Mechanism Choice in Scoring Auctions

Pasha Andreyanov, UCLA*
Job Market Paper
link: latest version

Abstract

A first-score auction requires weighing the price-bid against non-price characteristics of the firm. In this paper, I theoretically and empirically study the welfare implications of switching between the two leading designs of the scoring rule: linear ("weighted bid") and log-linear ("adjusted bid"), when the designer's preferences for quality and money are unknown. Motivated by the empirical application, I formulate a new model of scoring auctions, with two key elements: exogenous quality and a reserve price, and characterize the equilibrium for a rich set of scoring rules. The data is drawn from the Russian public procurement sector in which the linear scoring rule was applied from 2011 to 2013. I estimate the underlying distribution of firms' types nonparametrically and simulate the equilibria for both scoring rules with different weights. The empirical results show that for any log-linear scoring rule, there exists a linear one, yielding a higher expected quality and rebate. Hence, at least with risk-neutral preferences, the linear design is superior to the log-linear.

Keywords: procurement, scoring auction, exogenous quality, reserve price

*I am grateful to John Asker, Denis Chetverikov, Martin Hackmann, Elisabeth Honka, Hugo Hopenhayn, Moritz Meyer-Ter-Vehn, Tomasz Sadzik, Robert Zeithammer, the participants of the Theory and IO proseminars at UCLA, and also my brother, Boris Andreianov, for useful comments and suggestions.

1 Introduction

A first-score auction is often used in procurement to determine the best supplier among a variety of candidates. This mechanism weighs the price bid together with the non-price attributes of the firm using a publicly known scoring rule, and the participant with the highest score wins the contract at the price equal to its bid.

There are multiple shapes of the scoring rule, see Dini et al. (2006) and Molenaar and Yakowenko (2007) for an overview, but two of them are especially widespread. The linear ("weighted bid") rule is used, for example, in Delaware, Idaho, and Oregon, while the log-linear ("adjusted bid") rule is used in Alaska, Colorado, and Florida, for various public works contracts. The latter, also known as "price-quality ratio", is used extensively in Japan. It is unclear how a scoring rule is chosen in any given case, and the existing literature offers limited guidance on how to rank them. The theoretical literature was focused primarily on the optimal mechanism design and the differences between various auction formats, such as first-score and second-score, see Che (1993), Asker and Cantillon (2008) and Asker and Cantillon (2010), but not on the comparisons of the existing shapes of the scoring rule.

This paper is the first to compare the two leading designs of the scoring rule — linear and log-linear — in a first-score environment, and to rank them in terms of welfare using a structural approach. To achieve this goal, I formulate a novel model of scoring auctions, which allows for both linear and log-linear cases, and derive a tractable equilibrium characterization. I employ the equilibrium structure to analyze a novel dataset on bidding and quality assessments in Russian procurement auctions, and propose a computationally feasible procedure for nonparametric estimation and simulation of counterfactuals. Finally using the simulated equilibria, I compare the two designs in terms of welfare.

I leverage a novel dataset on bidding behavior and relevant quality measurements in scoring auctions using a public archive of Russian procurement auctions which took place from 2011 to 2013. Through this period, 46,387 contracts were assigned using the scoring auction format, across multiple industries and regions of the country. I select 3,228 auctions, based on criteria, such as uniformity of the scoring procedure and low market concentration. The data are split into ten parts based on the type of economic activity and the weight in the scoring formula.

The data exhibit two distinct features. First, all auctions in the sample use quality measurements such as a firm's experience, qualification of firm's personnel, or previous successful contracts. These characteristics cannot be exchanged for cost reduction in the short run, which I will exploit in the model by treating them as exogenous (nonstrategic). Second, the reserve price is chosen by a significant portion of bidders, a phenomenon that is often referred to as *bunching*.

Clearly, the model should reflect the patterns observed in the data. However, none of the existing models feature a reserve price, and the quality is always assumed to be a strategic choice variable. While the endogenous quality could be, to some degree, considered as a generalization of exogenous quality, the addition of the reserve price creates a distinct theoretical challenge due to its interaction with the scoring rule. Indeed, from the viewpoint of the score, the constraint is variable, so it might be binding for some firms, thus producing the bunching.

I build a novel model of a scoring auction in which quality is exogenous and the reserve price is explicitly set. The key to solving the model is to think of the firm's strategy as an unconstrained choice of the score, which is later censored at the reserve price constraint. Due to the feedback loop between the strategic shading and the expectations about the bunch, an additional level of endogeneity appears that was not featured in classic auction models. Despite its complexity, the model remains tractable with a wide set of scoring rules, which I call affine, nesting both linear and log-linear rules.

A construction of the equilibrium and a proof of its existence and uniqueness are the main theoretical results in this paper. At the core of the equilibrium is a one-dimensional strategy that captures the unconstrained choice of the score, pinned down by a simple (ordinary) differential equation. Moreover, this strategy is similar to the bidding strategy in a classic first-price auction, allowing the traditional nonparametric estimators to be applied.

To estimate the joint distribution of firm's cost and quality pairs from the observed bid and quality pairs, I use an algorithm similar to the two-step smoothing procedure in Guerre et al. (2000). In the first step, I use kernel smoothing and the optimality conditions derived from the model to estimate the amount of strategic shading from the observed scores.

In the second step, I use kernel smoothing to recover the joint distribution of bids and quality, while ignoring the bids equal to the reserve price. I use the boundary correction advocated in Hickman and Hubbard (2015) to estimate the density at the boundary, that is, the bids just below the reserve price. Finally, by reversing the shading, I recover the joint distribution of costs and quality wherever it is fully identified.

The last piece of the puzzle is the part of the distribution of costs and quality that maps into the reserve price, which I could not recover from the second step. Since there are multiple distributions that are consistent with the data, I pick a simple one. Namely, I extrapolate the missing part of the distribution from the boundary, that is, from the bids that were placed just below the reserve price. The final estimator therefore has a semiparametric flavor.

To relax some of the restrictions that the model puts on the data, I investigate the implications of auction-level heterogeneity. By assuming a multiplicative error term, I show that it has no impact on the estimation as long as it affects the reserve price in the same way as it affects the cost. This means that instead of the actual bids, I can simply use the normalized bids on a scale between 0 and 100, where 100 stands for the reserve price. For this to be empirically justified, I need evidence that the reserve price serves as a signal of the scale of the contract. I find such evidence in the regulation behind the data.

Finally, I am able to simulate the equilibria in both the default and the counterfactual scoring designs and, additionally, pick an arbitrary weight in the scoring formula. However, two problems prevent us from making decision on the mechanism choice right away. First, it is not absolutely clear which log-linear scoring rule to pick, as they vary by the weight in the scoring formula. Second, the welfare of the scoring auction is measured, at the very least, in two dimensions: expected quality and expected rebate, and so the two scoring rules might not necessarily be successfully ranked.

To overcome these problems, I propose the following approach. For each of the two scoring designs, I construct a frontier in the space of expected quality and rebate, spanned by the weight in the scoring formula, see Figure 1. The weight is a natural parameter to vary as it controls the trade-off between quality and rebate. By fixing a target level of expected quality, it is then possible to rank the two mechanisms in the remaining dimension.

My main empirical finding is that, based on the available data, for any log-linear scoring rule, there exists a linear one that is better in terms of rebate while having the same quality. In other words, the linear frontier is above the log-linear frontier, as seen on Figure 1. This it is true for every part of my data, see Figure 21 and Figure 22. The monetary loss associated with a switch to a log-linear rule with the same quality varies between 0.4% and 4%.

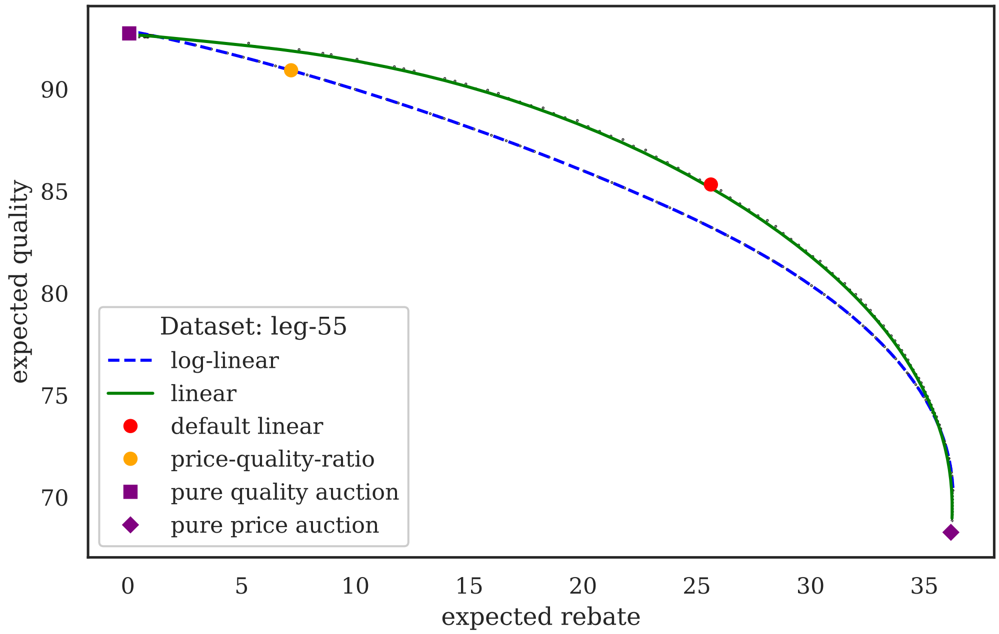

Figure 1: Welfare frontiers for the linear and log-linear families

This ranking may seem surprising, as I am comparing two equally plausible shapes of the scoring rule. However, there is a strong reason for such results. If the firms were to bid truthfully in a first-score auction, the mechanism would always pick the score-efficient firm. As a result, the (truthful) linear frontier would be the frontier of first-best mechanisms in the space of expected quality and rebate, and the (truthful) log-linear frontier would be below. The actual frontiers, of course, look slightly different due to the distortions associated with strategic shading and bunching. First, shading is pushing the truthful frontiers to the left, due to the informational rents paid to the firms. Second, bunching is pushing them towards the origin due to the inefficient screening. My empirical findings demonstrate that these distortions are not sufficient to change the ranking, at least with the available distributions.

Of course, these results implicitly rely on the risk-neutrality of the designer as I am using expected quality and expected rebate as measures of auction success. Mechanism design literature has long used expected revenue for ranking various classic auction mechanisms and I am simply following this tradition. It is possible, however, that when higher moments are considered, the ranking will change.

1.1 Related literature

The theoretical literature on scoring auctions is sparse compared to that of the classic auctions. Apart from Asker and Cantillon (2008) and Hanazano et al. (2016), only a few attempts have been made in — Branco (1997), Dastidar (2014) and Nishimura (2015) — to generalize the original Che (1993) framework. None of them, however, deals with the reserve price.

The closest model to ours can be found in Hanazano et al. (2016). The authors show how to solve for the symmetric equilibrium with the most general scoring rule. While losing much of the tractability of the earlier models, the only new scoring rule (used in practice) that this generality buys is, again, the log-linear. As in Che (1993) and Asker and Cantillon (2008), the quality in this paper is endogenous, and there is no reserve price.

We stress that our model is not a special case of the model of Hanazano et al. (2016); rather, we study different aspects of the scoring auction. While their focus is on nonlinearities in the scoring rule, our focus is on the interaction between the scoring rule and the reserve price. Also, our equilibrium characterization is more tractable due to the exogenous quality and the affine structure of the scoring rule.

Similar in spirit to our empirical finding of the superiority of the linear design are the theoretical results on the optimality of quasi-linear scoring rules in Che (1993) and Nishimura (2015). These papers, however, deal with a special environment, in which the distribution of firms' costs and quality is singular and they are already sorted in terms of productivity. The role of the auction mechanism then narrows to picking the efficient one. In our environment, the situation is more complex, as sorting is endogenous due to the full support of firm's cost and quality and, moreover, imperfect due to the bunching at the reserve price. Consequently, neither the linear nor the log-linear scoring rules can be generically considered as optimal, even from the perspective of the score.

The empirical literature comprises of studies of delegation and favoritism in scoring auctions: Dastidar and Mukherjee (2014), Adani et al. (2016), Huang (2016) and several reduced form studies of the choice between scoring and price only designs: Lewis and Bajari (2011), Koning and van de Meerendonk (2014) and Albano et al. (2008). Only a few papers contribute to the structural modeling of a scoring auction: Hanazano et al. (2016), Takahashi (2014) and Nakabayashi (2013), however, the questions investigated in these papers do not directly overlap with ours.

Finally, our model of an affine scoring rule could be used to model bid preferences, such as the ones studied in Marion (2007) and Krasnokutskaya and Seim (2011). Indeed, the intercept and the slope of bid in the scoring formula could be thought of as additive and multiplicative preferences. A firm enjoying the preferential treatment can be interpreted as a high-quality firm in a scoring auction. However, since our model is symmetric, we require that the firms treat each other's bid preferences as random, which might be rejected by the data if the pool of participants and the preferences are made public prior to the bidding stage.

The rest of the paper is organized as follows. In Section 2, I discuss the institutional background and the features of our data. In Section 3 and Section 4, I explain the theoretical and empirical framework for my analysis. In Section 5, we discuss potential heterogeneity problems and how we address them. In Section 6, I present the empirical findings, and I conclude in Section 7.

2 Data

We access an electronic archive of all public procurement contracts issued by the Russian federal government and municipalities between January 2011 and December 2013. Out of the 46,387 contracts that were assigned within the scoring auctions format, we select 3,228 contracts based on such criteria as similarity of contracts, uniformity of the scoring procedure and low market concentration. These auctions comprise a total award value of 219.6 million dollars[1].

The scoring auction is the most complex and regulated format among all used in Russian procurement. The contractee, which is typically not a private firm but a public body, that wishes to start a scoring auction has to follow the detailed instructions written by the Ministry of Economic Development. The auction process is supervised by a special commission and each step is made public. The whole process can take more than a month, depending on the size and complexity of the contract, see Table 1.

Step	Timing
1. Formation of the supervising commission and development of the notification and relevant tender documentation.	Prior to the publication.
2. Publication of the upcoming auction notification and relevant tender documentation. Starting bid collecting period.	-
3. Ending bid collecting period, opening quotes and publication of opening protocols.	1-4 weeks since the start of bid collection period.
4. Evaluation of bids and determining the winner. Publication of evaluation protocols and winner announcement.	Up to 10 days since the quotes we opened.
5. Signing the contract	Up to 10 days since the publication of auction protocols.

Table 1: Scoring auction timing and regulation

The archive consists of three types of files: notifications, protocols and tender documentation. The first two are standardized tables stored in the .xml format. The notification file contains an announcement of the upcoming auction together with a brief description of the contract, the industry code and the reserve price. The official site allows anybody to search the upcoming auctions by the information contained in the notification, thus reducing participation costs. The protocol file is created after the auction and contains all submitted bids together with the announcement of the winner. Finally, the documentation is a detailed description of the contract and the quality assessment, compiled by the contractee, typically in a form of one or several .doc files.

To link each contract with a certain type of economic activity, we use a 7-digit industry code called OKPD, extracted from the notifications. This code is an archaic classification of industries, currently replaced by a newer version OKPD2. We focus on five groups of economic activity: education, scientific research, legal services, technical services, and security.

[1]using the approximate exchange rate of 30 roubles per dollar in 2012

dataset name	industry name	OKPD code	price weight	total auctions	total bids	total awards in mln $	total reserve in mln $
Edu-80 Edu-55	Education	80*****	0.80 0.55	89 272	225 712	4.6 13.3	5.8 16.1
Sci-80 Sci-55	Scientific Research	73*****	0.80 0.55	162 298	487 854	11.7 30.9	17.8 38.0
Leg-80 Leg-55	Legal Services	75*****	0.80 0.55	929 55	3048 159	5.2 5.6	10.9 7.4
Tec-80 Tec-55	Technical Design	71*****	0.80 0.55	936 47	2917 129	99.5 7.3	158.8 9.3
Sec-80 Sec-55	Security	84*****	0.80 0.55	409 31	1181 94	33.4 8.1	42.0 9.6
Total				3228	9806	219.6	315.6

Table 2: Auction aggregate statistics

dataset name	average bidder quality score	average bidder price score	average bidder score	average winning rebate %	average winning quality %	average bidders
Edu-80 Edu-55	66.2 72.9	23.6 17.0	32.1 42.2	30.7 19.6	79.7 87.9	2.5 2.6
Sci-80 Sci-55	61.5 64.6	25.0 18.8	32.3 39.4	32.8 16.7	73.8 85.4	3.0 2.9
Leg-80 Leg-55	74.7 59.0	41.1 23.0	47.8 39.2	52.3 17.3	82.8 94.5	3.3 2.9
Tec-80 Tec-55	60.2 66.2	28.2 21.0	34.6 41.4	36.3 20.8	74.0 87.7	3.1 2.7
Sec-80 Sec-55	74.0 56.2	15.7 15.0	27.4 33.6	19.9 13.4	84.7 76.6	2.9 3.0

Table 3: Auction average statistics

The first group (Edu) consists of contracts for citizen education programs. These types of services are often purchased by the government when a factory is shut down or experiences a sharp decrease in labor demand. The type of quality that is required from the firm is measured in years of operation on the market and the number of people that went through the education program. The second group (Sci) consists of contracts for sociological, statistical, economic and other scientific research, with quality being typically some sort of experience, government accreditation, or the qualification (such as a doctoral degree) of personnel. The third group (Leg) consists of various legal services, for the most part related to the mandatory yearly inspection (audit) of firms. Here, quality is measured by the number of audits in the past, positive reviews from the clients, and also the number of personnel with a proper government accreditation. The fourth group (Tec) mostly consists of contracts for development of project documentation and estimates, with quality measurements very similar to (Leg). The fifth group (Sec) has contracts for maintaining security in various civil buildings and facilities. In this group, the quality is measured by the experience of the personnel and the availability of special equipment.

In each of our auctions, a linear scoring formula was used, which can be described as:

$$score = price\ weight \cdot price\ score + (1 - price\ weight\ ) \cdot quality\ score \tag{1}$$

$$price\ score = 100 \cdot (reserve\ price - price\ bid\ )/reserve\ price \tag{2}$$

where the price weight parameter is either 0.8 or 0.55. The quality score is assigned to each firm independently and strictly according to the evaluation criteria described in the contract documentation. These criteria vary by contract, but the final quality score is always between 0 and 100. Overall, we have 10 different datasets, which we summarize in Table 2.

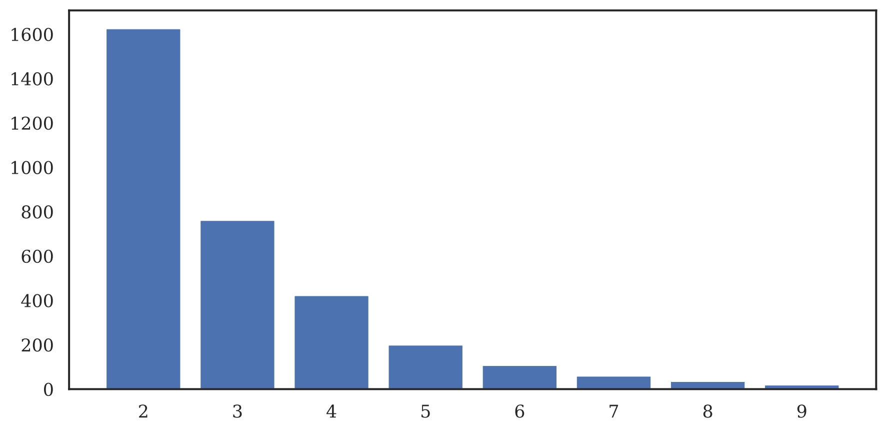

Figure 2: Aggregate participation histogram

The participation pattern is similar in all 10 datasets. Approximately 1/2 of the auctions have 2 bidders, 1/4 with 3 bidders, 1/8 with 4 bidders, etc., see Figure 2. The are a few dozens of auctions that have a single participant, which we ignore. Since it is a sealed-bid format, there is no public signal about the number of firms in the upcoming auction, so we will treat it is unknown.

The rest of the section is organized as follows. First, in Section 2.1, we discuss the institutional background of our data. Then, in Section 2.2, we discuss an example of how quality is assessed. Finally, in Section 2.3, we discuss certain distributional properties of our data.

2.1 Institutional background

In contemporary Russia, the procurement regulation is very similar to the one in the European Union and the US, with the only exception that it is more centralized. However, it did not always look the same, rather it was constantly evolving in accordance with a sequence of federal laws and presidential decrees, gradually shaping the national procurement system over the last 25 years, see Table 4.

Our data spans the middle of that time period, beginning with the earliest historical records in year 2011, available at the central e-procurement website, and ending with the major overhaul of procurement after the adoption of a new federal law in year 2014.

federal law	year	general regulation	scoring auction regulation
FZ-60, UP-305, FZ-97	1994	First attempts are made to replace central planning with competitive procurement	Scoring auctions are mostly unregulated, leaving room for corruption and preferential treatment
FZ-94	2006	Upcoming auction notifications become available online, on local websites.	The scoring auction format becomes limited to a small range of economic activities
	2011	Explicit procedures for determining the reserve price are introduced.	-
FZ-44	2014	A central auctioning platform is created, with access to all past (starting from 2011) and upcoming auction notifications and protocols.	The linear formula is replaced with the minimal bid formula. All time-related quality criteria are banned.

Table 4: Procurement regulation

In our data, one of the most important details is the way the reserve price is determined. More precisely, each time a government official sets a reserve price in a procurement auction, he has to follow the precise instructions recommended by the Russian Ministry of Economic Development. Typically, the reserve price is an average of three of more anonymous market offers from the firms in the industry, or the costs of similar contracts completed in the past.

Crucially, all the calculations and references have to be made public, together with the rest of the tender documentation, prior to the auction. The results of the auction could be even contested in court if the bidder believed the reserve price to be misleading. This indicates that the reserve price is also a signaling device that helps firms calibrate their bids.

2.2 Quality of the firm

The quality score is assigned to the firm based on several quality criteria, such as experience, number of successful contracts, qualification, etc. Even within an industry, there is a significant variation in the sets of criteria selected for each particular auction and in the evaluation scheme. This gives rise to an almost infinite variety of schemes for quality evaluation, leaving us no option but to take the final quality score at face value.

A typical formula for the quality score consists of two criteria:

$$\textit{quality score} = \textit{first criterion score} + \textit{second criterion score}.$$

For example, in the Education industry, these are often the number of customers (graduates) and the number of years that the firm has been operating on the market:

first criterion		*second criterion*	
less than 200 customers	20 points	less than 3 years	10 points
between 201 and 500 customers	30 points	between 4 and 7 years	20 points
more than 501 customers	60 points	between 8 and 10 years	30 points
		more than 11 years	40 points

Table 5: Quality criteria example 1

The quality evaluation scheme may consist of more than two criteria, or sometimes criteria have sub-divisions. In the following example from Scientific Research industry, quality is measured using two criteria: number of past contracts and qualification. However, the first criterion is split into three sub-criteria that correspond to different types of contracts:

first criterion						*second criterion*	
type 1 contracts		*type 2 contracts*		*type 3 contracts*		*qualified personnell*	
$\geqslant 5$	30 points	$\geqslant 2$	25 points	$\geqslant 1$	20 points	$\geqslant 4$	25 points
4	24 points	1	10 points	0	0 points	3	18 points
3	18 points	0	0 points			2	12 points
2	12 points					1	6 points
1	6 points					0	0 points
0	0 points						

Table 6: Quality criteria example 2

Note that in these examples, the evaluation instructions are fixed and do not depend on the characteristics of actual bidders. This is not always the case in our data. For certain auctions, the maximal points are assigned to the firm with the maximal among the current bidders number of customers (years), and all other bidders get proportionally less. We do not include these auctions in our datasets since they would require a more powerful structural model, able to deal with scoring formulas that depend on the characteristics of many bidders.

In our datasets, we have pooled together a large number of auctions with qualities of different nature. Since we take the final score at face value, we essentially assume that the utility of the designer (as a function of various quality criteria) is aligned with the quality evaluation scheme. For example, in Table 5, that would mean that 10 quality points are approximately equivalent to 3 years of experience, or 200 additional customers.

Due to the variety of the schemes and their non-linearity, we will abstain from exact interpretations of quality and treat it as an ordinal measure. In other words, we care about the fact that more quality is better than less quality, but not about what that quality means.

2.3 Spikes and bunches

In the examples of the previous section, the functions that map contracts or years into quality points are bounded from both sides. This means that two firms can be formally different in terms of their age and experience, but still have the same maximum (or minimum) quality points. This gives rise to two pronounced spikes in the quality score histogram, see Figure 3, with the right spike containing approximately 25% of the data (quality between 99 and 100).

Since our model is continuous, we will have to make sure that these distributional anomalies are smoothed out in the estimation phase. But this will happen automatically since we do not perform boundary correction in the q dimension. While this might seem like a bad idea because of over-smoothing at the boundary, it will be, in fact, beneficial for the model because sharp increases in the estimated density of quality might lead to non-existence of continuous equilibria.

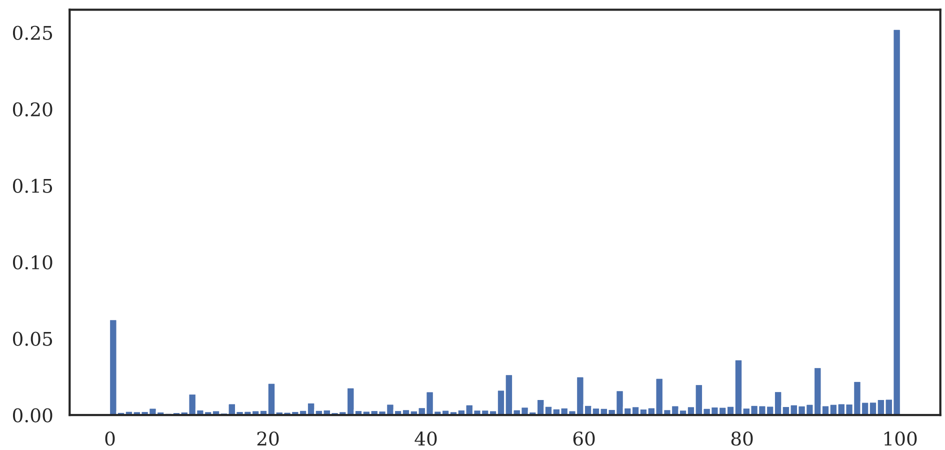

Figure 3: Aggregate histogram of quality scores.

Another distributional anomaly happens when bidders set their bid exactly at the reserve price, which we refer to as a bunch at the reserve price, see Figure 4. Approximately 7% of the data is concentrated in the range of price scores between 0 and 1, out of which 4.5% are exactly at price score 0. Though it is true that high quality firms tend to bid less aggressively, only a small fraction of the bunch is attributed to the spike in quality.

To see that the bunch is a separate phenomenon, we drop all the bids with quality greater than 99. The resulting histogram remains virtually unchanged, see Figure 5, and 3.2% of bidders still choose to bid exactly the reserve price. However, if we focus on the auctions with price weight 0.55, the pattern becomes more pronounced, see Figure 6.

This observation indicates that there may be some strategic factor at play that makes the bidders choose the reserve price with a positive probability. In the theoretical part of the paper, we will show that this feature is indeed a hallmarc of the scoring auction.

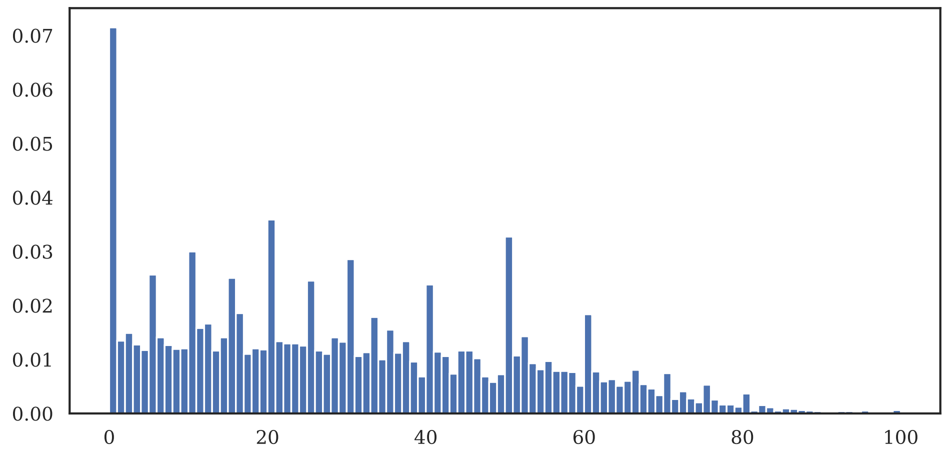

Figure 4: Aggregate histogram of price scores

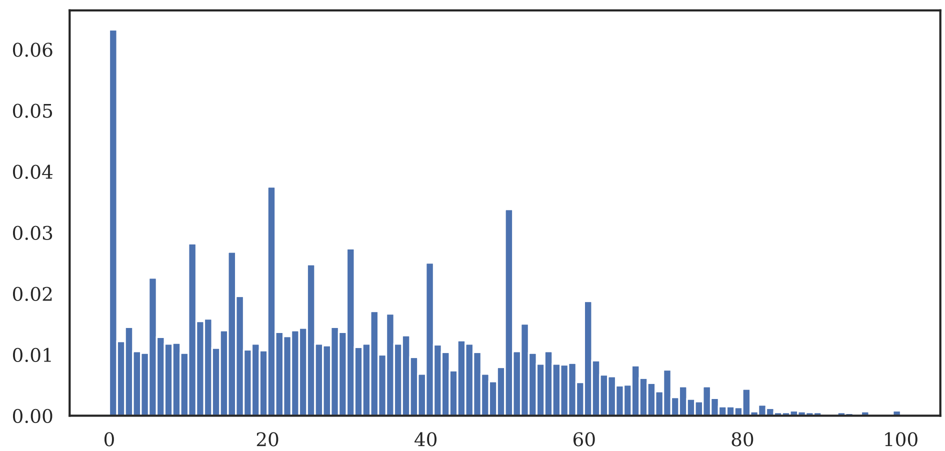

Figure 5: Aggregate histogram of price scores conditional on quality scores below 99

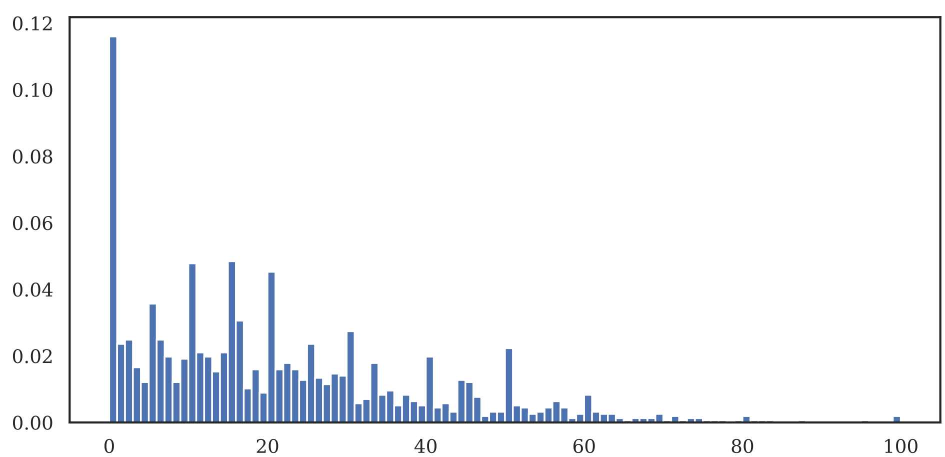

Figure 6: Aggregate histogram of price scores conditional on quality scores below 99 and price weight equal to 0.55

3 Model

A single contract is auctioned among N ex-ante identical, risk-neutral firms. Each firm simultaneously and secretly submits its bid b to the auctioneer. Each firm has quality q, which, together with the bid, is mapped into the firm's score, using a publicly known scoring rule. The firm with the highest score wins the contract at the price equal to its bid, which can not exceed a publicly known reserve price r.

A critical assumption in our model is that q is an exogenous and perfectly verifiable characteristic of the firm, observed by the auctioneer, but not the other firms. The cost c is private, and there is no technology that allows the firm to exchange it for quality or vice versa.

Assumption 1. *Each firm draws (c, q) independently, from the same distribution with a twice continuously differentiable cdf F, with full support on $[0, r] \times [\underline{q}, \overline{q}]$.*

We consider a broad class of scoring rules that we call *affine*, which nests the *quasi-linear* scoring rules studied in Che (1993) and Asker and Cantillon (2008).

Definition 1. *A scoring rule is* ***affine*** *if it can be represented by a function $s(b, q)$:*

$$s(b, q) = \alpha(q) + \beta(q)(r - b). \tag{3}$$

Three cases of affine scoring rules are of special interest:

$$\begin{array}{lll}
\textit{quasi-linear:} & \beta(q) = 1 & \Rightarrow \quad s(b, q) - r = \alpha(q) - b, \\
\textit{linear:} & \beta(q) = 1,\ \alpha(q) = wq & \Rightarrow \quad s(b, q) - r = wq - b, \\
\textit{log-linear:} & \alpha(q) = -r/q^w,\ \beta(q) = 1/q^w & \Rightarrow \quad -\log(-s(b, q)) = w \log q - \log b.
\end{array}$$

Assumption 2. *$\alpha(q), \beta(q)$ are twice continuously differentiable, $\alpha(q)$ is strictly increasing, and $\beta(q)$ is positive.*

Additionally, we assume that some firms will not come to the auction, and the participating firms hold identical beliefs about the turnout n.

Assumption 3. *The turnout is random and follows a discrete distribution with known probabilities p_n, where $1 \leqslant n \leqslant N$.*

We are looking for a symmetric, pure strategy BNE of this game.

3.1 Equilibrium Structure

We focus on the equilibrium scoring strategy, since the competition holds in the score dimension.

Our first goal is to introduce a function $\theta(c, q)$, that we will refer to as the firm's *type*, which can be used as a natural argument of the scoring strategy. We will refer to the level lines of this function as the *iso-types*.

Definition 2. *Denote the firm's* ***type*** *by $\theta(c, q)$, where*

$$\theta(c, q) = \alpha(q) + \beta(q)(r - c). \tag{4}$$

The type is similar to the firm's *production potential* in Che (1993) and the *pseudo-type* in Asker and Cantillon (2008). Without the reserve price, it would completely define the equilibrium behavior of the firm, due to the affine structure of the scoring rule. Indeed, if $\pi(s)$ is the equilibrium probability of winning associated with a score s, then the firm's profit can be written as:

$$\pi(s)(b-c) = \pi(s)(\underbrace{\alpha(q) + \beta(q)(r-c)}_{\text{firm's type}} - s)/\beta(q) = \pi(s)(\theta - s)/\beta(q),$$

which proves that θ is a sufficient statistic for the optimal choice of the score. This choice, however, might be infeasible in the presence of the reserve price.

Note that the iso-types depend only on the shape of the scoring rule, in other words they are exogenous. For the linear scoring rule $s = q - 2b$ these are parallel lines, and for a log-linear $s = -b/q$ these are rays originating at $c = 0, q = 0$, see Figure 7.

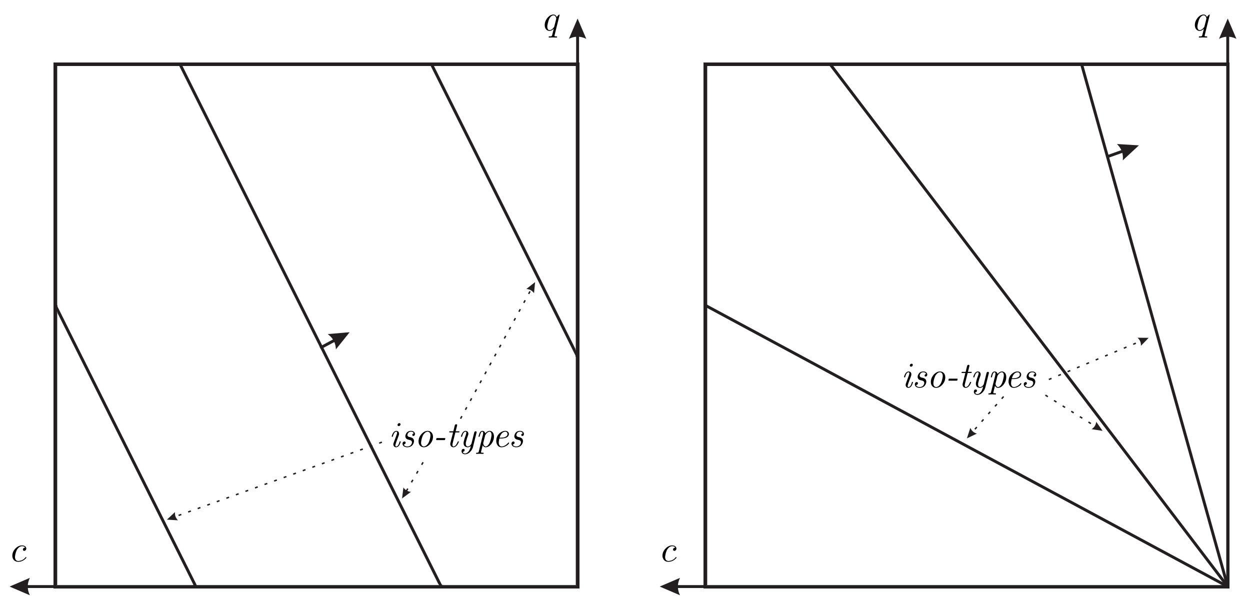

Figure 7: Iso-types for a linear and a log-linear scoring rule.

Our second goal is to define two scoring strategies that capture the firm's optimal choice of the score, with and without censoring imposed by the reserve price.

Definition 3. *Denote the equilibrium* ***scoring strategy*** *by* $\sigma^*(\theta, q)$.

This strategy represents the constrained choice of the score, and we will refer to its level lines as the *iso-scores*. By definition, it solves the following optimization problem:

$$\sigma^*(\theta, q) \in \arg\max_{s \geqslant \alpha(q)} \pi(s)(\theta - s). \tag{5}$$

The underlying bidding strategy can be easily recovered from the scoring strategy using formulas $\theta = \alpha(q) + \beta(q)(r-c)$ and $s = \alpha(q) + \beta(q)(r-b)$, which are simply accounting identities. Note that the situation when the bid is equal to r corresponds to the situation where the score is equal to $\alpha(q)$.

Definition 4. *Define* $\sigma(\theta) = \min_q \sigma^*(\theta, q)$ *as the equilibrium* ***uncensored scoring strategy****.*

This strategy represents the unconstrained choice of the score, and its level lines coincide with the iso-types. By construction, it solves the following optimization problem:

$$\sigma(\theta) \in \arg\max_s \pi(s)(\theta - s). \tag{6}$$

A natural conjecture is that the scoring strategy $\sigma^*(\theta, q)$ should be derived from $\sigma(\theta)$ via censoring at the $\alpha(q)$ threshold. This is not true in general, however, if we introduce a simple refinement of the equilibrium, this property can be established.

Assumption 4. $\sigma^*(\theta, q)$ *is continuous and a lower score is chosen when indifferent.*

Lemma 1. *Under Assumption 4, following equation holds:*

$$\sigma^*(\theta, q) = \max(\sigma(\theta), \alpha(q)). \tag{7}$$

Following equation (7), one can see that the iso-scores are kinked. Consequently, they are endogenous, as the position of the kink relies on the equilibrium through the $\sigma(\theta)$ function. We can nevertheless describe them qualitatively. When the score is a function of type, the iso-scores are aligned with the iso-types, but when the score is a function of quality, the iso-scores are horizontal. The set of firms that choose the reserve price, which we refer to as the *bunching region*, is outlined by the curve traced the kinks of the iso-scores, see Figure 8 for a stylized illustration.

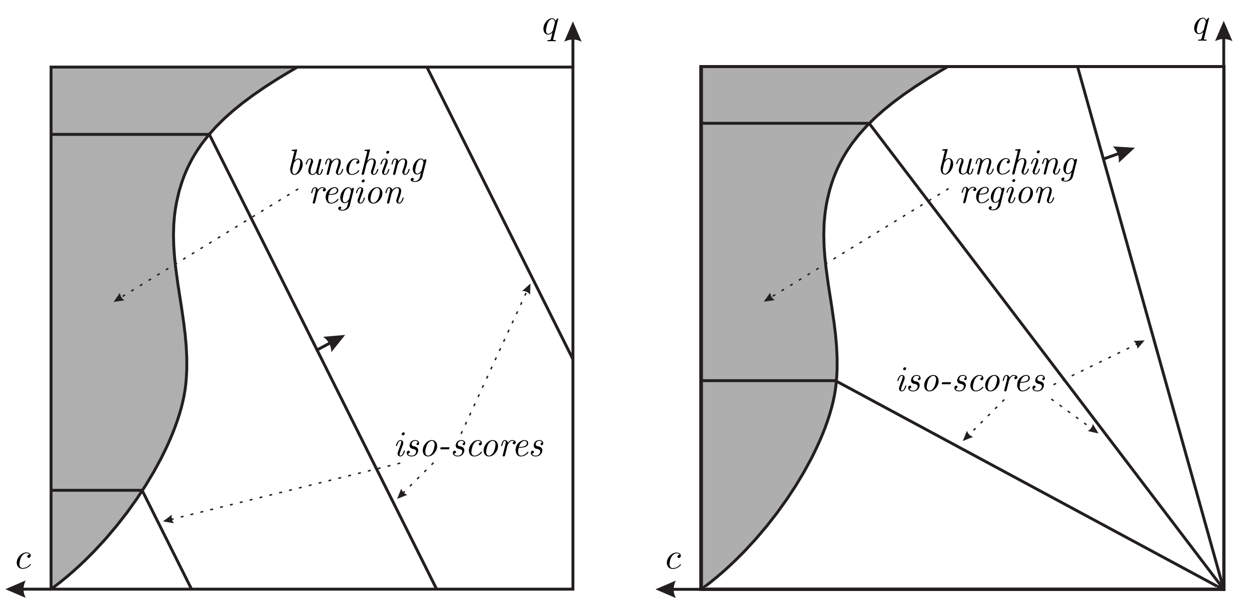

Figure 8: Iso-scores and the bunching region (grey) for a linear and a log-linear scoring rule.

The equilibrium strategy $\sigma^*(\theta, q)$ can be derived from the equilibrium strategy $\sigma(\theta)$ using equation (7). Consequently, we only have to characterize the $\sigma(\theta)$ strategy in order to pin down the equilibrium. To do that a few additional definitions and a regularity assumption are required.

Definition 5. *Denote the cumulative distribution function of* $(\theta, \alpha(q))$ *by* $\widetilde{F}$*, and its partial derivatives by* $\widetilde{F}_1'$ *and* $\widetilde{F}_2'$*. Define a function* $Z(\theta,\sigma) = (\sum p_n \widetilde{F}^{n-2}(\theta,\sigma))/(\sum p_n(n-1)\widetilde{F}^{n-2}(\theta,\sigma))$ *and a correspondence* $\delta(\theta) = \{\sigma \in [\underline{\theta}, \theta] : Z(\theta,\sigma)\widetilde{F}(\theta,\sigma) - (\theta-\sigma)\widetilde{F}_2'(\theta,\sigma) = 0\}$*, where* $[\underline{\theta}, \overline{\theta}]$ *is the support of* θ*.*

Assumption 5 (Regularity)**.** $\delta(\theta)$ *is a function, such that* $\delta(\theta) \geqslant \alpha(\overline{q})$ *for some* θ *in* $[\underline{\theta}, \overline{\theta}]$*.*

The role of this assumption will be discussed in detail in Section 3.2. We are now ready to give a characterization of the equilibrium.

Proposition 1. *Under Assumptions 1-5, there exists a unique symmetric BNE in the scoring auction, with a strictly monotone uncensored scoring strategy* $\sigma(\theta)$*, that solves:*

$$\sigma(\underline{\theta}) = \underline{\theta} \quad \text{and} \quad \sigma' = \frac{(\theta-\sigma)\widetilde{F}_1'(\theta,\sigma)}{Z(\theta,\sigma)\widetilde{F}(\theta,\sigma) - (\theta-\sigma)\widetilde{F}_2'(\theta,\sigma)} \quad \text{for all types in } [\underline{\theta}, \overline{\theta}]. \tag{8}$$

To illustrate the derivation of the differential equation (8), assume that there are two firms and there is no uncertainty about the turnout. Observe that the equilibrium probability of winning associated with a score $\sigma(\theta)$ is equal to the cumulative distribution function $\widetilde{F}(\theta, \alpha(q))$ evaluated at $(\theta, \sigma(\theta))$, which is also the area under the iso-score:

$$\pi(\sigma(\theta)) = Prob(\sigma^*(\hat{\theta}, \hat{q}) \leqslant \sigma(\theta)) = Prob(\sigma(\hat{\theta}) \leqslant \sigma(\theta),\ \alpha(\hat{q}) \leqslant \sigma(\theta)) = \widetilde{F}(\theta, \sigma(\theta)).$$

Indeed, if we switch to the $(\theta, \alpha(q))$ coordinates, the probability of winning will be to the left and below the (θ, σ) point. Moreover, the boundary of the bunching region will coincide with the $\sigma(\theta)$ curve, as they represent the same event when $\alpha(q)$ is equal to $\sigma(\theta)$. In these coordinates the bunching region is therefore the area between the 45 degree line and the $\sigma(\theta)$ curve. Note that while the $\sigma(\theta)$ is monotone, the boundary of the bunching region, obtained by tilting the curve to the left, might have an inflection point.

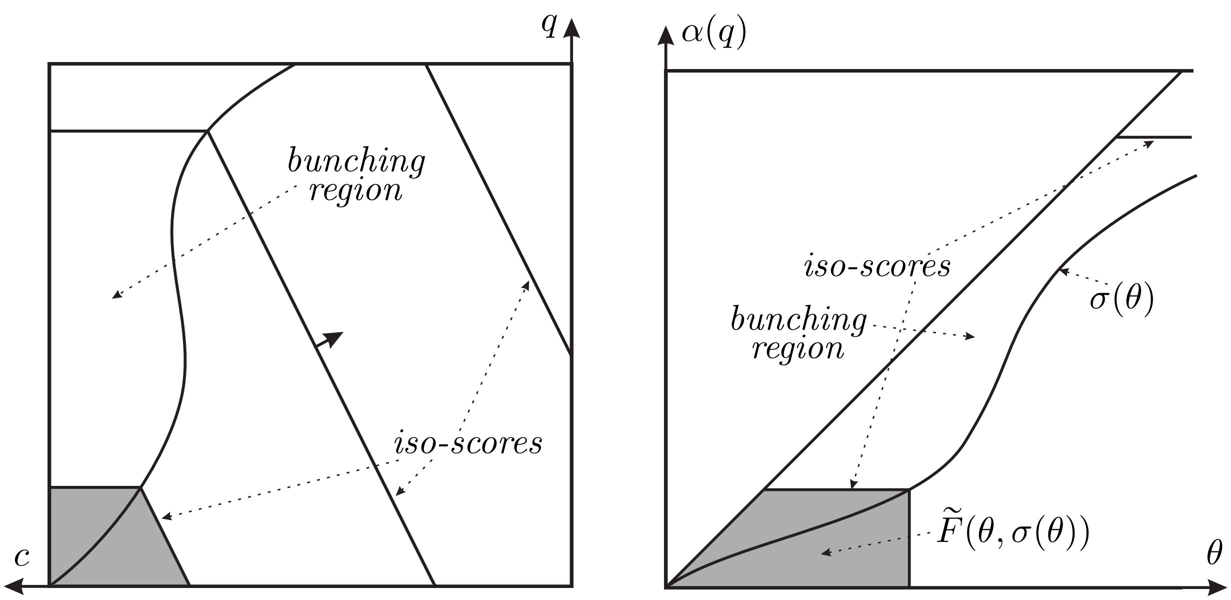

Figure 9: Iso-scores and the probability of winning (grey) for a linear scoring rule.

Bunching is typically considered incompatible with the symmetric equilibria in the first price auctions, as an infinitesimal deviation in the bid generates a sizable increase in the probability of winning. In the scoring auction, however, this is not the case as competition holds in the score, rather than the bid dimension. The bunch is spread across different scores due to the variation in q, and so the distribution relevant to the formation of best response is, in fact, continuous.

We can now derive the first order conditions associated with the optimal choice of the score.

First Order Conditions (2 firms):

$$\overbrace{\widetilde{F}(\theta,\sigma)d\sigma}^{\textit{marginal cost}} \approx \overbrace{(\theta-\sigma)\cdot(\widetilde{F}_1'(\theta,\sigma)d\theta + \underbrace{\widetilde{F}_2'(\theta,\sigma)d\sigma}_{\text{bunching effect}})}^{\text{marginal benefit}}. \tag{9}$$

The marginal cost of raising the score above the equilibrium can be written as $d\sigma$ times the probability of winning $\widetilde{F}(\theta,\sigma)$, while the marginal benefit is the profit margin $(\sigma-\theta)$ times the marginal increase in the probability of winning. The latter comes from two populations of firms, see Figure 10. The first are the firms outside the bunch, that have a marginally higher type θ, captured by the $\widetilde{F}_1'(\theta,\sigma)d\theta$ term. The second are the firms inside the bunch that have a marginally weaker constraint on the score, captured by the $\widetilde{F}_2'(\theta,\sigma)d\sigma$ term. We call the latter *bunching effect*, as it only appears in the presence of bunching.

After dividing both parts of equation (9) by $d\theta$, replacing $d\sigma/d\theta$ with σ', and solving for σ', we obtain a special case of the differential equation (8), where $Z(\theta,\sigma)=1$. We will derive it more rigorously and show that the second order conditions hold in Section 3.3.

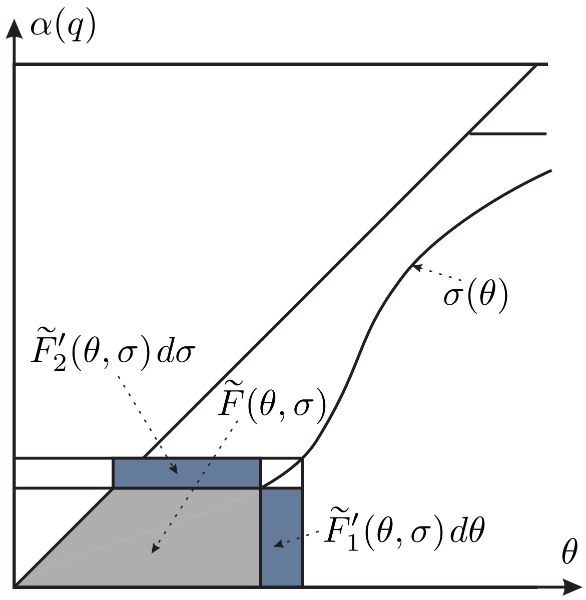

Figure 10: Marginal increase in the probability of winning.

While the distribution $\widetilde{F}$ itself is exogenous, the way it enters the first order conditions is clearly endogenous, as it is evaluated at (θ,σ). That is because the way expectations are formed about the strength of the competition depends on the amount of bunching, which in turn, depends on the amount of equilibrium shading. This creates a feedback loop between the equilibrium strategy $\sigma(\theta)$ and the equilibrium probability of winning $\pi(\sigma(\theta))$, which does not happen in a classic first-price auction, or in a scoring auction with no reserve price.

While the first order conditions are fairly complex due to the equilibrium expectations about the bunch, the firm's bidding problem is simple, as all the relevant information is captured by the observed distribution of scores. The firm then behaves simply as a monopolist operating over a downward sloping demand curve $\pi(s(b, q))$, with a price ceiling r:

$$\max_{b \leqslant r}(b-c)\pi(s(b,q)), \quad s(b,q) = \alpha(q) + \beta(q)(r-b).$$

We are interested in how the monopoly price (equivalently, the optimal bid) responds to an increase in c or q, assuming that the price ceiling does not bind. The first can be considered as an increase in the monopoly's marginal cost, which implies an increase in the monopoly price since b and c are complements in the profit function. The second can be considered as a positive demand shock that shifts the demand curve upwards. While it is natural to assume that the monopoly price will increase, this is not necessarily the case. Indeed, from the inverse elasticity rule we know that the defining factor is the elasticity of demand rather than its volume. If the optimal bid does not change with a small increase in q, the iso-bid will have an inflection point, as on Figure 11.

For the linear scoring rule, the shift of the demand curve can be offset by a proportional increase in the marginal cost, so that the profit margins remain the same. Consequently, the iso-bids will be parallel translations of each other, as on Figure 11.

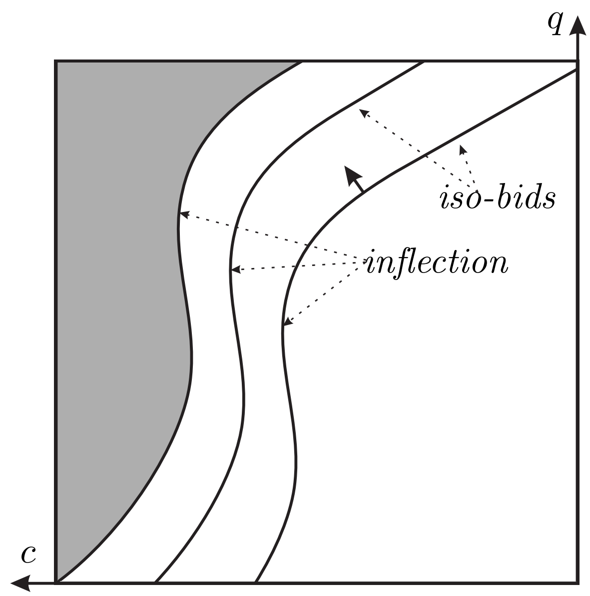

Figure 11: Iso-bids and bunching (grey area).

So far our analysis was for the case of two firms with no uncertainty about the turnout. To allow for random participation, we need to replace the probability of winning against a random firm $\widetilde{F}$ with the probability of winning against the strongest opponent firm. This probability is captured by a convoluted function $\sum p_n \widetilde{F}^{n-1}$, which finds its way into the first order conditions:

First Order Conditions (general):

$$\sum_{n=1}^{N} p_n \widetilde{F}^{n-1} d\sigma \approx (\theta - \sigma) \cdot (\widetilde{F}_1' d\theta + \widetilde{F}_2' d\sigma) \cdot \sum_{n=1}^{N} p_n (n-1) \widetilde{F}^{n-2}. \tag{10}$$

By introducing the $Z(\theta, \sigma)$ function we are able to isolate the effect of random participation in our differential equation:

$$\frac{(\theta - \sigma) \cdot (\widetilde{F}_1' + \widetilde{F}_2' \sigma')}{\sigma' \cdot \widetilde{F}} = Z = \frac{\sum_{n=1}^{N} p_n \widetilde{F}^{n-2}}{\sum_{n=1}^{N} p_n (n-1) \widetilde{F}^{n-2}}. \tag{11}$$

With a turnout fixed at n, Z is simply equal to $1/(n-1)$. Finally, after solving for σ', we obtain our differential equation:

$$\sigma' = \frac{(\theta - \sigma) \widetilde{F}_1'}{Z \widetilde{F} - (\theta - \sigma) \widetilde{F}_2'}. \tag{12}$$

Coupled with the initial value $\sigma(\underline{\theta}) = \underline{\theta}$, this differential equation can be solved locally under standard conditions, however, the solution might not necessarily exist globally for a number of reasons. In the next section we will focus on the conditions that are sufficient for the existence of the global solution.

3.2 Regularity, Existence and Uniqueness

The Initial Value Problem in Proposition 1 can be interpreted as a direct instruction to finding the equilibrium. However, one should be careful when constructing the $\sigma(\theta)$ trajectory as there are two cases how it may fail.

The first case is when $Z\widetilde{F} - (\theta - \sigma)\widetilde{F}_2'$ turns into zero thus preventing us from constructing a global solution. The solution then should be sought in the wider functional space of, possibly discontinuous, monotone functions. The second case is when $\widetilde{F}_1'$ turns into zero. Then the first order condition becomes degenerate and the trajectory switches, through a kink, to a different law of motion: $(\theta - \sigma)\widetilde{F}_2' = Z\widetilde{F}$, which can be interpreted as playing against the population of firms belonging entirely to the bunch. It is even possible for the trajectory to switch between the two laws of motion multiple times. The role of Assumption 5 is to make sure that the two previously mentioned cases do not occur, so that we can focus on the most regular scenario.

Recall the definition of the $\delta(\theta)$ correspondence:

$$\delta(\theta) = \{\sigma \in [\underline{\theta}, \theta] : Z(\theta, \sigma)\widetilde{F}(\theta, \sigma) - (\theta - \sigma)\widetilde{F}_2'(\theta, \sigma) = 0\}.$$

It is non-empty for all $\theta \in [\underline{\theta}, \overline{\theta}]$ by the *Intermediate Value Theorem*, and captures the set of points in the $(\theta, \alpha(q))$ coordinates for which the denominator in the right handside of the differential equation turns into zero. If, additionally, it is a function, it serves as a lower boundary to where the $\sigma(\theta)$ trajectory can go, because any continuous strategy will by pushed upwards when approaching $\delta(\theta)$. At the same time, when approaching the 45 degree line, the $\sigma(\theta)$ trajectory will be pushed to the right. Consequently, the trajectory always remains between these two functions:

$$\delta(\theta) \leqslant \sigma(\theta) \leqslant \theta.$$

In the theory of ordinary differential equations, the $\alpha(q) = \delta(\theta)$ and $\alpha(q) = \theta$ curves are referred to as *fences*, and together they form what is called an *anti-funnel*, see Figure 12. Existence and uniqueness of the trajectory that originates at $(\underline{\theta}, \underline{\theta})$ follows from the theory of fences and funnels, see Hubbard and West (1997) and Appendix A for more details.

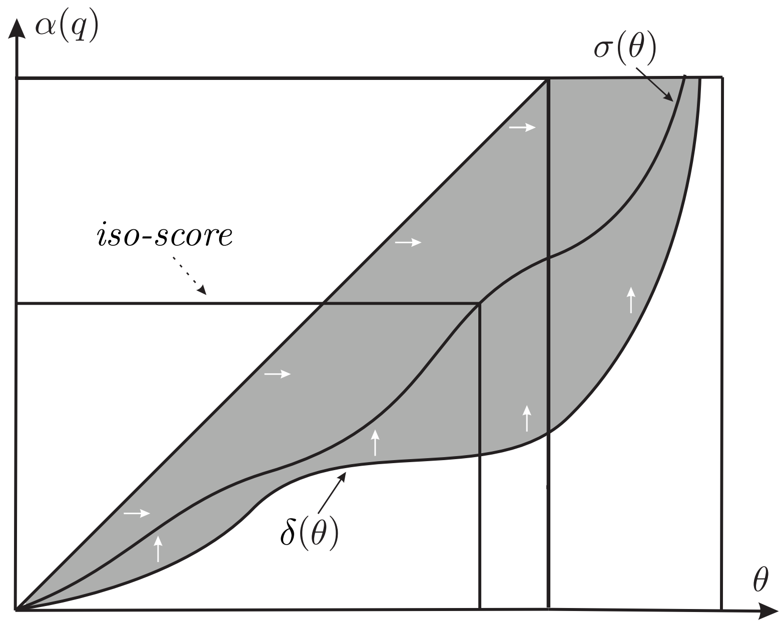

Figure 12: The anti-funnel (grey), the $\delta(\theta)$ and the $\sigma(\theta)$ curves.

The first part of Assumption 5 tells that $\delta(\theta)$ is a function, thus allowing us to establish the existence and uniqueness of the global solution to our differential equation. The second part tells that there exists a θ such that $\sigma(\theta) \geqslant \alpha(\overline{q})$, which means that the trajectory never passes through the point where $\widetilde{F}_1'$ turns into zero. Consequently, there is no switch in the law of motion and the differential equation correctly represents the first order conditions.

While the first order conditions are clearly necessary, it is not obvious whether they are sufficient for the equilibrium. To show this we will study a direct mechanism associated with the game in the next section.

3.3 Pseudo-type and the Direct Revelation Mechanism

We introduce a generalization of the *pseudo-type* in Asker and Cantillon (2008), which is a function, designed to be constant along the iso-score lines. Once the pseudo-type is defined, it can be thought of as a message in the direct mechanism, and the first order conditions can be derived using the revelation principle.

Definition 6. *Denote the firm's **pseudo-type** by $\rho(\theta, q)$, where*

$$\rho(\theta, q) = \sigma^{-1}(\sigma^*(\theta, q)). \tag{13}$$

By construction, the pseudo-type is constant along the iso-scores and, moreover, coincides with the type when the reserve price is not binding:

$$\rho(\theta, q) = \begin{cases} \theta, & \textit{if the reserve price is not binding,} \\ \alpha^{-1}(\sigma(\theta)), & \textit{if the reserve price is binding.} \end{cases}$$

The intuition behind the pseudo-type is the following. Since for certain types the reserve price is binding, the mechanism pools them with higher types (but same quality) all the way up to the value of $\sigma^{-1}(\alpha(q))$, which is the only type that would have chosen the reserve price willingly. The pseudo-type is therefore the type of the firm as perceived by the mechanism in equilibrium.

Definition 7. *Denote the* ***pseudo-type distribution*** *by $G(\rho)$ and the* ***residual pseudo-type distribution*** *by $\mathcal{G}(\rho)$, where*

$$\mathcal{G}(\rho) = \sum p_k G^{k-1}(\rho). \tag{14}$$

The distribution of the pseudo-type plays the same role as the distribution of values in a classic first-price auction. The area under the iso-score is equal to the probability of winning against a random opponent while choosing the score $\sigma(\rho)$, see Figure 13.

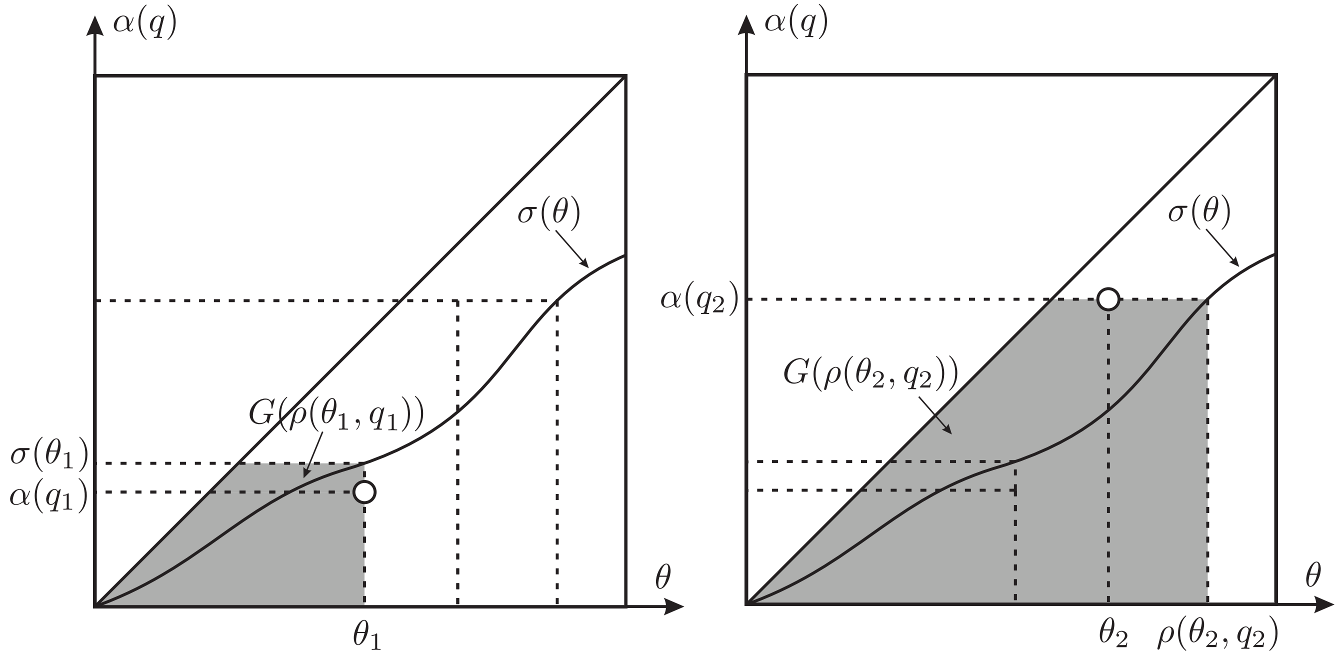

Figure 13: Pseudo-type and the probability of winning (grey).

We can now characterize our uncensored scoring strategy $\sigma(\theta)$ as an equilibrium strategy in a direct revelation mechanism where the firm submits the pseudo-type. The optimality conditions in this game are summarized below:

Direct Revelation Mechanism:

$$\sigma(\theta) \in \arg\max_{\rho} (\theta - \sigma(\rho))\mathcal{G}(\rho). \tag{15}$$

Indeed, the probability of winning against a single opponent by signaling ρ to the mechanism is equal to $G(\rho)$. The expected probability of winning is therefore $\mathcal{G}(\rho)$. The profit conditional on winning is equal to $b - c = (\theta - s)/\beta(q)$, where the score s is assigned based on the reported pseudo-type and therefore is equal to $\sigma(\rho)$.

Notice that the optimality conditions are exactly as in the first-price auction, if we interpret θ as value, $\sigma(\theta)$ as the bidding strategy and $\mathcal{G}$ as the value distribution. The classic first order conditions can be therefore written as:

First Order Conditions (general):

$$\overbrace{\sigma'(\rho)\cdot\mathcal{G}(\rho)}^{\text{marginal cost}} = \overbrace{(\theta-\sigma(\rho))\cdot\mathcal{G}'(\rho)}^{\text{marginal benefit}}. \tag{16}$$

Since $G(\theta)=\widetilde{F}(\theta,\sigma(\theta))$, the direct mechanism approach gives the same first order conditions as (10). But even without the first order conditions, the $\sigma(\theta)$ strategy can be characterized in terms of the pseudo-type distribution $\mathcal{G}$ by simply applying the *Envelope Theorem* to (15).

Proposition 2. *Under Assumptions 1-5, the equilibrium uncensored scoring strategy $\sigma(\theta)$ is a conditional expectation of the highest (among the other firms) pseudo-type below θ:*

$$\sigma(\theta)=\frac{1}{\mathcal{G}(\theta)}\int_{\underline{\theta}}^{\theta} z\,d\mathcal{G}(z)=\theta-\frac{1}{\mathcal{G}(\theta)}\int_{\underline{\theta}}^{\theta}\mathcal{G}(z)dz. \tag{17}$$

As in the classic first-price auction, the second order conditions follow from (and are equivalent to) the monotonicity of $\sigma(\theta)$, see, for example, Chapter 5 in Krishna (2010). To convince ourselves that the second order conditions are indeed satisfied in our model, we evaluate the change in profit when submitting a pseudo-type $\hat{\rho}$ while having pseudo-type ρ:

$$\begin{aligned}
(\theta-\sigma(\hat{\rho}))\mathcal{G}(\hat{\rho})-(\theta-\sigma(\rho))\mathcal{G}(\rho) &= \\
(\hat{\rho}-\sigma(\hat{\rho}))\mathcal{G}(\hat{\rho})+(\theta-\hat{\rho})\mathcal{G}(\hat{\rho})-(\rho-\sigma(\rho))\mathcal{G}(\rho)-(\theta-\rho)\mathcal{G}(\rho) &= \\
\int_{\rho}^{\hat{\rho}}\mathcal{G}(z)dz+(\theta-\hat{\rho})\mathcal{G}(\hat{\rho})-(\theta-\rho)\mathcal{G}(\rho) &= \\
\int_{\rho}^{\hat{\rho}}\mathcal{G}(z)dz-(\hat{\rho}-\rho)\mathcal{G}(\hat{\rho})-(\rho-\theta)(\mathcal{G}(\hat{\rho})-\mathcal{G}(\rho)) &\leqslant 0.
\end{aligned}$$

The change is negative since $\rho\geqslant\theta$ by the definition of the pseudo-type.

Corollary 1. *Under Assumptions 1-5, the first order conditions are sufficient.*

Formula (17) also demonstrates that $\sigma(\theta)<\theta$ for all $\theta>\underline{\theta}$. As a result, for every level of quality except for the lowest, the threshold cost $c(q)$ for which the firm is willing to pick the reserve price lies in the interior of the support:

$$c(q)=r-\frac{\sigma^{-1}(\alpha(q))-\alpha(q)}{\beta(q)}<r,\ \text{ for all } q>\underline{q}.$$

Consequently, as long as the distribution F has full support, there will be a positive mass of bids at the reserve price, conditional on every level of quality except for the lowest.

Corollary 2. *Under Assumptions 1-5, there is bunching at every level of $q>\underline{q}$.*

Coupled with $G(\theta)=\widetilde{F}(\theta,\sigma(\theta))$ and $\mathcal{G}(\rho)=\sum p_k G^{k-1}(\rho)$, equation (17) pins down the evolution of both $\sigma(\theta)$ and $\mathcal{G}(\theta)$. From this system of non-linear equations, we can derive a system of differential equations which is linear in σ', G' and $\mathcal{G}'$:

$$\sigma \cdot \mathcal{G}' + \sigma' \cdot \mathcal{G} = \theta \cdot \mathcal{G}', \tag{18}$$

$$\mathcal{G}' = G' \cdot \sum_{k=2}^{N} (n-1) p_n G^{n-2}, \tag{19}$$

$$G' = \widetilde{F}_1' + \widetilde{F}_2' \cdot \sigma'. \tag{20}$$

By eliminating G' and $\mathcal{G}'$ we arrive to the same differential equation as in Proposition 1.

3.4 Welfare

In this section, we will introduce three measures of welfare generated by a firm participating in an auction: firm's profit, quality and buyer's rebate $r - b$; that will be later used to rank the counterfactual scoring rules. It will be convenient to first analyze the interim version.

Definition 8. *For a firm with type θ and quality q:*

$$\begin{aligned} \textit{interim expected profit:} \quad & P(\theta, q) = (b - c)\mathcal{G}(\rho(\theta, q)), \\ \textit{interim expected quality:} \quad & Q(\theta, q) = q\mathcal{G}(\rho(\theta, q)), \\ \textit{interim expected rebate:} \quad & R(\theta, q) = (r - b)\mathcal{G}(\rho(\theta, q)). \end{aligned}$$

By applying the envelope theorem to (15), we can find the profit as a solution to the boundary value problem below:

$$\frac{\partial}{\partial \theta} P(\theta, q) = \mathcal{G}\left(\rho(\theta, q)\right)/\beta(q), \quad P(\alpha(q), q) = 0.$$

This allows us to write down the interim expected profit, quality and rebate in the $(\theta, \alpha(q))$ coordinates in terms of the residual distribution of the pseudo-type and the pseudo-type itself. The total expected surplus from a single participating firm is measured by

$$(r - c) \cdot \mathcal{G}(\rho(\theta(c, q), q)) = (\theta - \alpha(q)) \cdot \mathcal{G}(\rho(\theta, q))/\beta(q),$$

which is the area of the marked region in Figure 14. Similarly to a classic first-price auction, this area is split between firm's profit and seller's revenue (in our case, buyer's rebate) by the $\mathcal{G}(\rho(*, q))$ curve, which captures the probability of winning having type θ.

Corollary 3. *The interim expected profit, quality and rebate can be computed from the residual pseudo-type distribution $\mathcal{G}(\rho)$ and the pseudo-type $\rho(\theta, q)$ using formulas below:*

$$\begin{aligned} P(\theta, q) &= \int_{\alpha(q)}^{\theta} \mathcal{G}(\rho(z, q)) dz/\beta(q) \\ Q(\theta, q) &= q \cdot \mathcal{G}(\rho(\theta, q)) \\ R(\theta, q) &= (\theta - \alpha(q)) \cdot \mathcal{G}(\rho(\theta, q))/\beta(q) - P(\theta, q). \end{aligned}$$

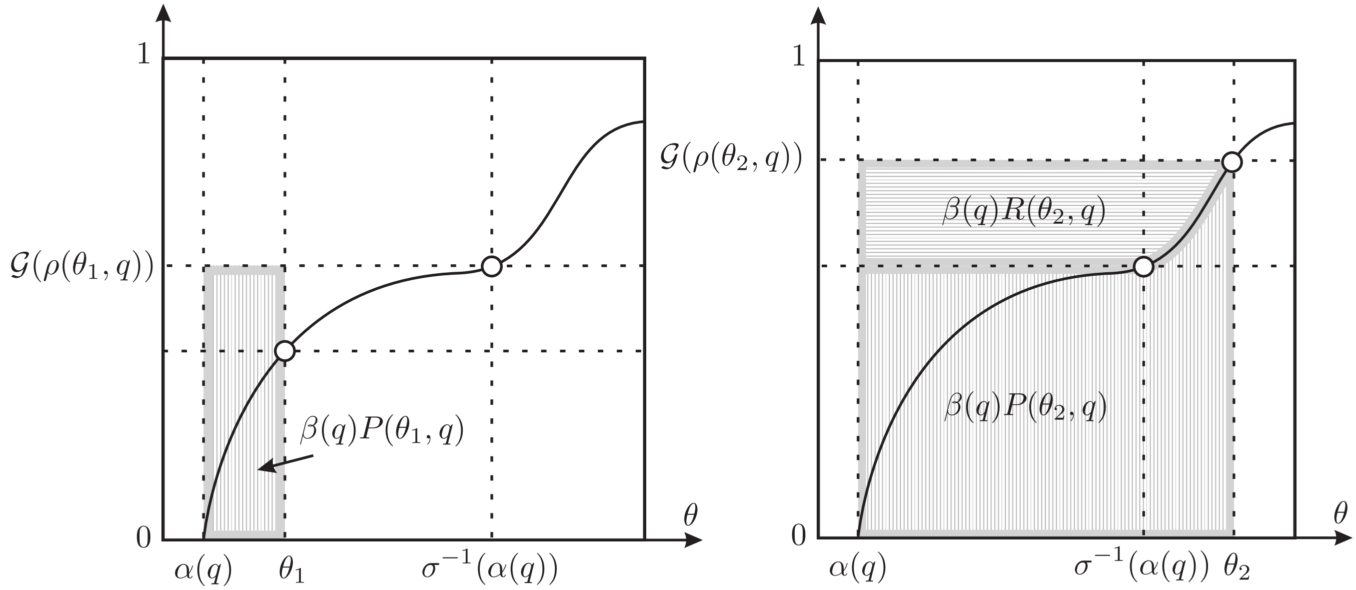

Figure 14: The firm's interim profit is the area under the $\mathcal{G}(\rho(*, q))$ curve, while the buyer's interim rebate is the area to the left from the $\mathcal{G}(\rho(*, q))$ curve.

Had the turnout been fixed, to compute the total ex-ante quality and rebate, we would average the interim ones with respect to the density of type and quality, and multiply by the number of participants. Due to random participation, the formulas are more tricky.

Conditional on the turnout n, the total rebate and quality extracted from firms with coordinates (θ, q) can be computed by renormalizing the interim ones by $G^{n-1}(\rho(\theta, q))/\mathcal{G}(\rho(\theta, q))$ and multiplying by n. These should be then averaged with the turnout probabilities p_n and integrated over the density of (θ, q).

Definition 9. *Define a function* $Y(\theta, q) = (\sum p_n n G^{n-1}(\rho(\theta, q)))/(\sum p_n G^{n-1}(\rho(\theta, q)))$.

Similarly to the $Z(\theta, \alpha(q))$ function, $Y(\theta, q)$ measures the extent of competition attributed to a firm with coordinates (θ, q). With fixed turnout it is simply equal to n.

Corollary 4. *The total (ex-ante) expected profit and quality can be computed from the interim ones, the* $Y(\theta, q)$ *function and the joint density of* $(\theta, \alpha(q))$ *using formulas below:*

$$\text{total expected quality:} \quad \int\int Q(\theta, q) Y(\theta, q) \tilde{f}(\theta, \alpha(q)) d\theta d\alpha(q),$$

$$\text{total expected rebate:} \quad \int\int R(\theta, q) Y(\theta, q) \tilde{f}(\theta, \alpha(q)) d\theta d\alpha(q).$$

4 Estimation

In this section, we treat our data as if there is no auction heterogeneity. We will show later in Section 5 that this is a reasonable approach, as long as the reserve price is a credible signal about the scale of the auction.

The theoretical analysis in Section 3.1 has shown that our model of scoring auctions shares important features with the classic first-price auction. Despite the fact that bidder's characteristic is two-dimensional, the core equilibrium strategy $\sigma(\theta)$ is one-dimensional, and, moreover, it is a solution to optimality conditions:

$$\sigma(\theta) \in \arg\max_{s} \pi(s)(\theta - s). \tag{21}$$

Since the probability of winning $\pi(s)$ is determined by the observed distribution of scores, nonparametric estimators of Guerre et al. (2000) and Li et al. (2002), which now can be considered as standard, can be easily generalized to work in our environment. However, we will make several important alterations that we explain below.

First, in order to capture the participation patterns in a symmetric model, we will estimate the strategy as a best response to the residual distribution of scores, that is the distribution of the maximum of scores among all but one bidder. Since this distribution can be thought of as generated by a composite bidder (the strongest among $n-1$ others), from the perspective of each single firm, the decision problem is as if it was competing against that composite bidder in a first-price auction. By that logic, it is absolutely justified to rely on the nonparametric techniques in Guerre et al. (2000) and Li et al. (2002) for the nonparametric estimation of the aforementioned distribution and the corresponding strategy.

Second, contrary to the classic first-price auction, it is not possible to use the estimated strategy $\sigma(\theta)$ to recover the whole distribution of (θ, q) from the observed distribution of (s, q), because only those bids that fall below the reserve price are coherent with that strategy. The bids that fall exactly at the reserve price therefore can not be used as part of the nonparametric estimation and should be dropped.

We first estimate the joint distribution of (b, q) using a bivariate kernel density estimator with boundary correction at $b = r$, advocated in Hickman and Hubbard (2015). We then transform it into the distribution of (s, q). Finally, by applying the inverse of the estimated $\sigma(\theta)$ strategy, we recover the part of the joint distribution of (θ, q) that does not map into the reserve price in equilibrium. The rest of the distribution is extrapolated from the boundary, namely, for every level of q, the density of a type that was forced to bid the reserve price will be set equal to the density of the type that chose the reserve price willingly, see Figure 15.

With the estimated distribution of (θ, q) at hand, we simulate the equilibria for the counterfactual scoring rules and calculate the expected quality and expected rebate using formulas derived in Section 3.4.

The rest of this section is split in three parts. In Section 4.1, we estimate the σ strategy. In Section 4.2, we recover the part of the distribution of (θ, q) that is identified. And, ultimately, in Section 4.3, we extrapolate the remaining part of the distribution, explain how to find the counterfactual equilibrium and compute welfare.

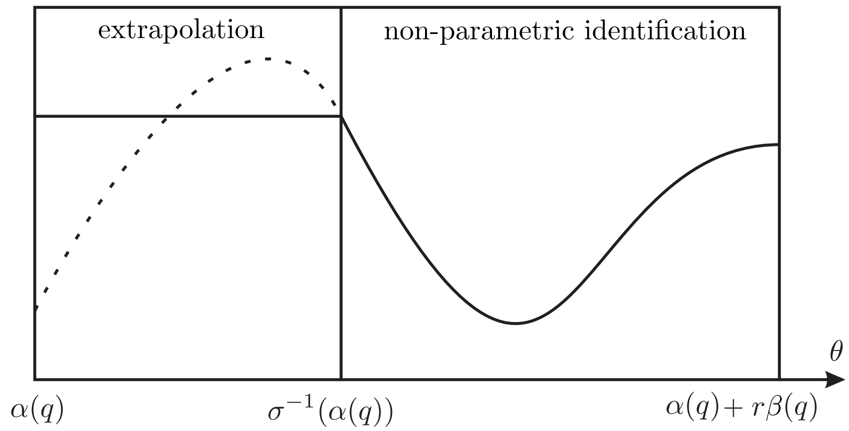

Figure 15: Estimation of the density of type, conditional on quality.

4.1 First-stage smoothing and strategy estimation

The first step in our estimation is to capture the equilibrium behavior imposed by the optimality conditions (21). The probability of winning $\pi(s)$ is measured by the distribution of the highest among $n-1$ firms score, which we refer to as a *residual distribution of scores.*

Definition 10. *Denote the **residual distribution of scores** by $\mathcal{H}$, and the corresponding density by $\hbar$.*

Following the logic of Guerre et al. (2000), we estimate the perceived distribution $\mathcal{H}$ against which each firm is playing in our symmetric model together with its density $\hbar$.

Definition 11. *Denote the auction index by j and the bidder index (in that auction) by i.*

Definition 12. *Denote the sample of scores by $S = \{s_{ij}\}$ and its size by M.*

We apply a classic kernel density estimator to the residual scores, that is the maximal score among all but one firm, for each firm i in each auction j:

$$\widehat{\hbar}(s) = \frac{1}{M} \sum_{s_{ij} \in S} K\left(\frac{\max_{k \neq i} s_{kj} - s}{h_s}\right),$$

with the kernel K and the bandwidth b_s borrowed from Li et al. (2002):

$$K(u) = \frac{35}{32}(1-u^2)^3 \mathbb{I}(|u| \leqslant 1), \quad h_s = 3.16 \cdot sd(s) \cdot m^{-1/5},$$

where $sd(s)$ is the standard deviation of the score in S. The choice of this bandwidth is based on the popular statistical recipe called the 'Silverman rule', which is calibrated for normal distributions, see Härdle (2012). The score distributions in our data look reasonably bell-shaped, see Figure 16, so we have no reason to distrust this approach.

Similarly, $\widehat{\mathcal{H}}$ can be obtained. From the practical point of view, however, we find it more natural to simply integrate the estimated density $\widehat{\hbar}$ numerically.

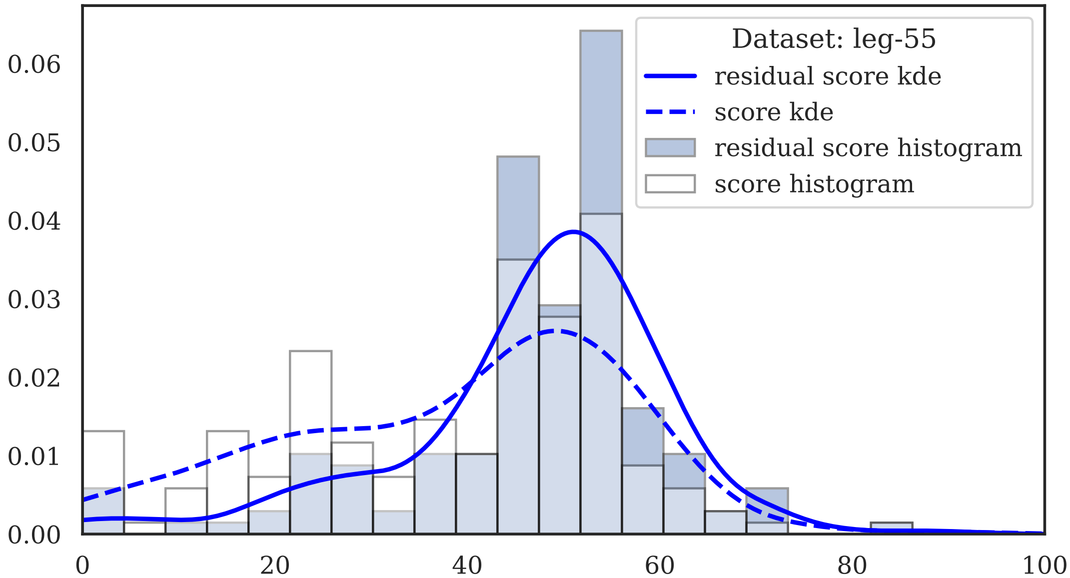

Figure 16: The estimators of densities of the score and residual score distributions.

The estimator of Guerre et al. (2000) is based on the first order conditions, and so it does not guarantee monotonicity of the strategy. However, a natural extremum estimator will have this property by construction:

$$\widehat{\sigma}(\theta) = \inf \arg\max_{s} \left(\widehat{\mathcal{H}}(s) \cdot (\theta - s) \right).$$

This estimator is convenient because it is aligned with our theoretical knowledge, and because extremum estimators are among the most studied and well-understood in the literature, see Newey and McFadden (1994) for an extensive overview.

4.2 Second-stage smoothing and boundary correction

The second step in our estimation is recovery of the joint distribution of (θ, q).

If we were to blindly follow the recipe in Guerre et al. (2000), we would use the estimated strategy $\hat{\sigma}$ to produce a sample of estimated pairs $(\hat{\theta}_{ij}, q_{ij})$, from which the density of the distribution of (θ, q) could be obtained using a 2-dimensional kernel density estimator. The reason why we cannot do this is that only the part of the distribution of (c, q) which maps into the interior bids, i.e. above the reserve price, can be correctly identified.

Definition 13. *Denote the sample of interior bids by D_I and its size by M_I:*

$$D_I = \{(b_{ij}, q_{ij}) | b_{ij} < r\}.$$

To estimate the part of the distribution of (c, q) that is nonparametrically identified, we find it easier to change the order or actions in the Guerre et al. (2000) algorithm. Namely, we will first apply a 2-dimensional kernel smoothing to D_I and only then use the estimated strategy $\hat{\sigma}$ to recover the relevant part of the distribution of (c, q). We will also put an additional effort to estimate the density at the boundary $b = r$, using *reflection method* advocated in Hickman and Hubbard (2015).

Definition 14. *Denote the reflected sample of interior bids by D_{IR}:*

$$D_{IR}(\tau) = \{(b_{ij}, q_{ij}) | b_{ij} = r - \tau(\tilde{b}_{ij}, q_{ij}), (\tilde{b}_{ij}, q_{ij}) \in D_I\},$$

where

$$\tau(y, q) = y + d(q)y^2 + 0.55 \cdot d^2(q)y^3, \tag{22}$$

$$d(q) = \frac{\log(\varphi(h_1, q)) - \log(\psi(h_0, q))}{h_1}, \tag{23}$$

$$\phi(h_1, q) = \frac{1}{M_I^2} + \frac{1}{M_I h_1 h_q} \sum_{D_I} K\left(\frac{h_1 - (b_{ij} - r)}{h_1}\right) K\left(\frac{q_{ij} - q}{h_q}\right), \tag{24}$$

$$\psi(h_0, q) = \max\{\frac{1}{M_I^2}, \ \frac{1}{M_I h_0 h_q} \sum_{D_I} K_0\left(\frac{b_{ij} - r}{h_0}\right) K\left(\frac{q_{ij} - q}{h_q}\right)\}, \tag{25}$$

and K_0 is the endpoint kernel:

$$K_0(u) = (6 + 18x + 12x^2) \cdot \mathbb{I}\{-1 \leqslant u \leqslant 0\}.$$

The idea behind the reflection method is that when the data is censored at a certain threshold, we can anticipate the unobserved part of the data assuming smoothness of the original distribution.

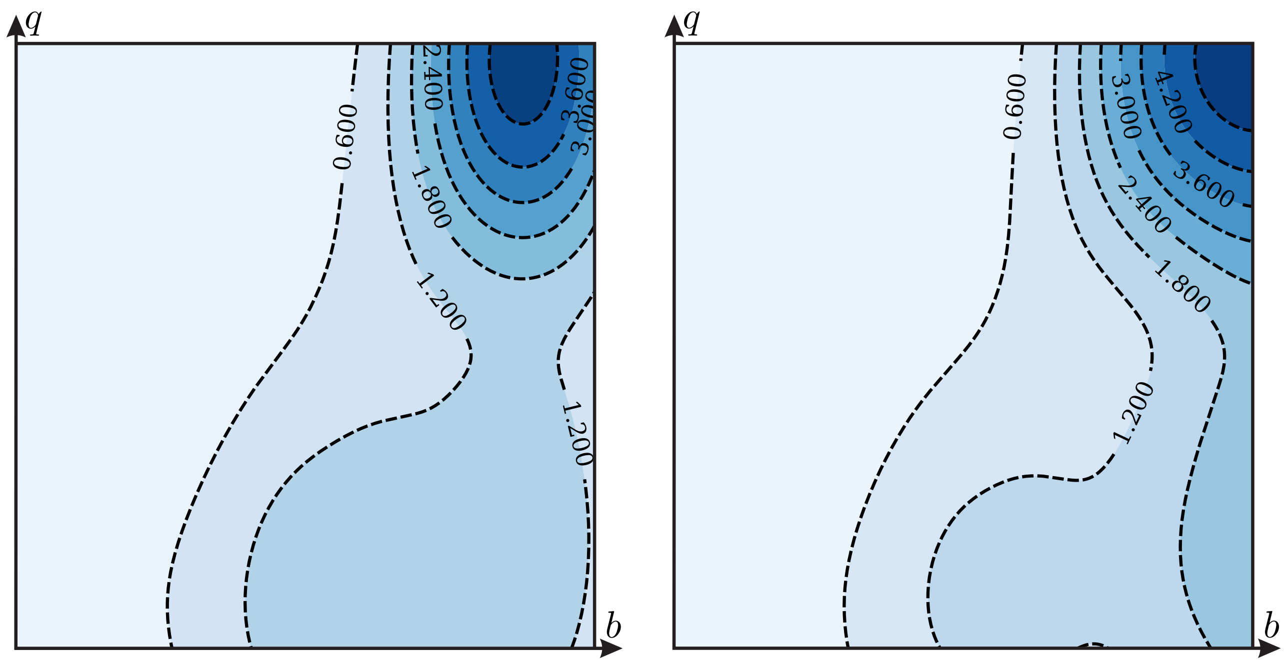

Figure 17: Estimator of $g(b, q)$ before and after boundary correction for Leg-55.

Definition 15. *Denote the densities of (b, q), and $(\theta, \alpha(q))$ by g and $\widetilde{f}$ respectively.*

Below is a boundary corrected kernel density estimator of the joint distribution of (b, q):

$$\widehat{g}(b, q) = \frac{1}{2M_I h_b h_q} \sum_{D_I \cup D_{IR}(\tau)} K\left(\frac{b_i - b}{h_b}\right) K\left(\frac{q_i - q}{h_q}\right),$$

where the main bandwidths are taken in accordance with the 'Silverman rule':

$$h_q = 3.16 \cdot sd(q) \cdot M_I^{-1/5}, \quad h_\theta = 3.16 \cdot sd(\theta) \cdot M_I^{-1/5}$$

and the bandwidths used in the construction of τ are:

$$h_1 = h_\theta \cdot M_I^{-1/20}, \quad h_0 = 1.48 \cdot h_1.$$

The practical consequence of not doing boundary correction is that the density would unnaturally decrease near the boundary ($b = r$), as seen on the left side of Figure 17. This would directly violate our knowledge of the equilibrium behavior, which predicts that firms bid essentially as if there is no reserve price and therefore the density should be smooth at the boundary.

We do not perform boundary correction for the quality dimension, on purpose, to oversmooth the spike in the neighborhood of $q = 100$, see Figure 3. If the estimated distribution is not smooth enough, there is a risk of non-existence of a continuous strategy.

Finally the relevant part of the joint distribution of (θ, q) can be obtained from the estimated joint distribution of (b, q) using the standard density transformation formula:

$$\textit{density of } (\theta, q) = \begin{cases} k \cdot \frac{\sigma'(\theta)}{\beta(q)} \cdot g(r - \frac{\sigma(\theta)-\alpha(q)}{\beta(q)}, q) & , \ 0 \leqslant \frac{\sigma(\theta)-\alpha(q)}{\beta(q)} \leqslant r \\ * & , \ \text{otherwise} \end{cases},$$

where $*$ stands for the unknown value of the density in the non-identified part of the distribution, and k stands for the constant to normalize the density. We will have to find a way to fill in the missing part of the distribution, which will be discussed in the next section.

4.3 Extrapolation and counterfactuals.

The last thing that impedes our ability to simulate equilibria in a counterfactual scoring rule is the part of the distribution of (θ, q) that is not identified. This fundamental problem stems from the censoring of the $\sigma(\theta)$ strategy due to the presence of the reserve price.

It is important to understand at this point that multiple densities $\widetilde{f}(\theta, q)$ will be consistent with the data. We will pick one such density due to its particular simplicity. Namely, we will say that the the density at the type $\theta < \sigma^{-1}(\alpha(q))$ is equal to the density at $\sigma^{-1}(\alpha(q))$, which is the closest type for which it is identified. The equilibrium strategy serves as a boundary in the $(\theta, \alpha(q))$ coordinates to the types that choose the reserve price, therefore the level lines of the distribution are horizontal to the left of the strategy, see Figure 18.

This gives us the following estimator for the whole range of $\theta \in [\alpha(q), \alpha(q) + \beta(q)r]$:

$$\widehat{\widetilde{f}}(\theta, a) = k \cdot \frac{\widehat{\sigma}'(\theta)}{\beta(\alpha^{-1}(a))\alpha'(\alpha^{-1}(a))} \cdot \widehat{g}(\min(r, r - \frac{\widehat{\sigma}(\theta) - a}{\beta(\alpha^{-1}(a))}), \alpha^{-1}(a)),$$

where k normalizes the density.

Once the density is estimated, for a new counterfactual scoring rule, we can solve the differential equation (8) using any suitable numerical integration method, such as, for example, Runge-Kutta, to obtain the new $\sigma(\theta)$ strategy and the new distribution $\mathcal{G}(\rho(\theta, q))$.

There are two ways how the numerical integration may fail, see Section 3.2. The first is if the trajectory reaches the boundary of the anti-funnel, then a continuous strategy does not

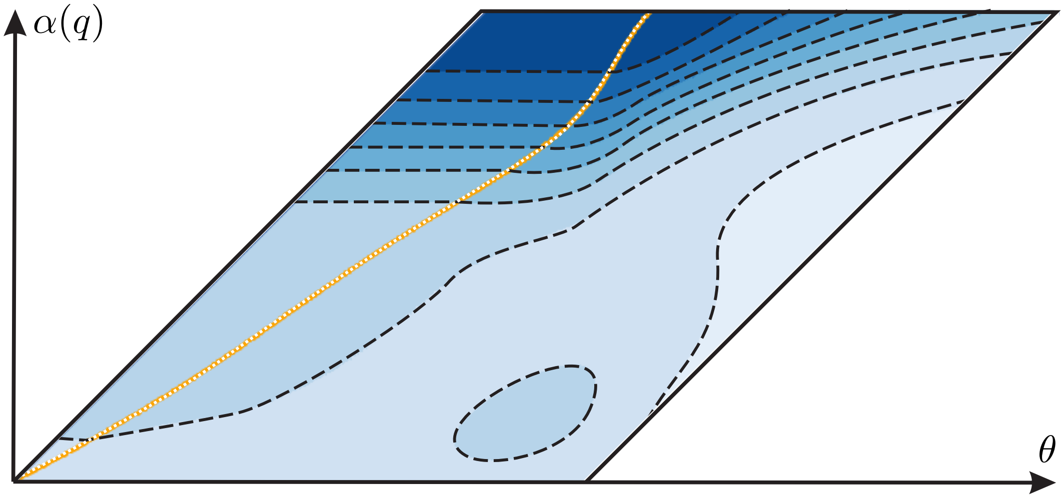

Figure 18: Estimator of the density of $(\theta, \alpha(q))$ after extrapolation for Leg-55.

exist, however, this never happens in our data. In the hypothetical situation where it happens, increasing the smoothness of the estimated distributions should generally help. The second is if the trajectory passes through the point where $\widetilde{F}_1' = 0$. In this case, starting from the level of quality q^* at which it happened, all bidders choose the reserve price. As a result, the interim rebate for all $q > q^*$ is zero and the interim quality can be calculated according to the distribution of probabilities of winning, based on the fact that highest quality wins.

Finally, the welfare properties of the scoring rule are captured by two functions: $Q(\theta, q)$ and $R(\theta, q)$, calculated using formulas in Corollary 3. The ex-ante versions of quality and rebate are then computed using the estimated density of $(\theta, \alpha(q))$, see Corollary 4.

5 Heterogeneity

Typically in the literature on structural estimation, an auction model is complemented with a separate model of heterogeneity, which instructs how to prepare the data before passing it to the main estimation routine. A common approach is to represent the value (in our case, the cost) of the bidder as either a sum or a product of two components: common and idiosyncratic. The common component is then filtered out using the observed auction characteristics. Such approach was taken, for example, in Haile et al. (2003) and Krasnokutskaya (2011).

In this paper, we take a different approach by relying on the reserve price as a control for heterogeneity. Precisely, we will assume that the bidder's cost is a product of an auction-specific component and a bidder-specific component:

$$c_{ij} = c_i \gamma_j, \quad r_j = r\gamma_j,$$

where γ_j represents the latent scale of the contract, and r_j is the reserve price in that auction. Crucially, the reserve price should follow the same scaling pattern.

Later in this section, we will show that, in this framework, the bids can be effectively homogenized by just normalizing them by the reserve price. But before that, it is important to understand why the reserve price is traditionally avoided as a control for the scale of the auction, and why our situation is different.

5.1 Reserve price as a control

In the classic value auctions, for which the bulk of empirical methods were developed, the reserve price is often set ad-hoc and is even sometimes entirely missing. That is because the benefits of fine-tuning the reserve price are very limited. In fact, a theoretical argument in Bulow and Roberts (1989) shows that it is more profitable to simply attract an additional bidder rather then set an optimal reserve price. Another explanation might be that costs of acquiring the statistical information needed for the optimal mechanism are prohibitively high. All of this makes the reserve price an unreliable instrument for economic analysis.

In procurement the situation is slightly different. First of all, since there is no natural upper bound to the bid (in value auctions bids are typically non-negative) the reserve price is always present, to cut the contractee's losses. Second, because the procurement contracts are often very complex and not all cost-relevant information can be conveyed in the tender documentation, the contractee might want to use the reserve price as a signaling device to help the firms estimate their costs. The reserve price then can be thought of as a publicly observed first bid. This signaling role is so important that it was, in fact, institutionalized by the Russian authorities, see Section 2.1

5.2 Homogenization of bids

Assume that, in any given auction j, the score, on top of satisfying all the assumptions of the model, has a particular shape:

$$s(b, q \mid r) = \tilde{s}(b/r, q),$$

which is true for the linear scoring formula that we have in our data. We can write down the equilibrium conditions for a representative auction with reserve price r in the Lagrangean form as in (26). It is easy to see that the equilibrium conditions stay unchanged if the same linear transformation $\psi_j(x) = \gamma_j x$, where $\gamma_j > 0$, is applied to all of the three variables c, b, and r:

$$\mathcal{H}\left(s(b, q \mid r)\right) \cdot (b - c) + \lambda(r - b) \to \min_{\lambda \geqslant 0} \max_{b}, \tag{26}$$

$$\mathcal{H}\left(s(\psi_j(b), q \mid \psi_j(r))\right) \cdot (\psi_j(b) - \psi_j(c)) + \lambda(\psi_j(r) - \psi_j(b)) \to \min_{\lambda \geqslant 0} \max_{b}. \tag{27}$$

Indeed, the same bidding strategy $b^*(c, q)$ is the solution to both (26) and (27), which means that the equilibrium bidding strategy in an auction that is scaled by γ_j, is also scaled by γ_j. In other words, the same equilibrium strategy $\sigma(\theta)$ can explain the behavior in multiple heterogeneous auctions as long as they are scalable to a single representative auction with a fixed reserve price, which is our model of heterogeneity.

Since in our data all the bids are divided by the reserve price inside the score, this effectively eliminates the impact of γ_j on the equilibrium behavior. Consequently, the model can be estimated as if there is no auction-specific heterogeneity.

Corollary 5. *Consider a model of heterogeneity, where for a firm i in auction j:*

$$s_j(b, q) = \tilde{s}(b/r_j, q), \quad c_{ij} = c_i\gamma_j, \quad r_j = \gamma_j r, \quad q_{ij} = q_i,$$

where (c_i, q_i) are drawn as in a symmetric model. Then the equilibrium strategies $\sigma(\theta)$ and $\sigma^*(\theta, q)$ do not depend on the scale of the auction γ_j.

While theoretically appealing, our approach to homogenization of bids has to yet prove itself successful in the data. To test it, we will randomly pick two bids (without return) from each auction and plot the scatterplot of logarithms of the two bids, and a scatterplot of the two bids normalized by the reserve price, see Figure 19 for the Leg-55, Sci-55 and Sec-80 datasets.

It can be seen with a naked eye that while the log-bids indeed require additional homogenization, it is far less obvious for the normalized bids, as the reserve price absorbs a significant portion of heterogeneity.

	before normalization		*after normalization*	
Dataset	Kendall	Pearson	Kendall	Pearson
Edu-80	0.86	0.97	0.27	0.39
Edu-55	0.85	0.97	0.23	0.29
Sci-80	0.81	0.94	0.21	0.29
Sci-55	0.79	0.94	0.01	0.03
Leg-80	0.68	0.91	0.2	0.13
Leg-55	0.82	0.94	0.21	0.13
Tec-80	0.81	0.94	0.33	0.5
Tec-55	0.86	0.97	0.27	0.39
Sec-80	0.89	0.99	0.36	0.52
Sec-55	0.78	0.97	0.11	0.4

Table 7: Correlation in bids before and after normalization.

The quality of homogenization varies across the datasets, see Table 7, but since there is no clear rule on how much correlation among bids is acceptable, we leave the data as it is.

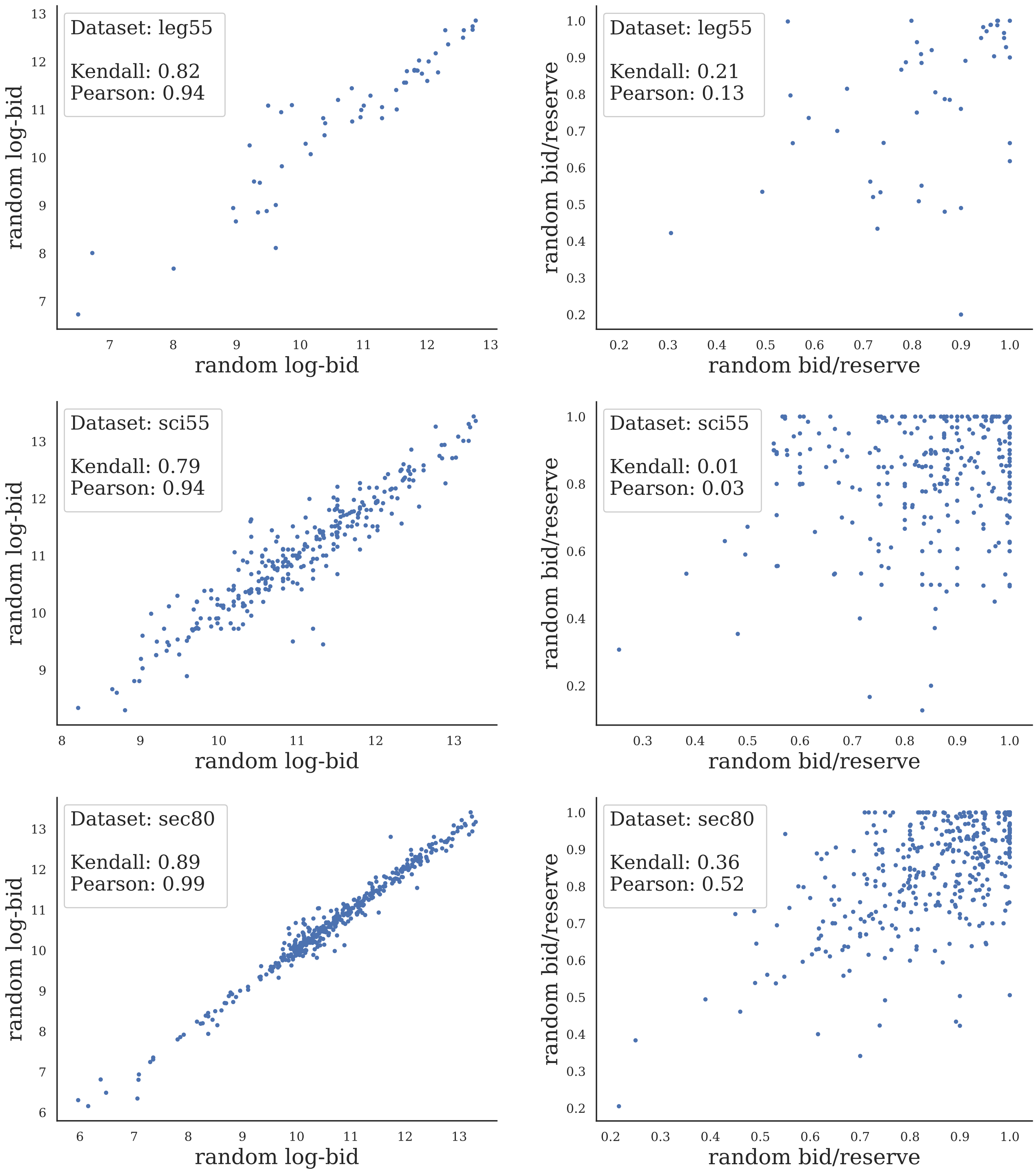

Figure 19: Normalization of bids.

6 Empirical results

For each of the ten datasets, and both scoring designs, I have traced the corresponding frontiers in the space of expected rebate and expected quality (both measured from 0 to 100), spanned by the weight in the scoring formula. For each dataset, I produced a figure that contains these frontiers, as well as the coordinates of the default linear scoring rule and its primary competitor: the price-quality-ratio ($s(b, q) = -b/q$) scoring rule.

I will focus below on the results for the Leg-55 dataset, characterized by the diagram in Figure 20. The diagram contains two arks, representing the welfare frontiers spanned by the two families: linear and log-linear.

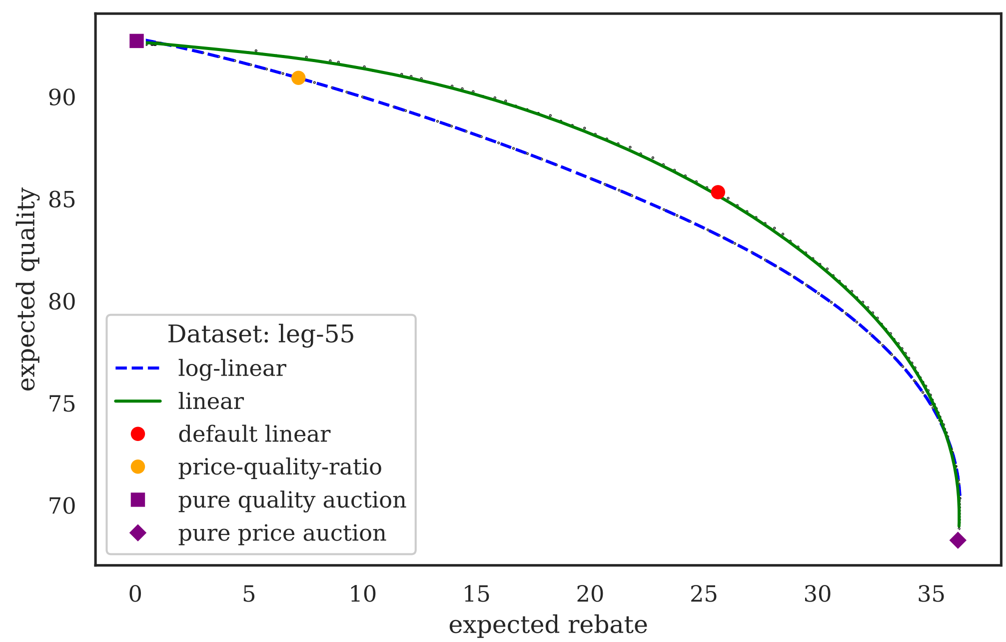

Figure 20: Welfare frontiers for the linear and log-linear families.

Several observations can be made.

(a) The arcs meet at the end points.

(b) The arcs are decreasing.

(c) The default linear rule is to the bottom-right from the price-over-quality rule.

(d) The linear design arc lies above the log-linear one.

All four observations are true for each of our datasets, see Figure 21. We will now explain the relevance of these findings for the main economic question addressed in the current paper.

Observation (a) should not be a surprise since the pure price auction and the pure quality auction can be considered as polar in the spectrum of scoring auctions produced by varying the weight in the scoring formula. The figure then confirms our basic intuition that for the weight approaching 0 or ∞ the two scoring designs become indistinguishable in terms of welfare.

Observation (b) is also anticipated, since the weight in the scoring formula is an intuitive control for the trade-off between quality and rebate. The figure confirms that no two members of the same scoring family can be ranked without the knowledge of designer's preferences. In other words, we can only hope to achieve an unambiguous improvement by switching between the linear and log-linear designs.

Observation (c) demonstrates the failure of the traditional approach to mechanism choice, when a default linear design is compared to a single member of the log-linear design. Indeed, a switch to price-over-quality would increase expected quality by approximately 10 points, but would decrease the expected rebate by approximately 30 points. Whether this is a beneficial exchange depends completely on how the contractee values quality in monetary terms, which can not be inferred from the auction data.

To quantify our findings, we compare the default linear scoring rule to the quality-equivalent member of the log-linear family. This can be interpreted as switching to the price-over-quality scoring rule, with a specially calibrated measure of quality to guarantee the same expected quality. Our results are summarized in Table 8, and can also be inferred from the frontiers in Figure 21.

dataset name	switch to price-over-quality	switch to log-linear with same expected quality
Edu-80	–26.42 % rebate and +11.99 % quality	-0.90 % rebate
Edu-55	–2.79 % rebate and +0.28 % quality	-2.10 % rebate
Sci-80	–27.59 % rebate and + 13.59 % quality	-1.16 % rebate
Sci-55	–5.90 % rebate and + 1.04 % quality	-3.12 % rebate
Leg-80	–14.50 % rebate and +4.25 % quality	-2.21 % rebate
Leg-55	–18.45 % rebate and +5.60% quality	-3.94 % rebate
Tec-80	–47.11 % rebate and +19.24 % quality	-1.31 % rebate
Tec-55	–4.11 % rebate and +0.61 % quality	-2.69 % rebate
Sec-80	–9.72 % rebate and + 6.10 % quality	-0.41 % rebate
Sec-55	–5.24 % rebate and +1.14 % quality	-1.80 % rebate

Table 8: Quantitative results

Finally, observation (d) is the main qualitative result in our paper. It appears that, for every member of the log-linear family, a member of the linear family dominates it in both welfare dimensions. In other words, even without knowing the designer's preferences for quality and rebate, we can tell that his utility, as long as it is a nondecreasing function of expected quality and rebate, is maximized at one of the linear scoring rules. Thus the linear design can be considered as superior to the log-linear one.

7 Conclusion

In this paper, I have investigated whether the linear scoring design is better than its main competitor, price-quality ratio, in terms of welfare, and received a strong positive answer.

My methodological contribution to the literature on scoring auctions is a class of models that are sufficiently flexible to feature both linear and log-linear scoring rules, yet they hold enough structure to produce a very tractable solution. These models are built on two core assumptions: exogenous quality and explicit reserve price. We have also developed a complete machinery for the estimation of these models and simulation of counterfactuals.

My empirical contribution is a qualitative result about the superiority of the linear scoring design. From Figure 20, as well as from every other simulation, it appears that every possible log-linear scoring rule is dominated by some linear scoring rule. Moreover, this finding does not depend on the exact knowledge of designer's preferences over quality and rebate, which makes it easy to be interpreted as policy advice.

Implicit in this assessment is the risk-neutrality of the designer's preferences, and so it is not surprising that the linear design succeeded over the log-linear. Had the preferences been risk-averse, the log-linear scoring rule could be advantageous.

Several avenues for future research are suggested.

Endogenous quality. My analysis of the affine scoring rule can be partially extended to the situation where quality is endogenous. This model can be thought of as part of a two-stage game, where the quality is chosen at the first stage, and is considered exogenous afterwards. However, to find an equilibrium in this game, more complicated tools, such as the ones developed in Hanazano et al. (2016), will be required.

Discontinuous strategies. Though I focus on the most regular case of a continuous strategy, situations may arise in practice where such equilibrium will not exist. In this case, the non-linear equation (17) may be considered as a heuristically derived characterization of the equilibrium. In fact, there is nothing in this equation, that particularly requires continuity.

Other scoring rules. While it might seem that the next logical step is to consider a scoring rule of an arbitrary shape, as in Hanazano et al. (2016), the empirical application calls for an even more general model where the scoring rule additionally depends on certain statistics of other firm's bids and qualities. This model would be useful for understanding the complicated scoring rules such as the "average bid" used in Italy and the "minimal bid" used in Russia.

8 Appendix A

Proof of Lemma 1

Assume that for some $\hat{\theta}, \hat{q}$, it is true that $\sigma^*(\hat{\theta}, \hat{q}) \neq \max(\sigma(\hat{\theta}), \alpha(\hat{q}))$, which means:

$$\sigma(\hat{\theta}) = \sigma^*(\hat{\theta}, \underline{q}) \leqslant \alpha(\hat{q}) < \sigma^*(\hat{\theta}, \hat{q}).$$

This can only be the case if $\sigma^*(\hat{\theta}, \hat{q})$ is a local maximum of a continuous function $\pi(s)(\hat{\theta} - s)$. Moreover, this function has to stay constant all the way between $s_1 = \sigma^*(\hat{\theta}, \underline{q})$ and $s_2 = \sigma^*(\hat{\theta}, \hat{q})$, otherwise the actual score $\sigma^*(\hat{\theta}, q)$ would be discontinuous in q somewhere in the range between $\underline{q}$ and $\hat{q}$, thus contradicting the continuity assumption. Staying constant, on the other hand, contradicts the assumption that a lower score is chosen when indifferent.

Proof of Proposition 1

The first step in the proof is to establish several important properties of the $\sigma(\theta)$ strategy. Clearly, $\sigma(\theta) = \min_q \sigma^*(\theta, q)$ is continuous, because $\sigma^*(\theta, q)$ is continuous in both arguments. It is also weakly increasing, by super-modularity of $\pi(s)(\theta - s)$, see Milgrom and Shannon (1994).

To see that $\sigma(\underline{\theta}) = \underline{\theta}$, note first that the actual distribution of scores $s = \alpha(q) + \beta(q)(r - b)$ is constrained by $\alpha(\underline{q}) = \underline{\theta}$ from below, as no firm is allowed to bid above the reserve price. The best (unconstrained) response $\sigma(\theta)$ therefore can not go below $\underline{\theta}$ as that would induce an upwards deviation for the lowest type. On the other hand, $\sigma(\theta)$ can not go above θ as that would guarantee a negative profit with positive probability for any type $\theta > \underline{\theta}$. Consequently, by continuity, the $\sigma(\theta)$ function at the lowest type $\underline{\theta}$ is confined to a single possible value $\underline{\theta}$.

To prove strict monotonicity, assume that the true range of the $\sigma^*(\theta, q)$ function is $[\underline{\theta}, \overline{s}]$. The $\pi(s)$ function is strictly increasing in this range by continuity of $\sigma^*(\theta, q)$ and $F(c, q)$. Therefore, the $\sigma(\theta)$ strategy is strictly increasing for all types that map into $(\underline{s}, \overline{s})$, see Milgrom and Shannon (1994). In other words, it can only be flat at two segments: one adjacent to $\underline{\theta}$ and one adjacent to $\overline{\theta}$. Typically, flat strategies are incompatible with symmetry in classic first-price auctions, however, due to the 2-dimensional firm characteristic and the wedge between σ and σ^*, eliminating such behavior requires slightly more work.

First, assume that there is a segment of types adjacent to $\underline{\theta}$ that maps into $\underline{\sigma}$. That would mean that, for all types in that segment such that $\theta > \underline{\theta}$, the actual score σ^* is equal to $\alpha(q) > \alpha(\underline{q}) = \underline{\theta}$ after truncation. Consequently, before truncation, the perceived probability of winning for all those types would be zero, which is worse than if they chose any other score above $\underline{\theta}$ but below their own type θ. Second, assume that there is a segment adjacent to $\overline{\theta}$, that maps into $\overline{s}$. That would mean that $\overline{s}$ is the maximum value of both σ and σ^*. Consequently, for a positive measure of (θ, q) the actual score would be $\overline{s}$, and, therefore there will exist at least one (θ, q) that would deviate. The $\sigma(\theta)$ strategy is therefore strictly monotone in the whole range of types.

Once we have invertibility of σ, we can construct the pseudo-type $\rho(\theta, q) = \sigma^{-1}(\sigma^*(\theta, q))$ and proceed with the direct mechanism approach in Section 3.3. Using the residual pseudo-type distribution $\mathcal{G}$ we can characterize $\sigma(\theta)$ as a solution to the optimality conditions below:

$$\sigma(\theta) \in \arg\max_{\rho}(\theta - \sigma(\rho))\mathcal{G}(\rho). \tag{28}$$

The necessary first order conditions can be written as a system of equations:

$$\sigma \mathcal{G}' + \sigma' \mathcal{G} = \theta \mathcal{G}', \quad Z\mathcal{G}'G = G'\mathcal{G}, \quad G' = \widetilde{F}_1' + \widetilde{F}_2'\sigma',$$

that can be uniquely solved, assuming that $\widetilde{F}_1' > 0$:

$$\sigma' = \frac{(\theta - \sigma)\widetilde{F}_1'}{ZG - (\theta - \sigma)\widetilde{F}_2'}, \quad G' = \frac{ZG\widetilde{F}_1'}{ZG - (\theta - \sigma)\widetilde{F}_2'}, \quad \mathcal{G}' = \frac{\mathcal{G}\widetilde{F}_1'}{ZG - (\theta - \sigma)\widetilde{F}_2'}.$$

Recalling that $G(\theta) = \widetilde{F}(\theta, \sigma(\theta))$, we obtain our differential equation:

$$\sigma' = \frac{(\theta - \sigma)\widetilde{F}_1'}{Z\widetilde{F} - (\theta - \sigma)\widetilde{F}_2'}.$$

The second order conditions are satisfied as $\sigma(\theta)$ strategy is a conditional expectation of the highest pseudo-type below θ, see Section 3.3 for a discussion. If the $\sigma(\theta)$ trajectory passes through the point where $F_1' = 0$, it will switch to a different law of motion: $(\theta - \sigma)\widetilde{F}_2' = Z\widetilde{F}$, however, under Assumption 5 this never happens, see Section 3.2 for a discussion.

It remains to show that coupled with the initial condition $\sigma(\underline{\theta}) = \underline{\theta}$, the differential equation has a unique solution in the range $[\underline{\theta}, \overline{\theta}]$, for which we will employ the theory of fences and funnels from Hubbard and West (1997).

An *anti-funnel* consists of two functions of time (in our case θ), called upper fence and lower fence. Both functions should be continuously differentiable and such that for each point on the fence, the slope of the vector field passing through is lower than the slope of the upper fence and greater than the slope of the lower fence at that point. The obvious choice of the upper fence is $\sigma = \theta$, because the vector field has zero slope there. For the lower fence a natural candidate is the $\sigma = \delta(\theta)$ curve:

$$\delta(\theta) = \{\sigma \in [0, \theta] : Z(\theta, \sigma)\widetilde{F}(\theta, \sigma) - (\theta - \sigma)\widetilde{F}_2'(\theta, \sigma) = 0\},$$

which is well defined under Assumption 5 and the vector field has an infinite slope there.

The coordinates (θ, σ) are inconvenient for the formal proof of uniqueness and existence, because the right handside of the differential equation is formally not defined at the lower fence. But this is only a notational problem. Assume for simplicity that $\underline{\theta} = 0$ and make a change of coordinates:

$$t = (\theta + \sigma)/2, \quad x = (\sigma - \theta)/2,$$

which corresponds to a 45 degrees clockwise rotation of the vector field around the $(0, 0)$ point. The new differential equation $x' = f(t, x)$ is now well-defined at the boundary of the new anti-funnel. Moreover, since all the functions α, β, F are assumed to be twice continuously differentiable, so does $\widetilde{F}$ and $Z(\theta, \sigma)$. Therefore, $f(t, x)$ is continuously differentiable, for all points inside and on the boundary of the anti-funnel. And since the new anti-funnel is a compact, $f(t, x)$ is uniformly bounded together with its first derivatives.

To establish existence we apply *Theorem 4.7.1 (Fence Theorem)* and *Theorem 1.4.4 (Anti-Funnel Theorem: Existence)* which only require that $f(t, x)$ satisfy continuity in t and Lipshitz condition in x inside the anti-funnel, which follows from continuous differentiability of f.

To establish uniqueness we apply *Theorem 4.7.5 (Second Uniqueness Criterion For Anti-Funnels)* which additionally requires that the anti-funnel is narrowing at the initial point (which is true in our case) and that inside the anti-funnel the following property holds:

$$\frac{\partial}{\partial x} f(t, x) \geqslant w(t), \quad \int_a^b w(t)dt > -\infty,$$

which follows from the fact that $\frac{\partial}{\partial x} f(t, x)$ is uniformly bounded there.

Proof of Proposition 2

We start from the direct mechanism derived in the proof of Proposition 1:

$$\sigma(\theta) \in \arg\max_{\rho}(\theta - \sigma(\rho))\mathcal{G}(\rho). \tag{29}$$

To apply the envelope argument, we use *Theorem 2* from Milgrom and Segal (2002). The premises of this theorem are satisfied since the profit function in (29) is continuously differentiable in θ, and its partial derivative with respect to θ is uniformly bounded. Since σ is strictly monotone (see proof of Proposition 1), for any θ there exists a q such that $\rho(\theta, q) = \theta$, and therefore the envelope conditions can be written as:

$$(\theta - \sigma(\theta))\mathcal{G}(\theta) - (\underline{\theta} - \sigma(\underline{\theta}))\mathcal{G}(\underline{\theta}) = \int_{\underline{\theta}}^{\theta} \mathcal{G}(z)dz. \tag{30}$$

Finally, since $\sigma(\underline{\theta}) = \underline{\theta}$ (see proof of Proposition 1), we obtain our formula:

$$\sigma(\theta) = \theta - \frac{1}{\mathcal{G}(\theta)} \int_{\underline{\theta}}^{\theta} \mathcal{G}(z)dz. \tag{31}$$

9 Appendix B

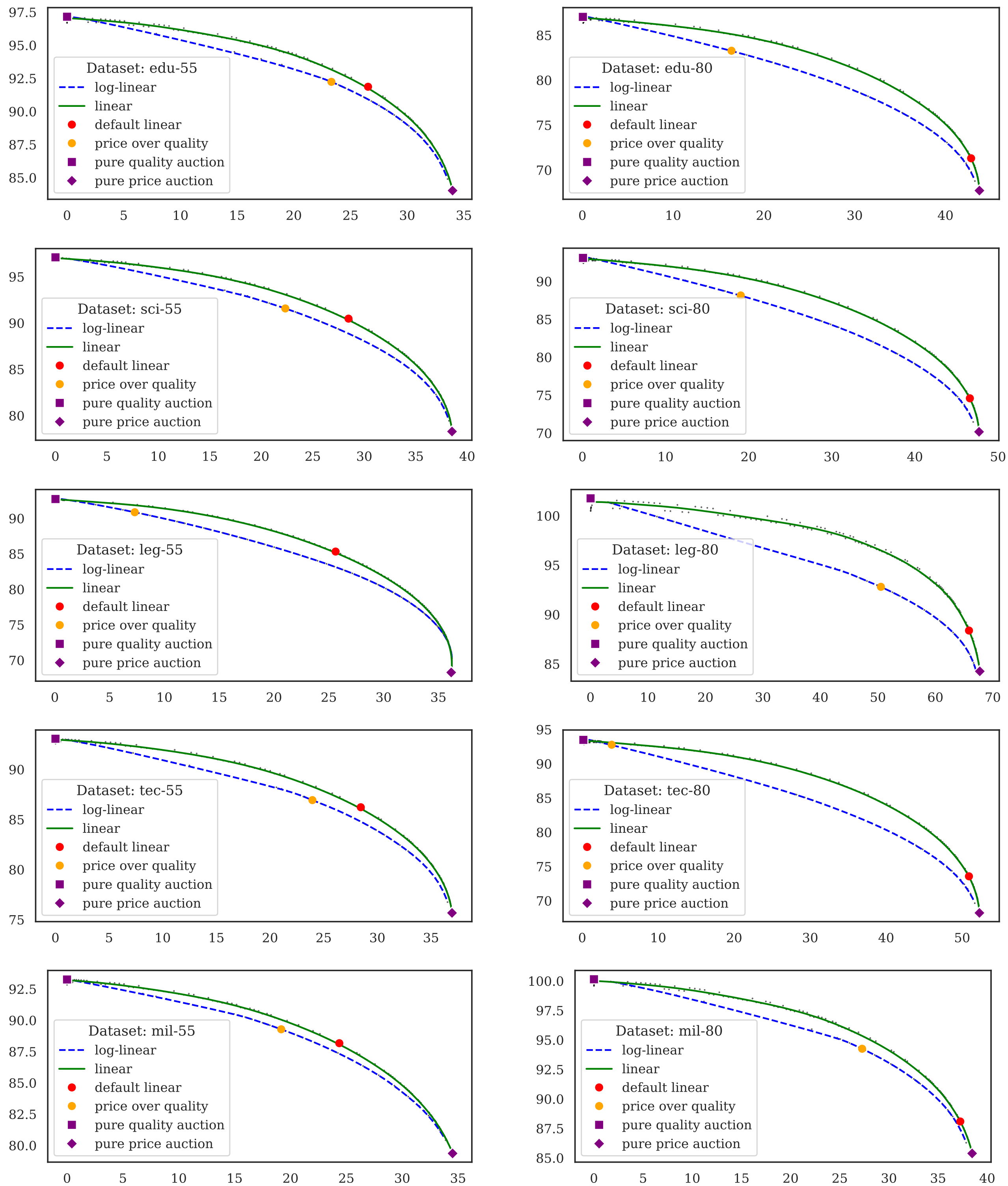

Figure 21: Welfare frontiers in the space of expected rebate and expected quality.

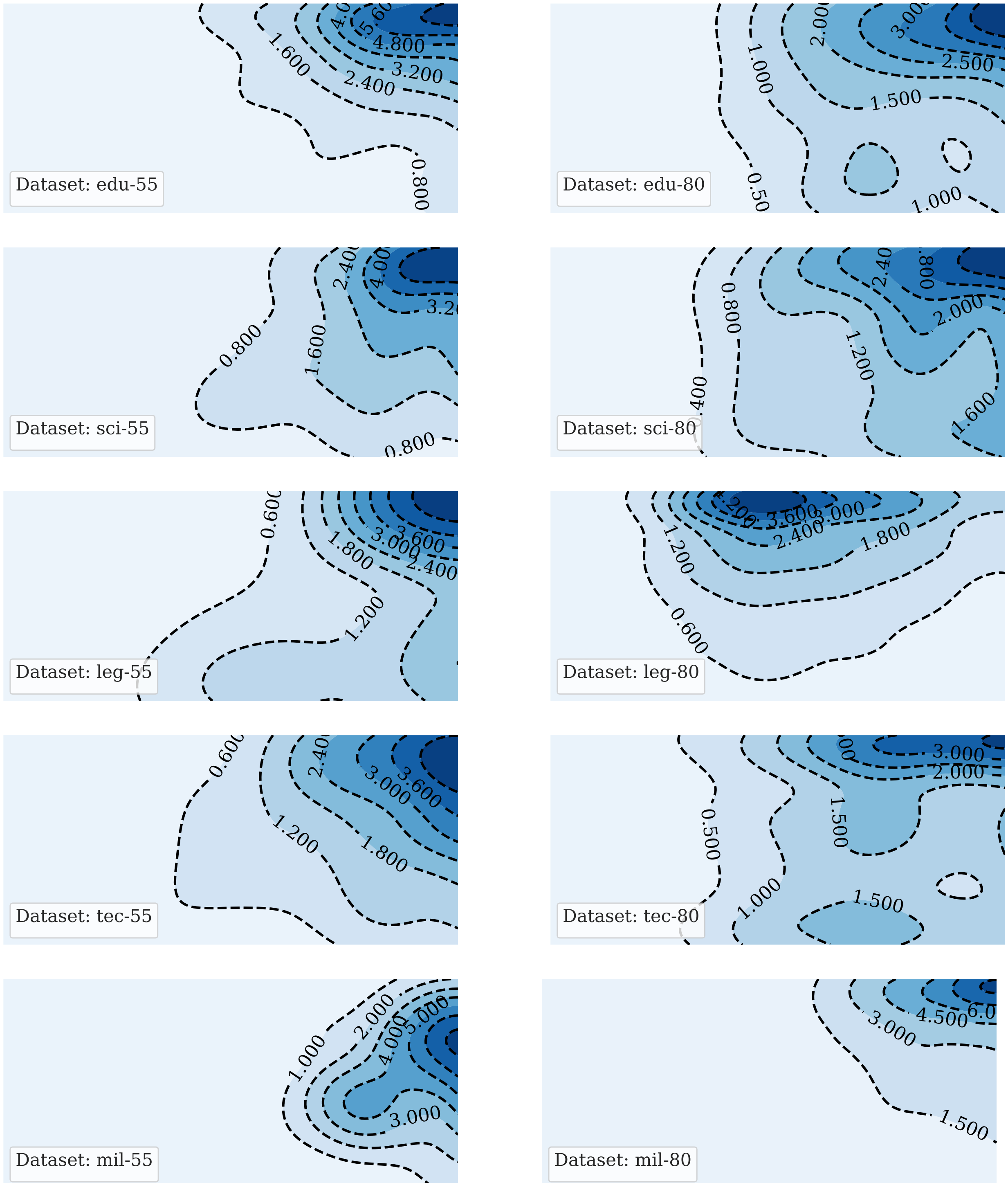

Figure 22: Joint distribution of bids and quality.

References

Adani, R. C. M., Valbonesi, P. et al. 2016. Favoritism in scoring rule auctions.Technical report, Dipartimento di Scienze Economiche" Marco Fanno".

Albano, G. L., Dini, F., Zampino, R. 2008. Bidding for complex projects: Evidence from the acquisitions of it services.

Asker, J., Cantillon, E. 2008. Properties of Scoring Auctions. The RAND Journal of Economics, 39, 69–85.

Asker, J., Cantillon, E. 2010. Procurement when price and quality matter. The RAND Journal of Economics, 41, 1–34.

Branco, F. 1997. The design of multidimensional auctions. The RAND Journal of Economics, 28, 63–81.

Bulow, J., Roberts, J. 1989. The simple economics of optimal auctions. Journal of Political Economy, 97, 1060–1090.

Che, Y.-K. 1993. Design Competition Through Multidimensional Auctions. The RAND Journal of Economics, 24, 668–680.

Dastidar, K. 2014. Scoring auctions with non-quasilinear scoring rules.

Dastidar, K. G., Mukherjee, D. 2014. Corruption in delegated public procurement auctions. European Journal of Political Economy, 35, 122–127.

Dini, F., Pacini, R., Valetti, T. 2006. Scoring Rules. In Handbook of Procurement, Cambridge University Press, Chap. 12.

Guerre, E., Perrigne, I., Vuong, Q. 2000. Optimal nonparametric estimation of first-price auctions. Econometrica, 68, 525–574.

Haile, P. A., Hong, H., Shum, M. 2003. Nonparametric tests for common values at first-price sealed-bid auctions. Working Paper 10105, National Bureau of Economic Research.

Hanazano, M., Hiroze, Y., Nakabayashi, J., Tsuruoka, M. 2016. Theory, identification, and estimation for scoring auctions.

Härdle, W. 2012. Smoothing techniques: with implementation in S. Springer Science & Business Media.

Hickman, B. R., Hubbard, T. P. 2015. Replacing Sample Trimming with Boundary correction in nonparametric estimation of first-price auctions. Journal of Applied Econometrics, 30, 739–762.

Huang, Y. 2016. Essays on procurement, scoring auction, and quality manipulation corruption. Ph.D. dissertation.

Hubbard, J., West, B. 1997. In Differential Equations: A Dynamical Systems Approach. Ordinary Differential Equations, Springer Science.

Koning, P., van de Meerendonk, A. 2014. The impact of scoring weights on price and quality outcomes: An application to the procurement of welfare-to-work contracts. European Economic Review, 71, 1 – 14.

Krasnokutskaya, E. 2011. Identification and estimation of auction models with unobserved heterogeneity. The Review of Economic Studies, 78, 293–327.

Krasnokutskaya, E., Seim, K. 2011. Bid Preference Programs and participation in highway procurement auctions. American Economic Review, 101, 2653–86.

Krishna, V. 2010. Auction Theory. Elsevier, 2nd edition.

Lewis, G., Bajari, P. 2011. Procurement Contracting With Time Incentives: Theory and Evidence. The Quarterly Journal of Economics, 126, 1173–1211.

Li, T., Perrigne, I., Vuong, Q. 2002. Structural estimation of the affiliated private value auction model. The RAND Journal of Economics, 33, 171–193.

Marion, J. 2007. Are bid preferences benign? the effect of small business subsidies in highway procurement auctions. Journal of Public Economics, 91, 1591 – 1624.

Milgrom, P., Segal, I. 2002. Envelope Theorems for Arbitrary Choice Sets. Econometrica, 70, 583–601.

Milgrom, P., Shannon, C. 1994. Monotone comparative statics. Econometrica: Journal of the Econometric Society, 157–180.

Molenaar, K., Yakowenko, G. 2007. Scoring Rules. In Alternative project delivery, procurement, and con- tracting methods for highways, American Society of Civil Engineers, Chap. 12.

Nakabayashi, J. 2013. Small business set-asides in procurement auctions: An empirical analysis. Journal of Public Economics, 100, 28 – 44.

Newey, W., McFadden, D. 1994. Large sample estimation and hypothesis testing. In Handbook of Econometrics, 4, Elsevier, Chap. 36.

Nishimura, T. 2015. Optimal design of scoring auctions with multidimensional quality.

Takahashi, H. 2014. Strategic design under uncertain evaluations: Theory and evidence from design-build auctions.