# Not Knowing the Competition: Evidence and Implications for Auction Design

Yunmi Kong[†]

December 29, 2019

**Abstract**

In a government auction program where first-price auctions generate significantly higher revenue than English auctions, I document evidence that bidders are uncertain about the number of auction entrants. Motivated by additional data evidence, I estimate a structural model of auctions in which rivals' participation is stochastic, allowing for bidders' risk aversion and asymmetry. Counterfactual simulations reveal that bidders' uncertainty about the number of entrants, combined with risk aversion, substantially softens the revenue impact of low competition in first-price auctions. This explains the observed revenue patterns and uncovers an empirically important reason for sellers to favor first-price auctions over English auctions.

This article incorporates content that was previously circulated under the title "Selective Entry in Auctions: Estimation and Evidence". I would like to thank the New Mexico State Land Office for generous assistance in collecting and interpreting the data. Special thanks are due to Stephen Wust, Joe Mraz, and Dan Fuqua. Chris Barnhill, Kevin Hammit, Tracey Noriega, Philip White, and Lindsey Woods also provided valuable industry insights. I would like to thank Xiaohong Chen, Jeremy Fox, Emmanuel Guerre, Kei Kawai, Vijay Krishna, Laurent Lamy, John Lazarev, Alessandro Lizzeri, Xun Tang, anonymous referees, and participants of the NYU CRATE lunch (2014), NYU Stern Friday seminar (2015), SITE Workshop (2015), World Congress of the Econometric Society (2015 Montreal), NYU alumni conference (2017), and Midwest Econometrics Group (2017 TAMU) for helpful comments and suggestions. I am deeply indebted to Isabelle Perrigne and Quang Vuong for their invaluable guidance and support. All remaining errors are mine. Financial support from NYU-CRATE is gratefully acknowledged.

[†]Department of Economics, Rice University. E-mail: yunmi.kong.01@gmail.com

# 1 Introduction

This article presents empirical evidence that uncertainty about the number of entrants combined with risk aversion, a very plausible feature of real-world auctions, is of first-order importance for auction design. In a government auction program where first-price auctions generate significantly higher revenue than English auctions, I document evidence that bidders are uncertain about the number of auction entrants. Motivated by additional data evidence, I estimate a structural model of auctions in which rivals' participation is stochastic from the bidder's point of view, allowing for bidders' risk aversion and asymmetry. Counterfactual simulations using the estimated model reveal that bidders' uncertainty about the number of entrants, combined with risk aversion, substantially softens the revenue impact of low competition in first-price auctions but has no effect in English auctions. This explains the observed revenue difference between auction formats and uncovers an empirically important reason for sellers to favor first-price auctions over English auctions. To my knowledge, this is the first study to show the empirical importance of uncertainty about entrants and risk aversion.

I study auctions run by the New Mexico State Land Office (NMSLO) in the Permian Basin between 2005 and 2014; the Permian Basin is one of the most prolific and economically important oil and gas basins in the world.[1] I document the following data patterns. (1) The first-price sealed-bid auction generates about 30 percent higher revenue than English auctions, controlling for characteristics of the auction item. (2) Bidders in the sealed-bid auction bid several multiples of the publicly announced reserve price even when they are the only bidder, whereas single-bidder English auctions yield exactly the reserve price. Pattern (2) reveals that bidders face uncertainty about the number of auction entrants; if a bidder knew for certain he was the only entrant, he would have bid the reserve price, which is the minimum acceptable bid. Meanwhile, I note that the majority of bidders are small, local independent operators, which makes it likely that at least some bidders are risk-averse.

[1] https://www.britannica.com/place/Permian-Basin

Whether bidders are risk-averse is important because we know from auction theory (McAfee and McMillan (1987) and Matthews (1987)) that, although uncertainty about entrants is revenue-neutral under risk-neutrality, it breaks revenue equivalence in the same direction as pattern (1) under reasonable forms of risk aversion. Intuitively, risk-averse bidders in first-price auctions insure themselves against the possibility of a large number of entrants by bidding relatively high, even when the expected number of entrants is low. Bidders also exhibit asymmetry, which, as shown in Maskin and Riley (2000a), can contribute to pattern (1) as well.

To understand the cause of the revenue difference between auction formats, I estimate an auction model in which bidders are uncertain about the set of entrants and allowed to be asymmetric and risk-averse. From a bidder's perspective, each potential competitor enters the auction with some known probability, resulting in uncertainty about the set of entrants. I allow these entry probabilities to differ by auction format, so that different entry rates can also be an explanation for the observed revenue difference. Exploiting the presence of two auction formats, I nonparametrically identify and estimate bidders' value distributions and utility functions. Estimates indicate that bidders are moderately risk-averse, at a level similar to what has been measured in previous studies of risk aversion.

Counterfactual simulations that isolate the effect of each model feature reveal that uncertainty about entrants combined with risk aversion is the main cause of the observed revenue difference between first-price and English auctions. This finding is robust to alternative specifications of the details surrounding auction entry, such as the number of potential bidders and whether entry is selective. Uncertainty and risk aversion yield this outcome by boosting revenue in first-price sealed-bid auctions while not affecting English auctions. The boost is largest when the number of realized entrants is low, softening the negative effect of low competition. Empirically, the magnitude of this boost is often similar to that of having one additional entrant. As such, this is a first-order consideration for sellers choosing between first-price and English auctions in low competition environments.

The qualitative patterns documented in this study are not particular to the NMSLO auctions; they are consistent with early intuition in natural resource policy that preceded – and did not have the benefit of – the auction theory used in this article. Surveying federal and state natural resource auctions, Mead (1967) remarks anecdotally that even when "a lack of bidder interest [...] results in one-bidder sales under sealed bid procedures, such sales may yield a price close to a competitive price," and where "competition is unreliable, sealed bidding is the more appropriate method since it introduces a measure of uncertainty about who may appear as a bidder." The findings are also important because, as Bulow and Klemperer (1996) have noted, low competition is a critical concern, and it is common in government auctions; Haile, Hendricks and Porter (2010), in their description of U.S. offshore oil and gas lease bidding, remark that "93% of tracts sold attracted one or two bids" in 1998-2002.

**Related literature** Empirical evidence for and consequences of bidders not knowing the number of entrants is a focus of this article. To be clear, this study is not alone in allowing uncertainty about the number of entrants, but it is unique in empirically documenting the implications thereof for choice of auction format and protection from very low revenue. Empirical studies that allow bidders' uncertainty about the number of entrants include the internet auction literature, which often models bidders' arrival as a Poisson process, and some of the literature on endogenous bidder entry. Note that this uncertainty makes a difference only when the bidding strategy depends on the number of entrants. For auctions modeled as second price auctions with private values, this uncertainty is irrelevant for bidding.

In this existing literature, uncertainty about entrants is not a focus per se but a byproduct of the entry model. Examples include Li and Zheng (2009), who analyze the effect of an increased number of potential bidders when entry is endogenous, and Bhattacharya, Roberts and Sweeting (2014), who examine the merit of different ways to organize bidders' entry. In other studies, including Athey, Levin and Seira (2011) and Krasnokutskaya and Seim

(2011), bidders enter endogenously but learn the number of competitors before bidding. Athey et al. (2011), in particular, also use data from two different auction formats and compare their performance, finding that first-price auctions attract more small bidders than English auctions. In all the aforementioned studies, bidders are risk-neutral.

Two methodological studies that allow both risk aversion and bidders' uncertainty about entrants are Li, Lu and Zhao (2015) and Gentry, Li and Lu (2017). Li et al. (2015) develop a test for the form of risk aversion based on symmetric bidders' entry into first-price versus ascending auctions, and Gentry et al. (2017) study identification in first-price auctions with risk-averse bidders and selective entry. Neither is concerned with a revenue comparison between different auction formats.

The theoretical foundations for my study come from McAfee and McMillan (1987) and Matthews (1987), who show that, under nonincreasing absolute risk aversion, expected revenue in a first-price auction is higher when bidders do not know the number of bidders $n$ than when they do. Per Riley and Samuelson (1981), revenue in the latter case is in turn higher than in an English (or second-price) auction, which is unaffected by uncertainty about $n$. This is in contrast to the risk-neutral case, in which first-price and English auctions with or without uncertainty about $n$ would all be revenue-equivalent, as Harstad, Kagel and Levin (1990) show. Menezes and Monteiro (2000) show that this revenue ranking of first-price over second-price auctions continues to hold under a model of endogenous entry. As these studies are concerned with general theoretical results for expected revenue, they do not disaggregate revenue patterns, discuss magnitudes, present empirical evidence, or specifically discuss the case of low competition.

Meanwhile, when laboratory experiments induce uncertainty about $n$, the revenue rankings that result are largely consistent with these risk-averse models and inconsistent with risk-neutral models. Dyer, Kagel and Levin (1989), Isaac, Pevnitskaya and Schnier (2012), and Aycinena and Rentschler (2018) perform such experiments.

More generally, although empirical studies typically assume risk neutrality, bidders are found to overbid relative to risk-neutral Nash in first-price auction experiments. Risk aversion has long been considered a candidate explanation for the overbidding, for instance in Cox, Roberson and Smith (1982) and Cox, Smith and Walker (1983). Bajari and Hortaçsu (2005) take data from a first-price auction experiment and compare four alternative structural models in their ability to recover bidders' value distributions from observed bids. They find that a risk-averse Bayes-Nash model performs better than both risk-neutral Bayes-Nash and behavioral models of bidding.

# 2 New Mexico's Oil and Gas Lease Auctions

## Overview

The New Mexico State Land Office (NMSLO) administers oil and gas leases on its trust lands. These leases grant the lessee the exclusive right to drill the leased land for a specified number of years. In return for the lease, the lessee pays the lessor an upfront lump sum "bonus", which can be considered the price of the lease, in addition to an annual rental and royalties on production. These leases are sold via monthly auctions where bidders bid on the amount of the bonus, of which the public reserve price is approximately $15.63 per acre. Eighty percent of the leases auctioned since 2005 are located in the Permian Basin. The Permian Basin has produced oil for 100 years and "is one of the most well-studied geologic regions of the world" "owing to its economic importance";[2] at the end of 2018, it was the second most productive oil field in the world.[3]

Conversations with agency staff as well as bidders reveal that valuations of a lease are idiosyncratic by bidder. Firms have different probabilities of drilling the tract within the five-year lease term, which depends on how they see the lease fitting into their overall portfolio and development strategy. They also differ in well and field design, recovery rates, aggressiveness

[2] https://www.britannica.com/place/Permian-Basin

[3] https://www.forbes.com/sites/rrapier/2018/12/27/why-the-permian-basin-may-become-the-worlds-most-productive-oil-field/

of hedging programs, cash flow, and alternative options for land acquisition to name some examples. All of these things factor into how they value a lease.

An interesting feature of the NMSLO auctions is that it uses both the first-price sealed-bid (S) and English or ascending oral (O) auction formats within the Permian Basin. As later sections explain, this feature will deliver extra insights regarding the practice of auctions beyond what could have been learned from single-format datasets. The first-price sealed-bid auctions are run via mail-in bids which are opened on auction day, whereas the English auctions are run via a live auctioneer one at a time. During 2005-2014, an average of 19 and 18 Permian Basin leases were sold each auction day via the S and O formats, respectively. The NMSLO records the dollar amount and bidder identity for every bid submitted in the first-price sealed-bid (S) auction. Only the transaction price and winner identity are recorded for the English (O) auction. Hence, the number of entrants for each lease is observed in the data for S but not for O. Figure 1 shows that overall, the number of entrants per auction is not high, with a mean and median of 3 sealed bids, and 46% of auctions receive 2 or fewer bids.

To learn whether there are bidders who participate exclusively in one auction format, I compare the bidder names observed in S to the winner names observed in O; this is an imperfect comparison because only the winners' names are observed in O. Specifically, I compute the fraction of O auctions that are won by bidders who also bid in the S auctions, and the fraction of S bids that come from bidders who also win in the O auctions. Table 1 provides year- and area-specific comparisons as well as a comparison based on the aggregate data (labeled "All"). The fraction of O wins from S bidders is always higher than the fraction of S bids from O winners, as expected when all bidders are observed in S but only the winning names are observed in O. The former fraction is above 90% in every year and area, and 98% in the aggregate comparison. Overall, the vast majority of auctions involve bidders that are recorded to have participated in both formats.

Also, I do not find that bidders withdraw or reduce their participation rate after winning auctions. To check this, I regress each bidder's S participation rate in quarter $t$ on the number of auctions it won in quarter $t-1$ in a fixed effects regression with bidder fixed effects, using 8,463 bidder-quarter observations. The regression coefficient on auctions won in the previous quarter is 0.0076, a small positive number, with clustered standard error 0.0009.

As for the size of land tracts covered by these leases, 320 acres (half a square mile) is by far the most common. To avoid excessive heterogeneity, I focus the analysis that follows on 320-acre leases in the Permian Basin sold during 2005-2014, excluding outlier leases that have a quality index in the top or bottom 5%. The quality index is a scalar summarizing the characteristics of the lease, which are described shortly; the formation of the quality index is detailed in Section 5. Meanwhile, Kong (2017) uses a subset of the NMSLO auctions in which adjacent leases are sold to study the identification of synergy between auctions. To abstract away from inter-auction effects, I exclude the entire subset analyzed by Kong (2017) from my analysis. Going forward, I refer to this set of leases as my estimation sample.[4]

## Differences between auction formats

Figure 2 maps the sections – one-square-mile (640-acre) blocks of land in the Public Land Survey System – that contain my estimation sample, color-coded by the auction format(s) used in each section. It shows that S leases and O leases are not spatially segregated but well mixed. Nonetheless, I do not rely on unconditional exogeneity of auction format but control for lease characteristics, including ex-post production volume and a measure of unobserved heterogeneity, to assess the effect of auction format on revenue, as explained below.

Table 2 provides a summary of sample size, auction revenue, and the observable characteristics of leases by auction format. Observable characteristics fall into three categories: lease terms, location of the tract (encompassing geological features), and time of auction

[4] Bhattacharya, Ordin and Roberts (2018) use NMSLO auctions in the Permian Basin to study the relationship between the auctions and drilling activity. They focus on the first-price sealed-bid auctions only and exclude premium (VB) leases from their data sample.

(industry, economic, local conditions of that time). In the first category, the royalty rate is indicated by the lease prefix; tracts deemed "regular" are assigned a "V0" lease prefix with 16.67% royalty, and tracts deemed "premium" are assigned a "VB" lease prefix with 18.75% royalty. The annual rental is either \$0.50 or \$1 per acre depending on geographic location.

I observe the location of the leased tract, which implies geological information. This includes the barrel-of-oil-equivalents (BOE) produced on the tract between 1970 and the auction date and the BOE produced after the auction date through 2014. I also observe whether a drill bit is recorded to have been drilled into the ground ("spudded") somewhere in the same section as the lease prior to its auction date. I come back to quantifying the unobservable information implied by geographic location shortly.

To represent the effects of time, the table includes oil prices (West Texas Intermediate) and gas prices (natural gas 1 month futures) at the time of auction. In addition, average price per acre in the previous month's auctions and average price per acre in the federal Bureau of Land Management's[5] lease sales in the same quarter are included to reflect local and industry conditions around the time.

According to the table, there is a noticeable revenue difference between the two auction formats, but there are also differences between the auction items sold. I next assess whether auction format matters for revenue after controlling for these lease characteristics. In particular, I include ex-post production volume directly as a control variable, so any remaining differences between formats cannot be attributed to production volume. Of course, leases may differ between formats in more general ways – including but not limited to costs, difficulty of development, etc. – due to unobserved geographical and geological features. Importantly, however, these features are not discontinuous but are expected to change gradually in geographic location; they are similar for nearby tracts.

Following Kong (2017), I exploit this property of land-based heterogeneity to construct a smooth, location-based "heatmap" index as follows. I take deflated sealed-bid data from the

[5]The BLM is a bureau that manages federal public lands, and is distinct from the State Land Office that manages state trust lands. Their auctions are quarterly.

NMSLO auctions and fit a smooth surface of the log bids per acre on geographic (north-south and east-west) coordinates using local quadratic regression. This procedure is performed once for each auction, excluding own-auction bids from the smoothing procedure, and the index for each tract is the value predicted by the fitted surface excluding own-auction bids. This heatmap index is meant to capture the effects of location-determined unobserved heterogeneity. The key assumption underlying this procedure is that unobserved quality is spatially continuous; the index for a given lease could be confounded if that lease lacks a quality shared by all nearby leases. Oil leases satisfy this assumption because unobserved quality is closely tied to geological features, which are continuous in space. For example, Hodgson (2018) models the probability of a successful well as a continuous function over space and explains that such spatial interpolation is representative of industry techniques for predicting geological features.

As a preliminary assessment of the heatmap index's ability to capture location-determined heterogeneity, Table 3 regresses post-auction production volume on the heatmap index and compares the index's explanatory power to that of the other covariates. A comparison of the resulting $R^2$ values indicates that the heatmap index explains more of the variance in production than all of the other covariates combined, confirming its ability to pick up heterogeneity not captured by the other covariates.

Table 4 column (1) displays a regression of auction revenue on auction format, controlling for the observable characteristics listed in Table 2 and year and calendar-month fixed effects. Column (2) additionally controls for the heatmap index. In column (1), all of the statistically significant coefficients – on the "premium" lease prefix dummy, ex-post production volume, oil price, and same quarter BLM price per acre – have the expected positive signs. In column (2), all coefficients retain the same signs, but the coefficients associated with geography and geology shrink considerably as the heatmap index soaks up location effects. The coefficient on the heatmap index is particularly encouraging with regard to its relevance; the index is in units of log dollars, so the interpretation is that the elasticity of auction price with respect to the index (converted to dollars) is 0.95. Moreover, the adjusted $R^2$ increases from 0.156 to 0.268 after inclusion of the heatmap index, demonstrating that it captures additional heterogeneity not represented by the other covariates. The coefficient on auction format remains relatively stable by contrast and statistically significant at the 1% level, corroborating evidence from the map that the observed revenue difference is not explained away by systematic disparities in geography or geology. Strikingly, using the sealed bid (S) format over the English (O) format is associated with a log revenue difference of 0.305 even after controlling for the observed and unobserved heterogeneity of auction items.

In light of the revenue difference in favor of the S format, it is interesting that the O format is used at all. The reason is likely multifaceted. Partially, it is a historical idiosyncracy of New Mexico, as the practice has been maintained for decades although no other state or federal leasing agency currently uses both formats. Partially, the O auction offers benefits that are not directly represented by auction revenue. For instance, O auctions facilitate the realization of synergies between adjacent leases in Kong (2017).

## Uncertainty about the number of entrants

I now investigate in more detail the patterns behind the S-O revenue difference documented above. Figure 3 plots a histogram of the price obtained in auction, separately for each auction format. One immediately noticeable difference between the two formats is that O has a large concentration at the bottom end of prices, whereas S does not. Those first two bars in the O histogram consist entirely of reserve price sales, where the tract sold exactly at the reserve price. By nature of the auction format, a reserve-price O sale indicates that only one bidder raised his hand at the English auction. So the mass at the bottom of the O histogram is caused by auctions with one entrant.

I find that 16% of S auctions also receive only one bid. So why doesn't the S histogram have a similar mass at the reserve price? It turns out that in the S auctions, one-bidder leases sell for on average 7.0 times the reserve price; the median is 4.6 times and the standard deviation of this multiple is 7.6. As shown in Table 2, no S auctions have a price equal to the reserve price. Clearly, the lone bidders in the S auction did not know beforehand that they would be the only bidder. If they had known for certain, they would have bid the minimum acceptable bid, so these leases would have sold at the reserve price, just as in the O auction. This demonstrates that bidders are uncertain about the number of entrants they will compete against. They choose their S bids based on a prior they have regarding the number of entrants $n$ rather than the $n$ realized ex-post. In the example of $n = 1$ discussed above, bidders appear as though they "overbid" ex-post because their priors assigned positive probability to the event that $n > 1$. More generally, when the realized $n$ is lower than expected, they will appear to have "overbid"; when the realized $n$ is higher than expected, they will appear to have "underbid."

However, uncertainty about the number of entrants in itself is insufficient to match the data patterns if bidders are risk neutral. Theory (see Harstad, Kagel and Levin (1990), Krishna (2010)) says uncertainty about the number of entrants is overall revenue-neutral for risk-neutral bidders. In the language used above, cases of "overbidding" when realized $n$ is low balance out cases of "underbidding" when realized $n$ is high in such a way that the uncertainty is revenue-neutral in the aggregate. Then S and O would be revenue-equivalent, in contrast to the evidence in Table 4. Therefore, I proceed to discuss evidence for relaxing two key assumptions underlying the revenue equivalence result: symmetry among bidders and risk neutrality.

## Asymmetry among bidders

As explained in Maskin and Riley (2000a) and Krishna (2010), asymmetry among bidders can break revenue equivalence in either direction. Also, Athey, Levin and Seira (2011) explain that asymmetry can lead to different entry rates into S versus O auctions.

I find evidence of asymmetry in the data. During 2005-2014, I observe 217 names bidding in the Permian Basin, of which 116 appear in my estimation sample as defined previously.

Extra-auction information is scarce for a large fraction of these names. Nonetheless, auction statistics yield information regarding their heterogeneity. Recalling that winning is the only statistic available for O, I observe that the average number of all Permian Basin wins per name during these ten years is 19.4. In line with the Pareto Principle, the roughly 15% of bidder names who exceed this average constitute roughly 85% of all wins. Following common practice in modeling asymmetric bidders, I first divide bidders into two subgroups, defining this top 15% of Permian names as "core" bidders and the remaining names as "fringe" bidders. In addition, I distinguish the most frequent bidder as a separate subgroup due to its dominant presence in these auctions. The most frequent bidder won roughly 30% of these auctions, whereas no other bidder's share exceeded single digit percentages.

Table 5 provides auction statistics from my estimation sample for each subgroup of bidders. In the S auction, entry rates differ much more than win rates conditional on entry, which look relatively similar between subgroups. With the exception of the top bidder, any given bidder bids in a small fraction of the S auctions offered. Meanwhile, the top bidder's share of all wins is higher in S than in O, whereas this is reversed for the other bidder subgroups. Only the top bidder's S-O difference in share of wins is statistically significant, as the last row of Table 5 shows.

The model in Section 3 allows for asymmetry between these three subgroups of bidders. This is a finite approximation of a richer asymmetry in which no two firms are the same in practice and some characteristics of firms – such as a reputation for more sophisticated hedging programs or larger cash flow – may be known among industry participants. In this study, the finite approximation works better than usual because bidders do not know who they are bidding against at the time of bidding; rather than bidding against a specific firm's bid distribution, they are genuinely bidding against a mix of multiple firms' bid distributions.

### Risk aversion

Based on the identity of the bidders in the data, it is highly plausible that many are risk-averse. I manually searched firm websites, Businessweek, Cortera, Hoovers, Linkedin, Manta, Wikipedia, Yahoo, and other sources for employee counts on the bidder names I see in my estimation sample. Among S bids from core bidders excluding the top bidder, 85% come from names either for which I could not find information or which had less than 50 employees. This figure is 82% among fringe bidders. Even the top bidder is a privately held, independent company based locally in New Mexico, though relatively large at about 300 employees. According to staff conversations, the majority of bidders in the NMSLO auctions are local independent operators rather than major integrated companies. In general, we consider smaller, privately held firms more likely to be risk averse. In the literature, George et al. (2005) find that internal ownership of a firm is associated with more risk aversion, and Hiebl (2013), surveying the literature on family firms, reports that family firms are found to be more risk-averse than non-family firms in a majority of studies on the topic.

As reviewed in the literature section, risk aversion is often found to be the best explanation of actual bidding behavior in experimental studies. Moreover, uncertainty about the number of entrants is known to interact with risk aversion in a manner consistent with the revenue ranking observed in my data. This makes intuitive sense; a risk-averse bidder is more sensitive than a risk-neutral bidder to the risk of more bidders turning out than expected and bids more aggressively as a result.

## 3 Model of auctions with asymmetry, risk aversion, and uncertainty about the set of entrants

Motivated by the empirical evidence, I present a model of auctions with uncertainty about the set of entrants, where bidders are allowed to be asymmetric and risk-averse. I assume independent private values (IPV). This is not to say common values are not present in the empirical application; as in most real world auctions, private and common value components likely coexist. I provide a few comments relating to this assumption. The Permian Basin is an area where knowledge of the geology is more complete due to a long history of production dating back to the 1920s. Permits for new seismic surveys are rarely requested in the basin, as these are done only in areas that are not well known and much of the basin has already been drilled in the past. Electric wireline logs from all drilled wells are made public by the New Mexico Oil Conservation Division. Moreover, the data sample covers years 2005-2014, coinciding with the boom in horizontal drilling and hydraulic fracturing, which also reduce production uncertainty.[6] This is important because the key feature of common value auctions is that I as a bidder think other bidders have value-relevant information - e.g. about how much oil is underground - that I do not possess. If all bidders have similar assessments of underlying oil, non-common components are the primary driver of value differences, and that is a private value auction. In terms of the limits of empirical methology, identification of English auction models and identification of risk aversion under the common value paradigm remain open questions at this time.

Regarding independent versus affiliated values, the predicted effects of affiliation go in the opposite direction of the observed revenue ranking. As shown in Milgrom and Weber (1982) and discussed in Krishna (2010), the O auction outperforms the S auction when values are affiliated, and uncertainty about $n$ further reduces S revenue due to the linkage principle.[7]

Finally, I assume the NMSLO's reserve price is not binding, as the agency considers it a "starting point" for serious bidders and tries not to offer tracts for which it might be binding.

[6]The following quote explains why horizontal drilling and hydraulic fracturing reduce production uncertainty. "Previously an oil company had to intersect a geologic structure to be successful so the affected areas were very small and the possibility of a dry hole was large. However with these new technologies the oil companies are now able to drill into the oil bearing source rock so there are no dry holes, and the target area is now in the form of a blanket (your oil bearing formation) as opposed to a point (a geologic structure)." Source: http://www.niobraranews.net.

[7]The linkage principle establishes a revenue ranking of auctions based on the derivative of the expected price paid upon winning with respect to one's own signal. The steeper is this derivative or "linkage", the greater is auction revenue. Milgrom and Weber (1982) show that when values are affiliated, linkage is greater in an English auction than in a first-price auction, and revealing information – such as the number of bidders – increases linkage relative to concealing it.

I discuss the implications of this assumption for my post-estimation analysis in Section 7.

## Setup

There are $M$ subgroups of bidders, denoted by $m \in \{1, ..., M\}$. The number of potential bidders $N_m$ in each subgroup is assumed to be common knowledge. Each subgroup has potentially different distributions $F_m(\cdot)$ of values $v$ on support $[\underline{v}, \bar{v}]$ and utility functions $U_m(\cdot)$. To allow for risk aversion, let $U_m(\cdot)$ be twice continuously differentiable with $U_m(0) = 0$, $U'_m(\cdot) > 0$ and $U''_m(\cdot) \leq 0$. From a bidder's perspective, each potential bidder from subgroup $m$ enters the auction with a subgroup-specific probability $p_m$. As a result, bidders do not know the number and composition of entrants that will actually participate in the auction.

This type of uncertainty can be generated by any number of entry models. Here, I present a model of bidding under the framework of nonselective entry. The endogenous entry model of Levin and Smith (1994) – used by Li and Zheng (2009), Athey, Levin and Seira (2011), and Krasnokutskaya and Seim (2011) – fits into this framework, but I remain agnostic as to the specific entry model generating the observed entry probabilities. As a consequence, this study will not attempt to make predictions that depend on a specific entry model; rather, it will focus on the effect of bidders' uncertainty about the set of entrants. In Section 7, I check the robustness of my empirical findings to the selective entry framework.

In the remainder of this section, I discuss Bayesian Nash equilibrium for bidding in the first-price sealed-bid auction and the English auction.

## First-price sealed-bid auction (S)

Suppose every potential bidder from subgroup $m$ enters with probability $p_m$ and employs a monotonic bidding strategy denoted by $b_m(\cdot; p)$, where $p = (p_1, ..., p_M)$ is the profile of entry probabilities. Define

$$J_m(b|p) \equiv (1 - p_m) + p_m F_m(b_m^{-1}(b; p)),$$

which is the probability that a potential bidder from subgroup $m$ either does not enter or enters and bids less than $b$. The probability of winning $W$, for an entrant from subgroup $m$ who bids $b$, is the probability that all potential bidders either do not bid or bid less than $b$,

$$W_m(b|p) \equiv J_m(b|p)^{N_m - 1} \prod_{i \neq m} J_i(b|p)^{N_i}.$$

Note that uncertainty regarding the number of entrants is built into $W$, as each potential competitor bids with probability $p_m$. Now, the expected profit of a subgroup $m$ entrant who draws value $v$ and bids $b$ is

$$\pi_m(v, b|p) \equiv U_m(v - b) W_m(b|p). \tag{1}$$

Then, the bidder chooses bid $b$ to maximize $\pi_m(v, b|p)$, which yields a first-order condition for bidding:

$$\frac{U'_m(v - b)}{U_m(v - b)} = \frac{W'_m(b|p)}{W_m(b|p)}. \tag{2}$$

Intuitively, the bidder chooses a bid that balances its marginal effect on his utility conditional on winning against its marginal effect on the probability of winning. Replacing $W_m(b|p)$ with its definition and imposing the equilibrium condition $b = b_m(v; p)$ results in a differential equation for the equilibrium bidding strategy of subgroup $m$. Lebrun (1999), Maskin and Riley (2000a), and Maskin and Riley (2000b) among others have derived systems of differential equations characterizing equilibrium bidding with asymmetric bidders. In general, closed-form solutions cannot be obtained for the bid functions. A monotone pure-strategy equilibrium exists for this bidding game, as can be shown by relating this game to the framework in Reny and Zamir (2004).[8] Maskin and Riley (2003) prove that equilibrium is unique when all bidders have the same utility function, but uniqueness has not been proven for asymmetric utility functions.

Anticipating the empirical analysis, it is useful to rewrite the first-order condition with simplified notation. Defining $\lambda(x) = \frac{U(x)}{U'(x)}$, we have

$$\lambda_m(v - b) = \frac{W_m(b|p)}{W'_m(b|p)}. \tag{3}$$

### English (ascending oral) auction (O)

Now consider bidding in the English (O) auction. The profile of entry probabilities is denoted $p^o = (p^o_1 ..., p^o_M)$, where the "o" superscript indicates the O auction throughout. If only one bidder enters the auction, that bidder wins and pays the reserve price $r$. Under the private value paradigm, the English auction is equivalent to the second price auction. Even with uncertainty, risk aversion, and asymmetry, it is still a weakly dominant strategy for each bidder to bid his value. Although there can be other equilibria, this is the unique sequential equilibrium.

## 4 Identification

Before describing my estimation procedure, I briefly discuss identification of the model. The auction model with risk aversion and asymmetric bidders is not identified from either the English (O) or first-price sealed-bid (S) auction alone absent additional conditions. This relates to the nonidentification result in Guerre, Perrigne and Vuong (2009), which states that the first-price auction model with risk-averse bidders is not identified from the observed distribution of bids. The same is obvious for the English auction, as English auction bids

[8] Reny and Zamir (2004)'s model accommodates first-price auctions with risk-averse, asymmetric bidders and a binding reserve price. The bidding game here can be reformulated as a first-price auction with a binding reserve price $\underline{v}$, where bidders draw $v < \underline{v}$ with probability $1 - p_m$ and draw $v \sim F_m(\cdot)$ otherwise.

contain no information about bidders' utility functions. However, if the O auction and S auction share the same $F_m(\cdot)$ and $U_m(\cdot)$ - as they do here, sharing the same pool of bidders - we gain identifying power by using data from both auction formats simultaneously as follows.

As a preliminary step, the entry probabilities $p$ and $p^o$, as Bernoulli probabilities, are identified from observed moments of the number of entrants as a fraction of potential bidders. Then, value distributions are identified from O data by appealing to ideas from Athey and Haile (2002); their Theorems 2 and 3 establish that asymmetric value distributions are identified nonparametrically from transaction prices and winner identities alone. Next, taking these value distributions, a bidder's utility function is identified nonparametrically from her S auction bids as the function that satisfies the first-order condition for bidding in (3), as in Lu and Perrigne (2008). This identification argument provides a basis for estimating the auction model of Section 3, which is the topic of the next section.

# 5 Estimation

In this section, I develop a multi-step estimation procedure to estimate $F_m(\cdot)$ and $U_m(\cdot)$. The main steps of estimation consist of a sieve maximum likelihood estimator for $F_m(\cdot)$ followed by a constructive estimator for $U_m(\cdot)$ that exploits the first-order condition for bidding in the sealed bid (S) auction. I also discuss how I estimate the number of potential bidders $N_m$ and entry probabilities $p_m$ and $p_m^o$. Before describing this estimation procedure in detail, I first address heterogeneity across auction items.

### Auction-level heterogeneity

With many covariates or lease characteristics $z$, nonparametric estimators suffer from the curse of dimensionality, meaning there is not enough data to condition estimates on every combination of covariate values. In order to overcome this problem, I follow the literature in taking a single index approach, in which the primitives depend on the vector $z$ only through a scalar index $z'\beta$, i.e. $F_m(v|z) = F_m(v|z'\beta)$. Note that I use the index only for conditioning; this is different from the approach taken in Haile, Hong and Shum (2003), who subtract the index from bids and work with the resulting residuals. I estimate $\beta$ by regressing the log of submitted sealed bids on all the covariates $z$ listed in Table 4, excluding auction format. Table 2 provides summary statistics of the covariates. The heatmap index described in Section 2 is included to account for location-based unobserved heterogeneity.

Figure 4 shows the observed distribution of $z'\hat{\beta}$, separately for the S and O auction. The distributions are similar; the distribution for S has mean 10.53 and standard deviation 0.54, and the distribution for O has mean 10.49 and standard deviation 0.56. Both have a median of 10.48.

## Number of potential bidders

Although I observe the identities of all bidders in the S auction and all winners in the O auction for every auction, the number of potential bidders is not directly observed. The list of leases to be auctioned and their descriptions are published online for free, and there is no process for registering interest for a particular lease or auction date prior to bidding. A common measure of potential bidders in the literature is the number of unique bidder names $\tilde{N}$ observed in auctions of the same time period and area. As the number of items auctioned grows large relative to the number of names, $\tilde{N}$ would converge to the true $N$. However, in this dataset the number of items auctioned is not large enough to avoid an undesirable feature of $\tilde{N}$: even if the true $N$ is fixed, variation in the quantity of auction items generates substantial variation in observed $\tilde{N}$. When each potential bidder enters with some probability, by construction the number of unique bidder names observed in the data increases with the number of items auctioned.

With this in mind, I estimate my model under two different specifications for counting the number of potential bidders. In the primary specification, I count the number of unique names from subgroup $m$ bidding in the Permian Basin on each auction date, and let $N_m$ be the largest of this count over all auction dates. This has the advantage of being immune to the spurious fluctuations explained above, but has the disadvantage of being insensitive to any genuine fluctuations in $N$. In the alternative specification, I count $N_m$ separately for each year and sub-area (Chaves county, Eddy county, Upper Lea county, Lower Lea county) of the Permian Basin as the number of unique bidder names from subgroup $m$ observed in that year and sub-area. This has the advantage of being sensitive to genuine fluctuations in $N$, but has the disadvantage of also fluctuating spuriously with the number of items offered in that year and sub-area. Table 6 provides a summary of the $N_m$ thus counted. Areas outside the named counties do not have enough observations to compute the alternative specification, so those auctions (7% of auctions in the estimation sample) are dropped in the alternative specification.

## Estimation of entry probabilities

As an intermediate step to recovering model primitives $F_m(\cdot)$ and $U_m(\cdot)$, I need to estimate the entry probability for each subgroup $m$. In the first-price sealed-bid (S) auctions, the number of entrants $n_m$ is observed, so

$$p_m(z'\beta) = \mathbb{E}\left[\frac{n_m}{N_m} \mid z'\beta\right].$$

I estimate $p_m(z'\beta)$ via a kernel regression of $n_m/N_m$ on $z'\beta$.

Meanwhile in the English (O) auctions, the number of entrants $n$ is generally not observed, though the events $n = 0$ and $n = 1$ are detectable as non-sales and reserve price sales, respectively. As a result, estimating entry probability $p^o_m(z'\beta)$ requires much more data than for S auctions and can be difficult in practice if $n = 0$ is a rare event, as it is in this dataset. As a practical alternative, I take an approach which exploits the $p_m(z'\beta)$ estimated from S auctions while still allowing entry rates to systematically differ by auction format. In particular, I specify $p^o_m(z'\beta) = \gamma_m p_m(z'\beta)$, where $\gamma_m$ is a subgroup-specific parameter to be estimated as part of the maximum likelihood step explained below. This specification relies on S data, where $n_m$ is directly observed, to learn how the number of bidders responds to auction items of higher quality ($z'\beta$), while using O data to estimate systematic entry rate differences between O and S.

## Estimation of value distributions and O entry parameters

Recall that not all bids, but only transaction prices, are observed for the O auction. I propose a sieve maximum likelihood estimator to estimate bidders' value distributions and O entry parameters $\gamma_m$ from these observables. Specification of the likelihood relies only on the transaction price equaling the second highest value among entrants.

Specifically, we see for each item $k = 1, ..., K$ the transaction price and the identity of the winning bidder. So the log likelihood of the observed data is the log likelihood of observed transaction prices $t$ and winner's subgroups $m$ given entry probabilities $p^o_m(\cdot) = \gamma_m p_m(\cdot)$ and auction covariates $z$.[9] Namely,

$$\sum_{k=1}^{K} \log(\Pr(\text{2nd highest } v \text{ among entrants} = t_k \text{ \& winner's subgroup} = m_k | \gamma_m, z'_k\beta)). \quad (4)$$

The idea is to find the $F_m(v|z'\beta)$ and $\gamma_m$ that maximize this log likelihood. Before writing the mathematical expression for (4), some notation is useful. Let $H_m(t|p^o_m, z'\beta) \equiv (1 - p^o_m) + p^o_m F_m(t|z'\beta)$, the probability that a bidder either does not bid or does bid with $v \leq t$. Also, define its derivative $h_m(t|p^o_m, z'\beta) \equiv p^o_m f_m(t|z'\beta)$. Using $H_m$ and $h_m$ as shorthand for $H_m(t|p^o_m, z'\beta)$ and $h_m(t|p^o_m, z'\beta)$, respectively, the likelihood of observing transaction price $t \in [\underline{v}, \bar{v}]$ and a winning bidder from subgroup $m$ is

[9] The transaction price is a continuous variable, so Pr(2nd highest $v$ among entrants $= t$) represents a density.

$$N_m(1-H_m)\left[\left((N_m-1)h_m H_m^{N_m-2}\prod_{i\neq m} H_i^{N_i}\right)+\sum_{i\neq m}\left(N_i h_i H_m^{N_m-1}H_i^{N_i-1}\prod_{j\neq i,m} H_j^{N_j}\right)\right]. \quad (5)$$

This is the likelihood that the second highest value among entrants is $t$ and the winner is from subgroup $m$ when there are $N_m$ potential bidders from each subgroup who "draw" their value from $H_m(\cdot|p_m^o, z'\beta)$, which has built in the fact that each potential bidder bids with probability $p_m^o$. In addition, the likelihood of an observation where no one bids is $\prod_{m=1}^{M}(1-p_m^o)^{N_m}$, and the likelihood that a lease sells at the reserve price to a bidder from group $m$ is $N_m p_m^o(1-p_m^o)^{N_m-1}\prod_{i\neq m}(1-p_i^o)^{N_i}$.

To use sieve estimation, $F_m(v|z'\beta)$ is approximated using a bivariate Bernstein polynomial after normalizing $v$ and $z'\beta$ to have support in $[0,1]$, the domain of Bernstein polynomials.[10]

$$F(v|z'\beta) = B_{a,b}(v, z'\beta)$$

$$\equiv \sum_{i=0}^{a}\sum_{j=0}^{b}\alpha_{i,j}\begin{pmatrix} a \\ i \end{pmatrix} v^i(1-v)^{a-i}\begin{pmatrix} b \\ j \end{pmatrix}(z'\beta)^j(1-z'\beta)^{b-j}.$$

Polynomial approximation imposes that $F(v|z'\beta)$ be continuous not just in $v$ but also in $z'\beta$. A useful property of Bernstein polynomials is that functional restrictions, which are potentially complex, can be imposed via simple linear restrictions on the Bernstein coefficients $\alpha_{ij}$. Specifically, as $F_m(v|z'\beta)$ is a cdf, $B(v, z'\beta)$ should be weakly increasing in $v$; with Bernstein polynomials, this is imposed via the simple restriction that $\alpha_{i,j} \leq \alpha_{i',j}$ if $i < i'$. Properties of Bernstein polynomials are detailed in Lorentz (1986); other work using Bernstein polynomials in the estimation of auction models includes Komarova (2017), Kong (2017), and Compiani, Haile and Sant'Anna (2018).

Now the likelihood (5) can be expressed in terms of polynomial $B(\cdot,\cdot)$ by replacing $H_m$

[10]These normalized values are used only inside the Bernstein polynomial; when performing all other calculations, true values are used.

with $(1 - p_m^o) + p_m^o B(t, z'\beta)$ and $h_m$ with $p_m^o B_1(t, z'\beta)$. Finally, $F_m(v|z'\beta)$ and $\gamma_m$ can be estimated by finding the parameters $\alpha_{i,j}$ and $\gamma_m$ that maximize the log likelihood of the O data.

## Estimation of utility functions

I estimate nonparametric utility functions $U_m(\cdot)$ for each subgroup by exploiting the first-order condition for bidding in S, introduced in the model section and restated below:

$$\lambda_m(v - b) = \frac{W_m(b|p)}{W'_m(b|p)},$$

where $\lambda_m(\cdot) \equiv U_m(\cdot)/U'_m(\cdot)$ and $W_m(b|p)$ is the probability of winning for an entrant from subgroup $m$ who bids $b$ conditional on entry probabilities $p$. As entry probabilities $p_m(z'\beta)$ are functions of $z'\beta$, conditioning on $z'\beta$ is sufficient to condition on $p$. In short, I estimate $\lambda_m(\cdot)$ as the nonparametric function that maps $v - b$ to $W_m(b|z'\beta)/W'_m(b|z'\beta)$, following the logic of Lu and Perrigne (2008). Detailed steps follow.

Define $G_m(b|z'\beta)$ to be the distribution of S bids from subgroup $m$ conditional on $z'\beta$. As bid functions are monotonic, the $\alpha$-quantile of values should map to the $\alpha$-quantile of bids. Abstracting from conditioning on $z'\beta$ for the remainder of the section, the first-order condition can be written as $\lambda_m(F_m^{-1}(\alpha) - G_m^{-1}(\alpha)) = W_m(G_m^{-1}(\alpha))/W'_m(G_m^{-1}(\alpha))$ as a consequence. Having estimated the value distribution $F_m(\cdot)$ as explained before, I estimate $G_m(\cdot)$ as the smoothed empirical cdf of observed S bids. As for the probability of winning $W_m(\cdot)$, it is equivalent to the distribution of the highest competing bid facing a subgroup-$m$ bidder. These are directly observed, so I estimate $W_m(\cdot)$ as the smoothed empirical cdf of highest competing bids. Having thus estimated $F_m(\cdot)$, $G_m(\cdot)$, and $W_m(\cdot)$, I use the first-order condition above to construct $\hat{\lambda}_m(\cdot)$ as the function that maps $\hat{F}_m^{-1}(\alpha) - \hat{G}_m^{-1}(\alpha)$ to $\hat{W}_m(\hat{G}_m^{-1}(\alpha))/\hat{W}'_m(\hat{G}_m^{-1}(\alpha))$, for every quantile $\alpha$. Then, as $\lambda_m(x) \equiv U_m(\cdot)/U'_m(\cdot)$, I have $\hat{U}_m(y) = exp \int_1^y 1/\hat{\lambda}_m(t)dt$; this estimates $\hat{U}_m(\cdot)$ to scale, with the scale normalization

$U_m(1) = 1$. Other normalizations can be chosen as well; the scale is easily adjustable.

## 6 Estimation results

Figure 5 depicts the density of the estimated value distributions $F_m(\cdot)$ for each subgroup of bidders, for both specifications of $N$ described in Section 5; specification 2 is shown by the dotted lines. Appendix Figure 11 provides the associated cumulative distribution functions and bootstrap intervals thereof. In specification 1, the top bidder (subgroup 1) has a higher mode than the fringe bidders, and the value distribution of core bidders (subgroup 2) is bimodal. In specification 2, the fringe bidders' (subgroup 3) distribution changes the most; as seen previously in Table 5, fringe bidders constitute the smallest fraction of bids and wins. As a result, there is less relevant data, and estimates for the fringe bidders are consistently least precise among subgroups for all results that follow. Overall, there does not seem to be a clear dominance relation among the subgroups' value distributions.

Table 7 shows the estimated O entry parameters $\gamma_m$. As discussed in Section 5, $\gamma_m$ is the ratio of the O entry rate to the S entry rate for subgroup $m$. Figure 6 provides 95% bootstrap intervals based on 1000 bootstrap samples and the primary specification of $N$. Subgroup 1 enters O at a lower rate than it enters S, whereas subgroup 2 enters O at a higher rate. The bootstrap interval for subgroup 3 contains $\gamma_3 < 1$, but overall it lies mostly in $\gamma_3 > 1$. In the next section, I account for these entry rate differences between auction formats as I analyze the revenue patterns observed in the data.

The other estimated primitive is the utility function $U_m(\cdot)$, which I estimate nonparametrically as explained in Section 5. For purposes of presenting the level of risk aversion concisely and comparing to the risk neutral benchmark, I also estimate $U_m(\cdot)$ parametrically using the constant relative risk aversion (CRRA) specification. Risk neutrality corresponds to a CRRA parameter of zero, and positive parameter values indicate risk aversion. Table 7 presents the estimated risk parameters for each subgroup, and Figure 6 provides 95% bootstrap intervals based on the first specification of $N$. The bootstrap interval for subgroup 2 lies entirely in the positive risk aversion domain. The bootstrap interval for subgroup 1 contains risk neutrality but lies mostly in the positive domain. The bootstrap interval for subgroup 3 is much wider than the others and contains risk neutrality. As a result, the null hypothesis of risk neutrality cannot be rejected for subgroups 1 and 3. Meanwhile, regarding the levels of risk aversion, the estimates are in the range of what is found in the literature; Holt and Laury (2002) measure CRRA parameters centered around 0.3-0.5 in laboratory experiments, and Lu and Perrigne (2008) measure roughly 0.59 using data from the U.S. Forest Service timber auctions.

To assess model fit, I compare the observed data to data simulated from the estimated model, according to the observed distribution of $z'\beta$ shown in Figure 4. The sample size of the simulated data is 100 times that of the observed sample. First, I compare the observed versus simulated number of entrants $n$. In the O auction data, we observe $n$ only when it is 0 or 1. The observed probabilities of $n = 0$ and $n = 1$ in the estimation sample are 3.8% and 11.9%, respectively, and the simulated analogs are 2.6% and 12.6%. Figure 7 similarly shows the observed versus simulated distributions of $n$ for the S auction, which are also close. Next, I compare the observed versus simulated distribution of O prices - as only prices are observed for O - and S bids in Figure 8. As $n = 1$ leads to a mass of reserve price sales in O, Figure 8 excludes this mass and shows O prices for $n > 1$. Meanwhile, when it comes to simulating S bids, rather than simulating asymmetric bid functions for every value of $z'\beta$, which is a continuous variable, I simulate the representative case of median $z'\beta$ and compare to observed bids conditional on median $z'\beta$, where this conditioning is done via kernels. Overall, the estimated model fits the data very well.

# 7 Importance of uncertainty and risk aversion

## Explaining the observed revenue patterns

In Section 2, we observed that the first-price sealed-bid (S) auction generates more revenue than the English (O) auction and that lone bidders in S auctions bid much more than the reserve price. According to auction theory, the model I estimated has multiple features that could contribute to the observed revenue difference between S and O: (1) different entry rates into the two auction formats, (2) bidders' uncertainty about who will bid combined with risk aversion, and (3) asymmetry between bidders. Fixing quality index $z'\beta$ at the median value, I investigate the contribution of each feature to the observed revenue patterns.

Table 8 summarizes the effect of each feature stated above on simulated auction revenue. The first row simulates the full model. The second row shuts down feature (1) and shows what revenue would be if O had the same entry rates as S. The third row shuts down both features (1) and (2) by additionally making bidders know the set of entrants and be risk-neutral. Simulating S revenue in the third row is not trivial because bid functions must be numerically simulated for every possible combination of $(n_1, n_2, n_3)$, where $n_m$ is the number of entrants from subgroup $m$. This is necessary because every combination of entrants leads to a different bid function when the set of entrants is known by bidders as they bid. I modify the algorithm of Fibich and Gavish (2011) to simulate each of those asymmetric bid functions.[11]

We see, from examining each row, that feature (2) is the main contributor to S earning more revenue than O. Feature (3) by itself would cause only a minor revenue difference in this setting, and feature (1) works in the opposite direction of the overall revenue difference. Both specifications of the number of potential bidders $N$, discussed in Section 5, lead to the same qualitative conclusion. The rest of this section delves deeper into the Table 8 numbers. For illustrative purposes, I use estimation results from the primary specification of $N$.

[11] The appendix describes how I modify the algorithm for type-symmetric bidding.

First, consider feature (1). As a result of the entry rate differences shown in Table 7, the O auction attracts on average 2.85 bidders, compared to the S auction's average of 2.7. This explains why O revenue in Table 8's first row is higher than in the second row, where feature (1) is shut down.

Next, consider the effect of feature (2). Figure 9 enables a more in-depth comparison of Table 8's second and third rows by disaggregating expected revenue by the realized number of entrants $n$. The yellow bars depict current S revenue, corresponding to Table 8's second row, column S. The green bars show what S revenue would have been if risk-neutral bidders knew the set of entrants when they bid, as is typically assumed in the empirical auction literature. This corresponds to Table 8's third row, column S. Finally, the dark blue bars depict what O revenue would have been if entry rates were the same as in S, so that entry rates are not a factor in Figure 9's revenue comparison. This corresponds to Table 8's second row, column O. As a reference, the red dotted line marks the value of the item to the median core bidder.

The dark blue and green bars confirm that a low number of entrants is very damaging to auction revenue; cases with just one entrant are particularly devastating. In light of this, the yellow bars are intriguing. When the number of realized entrants is low, the yellow bars display a large revenue boost similar in size to having one additional entrant in the green bars. In Section 2, we saw that one-bidder S auctions earn much more than the reserve price; this evidence maps to the first yellow bar versus the first blue bar. Moreover, the yellow bar for one entrant comes close to the green and blue bars for two entrants, consistent with Mead (1967)'s remark that "one-bidder sales under sealed bid procedures [...] may yield a price close to a competitive price." Meanwhile, at high realizations of $n$, the yellow bars end up "underbidding," as bidders' priors about $n$ assign high probabilities to lower $n$. So the yellow bars can drop lower than the green bars, but here the drop is not large enough to wash out the gains made at lower realizations of $n$. This is due to the interaction of uncertainty with risk aversion. In total, Table 8 and Figure 9 reveal that uncertainty combined with risk aversion can provide meaningful protection against the effects of low competition in a real auction environment. This protection applies to the S auction but not to the O auction, where bidders have a weakly dominant strategy for bidding unaffected by the number of entrants or risk aversion. Therefore, this is a first-order reason for sellers to favor first-price sealed-bid auctions over English auctions in low competition environments.

What is the overall revenue gained by running a first-price sealed-bid (S) auction instead of an English (O) auction in the NMSLO setting? To quantify this, I take each of the items auctioned by S in the estimation sample and simulate 100 O auctions conditional on that item's quality index $z'\beta$. This format change also accounts for the accompanying change in entry rates; bidders now enter the auctions at O entry rates, not S entry rates. As shown in Table 2, the observed average revenue per S auction is \$128,822. Simulations that convert these auctions to O predict that this revenue would drop to \$101,848, or by \$26,974 per item. The simulated number is quite similar to the observed average revenue per O auction, \$102,481, which is consistent with Figure 4, where we saw that the distribution of $z'\beta$ is similar in the two auction formats. Therefore, \$26,974 is also an approximation of the revenue that would be gained per item by converting the O auctions to S. Of course, a caveat when interpreting these simulations is that they are conducted assuming the number of potential bidders would remain at current levels; they do not account for any changes in the number of potential bidders that could result from converting all auctions to S or to O.

## Robustness of findings

### Robustness to selective entry

In estimation, the one restriction I placed on the entry model is that it be nonselective. Now I examine the robustness of my findings to misspecification of the entry framework, or to the possibility that the true entry model is selective. To clarify, nonselective entry means the distribution of values conditional on entry does not depend on the entry rate. On the other hand, if entry is selective, the distribution of values conditional on entry does vary with the entry rate because bidders with the highest values are most likely to enter. Therefore, if the true entry model is selective, different entry rates between S and O can play a bigger role in explaining the revenue patterns we observe. Specifically, we saw in Table 7 that the S auction overall has a lower entry rate than O. Then under selective entry, entrants' values in S would on average be higher than in O, potentially contributing to higher revenue in S.

The severity of misspecification would increase with the degree of selectiveness. Therefore, I subject my findings to the most stringent test of robustness by re-estimating the auction model under the most selective form of entry, the perfectly selective entry model of Samuelson (1985). If my findings persist under perfectly selective entry, they should persist under any intermediate models between perfectly selective and nonselective entry, such as the entry models of Marmer, Shneyerov and Xu (2013) and Gentry and Li (2014). In a perfectly selective entry model, an entry rate of $p$ implies that exactly those bidders in the top $p$ portion of the unconditional value distribution enter the auction. As a result, two aspects of the estimation procedure change. First, the value distributions estimated in Section 5 are now interpreted as value distributions conditional on entry at O entry rates. Second, when estimating utility functions using S bids as in Section 5, the relevant value distribution must be adjusted according to S entry rates. For example, if the entry rate in S is $1/\gamma_m$ times that of O, with $\gamma_m \geq 1$, the top $1/\gamma_m$ of the value distribution estimated from O would be the relevant distribution for S. In Table 7 we saw that $\gamma_m > 1$ for subgroups 2 and 3. Only subgroup 1 has $\gamma_1 < 1$, entering S at a higher rate than O. For this subgroup, selective entry actually works in favor of my original findings by reducing the contribution of different entry rates to the revenue dominance of S. In this robustness test, I assume subgroup 1 has nonselective entry for two reasons. First, this biases the test against my findings, making it if anything more robust. Second, as perfectly selective entry left-truncates the value distribution, we cannot infer the truncated portion, which is what we would need to do for $\gamma_m < 1$.

After re-estimating the auction model under selective entry, I repeat the simulations of

Table 8 in Table 9. Comparing Table 9 to Table 8, O revenue in the second row shows the effect of selective versus nonselective entry; when O revenue is simulated at S entry rates, the number of entrants decreases in both Tables, but in Table 9 the values conditional on entry are higher, leading to higher revenue than in Table 8. Moving from the first to second row in Table 9 narrows the S-O revenue gap, so as predicted, selective entry increases the contribution of different entry rates to the observed revenue patterns. Nonetheless, the bulk of the S-O revenue difference is explained by feature (2) as before. Thus, I confirm that this finding is robust to selective entry.

#### Robustness to null hypothesis of risk neutrality for a subset of bidders

In the empirical analysis of auctions, bidders are commonly assumed to be risk-neutral by default. In Figure 6, the null hypothesis of risk neutrality could not be rejected for subgroups 1 and 3, whereas it was rejected for subgroup 2. At a general level, this leads to an interesting question of whether the findings above can be relevant to settings where only a subset of bidders are risk-averse and the remainder are risk-neutral. Specifically for the NMSLO, we may ask whether the findings are robust if subgroups 1 and 3 are in fact risk-neutral. I conduct a test of robustness in this regard by exploiting bootstrap samples in which subgroups 1 and 3 have CRRA parameter estimates close to zero (risk neutrality). I use these bootstrap samples and corresponding model estimates to repeat the simulations of Table 8. Table 10 displays the results for bootstrap samples in which the CRRA parameters were closest to zero. As before, feature (2) is the primary explanation for the observed revenue difference between formats. The qualitative conclusion is the same as in Table 8 and demonstrates that not all bidders need to be risk-averse for these forces to be important.

#### Discussion of the non-binding reserve price assumption

My model assumes that there is no binding reserve price. What would the implications of this assumption be for my analysis if the reserve price $r$ were binding in practice? The first implication would be that the $F_m(\cdot)$ I estimate would in fact be the left-truncated distribution $F_m^*(\cdot|v \geq r)$, where the star superscript indicates the true unconditional distribution. The second implication would be that the estimated non-entry probability $(1 - p_m)$ would in fact be $(1 - p_m^*)F_m^*(r)$, the true non-entry probability multiplied by the constant $F_m^*(r)$. For my post-estimation analysis, the first implication is without loss as long as I do not simulate changes to the reserve price; knowledge of the distribution to the left of the reserve is irrelevant for simulating bids because bidders with those values do not bid. The second implication is without loss as long as I do not simulate new entry equilibria that are different from the observed ones for each auction format; if I know the effective entry rate $(1 - p_m^*)F_m^*(r)$, the separate knowledge of $(1 - p_m^*)$ and $F_m^*(r)$ is irrelevant for simulation. Therefore, the assumption is without loss for my post-estimation analysis, as I neither simulate alternative reserve prices nor simulate new entry equilibria.

## Ignoring uncertainty and risk aversion

The observation that lone bidders bid more than the reserve price is easily explained by their uncertainty about who will bid, but it is a puzzle for "standard" models that ignore uncertainty. Beyond this observation, the data contain additional phenomena that are easily explained by my model but are awkward to rationalize for the standard model. Consider the most common composition of S bidders observed in the estimation sample, an $n = 2$ auction with one bidder from subgroup 1 (recall that there is only one bidder in subgroup 1) and one bidder from subgroup 2. In S auctions with this composition of bidders, the maximum sealed bid observed from the subgroup 1 bidder is about $180,000. In a standard model where bidders know the set of entrants as they bid, we know from auction theory that there is little reason for the other bidder to bid much more than $180,000 in these auctions. However, the subgroup 2 bidder does in fact bid more 7% of the time, for a maximum bid of about $400,000. More generally stated, the data show bidders choosing bids in regions where the only other entrant's bid distribution is very sparse. The phenomenon becomes even more pronounced after controlling for heterogeneity across auction items.

In the standard risk-neutral model, we know from Guerre, Perrigne and Vuong (2000) that the value rationalizing each first-price auction bid is identified as $v_i = b_i + G_{-i}(b_i)/g_{-i}(b_i)$, where the $G_{-i}(\cdot)$ and $g_{-i}(\cdot)$ indicate the bid distribution and density of the other subgroup. Therefore, the standard model would have to rationalize the phenomenon above - bids in regions with very small $g_{-i}(b)$ - with very large $v$. Figure 10, displaying the value distribution (CDF) estimated from the standard model, confirms this prediction.

In my model, the observed phenomenon is not a puzzle because bidders' priors about $n$ place positive probability on there being other entrants, and this uncertainty is amplified by risk aversion. The difference between the standard model's estimate and mine in Figure 10 demonstrates that the consequences of ignoring uncertainty and risk aversion in empirical analysis can be large.

# 8 Conclusion

This article shows that uncertainty about the number of entrants and risk aversion are of first-order empirical importance in auction design. It does this by combining a rare, dual-format auction dataset with a close examination of reduced-form data patterns and nonparametric structural analysis. The structural model is data-motivated and rich, allowing for asymmetry among bidders, risk aversion, and uncertainty about entrants; the latter two features are seldom accounted for in the empirical auction literature. A key empirical finding of the study is that, in first-price sealed-bid auctions, uncertainty combined with risk aversion moderates the damage from a low number of entrants, often bolstering revenue by an amount similar to having one additional entrant. The logic behind this effect is intuitive, it explains the observed data patterns well, and its importance in explaining these patterns is shown to be robust. This effect does not apply to English auctions. An immediate policy implication is to favor first-price sealed-bid auctions over English auctions in low-entry environments.

# References

Athey, S. and P. A. Haile, “Identification of Standard Auction Models,” *Econometrica* 70 (2002), 2107–2140.

Athey, S., J. Levin and E. Seira, “Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions,” *The Quarterly Journal of Economics* 126 (2011), 207–257.

Aycinena, D. and L. Rentschler, “Auctions with Endogenous Participation and an Uncertain Number of Bidders: Experimental Evidence,” *Experimental Economics* (2018), 1–26.

Bajari, P. and A. Hortaçsu, “Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data,” *Journal of Political Economy* 113 (2005), 703–741.

Bhattacharya, V., A. Ordin and J. W. Roberts, “Bidding and Drilling Under Uncertainty: An Empirical Analysis of Contingent Payment Auctions,” Working Paper, 2018.

Bhattacharya, V., J. W. Roberts and A. Sweeting, “Regulating Bidder Participation in Auctions,” *The RAND Journal of Economics* 45 (2014), 675–704.

Bulow, J. and P. Klemperer, “Auctions Versus Negotiations,” *The American Economic Review* 86 (1996), 180–194.

Compiani, G., P. Haile and M. Sant’Anna, “Common Values, Unobserved Heterogeneity, and Endogenous Entry in U.S. Offshore Oil Lease Auctions,” Working Paper 24795, National Bureau of Economic Research, 2018.

Cox, J. C., B. Roberson and V. L. Smith, “Theory and Behavior of Single Object Auctions,” *Research in Experimental Economics* 2 (1982), 1–43.

Cox, J. C., V. L. Smith and J. M. Walker, “Tests of a Heterogeneous Bidders Theory of First Price Auctions,” *Economics Letters* 12 (1983), 207–212.

Dyer, D., J. H. Kagel and D. Levin, "Resolving Uncertainty about the Number of Bidders in Independent Private-Value Auctions: An Experimental Analysis," *The RAND Journal of Economics* 20 (1989), 268–279.

Fibich, G. and N. Gavish, "Numerical Simulations of Asymmetric First-Price Auctions," *Games and Economic Behavior* 73 (2011), 479–495.

Gentry, M. and T. Li, "Identification in Auctions with Selective Entry," *Econometrica* 82 (2014), 315–344.

Gentry, M. L., T. Li and J. Lu, "Identification and Inference in First-Price Auctions with Risk-Averse Bidders and Selective Entry," Working Paper, Available at SSRN 2587812, 2017.

George, G., J. Wiklund and S. A. Zahra, "Ownership and the Internationalization of Small Firms," *Journal of Management* 31 (2005), 210–233.

Guerre, E., I. Perrigne and Q. Vuong, "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica* 68 (2000), 525–574.

———, "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica* 77 (2009), 1193–1227.

Haile, P., K. Hendricks and R. Porter, "Recent U.S. Offshore Oil and Gas Lease Bidding: A Progress Report," *International Journal of Industrial Organization* 28 (2010), 390–396.

Haile, P. A., H. Hong and M. Shum, "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," Working Paper 10105, National Bureau of Economic Research, 2003.

Harstad, R. M., J. H. Kagel and D. Levin, "Equilibrium Bid Functions for Auctions with an Uncertain Number of Bidders," *Economics Letters* 33 (1990), 35–40.

Hiebl, M. R., "Risk Aversion in Family Firms: What Do We Really Know?," *The Journal of Risk Finance* (2013).

Hodgson, C., "Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry," Working Paper, 2018.

Holt, C. A. and S. K. Laury, "Risk Aversion and Incentive Effects," *The American Economic Review* 92 (2002), 1644–1655.

Isaac, M., S. Pevnitskaya and K. S. Schnier, "Individual Behavior and Bidding Heterogeneity in Sealed Bid Auctions where the Number of Bidders is Unknown," *Economic Inquiry* 50 (2012), 516–533.

Komarova, T., "Extremum Sieve Estimation in k-out-of-n Systems," *Communications in Statistics - Theory and Methods* 46 (2017), 4915–4931.

Kong, Y., "Sequential Auctions with Synergy and Affiliation Across Auctions," Working Paper, 2017.

Krasnokutskaya, E. and K. Seim, "Bid Preference Programs and Participation in Highway Procurement Auctions," *The American Economic Review* 101 (2011), 2653–2686.

Krishna, V., *Auction Theory*, 2nd edition (San Diego: Academic Press, 2010).

Lebrun, B., "First Price Auctions in the Asymmetric N Bidder Case," *International Economic Review* 40 (1999), 125–142.

Levin, D. and J. L. Smith, "Equilibrium in Auctions with Entry," *The American Economic Review* 84 (1994), 585–599.

Li, T., J. Lu and L. Zhao, "Auctions with Selective Entry and Risk Averse Bidders: Theory and Evidence," *The RAND Journal of Economics* 46 (2015), 524–545.

Li, T. and X. Zheng, "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions," *The Review of Economic Studies* 76 (2009), 1397–1429.

Lorentz, G. G., *Bernstein Polynomials*, 2nd edition (New York: AMS Chelsea Publishing, 1986).

Lu, J. and I. Perrigne, "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data," *Journal of Applied Econometrics* 23 (2008), 871–896.

Marmer, V., A. Shneyerov and P. Xu, "What Model for Entry in First-Price Auctions? A Nonparametric Approach," *Journal of Econometrics* 176 (2013), 46–58.

Maskin, E. and J. Riley, "Asymmetric Auctions," *The Review of Economic Studies* 67 (2000a), 413–438.

———, "Equilibrium in Sealed High Bid Auctions," *The Review of Economic Studies* 67 (2000b), 439–454.

———, "Uniqueness of Equilibrium in Sealed High-Bid Auctions," *Games and Economic Behavior* 45 (2003), 395–409.

Matthews, S., "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View," *Econometrica* 55 (1987), 633–646.

McAfee, R. P. and J. McMillan, "Auctions with a Stochastic Number of Bidders," *Journal of Economic Theory* 43 (1987), 1–19.

Mead, W. J., "Natural Resource Disposal Policy - Oral Auction Versus Sealed Bids," *Natural Resources Journal* 7 (1967), 194–224.

Menezes, F. M. and P. K. Monteiro, "Auctions with Endogenous Participation," *Review of Economic Design* 5 (Mar 2000), 71–89.

Milgrom, P. R. and R. J. Weber, “A Theory of Auctions and Competitive Bidding,” *Econometrica* 50 (1982), 1089–1122.

Reny, P. J. and S. Zamir, “On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-Price Auctions,” *Econometrica* 72 (2004), 1105–1125.

Riley, J. G. and W. F. Samuelson, “Optimal Auctions,” *The American Economic Review* 71 (1981), 381–392.

Samuelson, W. F., “Competitive Bidding with Entry Costs,” *Economics Letters* 17 (1985), 53–57.

# Figures and Tables

Figure 1: Histogram of number of bids received in S auctions, Permian Basin 2005-2014

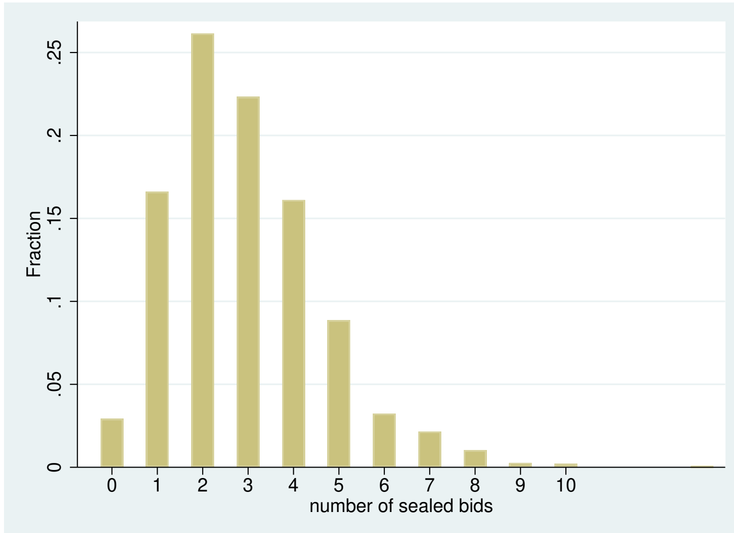
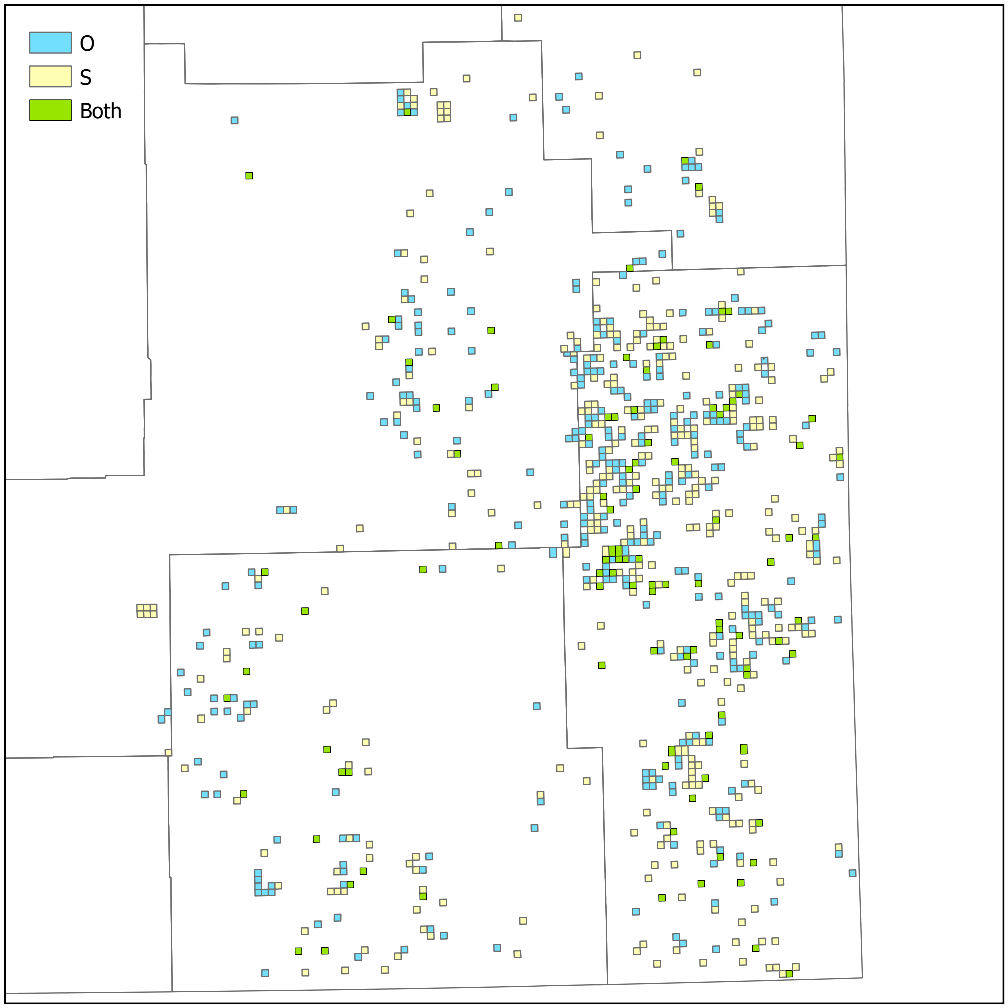



Table 1: Bidder pool comparison, Permian Basin 2005-2014

| Fraction by year | All | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| O wins from S bidders | 0.98 | 0.95 | 0.97 | 0.92 | 0.95 | 0.95 | 0.98 | 0.98 | 0.92 | 0.97 | 0.95 |
| S bids from O winners | 0.97 | 0.90 | 0.89 | 0.81 | 0.86 | 0.86 | 0.91 | 0.96 | 0.89 | 0.86 | 0.94 |

| Fraction by county | All | Chaves | Eddy | Upper Lea | Lower Lea |
|---|---|---|---|---|---|
| O wins from S bidders | 0.98 | 0.96 | 0.95 | 0.98 | 0.97 |
| S bids from O winners | 0.97 | 0.84 | 0.86 | 0.95 | 0.89 |

Figure 2: Map of sections by auction format, estimation sample




Each colored square is a section. The larger polygons are counties.

Table 2: Summary statistics by auction format, estimation sample

| | S | O |
|---|---|---|
| Number of auctions | 674 | 556 |
| Average across auctions: | | |
| Revenue ($) | 128,822 | 102,481 |
| Unsold | 3.0% | 3.8% |
| Sold at reserve price | 0% | 12% |
| Lease prefix VB | 45% | 36% |
| Annual rental per acre ($) | 0.73 | 0.72 |
| Production, 1970-auction date (BOE) | 72,807 | 125,839 |
| Production, auction date-2014 (BOE) | 126,370 | 102,919 |
| Well spudded before in same section | 48% | 43% |
| Natural gas 1 month futures in auction month ($) | 5.8 | 6.0 |
| WTI oil prices in auction month ($) | 79.6 | 81.3 |
| Average price/acre in BLM sale, same quarter ($) | 878 | 1,115 |
| Average price/acre in last month's auction ($) | 290 | 338 |

Table 3: Preliminary assessment of the heatmap index

| | (1) | (2) | (3) |
|---|---|---|---|
| | ln(production) | ln(production) | ln(production) |
| heatmap index | 2.652*** | | 2.422*** |
| | (0.271) | | (0.268) |
| auction format S | | 0.205 | 0.142 |
| | | (0.207) | (0.194) |
| lease prefix VB | | 0.570** | -0.078 |
| | | (0.236) | (0.233) |
| ln(production) 1970-auction date | | 0.038 | 0.030 |
| | | (0.036) | (0.034) |
| section drilled before | | 0.406* | 0.364* |
| | | (0.223) | (0.208) |
| ln(gas futures) | | 0.921 | 0.327 |
| | | (0.809) | (0.774) |
| ln(WTI oil price) | | 0.105 | 0.662 |
| | | (0.726) | (0.684) |
| same quarter BLM price/acre | | -0.380 | -0.193 |
| | | (0.263) | (0.246) |
| last month price/acre | | -0.397* | -0.288 |
| | | (0.209) | (0.195) |
| Constant | Y | Y | Y |
| Year FE | N | Y | Y |
| Calendar-month FE | N | Y | Y |
| Observations | 1189 | 1189 | 1189 |
| $R^2$ | 0.141 | 0.114 | 0.213 |
| Adjusted $R^2$ | 0.141 | 0.092 | 0.193 |

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

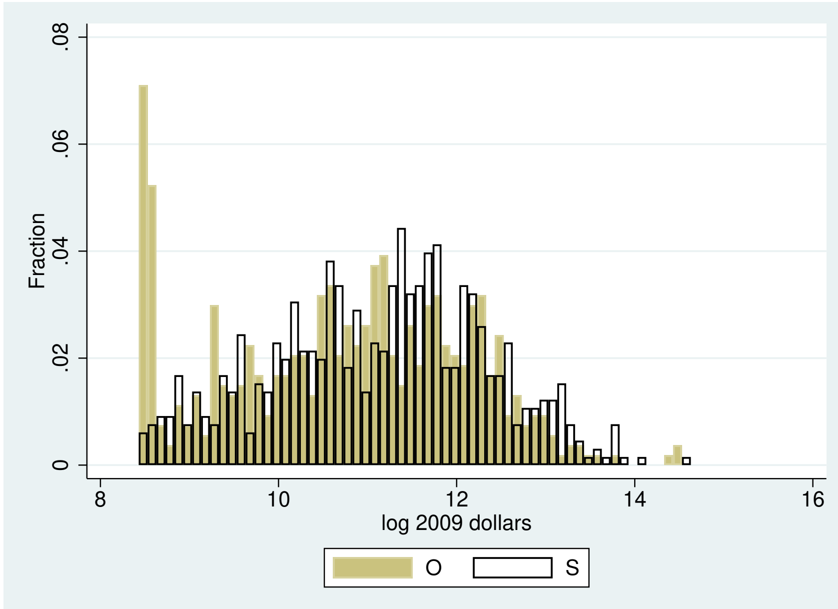
Table 4: Auction format and auction revenue, estimation sample

| | (1) auction revenue | (2) auction revenue |
|---|---|---|
| auction format S | 0.321*** | 0.305*** |
| | (0.073) | (0.068) |
| lease prefix VB | 0.404*** | 0.174** |
| | (0.077) | (0.072) |
| ln(production) 1970-auction date | 0.003 | 0.002 |
| | (0.011) | (0.010) |
| ln(production) auction date-2014 | 0.069*** | 0.025** |
| | (0.010) | (0.011) |
| section drilled before | 0.093 | 0.095 |
| | (0.074) | (0.069) |
| ln(gas futures) | -0.112 | -0.305 |
| | (0.226) | (0.205) |
| ln(WTI oil price) | 0.497** | 0.721*** |
| | (0.208) | (0.196) |
| same quarter BLM price/acre | 0.185** | 0.242*** |
| | (0.074) | (0.067) |
| last month price/acre | 0.029 | 0.055 |
| | (0.093) | (0.085) |
| heatmap index | | 0.953*** |
| | | (0.070) |
| Constant | Y | Y |
| Year FE | Y | Y |
| Calendar-month FE | Y | Y |
| Observations | 1189 | 1189 |
| $R^2$ | 0.177 | 0.287 |
| Adjusted $R^2$ | 0.156 | 0.268 |

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Figure 3: Histogram of ln(price), estimation sample




Prices are in 2009 dollars, deflated by the GDP implicit price deflator.

Table 5: Statistics by bidder subgroup, estimation sample

| | top | core | fringe |
|---|---|---|---|
| count of names | 1 | 32 | 83 |
| average number of entrants in a given S auction | 0.84 | 1.63 | 0.41 |
| average S entry rate | 0.84 | 0.05 | 0.005 |
| average S win rate conditional on bidding | 0.40 | 0.38 | 0.32 |
| subgroup's share of S wins | 0.35 | 0.54 | 0.11 |
| subgroup's share of O wins | 0.30 | 0.58 | 0.13 |
| p-value for test of null hypothesis that subgroup's S share = O share | 0.05 | 0.28 | 0.25 |

Figure 4: Observed distribution of $z'\beta$ in S and O auctions

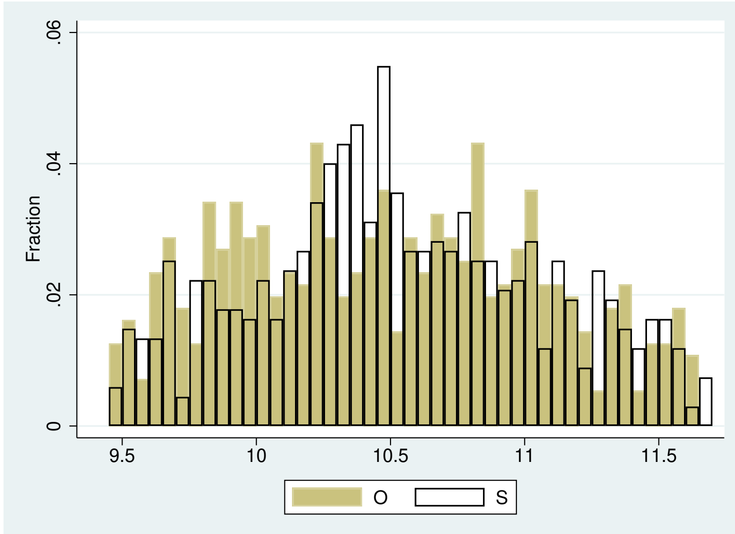



Table 6: Number of potential bidders by subgroup

| | Specification 1 | Specification 2 | | |
|---|---|---|---|---|
| | | average across years and sub-areas | min | max |
| $N_1$ | 1 | 1 | 1 | 1 |
| $N_2$ | 15 | 13 | 4 | 21 |
| $N_3$ | 13 | 10 | 2 | 26 |

Figure 5: Density of estimated value distributions at median $z'\beta$

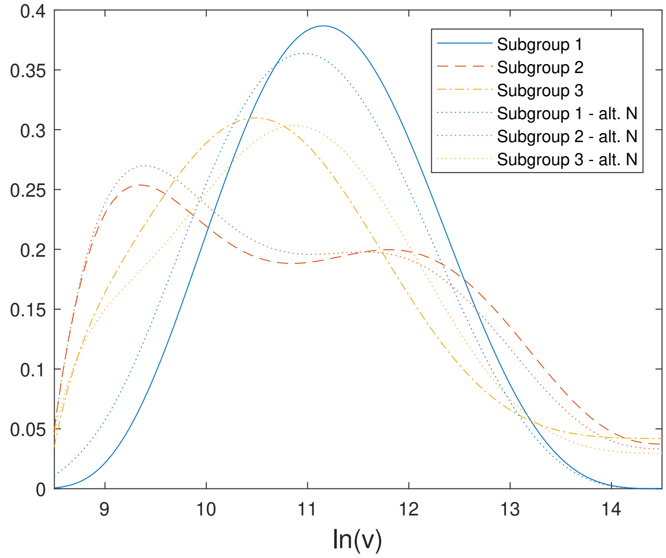



Table 7: Estimated O entry parameter $\gamma_m$ and risk aversion levels

| | $N$ - specification 1 | | $N$ - specification 2 | |
|---|---|---|---|---|
| Subgroup | $\gamma_m$ | CRRA risk | $\gamma_m$ | CRRA risk |
| 1 | 0.87 | 0.31 | 0.96 | 0.24 |
| 2 | 1.14 | 0.77 | 1.32 | 0.78 |
| 3 | 1.28 | 0.60 | 1.15 | 0.39 |

Figure 6: 95% bootstrap intervals for entry parameter $\gamma_m$ (left) and CRRA parameter (right)

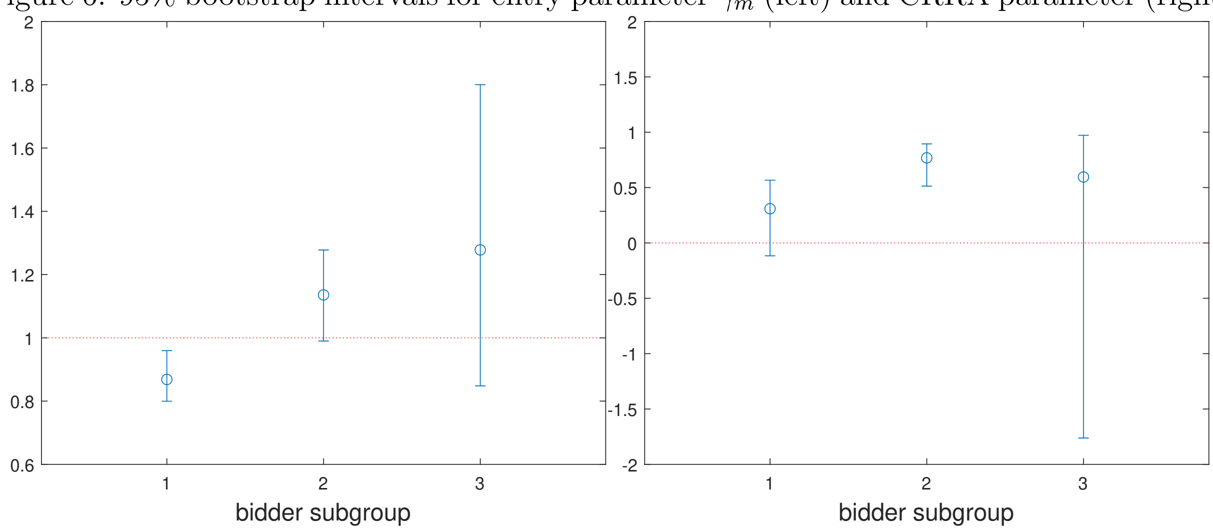


Figure 7: Observed vs. simulated number of S entrants

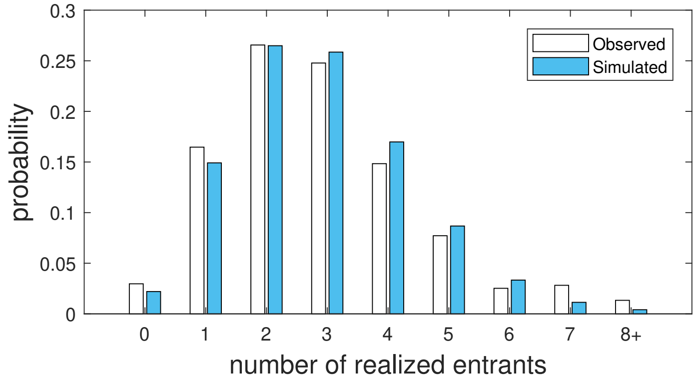



Figure 8: Observed vs. simulated distribution of O prices (left) and S bids (right)

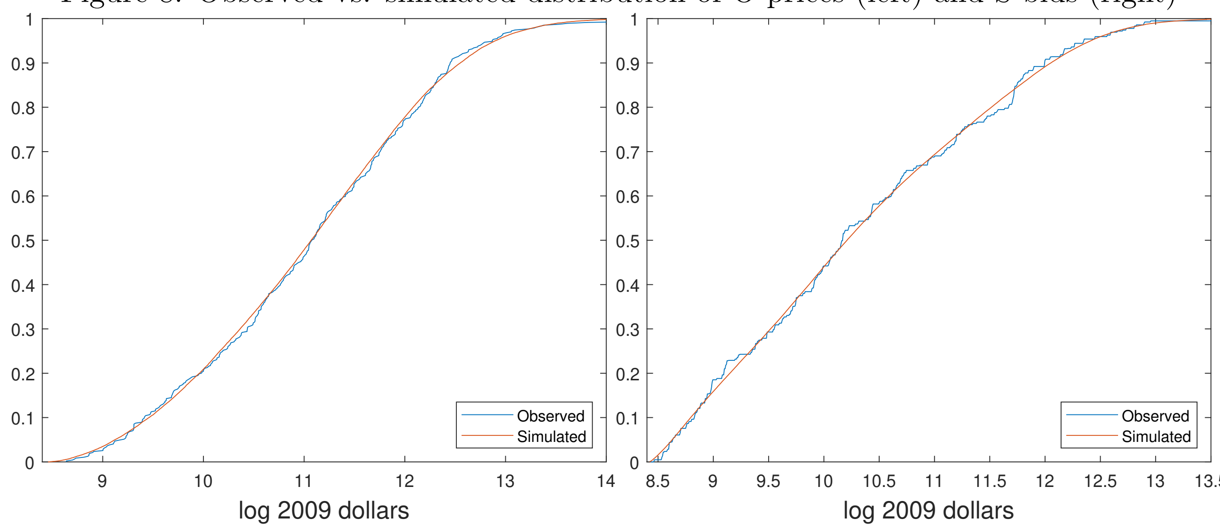


Table 8: Simulated auction revenue at median $z'\beta$

| | $N$ - Specification 1 | | $N$ - Specification 2 | |
|---|---|---|---|---|
| | O | S | O | S |
| (1),(2),(3) | 87,736 | 118,718 | 93,707 | 121,332 |
| (2),(3) | 81,473 | 118,718 | 75,251 | 121,332 |
| (3) | 81,473 | 80,848 | 75,251 | 73,370 |

Definitions:

(1): Different entry rates into the two auction formats

(2): Bidders' uncertainty about set of entrants combined with risk aversion

(3): Asymmetry between bidders

Figure 9: Simulated revenue disaggregated by realized $n$

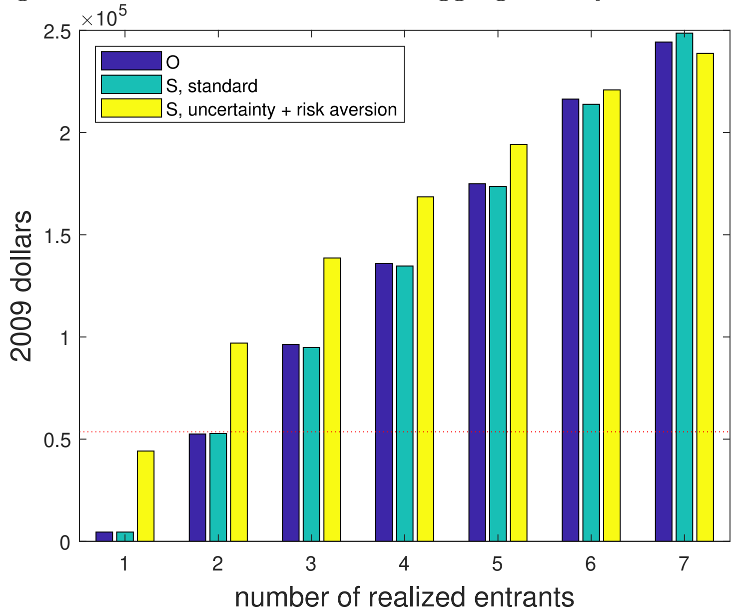



Table 9: Simulated auction revenue from estimating model based on selective entry

| | O | S |
|---|---|---|
| (1),(2),(3) | 88,021 | 116,692 |
| (2),(3) | 90,373 | 116,692 |
| (3) | 90,373 | 90,666 |

Table 10: Simulated revenue based on bootstrap samples with risk-neutral subgroups 1 and 3

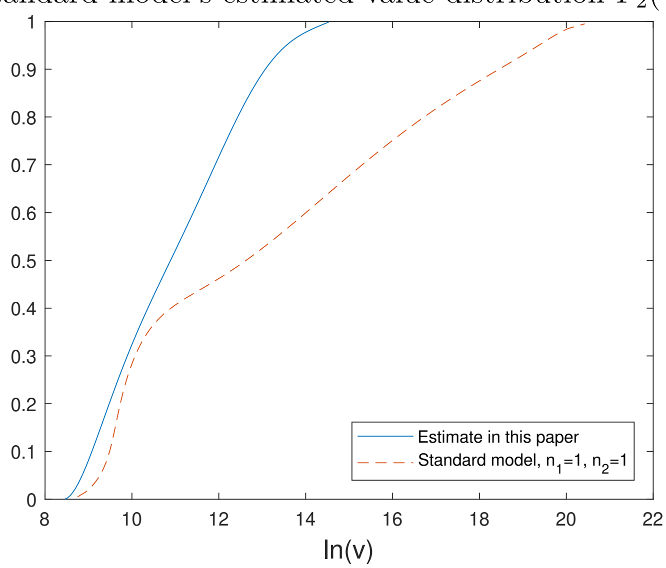
| | Bootstrap sample 1 | | Bootstrap sample 2 | |
|---|---|---|---|---|
| | O | S | O | S |
| (1),(2),(3) | 83,590 | 103,520 | 78,911 | 88,573 |
| (2),(3) | 81,770 | 103,520 | 73,732 | 88,573 |
| (3) | 81,770 | 80,670 | 73,732 | 74,681 |

CRRA parameter estimates for subgroups 1, 2, 3:

Bootstrap sample 1: 0, 0.68, 0.02

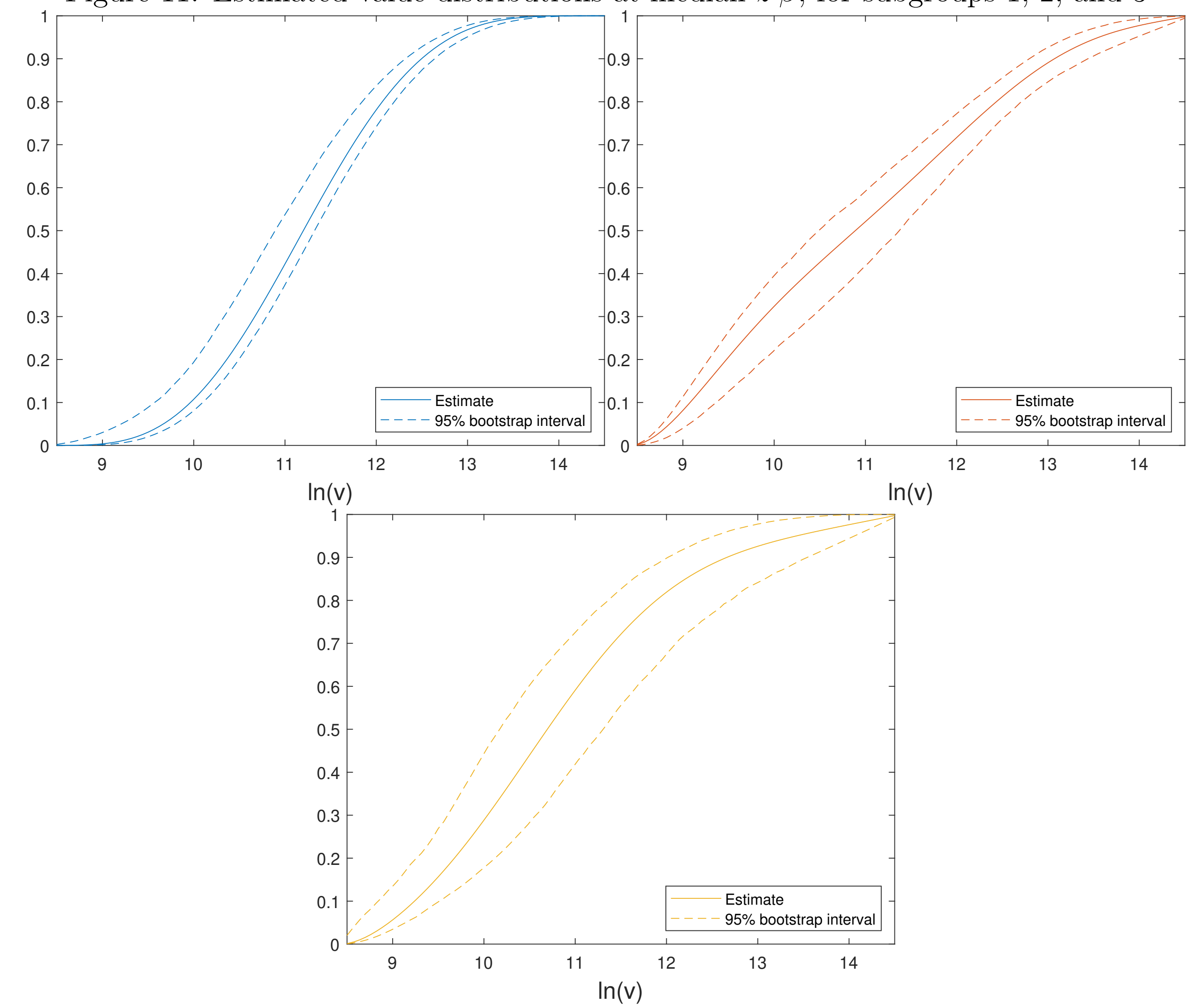
Bootstrap sample 2: 0, 0.70, 0.02

Figure 10: Standard model's estimated value distribution $\hat{F}_2(\cdot)$, median $z'\beta$



# Appendix

Figure 11: Estimated value distributions at median $z'\beta$, for subgroups 1, 2, and 3




**Fibich and Gavish (2011) algorithm modified for type-symmetric bidding** The asymmetric S bids in row (3) of Tables 8, 9, and 10 are simulated using the boundary-value method as explained in Section 5 of Fibich and Gavish (2011). To reflect the type-symmetric structure of my model, I modify the fixed-point iterations presented in equation (27a) of that article; instead of iterating through $n$ asymmetric bidders as presented there, I iterate through $n$ asymmetric subgroups of bidders with $N_j$ bidders from each subgroup $j \in \{1, ..., n\}$, updating bid functions at the subgroup-level rather than at the individual level. The modified fixed-point iteration equations are given below. Let $N \equiv \sum_{j=1}^{n} N_j$.

$$\frac{dv_i^{(k+1)}}{dv_n} = \frac{F_i(v_i^{(k)})f_n(v_n)}{f_i(v_i^{(k)})F_n(v_n)} \frac{A_i(v_i^{(k+1)} - b^{(k)}) - (2-N)B_i}{D_i}, \; i = 1, ..., n-1,$$

$$\frac{db^{(k+1)}}{dv_n} = \frac{f_n(v_n)}{F_n(v_n)} \frac{(N-1)B_n(v_n - b^{(k+1)})}{(v_n - b^{(k)})D_n},$$

where

$B_i \equiv (v_i^{(k)} - b^{(k)})^{N_i - 1} \prod_{j<i} (v_j^{(k+1)} - b^{(k)})^{N_j} \prod_{j>i} (v_j^{(k)} - b^{(k)})^{N_j},$

$D_i \equiv N_i B_i + (v_i^{(k)} - b^{(k)}) \left[ \sum_{j<i} \frac{N_j B_i}{(v_j^{(k+1)} - b^{(k)})} + \sum_{j>i, j<n} \frac{N_j B_i}{(v_j^{(k)} - b^{(k)})} + \frac{(1 - \sum_{j<n} N_j) B_i}{(v_n - b^{(k)})} \right],$

$A_i \equiv \frac{(N_i - 1)B_i}{(v_i^{(k)} - b^{(k)})} + \sum_{j<i} \frac{N_j B_i}{(v_j^{(k+1)} - b^{(k)})} + \sum_{j>i} \frac{N_j B_i}{(v_j^{(k)} - b^{(k)})},$

and

$B_n \equiv (v_n - b^{(k)})^{N_n} \prod_{j<n} (v_j^{(k+1)} - b^{(k)})^{N_j},$

$D_n \equiv \sum_{j<n} \frac{N_j B_n}{(v_j^{(k+1)} - b^{(k)})} + \frac{(1 - \sum_{j<n} N_j) B_n}{(v_n - b^{(k)})}.$