined as the bids divided by the book value of each auction.

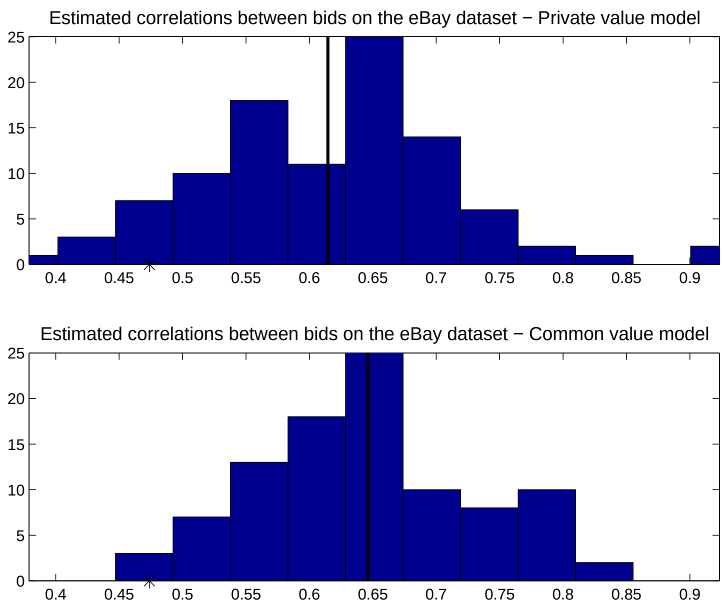

FIGURE 4. Histograms of 1000 estimated correlations between observed bids for each simulated dataset in each model. The vertical lines indicate the mean of the estimates, and the star marks out the estimated correlation in the actual eBay dataset.

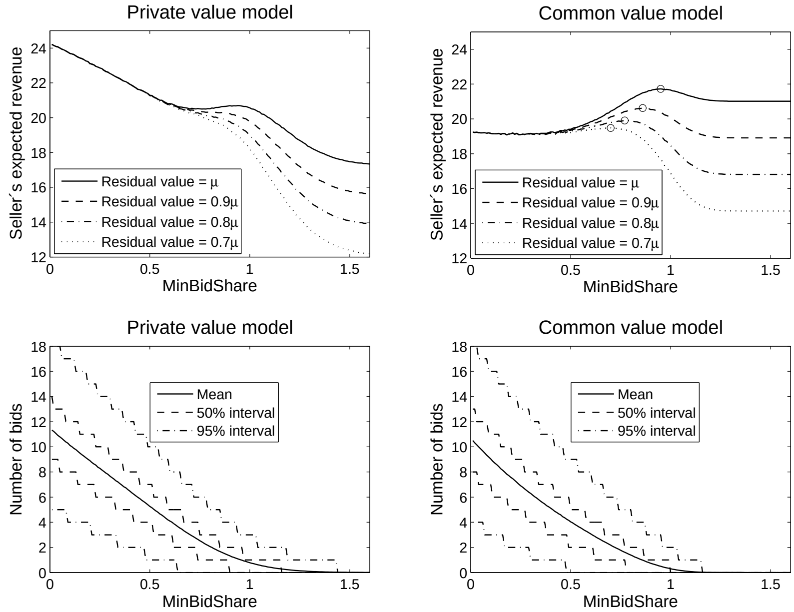

FIGURE 5. Comparing the optimal minimum bids between the models. The figure shows the posterior distribution of the sellers' expected revenue (upper graphs) and the number of bidders with signals $x \geq x^{\star}$ (lower graphs) as a function of the minimum bid divided by the book value. The circles in the upper right figure mark out the optimal MinBidShare value for each residual value in the CV model. The optimal minimum bid in the IPV model is zero, irrespective of the residual value.

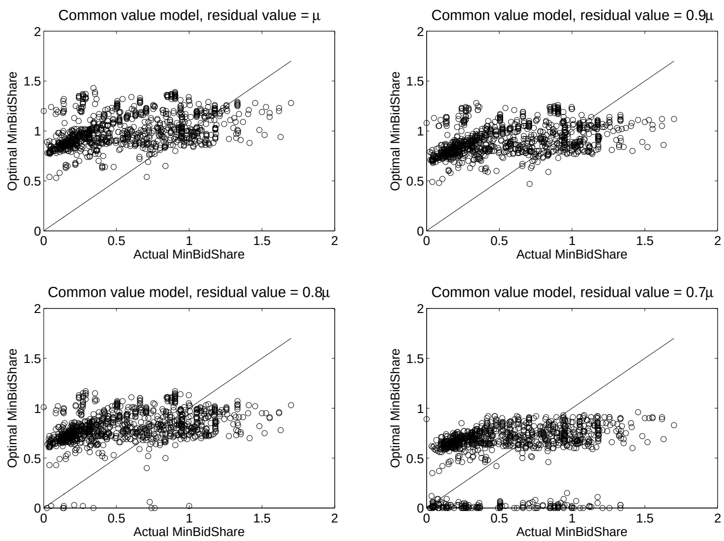

FIGURE 6. The actual and optimal MinBidShare value is compared in each auction to a 45−degree reference line in the CV model. The optimal MinBidShare value in the IPV model is zero in most auctions and a maximum of 0.02 across all residual values.

BAYESIAN INFERENCE IN STRUCTURAL SECOND-PRICE AUCTIONS WITH BOTH PRIVATE-VALUE AND COMMON-VALUE BIDDERS

BERTIL WEGMANN

Abstract. Auctions with asymmetric bidders have been actively studied in recent years. Tan and Xing (2011) show the existence of monotone pure-strategy equilibrium in auctions with both private-value and common-value bidders. The equilibrium bid function is given as the solution to an ordinary differential equation (ODE). We approximate the ODE and obtain a very accurate, approximate inverse bid as an explicit function of a given bid. This results in fast and numerically stable likelihood evaluations, which is an extremely valuable property for inference. We propose a model where the valuations of both common-value and private-value bidders are functions of covariates. The probability of being a common-value bidder is modeled by a logistic regression model with Bayesian variable selection. The model is estimated on a dataset of eBay coin auctions. We analyze the model using Bayesian methods implemented via a Metropolis-within-Gibbs algorithm. The posterior inference of the common-value part of the model is similar to the ones obtained from a model with only common-value bidders, whereas the private-value part of the model is more affected by the introduction of common-value bidders. There is on average a slightly larger probability for a bidder to be a common-value bidder, but this probability depends very little on the auction-specific covariates.

Keywords: Asymmetry, Bid function approximation, Common-value bidders, Gaussian model, Internet auctions, Markov Chain Monte Carlo, Ordinary differential equation, Private-value bidders.

1. Introduction

In empirical studies of auctions, researchers have typically been working within either the independent private-value or the pure common-value paradigms, see e.g. Bajari and Hortacsu (2003, henceforth BH), Sareen (1999), and Paarsch (1992) for good examples. Within the private-value paradigm the value of the object is known to the private-value bidder, whereas the common-value bidder needs to estimate the value that is common to all bidders. In recent years, auctions with asymmetric private-value and/or common-value bidders have been an actively studied research area. Campo, Perrigne, and Vuong (2003) propose a method to estimate asymmetric first-price auctions with affiliated private-value bidders, Maskin and Riley (2000) distinguish between *weak* and *strong* private-value bidders in their study of asymmetric auctions, and Chang and Tan (2010) studied two common-value bidders in which their private information are asymmetrically distributed. While these studies have focused on asymmetry in either the private-value or common-value paradigm, Goeree and Offerman (2002,2003) and Jackson (2009) studied auctions with values involving both private-value and common-value components, and Reny and Zamir (2004) prove the existence of equilibria in general asymmetric first-price auctions with interdependent values.

Tan and Xing (2011, henceforth TX) show the existence of a monotone pure-strategy equilibrium in auctions with both private-value and common-value bidders. In the solution of their model the private-value bidders bid their values, while the equilibrium bid for symmetric

Department of Statistics, Stockholm University, SE-106 91 Stockholm. E-mail: bertil.wegmann@stat.su.se.

common-value bidders is the solution to an ordinary differential equation (ODE) that depends on the parameters in the private-value distribution. To solve the ODE one needs to resort to numerical methods like the Runge-Kutta methods. This is very time-consuming and can not be used for likelihood-based inference, where the bid function needs to be evaluated over and over again for each bid in each auction.

We solve the ODE by approximating the bid function to obtain the inverse bid as an explicit function of a given bid. This is extremely valuable for a fast and numerically stable likelihood evaluation. We document that the accuracy of the approximation is very good by comparing it graphically to the exact bid function in TX. Following BH and Wegmann and Villani (2011, henceforth WV), we define similar Gaussian valuation distributions for both private-value and common-value bidders.

Evaluating the likelihood is substantially simpler if we condition on the valuation type of each bidder. We therefore augment the model with valuation indicators that record if each bid belongs to a private-value or a common-value bidder. This information is of course not observed, but a Bayesian approach allows us to treat these missing observations as unknown parameters to be estimated. We use a Metropolis-within-Gibbs algorithm to simulate from the joint posterior of the model parameters and the valuation indicators.

We contrast our model on the eBay coin auction dataset in WV with bids and auction-specific covariates. Contrary to BH and WV, our model does not explicitly take into account the seller's reservation price. We therefore restrict the empirical analysis to the 464 auctions with a negligible reservation price. Our empirical results show that the inference for the common-value part of the model is essentially unchanged by the presence of private-value bidders, whereas the estimates of the private-value part of the model is strongly affected by the introduction of common-value bidders. We find the results for the common-value distribution to be reasonable, where the book value and the condition of the auctioned object are the main determinants for the valuations. This is also true for the book value in the private-value distribution, but the results for the condition of the object is probably biased. Finally, it seems that the probability of being a common-value bidder depends very little on auction-specific covariates, but, on average, there is a slightly larger probability of being a common-value bidder for a given bid.

2. A model with both private-value and common-value bidders

2.1. Model setup and equilibrium bidding strategies. Assume that m common-value and n private-value risk-neutral bidders compete for a single object in a second-price auction. The number of private-value and common-value bidders in each auction is common knowledge to the bidders. Each private-value bidder knows his or her valuation, q, which is drawn independently from the same distribution. The value of the object, v, for common-value bidders is unknown and the same at the time of bidding, but a prior distribution for v is shared by the bidders. To estimate v, each bidder relies on his or her own information of the object modeled as a signal x, which is drawn independently from the same distribution, $x|v$. Further, the common value and the signals, $v, x_1, \ldots, x_m$, are independent of all private values $q_1, q_2, \ldots, q_n$, but have the same support $[\underline{q}, \overline{q}]$.

Because the auction involves symmetric private-value and common-value bidders, we consider a symmetric bid equilibrium for an arbitrary private-value and common-value bidder without loss of generality. Let $f_q(q)$ denote the probability density function of q, and let $F_q(q)$ denote the cumulative distribution function of q for the private-value bidders. It is a dominant equilibrium strategy for private-value bidders to bid their values in second-price auctions (Vickrey, 1961). This implies that a private-value bidder's equilibrium bid is not affected by the presence of common-value bidders, which gives the bid function for a private-value bidder as

$$b(q) = q. \tag{2.1}$$

Let $g_v(v)$ denote the probability density function of v, let $g_{x|v}(x|v)$ denote the conditional probability density function of $x|v$, and let $G_{x|v}(x|v)$ denote the conditional cumulative distribution function of $x|v$ $(x|x)$ for the common-value bidders. The equilibrium bidding strategy for a common-value bidder is characterized by different expected common values conditional on two possible events A and B: the highest bid of all other bidders is placed by a private-value bidder (A) or a common-value bidder (B) (TX). If A occurs the expected common value, $E[v|A]$, is given by

$$L(x) = \frac{\int_v v \cdot G_{x|v}^{m-1}(x|v) \cdot g_{x|v}(x|v) \cdot g_v(v)\ dv}{\int_v G_{x|v}^{m-1}(x|v) \cdot g_{x|v}(x|v) \cdot g_v(v)\ dv}, \tag{2.2}$$

and if B occurs the expected common value, $E[v|B]$, equals

$$H(x) = \frac{\int_v v \cdot G_{x|v}^{m-2}(x|v) \cdot g_{x|v}^2(x|v) \cdot g_v(v)\ dv}{\int_v G_{x|v}^{m-2}(x|v) \cdot g_{x|v}(x|v) \cdot g_v(v)\ dv}. \tag{2.3}$$

Let $g_{x|x_i}(x|x)$ be the probability density function of the signal x for any other bidder than i conditional on bidder is signal x, and let $G_{x|x_i}(x|x)$ be the corresponding conditional cumulative distribution function. The equilibrium bid function for an arbitrary bidder i with signal x, satisfies the solution to the following ordinary differential equation (ODE) (TX):

$$\frac{db}{dx} = \frac{(m-1)g(x|x)F(b)(H(x)-b)}{nG(x|x)f(b)(b-L(x))}, \tag{2.4}$$

with the boundary condition i) $b(\underline{x}) = H(\underline{x}) = L(\underline{x})$ *and ii)* $b'(\underline{x}) = \alpha L'(\underline{x}) + (1-\alpha)H'(\underline{x})$, where $\alpha = \frac{n}{n+m-1}$.

TX show that there exists an increasing solution, $b(x)$, such that $L(x) < b(x) < H(x)$ for all $x \in (\underline{x}, \overline{x}]$. The bid function can be evaluated by the Runge-Kutta method, which is a reasonably simple and robust algorithm for numerical solutions of ODEs. In the next section we briefly describe how this method can be implemented into our model, and compare the exact bid function to a useful approximate solution of the bid function.

Let θ_j be the probability of the event that the ith bidder is a common-value bidder in auction j. Further, let q_j denote the private value in auction j, let v_j denote the common value in auction j, and let x_{ij} denote the signal of the ith bidder in auction j. Similar to Wegmann (2011), we specify Gaussian models for both private and common values as

$$\begin{aligned}
q_j &\overset{iid}{\sim} N(\mu_{pj}, \sigma_{pj}^2), \quad i = 1, ..., n_j, \\
v_j &\sim N(\mu_{cj}, \sigma_{cj}^2), \quad j = 1, ..., D, \\
x_{ij}\,|v_j &\overset{iid}{\sim} N(v_j, \kappa\sigma_{cj}^2), \quad i = 1, ..., m_j, \\
\mu_{pj} &= z'_{\mu pj}\beta_{\mu p}, \\
\mu_{cj} &= z'_{\mu cj}\beta_{\mu c} \\
\ln \sigma_{pj}^2 &= z'_{\sigma pj}\beta_{\sigma p}, \\
\ln \sigma_{cj}^2 &= z'_{\sigma cj}\beta_{\sigma c}, \\
\ln \frac{\theta_j}{1-\theta_j} &= z'_{\theta j}\beta_{\theta},
\end{aligned} \tag{2.5}$$

where $z_j = (z'_{\mu pj}, z'_{\sigma pj}, z'_{\mu cj}, z'_{\sigma cj}, z'_{\theta j})'$ are auction-specific covariates in the regression models for $(\mu_{pj}, \sigma^2_{pj}, \mu_{cj}, \sigma^2_{cj}, \theta_j)$ in auction j. Note that we model θ as a function of auction-specific covariates. Bidder-specific covariates can equally well be used, but we do not have such information in our eBay dataset. Note also that we model θ using a logit link and the variances using a log-link, but any link functions can be used in our inferential methodology.

The Gaussian model for common values implies that

$$x \,|x_i \overset{iid}{\sim} N\left(\mu_{cj}, \frac{(\kappa+1)(\kappa+2)}{\kappa}\sigma^2_{cj}\right) \tag{2.6}$$

for any common-value bidder other than the ith bidder in an arbitrary auction.

2.2. **The approximate bid function for a common-value bidder.** A single evaluation of the posterior density (likelihood function) requires advanced numerical solutions of the ODE in (2.4) to obtain $b\left(x|\mu_p, \sigma^2_p, \mu_c, \sigma^2_c, \kappa, m, n\right)$. This is a very costly procedure that needs to be repeated for each bid in each auction in the dataset. Therefore, the likelihood evaluation is not fast enough to use for routine inference. Instead, we approximate the bid function to obtain the inverse bid function $\phi(b)$ as an explicit function of b, which implies that numerical integration is not necessary. This leads to much faster and numerically more stable likelihood evaluations. We first derive the approximate bid function and examine the accuracy of the approximation. Then, we state under which conditions our approximation approach holds to account for other valuation structures and auction setups.

In our approximation of the bid function, we use the accurate linear approximations of $L(x)$ and $H(x)$ in WV, given by

$$\hat{H}(x) = c + \omega\mu + (1-\omega)x, \tag{2.7}$$

where $c = -\frac{\sqrt{\kappa}\sigma\gamma\theta(m-2)}{\gamma(m-2)+1+\frac{\kappa}{2}}$, $\omega = \frac{\frac{\kappa}{2}}{\gamma(m-2)+1+\frac{\kappa}{2}}$, $\theta = 1.96$ and $\gamma = 0.1938$, and

$$\hat{L}(x) = \tilde{c} + \tilde{\omega}\mu + (1-\tilde{\omega})x, \tag{2.8}$$

where $\tilde{c} = -\frac{\sqrt{\kappa}\sigma\gamma\theta(m-1)}{\gamma(m-1)+\frac{1}{2}+\frac{\kappa}{2}}$, and $\tilde{\omega} = \frac{\frac{\kappa}{2}}{\gamma(m-1)+\frac{1}{2}+\frac{\kappa}{2}}$. Then, $\hat{L}'(x) = (1-\tilde{\omega})$ and $\hat{H}'(x) = (1-\omega)$. Following the intuition in TX, we propose an approximation of the bid function in (2.4) as a weighted average between $\hat{L}(x)$ and $\hat{H}(x)$,

$$\hat{b}(x) = \frac{\delta\hat{L}(x) + \hat{H}(x)}{\delta+1}, \tag{2.9}$$

where $\delta > 0$ is constrained to be independent of the signal x. Note that, by requiring δ to be positive the approximate bid function fulfills the crucial inequality $L(x) < b(x) < H(x)$ in TX. Taking the first derivative of $\hat{b}(x)$ and replacing b, $L(x)$, and $H(x)$ in $\left(\frac{H(x)-b}{b-L(x)}\right)$ of (2.4) by $\hat{b}(x)$, $\hat{L}(x)$, and $\hat{H}(x)$, respectively, the following equality holds:

$$\frac{db}{dx} = \frac{(m-1)g(x|x)F(b)}{nG(x|x)f(b)} \cdot \delta = \frac{\delta\hat{L}'(x) + \hat{H}'(x)}{\delta+1}. \tag{2.10}$$

Now, in order to solve for δ we approximate x in (2.10) by using the boundary conditions $i)$ and $ii)$ in (2.4). First, let x_0 be the lower bound of the signals given by the solution to $\hat{b}(x_0) = \hat{H}(x_0) = \hat{L}(x_0)$ in $i)$. Second, the solution to $ii)$ gives the slope $\hat{b}'_0(x_0)$ of a linear bid function $\hat{b}_0(x_0)$ through x_0. Hence, the approximated value of x for each bid b can be obtained by

$$\tilde{x} = x_0 + \frac{1}{\hat{b}_0'(x_0)}\left(b - \hat{b}_0(x_0)\right) = x_0 + \frac{b - \hat{b}_0(x_0)}{\alpha(1-\tilde{\omega}) + (1-\alpha)(1-\omega)}. \tag{2.11}$$

Replacing x by $\tilde{x}$ in (2.10), we solve the ODE by solving the quadratic equation for $\delta > 0$, which gives

$$\delta = \frac{1 - \tilde{\omega} - u(b) + \sqrt{(1-\tilde{\omega}-u(b))^2 + 4(1-\omega)u(b)}}{2 \cdot u(b)}, \tag{2.12}$$

where $u(b) = \frac{(M-1)g(\tilde{x}|\tilde{x})F(b)}{NG(\tilde{x}|\tilde{x})f(b)} > 0$ is a function of a bid b. Note that $4(1-\omega)u(b) > 0$ implies that the solution in (2.12) is unique for $\delta > 0$ that satisfies $\hat{L}(x) < \hat{b}(x) < \hat{H}(x)$. Solving for x in (2.9), gives the approximate inverse bid function

$$\phi(b) = \frac{\delta(b - \tilde{c} - \tilde{\omega}\mu) + (b - c - \omega\mu)}{\delta(1-\tilde{\omega}) + (1-\omega)}. \tag{2.13}$$

We compare the approximate inverse bid function in (2.13) to the exact solution obtained with the Runge-Kutta methods, see Dormand and Prince (1980). In our model, the lower bound of the signals is $-\infty$, which is untenable for the implementation of the Runge-Kutta methods. However, we find that sufficiently low values for x in (2.4, i)) only distort the solution of the exact inverse bid function very little. We typically use an initial value, x_0, where $\hat{H}(x) - \hat{L}(x) = 0.1$ and $\hat{b}_0(x_0) = \frac{\hat{L}(x)+\hat{H}(x)}{2}$, together with a maximum step length of 0.01 in the Runge-Kutta algorithm.

Figures 1, 2 and 3 compare the exact bid function to the approximate case graphically. The upper left subgraph of Figure 1 illustrates the comparison for a representative auction, based on the mean of the covariates in the eBay dataset and the posterior mean of the model parameters in Table 1. This gives $\mu_p = 23, \sigma_p = 16, \mu_c = 31, \sigma_c = 6.5, \kappa = 1.4, n = 2$, and $m = 3$. The approximation of the bid function is very good in all figures. Figure 1 shows that the accuracy of the approximation is about the same no matter how many private-value and common-value bidders that are present in an auction. In Figure 2 the approximation is notably better when the fairly large coefficient of variation (CV) of the private-value model decreases in the representative auction. The accuracy of the approximation seems to be most sensitive to changes in the parameters of the common-value model. The approximation deteriorates slightly for larger values of κ and σ_c in Figure 3.

Our approximation approach can be used for other auction setups and valuation structures. The conditions under which it can be used hinges upon the approximations of $L(x)$ and $H(x)$, and that $g(x|x_i)$ is a probability density function given the density functions $g(v)$ and $g(x|v)$. The conditions for the approximations of $L(x)$ and $H(x)$ are given in WV, and hold for the Gaussian common-value model in our model setup. In addition, WV verified that the approximations of $L(x)$ and $H(x)$ hold for a lognormal model. It can be verified that $g(x|x_i)$ is also lognormal in this case.

3. Bayesian Inference

We use Bayesian methods to estimate the model. The next three subsections describe the likelihood function, the prior distribution and the Markov Chain Monte Carlo (MCMC) algorithm that we use to sample from the joint posterior distribution of the model parameters and the private/common-value indicators of the bids.

3.1. **The likelihood function.** BH made the assumption that bids in parallel auctions are independent, and showed that last-minute bidding is a symmetric Nash equilibrium on eBay. This allows us to model eBay auctions as independent second-price auctions, which clearly simplifies the analysis. The probability distribution of each bid in the jth auction is a two-component mixture distribution with probability θ_j that the bid comes from a common-value bidder and with probability $(1-\theta_j)$ that the bid comes from a private-value bidder. It is usually straightforward to estimate a two-component mixture, at least with simulation methods. Our auction model poses a special challenge, since the common-value bid function depends on the number of private-value and common-value bidders, n and m, in the same auction, see Section 2. The only way of writing up the likelihood function is therefore to condition on m and subsequently to sum over m. Alternatively, we can augment the model with indicators, one for each bid, that determine whether a given bid comes from a common-value or a private-value bidder. Let $I_{ij} = 1$ if the ith bidder in the jth auction is a common-value bidder, and let I_j be a vector containing all valuation indicators for bids in the jth auction. Finally, let $\{I = I_1, I_2, \ldots, I_D\}$ denote the collection of all indicators. We use an MCMC algorithm to simulate from the joint posterior of the model parameters and the valuation indicators, see Section 3.3 below.

The likelihood function of bids is complicated by the fact that the highest bid is usually not observed because of eBay's *proxy bidding system* (see BH and WV for details). The bid distribution for a common-value bidder in a single auction is derived from the distribution of the signal as

$$g_b\left(b|\beta_{\mu p}, \beta_{\sigma p}, \beta_{\mu c}, \beta_{\sigma c}, z, \kappa, v, m\right) = g_{x|v}\left[\phi(b)|v, \kappa, \beta_{\sigma c}, z\right]\phi'(b), \tag{3.1}$$

where $\phi(b) = x$ is the inverse bid function. Let $\mathbf{b} = (b_1^\star, b_{p2}, b_{p3}, \ldots, b_{pn}, b_{c2}, b_{c3}, \ldots, b_{cm})$ be the vector of observed bids in an arbitrary auction, where $b_1^\star$ is the highest bid from either a private-value or common-value bidder that is lower than the highest bid in the auction, $b_{p2} > b_{p3} > \cdots > b_{pn}$, and $b_{c2} > b_{c3} > \cdots > b_{cn}$. Then the likelihood function for that auction is given by

$$\begin{aligned} &f_{\mathbf{b}}\left(b_1^\star, b_{p2}, b_{p3}, \ldots, b_{pn}, b_{c2}, b_{c3}, \ldots, b_{cm}|\mu_p, \sigma_p^2, \mu_c, \sigma_c^2, \kappa, I, \theta, z, v\right) = \\ &\left\{1 - F_q\left(b_{p2}|\mu_p, \sigma_p^2\right)\right\}^{I(b_{p1}>b_{c1})} \cdot \left\{f_q\left(b_1^\star|\mu_p, \sigma_p^2\right)\right\}^{I(b_{p1}<b_{c1})} \cdot \textstyle\prod_{i=2}^{n} f_q\left(b_{pi}|\mu_p, \sigma_p^2\right) \times \\ &\left\{1 - G_{x|v}\left[\phi\left(b_{c2}\right)|v, \kappa, \sigma_c^2\right]\right\}^{I(b_{p1}<b_{c1})} \cdot g_b\left(b_1^\star|v, \kappa, \sigma_c^2\right)^{I(b_{p1}>b_{c1})} \cdot \textstyle\prod_{i=2}^{m} g_b\left(b_{ci}|\mu_p, \sigma_p^2, v, \kappa, \sigma_c^2, m\right), \end{aligned} \tag{3.2}$$

3.2. **The prior distribution.** We need to assign a prior distribution to all unknown quantities: $I_j, v_j, \kappa, \beta_\theta, \beta_{\mu p}, \beta_{\sigma p}, \beta_{\mu c}$, and $\beta_{\sigma c}$. The prior distribution of the indicators in a given auction are iid Bernoulli distributed with probability θ_j. The indicator vectors are assumed to be a priori independent between auctions. The prior distribution of the common values, $v_1, \ldots v_D$, are independent $N(\mu_j, \sigma_j)$ variables according to the model. Following WV, we use an inverse Gamma prior for κ, $\kappa \sim IG\left(\bar{\kappa}, g\right)$, where $\bar{\kappa} = 0.25$ and $g = 4$. The five regression coefficient vectors are assumed to follow g-priors (Zellner, 1986) a priori. Specifically, the prior for β_μ, conditional on β_σ, is given by

$$\beta_{\mu\cdot}|\beta_{\sigma\cdot} \sim N[0, c(z'_{\mu\cdot}Dz_{\mu\cdot})^{-1}],$$

where $D^{1/2} = \text{Diag}[\exp(-z'_{\sigma_1}\beta_\sigma/2), \ldots, \exp(-z'_{\sigma_n}\beta_\sigma/2)]$, and $c > 0$ is a scaling factor that we set equal to the number of auctions in the data, making the information in the prior equivalent to a sample of one observation. Marginal g-prior for β_σ is given by

$$\beta_\sigma \sim N[0, c(z'_\sigma z_\sigma)^{-1}].$$

The same equivalence of information in the prior for β_θ does not apply if the g-prior is used for these parameters. Instead, we use the following prior for β_θ:

$$\beta_\theta \sim N[0, c_\theta I],$$

where I is the identity matrix and $c_\theta = 100$. Finally, the prior inclusion probability, π, of a given covariate, in the logistic regression model for θ_j, is set to 0.5.

3.3. **Metropolis-within-Gibbs sampling.** We use a Gibbs algorithm with the following eight parameter updating steps:

1. $\{I_i\}_{i=1}^D$ 2. β_θ 3. $\{v_i\}_{i=1}^D$ 4. κ 5. $\beta_{\mu p}$ 6. $\beta_{\sigma p}$ 7. $\beta_{\mu c}$ 8. $\beta_{\sigma c}$

We use a Metropolis-Hastings update for the valuation indicators in Step 1. The proposals for the I_j-vector is obtained by changing the current value of each element I_j with a certain pre-determined probability. This probability of changing a given element of I_j is here set to $\min\left(1, \frac{1.5}{n_j+m_j}\right)$. Note that we propose all elements of I_j jointly. Since the indicator vectors in different auctions are independent conditional on the model parameters, we can generate the indicator vectors $I_1, ..., I_D$ simultaneously in a single updating step.

The regression coefficients in the logistic regression in Step 2 are simulated using the efficient finite step Newton algorithm in Villani, Kohn and Giordani (2009). This allows us to do Bayesian variable selection among the covariates in the logistic regression model for θ with the indicators I_{ij} as the response variables. Variable selection introduces point masses at zero in the prior distribution to allow for the possibility that a subset of the covariates is absent in θ, see Smith and Kohn (1996). The gradient and Hessian are given in closed form. In the algorithm, Newton's method is used to iterate a small number of steps from the current point towards the posterior mode. The proposal is subsequently drawn from the multivariate t-distribution with mean equal to the terminal Newton point and covariance matrix equal to the negative inverse Hessian at this point. The degrees of freedom is set to 10. We find that it is sufficient to iterate only one Newton step towards the posterior mode. The algorithm is very efficient. We get very good convergence in the MCMC chain with a mean acceptance rate of 85%.

The common values, $v_1, ..., v_D$, in Step 3 are simulated by a Random Walk Metropolis (RWM) step, and the same applies to κ in Step 4. The scaling factors in the RWM are all chosen to roughly match the optimal Metropolis acceptance probabilities in Gelman et. al. (2004). Note also here that the common values are independent, conditional on the model parameters, so we can generate them all at once. We experimented with two different ways of simulating the regression coefficient vectors in Steps 5-8. First, we considered the same finite Newton step algorithm with variable selection as in Step 2, but our implementation of this method resulted in an MCMC chain that got stuck for long spells. Instead, we use simple RWM updates in Steps 5-8. RWM is a much less efficient algorithm, however, and we hope to shed more light on why we were not successful with the finite Newton algorithm.

4. Posterior results on eBay coin auction data

4.1. **Estimation results.** The model is estimated using the data from eBay coin auctions in WV. Their dataset involves *minimum bids* (*reservation prices*) that are set by the seller in each auction. Since we do not model auctions with minimum bids, we use the subset of the data with auctions where the minimum bid is less than half of the book value of the auctioned object. Therefore, the minimum bids can be assumed to have a negligble effect on the bidding process. This gives us data from 464 auctions, which includes auction-specific covariates and bid sequences from auctions that ended between November 7 and December 19, 2007 and December 27, 2007 to January 29, 2008. The book value of the object is defined by Book. The other covariates are dummy variables that are coded to be 1 if a certain characteristic is present in a given auction: sellers with large selling volumes, so called PowerSellers (Pow), whether or not the object has a major damage (MajBlem), and if the object is sold in its original sealed envelope (UnOpen).

Table 1 reports the posterior results of the estimated model, whereas Table 2 and 3 show the results for the estimated model with only private-value and only common-value bidders, respectively. The uncertainty of the parameter estimates is illustrated with 2.5 and 97.5 percentiles of the posterior for each parameter. The numerical efficiency of the MCMC algorithm is described by the inefficiency factor (IF), defined as the number of posterior draws needed to obtain the equivalent of one independent draw.

In general, the empirical results for the common-value distribution are essentially the same as the ones obtained from a model with only common-value bidders. It is interesting to note, however, that the variability in the signals of the common-value bidders decreases across auctions by the presence of private-value bidders, since the posterior mean estimate of κ is notably lower in this case. The estimates of the private-value distribution are strongly affected by the introduction of common-value bidders. The uncertainty of the parameters in the common-value distribution is substantially smaller than for the corresponding parameters in the private-value distribution. We find the signs and magnitudes of the estimated coefficients to be reasonable in the common-value distribution, where the book value and the condition of the auctioned object are the main determinants of the mean valuations. In the private-value distribution, μ_p is less affected by the book value and the condition of the object, but the coins in sealed envelopes ($\mathsf{UnOpen} = 1$) are still higher valued.

Figure 4 illustrates substantially more auctions with larger values of μ_c than μ_p. The valuations for the private-value bidders are more concentrated to low-value auctions. This might be explained by different bidding characteristics between private-value and common-value bidders. It is interesting to observe that the valuations of common-value bidders seem to be much more responsive to imperfections of the objects or missing coins in a package, compared to the private-value bidders (MajBlem has a strong negative effect in μ_c, but is not significant in mu_p). One interpretation of this finding is that private-value bidders have stronger attachment to the particular auctioned coin than the common-value bidders who may be more focused on subsequent resale of the object. This line of reasoning can also explain why common-value bidders react stronger than private-value bidders to changes in the book value of the object.

The major differences in σ between the distributions are substantially larger magnitudes in the posterior mean estimates of the constant and the coefficient of Book_d in σ_p. This suggest that the more heterogenous private-value bidders search the eBay marketplace for coins that match their collection, while the unknown market value for the common-value bidders is more precise.

The last but one column in Table 1 displays the posterior probability of including a given covariate in the model for θ. The posterior inclusion probabilities are all very small. Hence, the probability θ of being a common-value bidder in a given auction depends very little on certain auction-specific covariates. The probability of being a common-value bidder is equal to 0.58 in a *representative auction* with a covariate vector equal to the mean of the sample covariates, and the model parameters fixed at their posterior mean estimates. By taking the mean of each valuation indicator I in all auctions, we determine the probability of being a common-value bidder for each bid. Figure 5 presents a histogram of the probabilities of being a common-value bidder. In many cases the probability is either high, typically $0.6-0.8$, or low, where the probabilities are relatively uniformly distributed between $0-0.5$. Note that we only include covariates on the auction level. It would clearly be interesting to also model θ using information on the bidders as covariates. While such information is available on eBay, for example whether or not the bidder is a power seller or the bidder's feedback score, our current dataset does not include this information.

Finally, Table 4 presents posterior results for a Poisson regression with the total number of private-value and common-value bidders as response variable in each auction. This entry model is needed when predicting e.g. the price in future auctions, see WV. The main determinant of the entry process is $\log(\mathsf{Book})_d$. A larger book value attracts more bidders to the auction. The expected number of bidders is equal to 5.6 in the representative auction.

5. Conclusions

We model second-price auctions with both private-value and common-value bidders. The mean and variance of the valuation distributions for both types of bidders are modeled as functions of covariates. The probability that a given bidder is a common-value bidder is modeled by a logistic regression model with auction-specific covariates, and we use an MCMC algorithm that allows for Bayesian variable selection among the covariates.

In equilibrium, the private-value bidders bid their values and the symmetric common-value bidders bid according to the bid function in TX, which is the solution to an ODE. To solve the ODE one needs to resort to numerical algorithms like the Runge-Kutta methods. This is very time-consuming and is a major impediment to widespread use in practical likelihood-based work, since a single likelihood evaluation requires evaluation of the inverse bid function for every bid in every auction. We approximate the highly complicated bid function and document that the accuracy of the approximation is very good. The approximation is given as an explicit inverse bid function of a bid b and takes virtually no time to evaluate and is very stable numerically.

We estimate the model using Bayesian methods implemented via a Metropolis-within-Gibbs algorithm. The model is illustrated on a subset of the eBay coin auction dataset in WV. We find the posterior results for the common-value distribution to be reasonable, where the book value and the condition of the auctioned object are the main determinants for valuations. This is also true for the book value in the private-value distribution, but the condition of the object does not affect the valuations significantly. Finally, we find that the probability of being a common-value bidder depends very little on auction-specific covariates, but, on average, there is a slightly larger probability for a given bid to be placed by a common-value bidder.

References

[1] Bajari, P., and Hortacsu, A. (2003), The Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions, *The Rand Journal of Economics*, **3(2)**, 329-355.

[2] Campo, S., Perrigne I., and Vuong, Q. (2003), Asymmetry in First-Price Auctions with Affiliated Private Values, *Journal of Applied Econometrics*, **18**, 179-207.

[3] Cheng, H., and Tan, G. (2010), Asymmetric Common Value Auctions with Applications to Private Value Auctions with Resale, *Economic Theory*, $\mathbf{45(1-2)}$, 253-290.

[4] Dormand, J. R., and Prince, P. J. (1980), A family of Embedded Runge-Kutta Formulae, *Journal of Computational and Applied Mathematics*, **6**, 19-26.

[5] Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2004), *Bayesian Data Analysis.* Chapman & Hall/CRC.

[6] Goeree, J., and Offerman, T. (2003), Competitive Bidding in Auctions With Private and Common Values, *Economic Journal*, **113**, 598-614.

[7] —— (2002), Efficiency in Auctions with Private and Common Values: An Experimental Study, *American Economic Review*, **92(3)**, 625-643.

[8] Jackson, M., (2009), Non-existence of Equilibrium in Vickrey, Second-price, and English auctions, *Review of Economic Design*, **13**, 137-145.

[9] Maskin, E., and Riley, J. (2000), Asymmetric Auctions, *Review of Economic Studies*, **67(3)**, 413-38.

[10] Paarsch, J. H. (1992), Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions, *Journal of Econometrics*, $\mathbf{51(1-2)}$, 191-215.

[11] Reny, P., and Zamir, S. (2004), On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-price Auctions, *Econometrica*, **72**, 1105-1125.

[12] Sareen, S. (1999), Posterior Odds Comparison of a Symmetric Low-Price, Sealed-Bid Auction within the Common-Value and the Independent-Private-Values Paradigms,*Journal of Applied Econometrics*, **14(6)**, 651-676.

[13] Smith, M., and Kohn, R. (1996), Nonparametric Regression using Bayesian Variable Selection, *Journal of Econometrics*, **75(2)**, 317-343.

[14] Tan, X., and Xing, Y. (2011), Auctions with both Common-value and Private-value Bidders, *Economics Letters*, **111**, 99-103.

[15] Vickrey, W. (1961), Counterspeculation, Auctions, and Competitive Sealed Tenders, *The Journal of Finance*, **16(1)**, 8-37.

[16] Villani, M., Kohn, R., and Giordani, P. (2009), Regression Density Estimation using Smooth Adaptive Gaussian Mixtures, *Journal of Econometrics*, **153(2)**, 155-173.

[17] Wegmann, B. (2011), Bayesian Comparison of Private and Common Values in Structural Second-Price Auctions, manuscript, Department of Statistics, Stockholm University.

[18] Wegmann, B., and Villani, M. (2011), Bayesian Inference in Structural Second-Price Common Value Auctions, *forthcoming in Journal of Business and Economic Statistics.*

[19] Zellner, A. (1986), On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, 233-243. North-Holland/Elsevier.

Table 1. Posterior inference for the structural model with both private-value and common-value bidders

Parameter	Covariate	Mean	Std	2.5	97.5	Incl Prob	IF
κ	–	1.427	0.241	1.023	1.996	–	106.887
μ_p	Const	23.246	1.749	19.687	26.552	–	148.112
	Book_d	0.338	0.061	0.217	0.451	–	138.304
	Book*Pow	-0.037	0.047	-0.134	0.057	–	179.345
	Book*Unopen	0.319	0.109	0.091	0.524	–	123.107
	Book*MajBlem	0.063	0.125	-0.159	0.314	–	156.576
$\log(\sigma_p^2)$	Const	5.603	0.084	5.445	5.767	–	108.826
	$\log(\text{Book})_d$	1.677	0.094	1.505	1.876	–	100.063
	log(Book)*Pow	0.021	0.027	-0.028	0.077	–	189.796
	log(Book)*Unopen	0.245	0.051	0.148	0.340	–	184.672
	log(Book)*MajBlem	0.127	0.050	0.028	0.205	–	198.927
μ_c	Const	30.687	0.626	29.400	31.879	–	135.916
	Book_d	0.718	0.021	0.679	0.759	–	102.890
	Book*Pow	0.050	0.026	0.008	0.088	–	198.639
	Book*Unopen	0.228	0.051	0.138	0.342	–	167.643
	Book*MajBlem	-0.260	0.043	-0.347	-0.177	–	174.663
$\log(\sigma_c^2)$	Const	3.736	0.088	3.585	3.896	–	141.139
	$\log(\text{Book})_d$	1.518	0.095	1.335	1.705	–	148.827
	log(Book)*Pow	0.076	0.049	-0.012	0.158	–	200.298
	log(Book)*Unopen	0.310	0.047	0.204	0.386	–	190.877
	log(Book)*MajBlem	0.050	0.041	-0.012	0.122	–	199.356
$\log \frac{\theta}{1-\theta}$	Const	0.335	0.057	0.224	0.448	–	18.257
	$\log(\text{Book})_d$	-0.019	0.128	-0.270	0.236	0.133	–
	log(Book)*Pow	-0.078	0.186	-0.446	0.300	0.137	–
	log(Book)*Unopen	0.103	0.367	-0.669	0.796	0.122	–
	log(Book)*MajBlem	-0.041	0.088	-0.209	0.133	0.153	–

Note: $c = n, \bar{\kappa} = 0.25, g = 4, \pi = 0.5$, and $x_d = x - \bar{x}$. The last column displays the inefficiency factors for covariates with at least 0.3 in inclusion probability.

TABLE 2. Posterior inference for the structural model with only private-value bidders

Parameter	Covariate	Mean	Std	2.5	97.5	IF
μ_p	Const	27.626	0.250	27.130	28.117	50.662
	Book_d	0.593	0.009	0.575	0.611	16.288
	Book*Pow	0.024	0.010	0.008	0.050	194.598
	Book*Unopen	0.208	0.026	0.158	0.261	126.183
	Book*MajBlem	-0.218	0.030	-0.278	-0.165	151.667
$\log(\sigma_p^2)$	Const	4.208	0.035	4.143	4.277	53.452
	$\log(\text{Book})_d$	1.520	0.043	1.436	1.601	29.450
	log(Book)*Pow	0.033	0.013	0.009	0.059	196.666
	log(Book)*Unopen	0.307	0.030	0.252	0.365	165.406
	log(Book)*MajBlem	0.032	0.029	-0.031	0.087	194.107

Note: $c = n$ and $x_d = x - \bar{x}$. The last column displays the inefficiency factors.

TABLE 3. Posterior inference for the structural model with only common-value bidders

Parameter	Covariate	Mean	Std	2.5	97.5	IF
κ	–	3.840	0.405	3.113	4.692	18.702
μ_c	Const	34.731	0.373	33.987	35.445	46.147
	Book_d	0.758	0.014	0.731	0.787	24.194
	Book*Pow	0.036	0.013	0.012	0.061	197.874
	Book*Unopen	0.398	0.039	0.330	0.479	135.561
	Book*MajBlem	-0.209	0.043	-0.293	-0.122	173.153
$\log(\sigma_c^2)$	Const	4.122	0.038	4.047	4.197	58.419
	$\log(\text{Book})_d$	1.490	0.042	1.407	1.573	60.226
	log(Book)*Pow	0.040	0.013	0.011	0.068	198.256
	log(Book)*Unopen	0.306	0.026	0.258	0.359	176.147
	log(Book)*MajBlem	0.080	0.064	-0.016	0.179	201.706

Note: $c = n, \bar{\kappa} = 0.25, g = 4$, and $x_d = x - \bar{x}$. The last column displays the inefficiency factors.

Table 4. Poisson regression with the total number of private-value and common-value bidders in each auction as response variable

Response	Covariate	Mean	Std	2.5	97.5	Incl Prob	IF
$n+m$	Const	1.724	0.023	1.679	1.769	1.000	2.095
	$\log(\text{Book})_d$	0.131	0.029	0.075	0.188	0.999	1.296
	Pow	-0.006	0.022	-0.123	0.047	0.150	-
	Unopen	0.074	0.078	0.003	0.248	0.587	1.359
	MajBlem	-0.120	0.137	-0.449	0.027	0.580	1.263

Note: $c = n, \pi = 0.5$, and $x_d = x - \bar{x}$. The last column displays the inefficiency factors for covariates with at least 0.3 in inclusion probability.

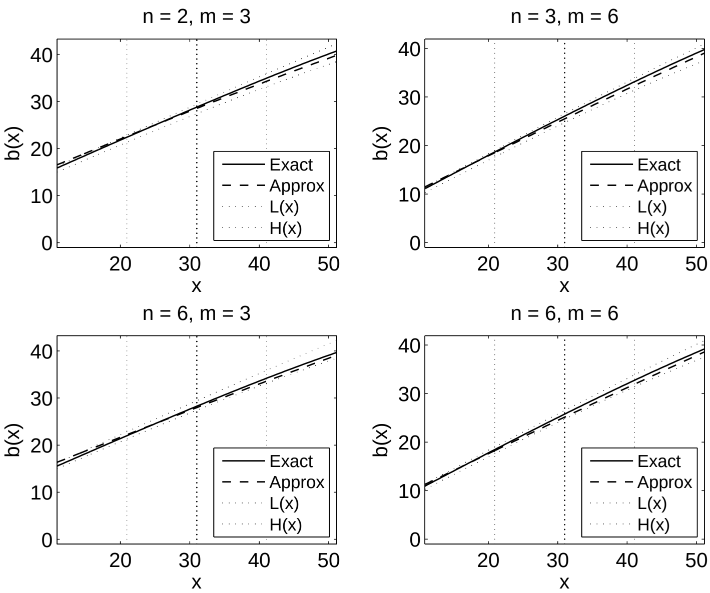

Figure 1. Examining the accuracy of the approximate bid function for different number of private-value and common-value bidders. The parameter values in the private-value and common-value models are $\mu_p = 23, \sigma_p = 16, \mu_c = 31, \sigma_c = 6.5$, and $\kappa = 1.4$. The vertical lines represent the mean $\pm$ 1 standard deviation in the unconditional distribution of the signals, x. The bid function is shown for signals that range between $\pm$ 2 standard deviations from the mean.

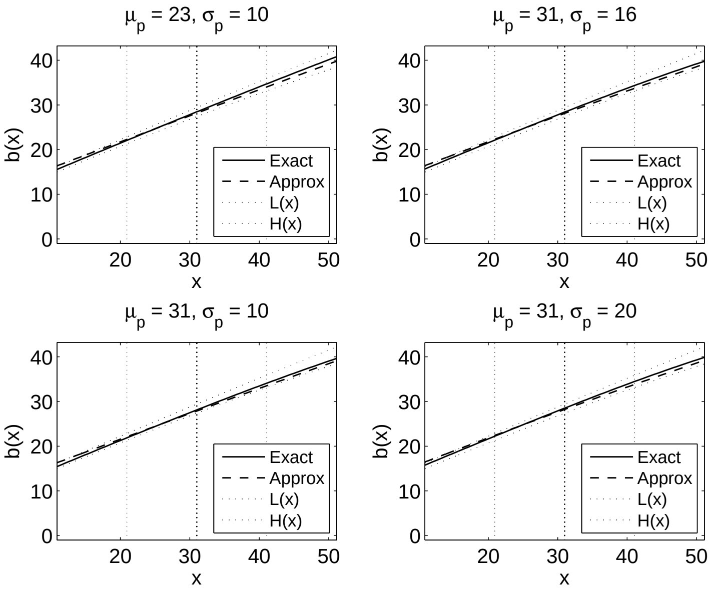

FIGURE 2. Examining the accuracy of the approximate bid function for different parameter configurations of the private-value model. The parameter values in the common-value model and the number of different bidders are $\mu_c = 31, \sigma_c = 6.5, \kappa = 1.4, n = 4$, and $m = 3$. The vertical lines represent the mean $\pm$ 1 standard deviation in the unconditional distribution of the signals, x. The bid function is shown for signals that range between $\pm$ 2 standard deviations from the mean.

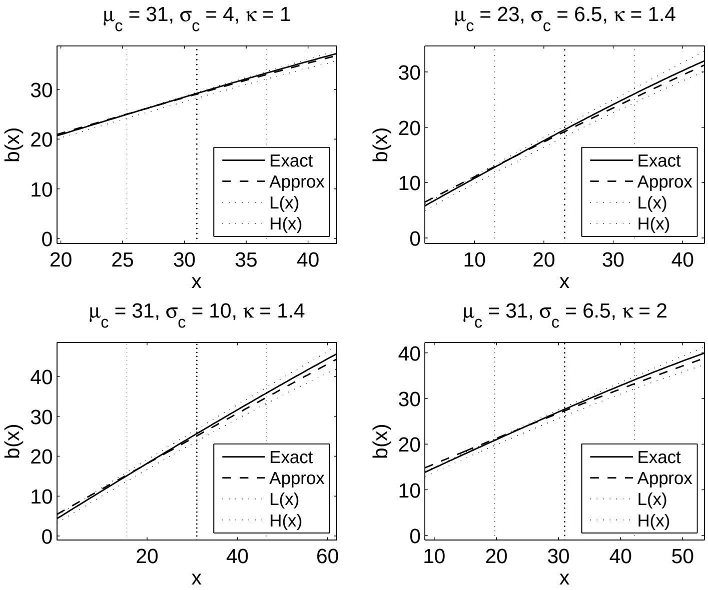

FIGURE 3. Examining the accuracy of the approximate bid function for different parameter configurations of the common-value model. The parameter values in the private-value model and the number of different bidders are $\mu_p = 23, \sigma_p = 16, m = 4$, and $n = 2$. The vertical lines represent the mean $\pm$ 1 standard deviation in the unconditional distribution of the signals, x. The bid function is shown for signals that range between $\pm$ 2 standard deviations from the mean.

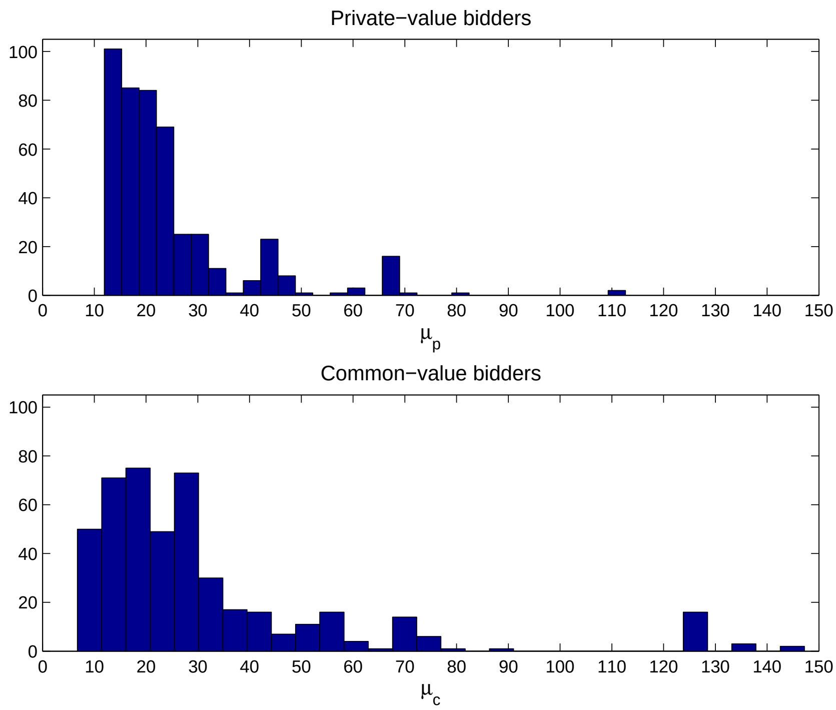

FIGURE 4. Histograms of μ across auctions.

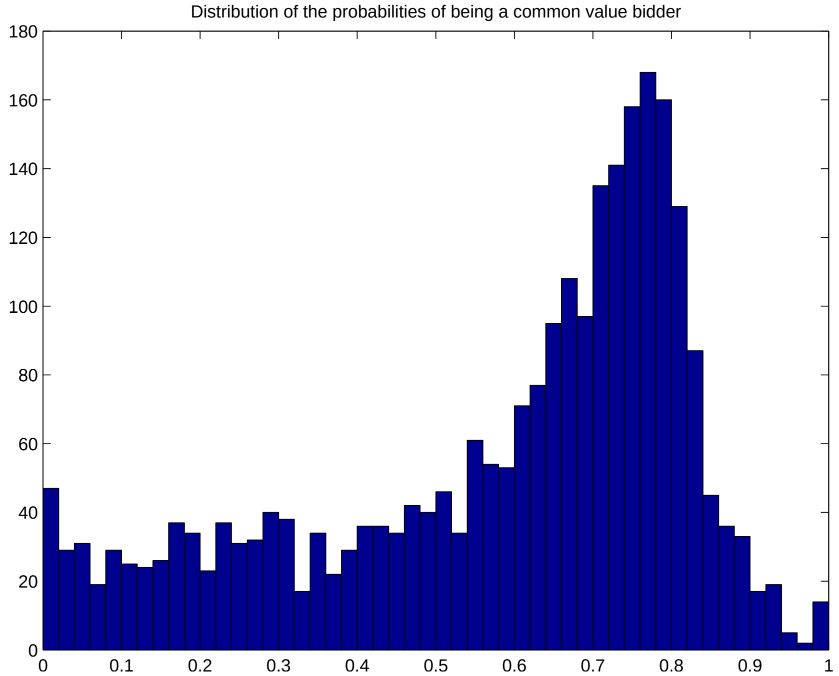

FIGURE 5. Histogram of the probabilities of being a common-value bidder for each bid across auctions.