CIRJE-F-1018

Auctions For Complements –An Experimental Analysis

Daniel Marszalec
The University of Tokyo

July 2016

Auctions For Complements - An Experimental Analysis*

Daniel Marszalec†

July 8, 2016

Abstract

Abstract: I evaluate the performance of four static sealed-bid package auctions in an experimental setting with complementarities. The valuation model comprises two items, and three bidders: two 'local bidders demand one item only, while the third (global) bidder only wants both. The rules I compare include the Vickrey and first-price auctions, Vickrey Nearest Rule and the Reference Rule. Auction-level tests find the first-price auction revenue dominant overall, while the Vickrey auction performs worst; the other two rules rank intermediate. Bidder-level tests of the experimental data reject the competitive equilibrium bidding functions: overbidding is widespread in all four auctions, and bidders are averse to submitting boundary bids. I also observe behaviour consistent with collusive bidding in the Vickrey auction. Contrary to theoretical predictions, the Vickrey auction performs worst on efficiency, primarily for this reason.

JEL Classification: D44, C90

Keywords: Auctions, Experimental Economics, Core-Selecting Auctions, Collusion

*I would like to thank Vince Crawford and Paul Klemperer, as well as Ian Jewitt,Thees Spreckelsen and Alex Teytelboym for their comments and criticisms. Financial support from the British Academy, the ESRC, the Royal Economic Society, and CIRJE is also gratefully acknowledged.

†Graduate School of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan. Email: dmarszalec@e.u-tokyo.ac.jp

The growth in popularity of auctions has seen them applied to an ever wider range of markets, including markets with multiple packages and complementarities. A stylized example of such a situation is an auctioneer selling a jacket and a pair of trousers. Some buyers may only want the jacket, others may only need the trousers, but some customers may want a complete suit, and thus prefer to buy both. The fact that the two garments match creates additional value for the buyer who wants both - this is the complementarity. More complex demand patterns of a similar kind are present in the auctions for mobile telephony spectrum, contracts for serving bus routes or airport take-off and landing slots.[1] To deal with this increased complexity, a new class of mechanisms, called core-selecting auctions, have been developed and implemented, though our understanding of their incentive properties is still incomplete. I conduct a bidding experiment to evaluate the performance of two static core-selecting auctions (the Vickrey Nearest and the Reference Rule) against two older alternatives (the Vickrey and first-price auctions).

The motivation for picking the Vickrey and first-price auctions is that they cover two extremes in terms of bidder incentives. In the Vickrey auction truthful bidding is a dominant strategy, while the first-price auction gives strong incentives for bidding below value. Both auctions also embody well-known theoretical weaknesses, which have limited their use in practice: the Vickrey auction may generate low revenue, and the first-price auction can be inefficient. A key motivation behind the use of core-selecting rules is that they should generate outcomes which are the "best of both worlds," with efficiency better than in first-price, and revenue higher than in Vickrey auctions.[2] To achieve this aim, the core-selecting rules partially de-couple bidders' payments from their own bids (to encourage close to truthful bidding), while requiring that the payments lie in the core (thereby reducing the likelihood of low-revenue outcomes).

My main finding is the strong performance of the simplest of the four rules, the first-price auction: it is revenue-dominant without losing efficiency. I cannot reject rev-

[1]On mobile spectrum, see Danish Business Authority (2012), ComReg (2012) and Ofcom (2012). The auction of London bus routes is discussed in Cantillon and Pesendorfer (2006). An auction solution to allocating landing slots is discussed in Federal Aviation Administration (2008).

[2]Sun and Yang (2006, 2009) have also proved that in the setting of my paper, there exists a dynamic incentive-compatible mechanism which finds the competitive equilibrium. In the present experiment, I only consider one-shot sealed-bid auctions, and thus do not include this mechanism in my comparison.

enue equivalence between the remaining three auctions. The Vickrey auction is least efficient, and no significant efficiency difference emerges between the first-price and the core-selecting rules.

At the bidder level, I test the experimental data against the Bayesian Nash equilibrium bidding functions for all four rules, as derived by Ausubel and Baranov (2010). The theory is not supported by my experiment, and overbidding is frequent in each auction. In the core-selecting auctions, when bidders' behaviour diverges from equilibrium, they do not revert to a truth-telling rule-of-thumb. Instead they attempt to game the rule to their advantage, albeit unsuccessfully. I also find evidence of attempted collusion in the Vickrey auction, which can explain the low revenue and efficiency of this auction. In the first-price auction when bidders deviate from theoretical equilibrium, they do so in predictable ways that do not undermine efficiency or revenue.

The simplest auction is thus most robust in my experiment, and the attractive properties of the core-selecting rules are not fully borne out when bidders' behaviour deviates from expectation. Recently, many real-world package auctions have used complex core-selecting designs, without giving much attention to first-price rules. Against this backdrop, my results invite a re-consideration of the merits of the humble first-price package auction as a viable and easy to understand alternative, which warrants further research.

Recent experimental auction literature has focused on dynamic auctions, such as the combinatorial-clock, and simultaneous ascending auctions.[3] This strand of research has been primarily concerned about efficiency properties of those auctions, and how bidders select packages in settings with complex valuation patterns. However, many practical implementations of such dynamic designs feature a one-shot static auction as their final phase: for example, the Danish, Irish and UK spectrum auctions in 2012, all used a Vickrey-Nearest type rule to determine the final prices and allocations of licences, after a dynamic auction had been used to determine the relevant packages.[4] My work is naturally seen as investigating how these static rules perform, given that a selection of packages has already been set. At the time of writing, there had been no prior experimental work in

[3]Kagel, Lien and Milgrom (2010 and 2014), and Kazumori (2010) are good examples of this.
[4]See ComReg (2012), Danish Business Authority (2012), and Ofcom (2012).

this area.

The rest of the paper is structured as follows. The auction rules and valuation model are introduced in Section 1, and the precise formulation of the hypotheses which I test are discussed in Section 2. The experimental setup is presented in Section 3, and Section 4 performs a quality check of the data. Auction level results and hypothesis tests are presented in Section 5, while bidder-level analysis is conducted in Section 6. Section 7 discusses the interpretation of the results, and Section 8 concludes.

1 Auction Setup and Rule Descriptions

My model consists of three bidders and two items, sold simultaneously. I label the items as '1' and '2', and assume that two of the bidders have a positive valuation on one item only. These are the 'local' bidders, and I label them as L1 and L2, corresponding to which item they value positively. The third bidder, G - the 'global' bidder - has a positive value only on the bundle of 1 and 2 together, and zero value on 1 and 2 individually. Each bidder is only permitted to bid on the bundle they value positively, so the auctioneer always receives three bids.

To model complementarity, I assume that the local bidders' values are drawn from a uniform distribution on [0,100], while the global bidder's value is drawn from a uniform distribution on [0,200]. I will use b_{L1} to denote the bid of bidder L1, b_{L2} for the bid of bidder L2, and b_G for the bid of global bidder G. The auction rule itself is described by $P(b_{L1}, b_{L2}, b_G)$, a payment vector conditional on the bid-triplet (b_{L1}, b_{L2}, b_G). Individual payments assigned by an auction mechanism to the three bidder types are labelled as p_{L1}, p_{L2} and p_G, such that $P(b_{L1}, b_{L2}, b_G) = (p_{L1}, p_{L2}, p_G)$.

Prior to calculating the bidders' payments, the auctioneer solves a winner-determination problem: he picks a feasible bid-maximising allocation such that each item gets assigned to at most one bidder. In the present setting there are only two sensible allocations.[5] If the sum of local bids is higher, the L-types win one item each; otherwise the G-type wins

[5]More allocations are feasible, but not really 'sensible': for example, only selling one item is feasible, but not sensible. Aggregate revenue could be increased by offering the unsold item at a price $\varepsilon > 0$. If a bidder's value on this item is positive, we have a Pareto improvement.

both.[6] The winner-determination procedure is common to all the rules I analyse.

1.1 The Vickrey Auction

The multi-unit Vickrey Auction, an extension of the standard Vickrey-Clark-Groves mechanism to the auction context, has the main aim of inducing truthful value revelation amongst the bidders. This, in turn, enables the implementation of an efficient value-maximising allocation. Irrespective of bidder type, in the Vickrey auction the price paid by each winning bidder is determined solely by the bids of the other two bidders. This price is calculated such that each bidder receives a payoff equal to the incremental surplus they bring to the auction.

For a numerical example, let $(b_{L1}, b_{L2}, b_G) = (48, 40, 60)$. Bidders L1 and L2 win an item each, as the sum of their bids exceeds J's bid. The surplus that bidder L1 brings to the system is 28: without L1's bid, the auctioneer only faces the bids of $b_j = 60$ and $b_{i2} = 40$, whereby G would win both items, and the surplus - evaluated at the bidders' bids - would be 60. With L1's bid of 48, L1 and L2 win instead, and the total surplus is 88 - an increase of 28. To give L1 a surplus of 28, the payment must solve the equation $b_{L1} - p_{L1} = 28 \implies p_{L1} = 48 - 28 = 20$. By similar calculations, L2's payment is $p_{L2} = 12$.

To generalise the above reasoning, and after imposing a non-negativity constraint on prices, the Vickrey auction payments can be written as:

$$P^{VA}(b_{L1}, b_{L2}, b_G) = \begin{cases} (VP_{L1}, VP_{L2}, 0) & if \quad b_{L1} + b_{L2} \geq b_G \\ (0, 0, b_{L1} + b_{L2}) & if \quad b_{L1} + b_{L2} < b_G \end{cases} \tag{1}$$

$$where \quad : \quad \begin{array}{l} VP_{L1} = \max[(b_G - b_{L2}), 0)] \\ VP_{L2} = \max[(b_G - b_{L1}), 0)] \end{array}$$

There are two well-known problems with the Vickrey auction, which limit its practical usefulness: the possibility of low revenue, and susceptibility to collusion. From equation (1) we see that in the case when $b_{L1} + b_{L2} > b_G$ with $0 < b_{L1} < b_G$ and $0 < b_{L2} < b_G$,[7]

[6] Ties are broken randomly.

[7] This case corresponds to the situation where L1 and L2 together out-bid G, but neither of the local bids, on their own, would be sufficient to out-bid the global bidder.

the Vickrey auction 'leaves money on the table', in that $p_{L1} + p_{L2} < b_G$: the seller has a seen a bid that exceeds the sum of payments he receives from the winning bidders. This is equivalent to saying that the Vickrey auction outcomes frequently lie outside the core.

The core is defined as a set of allocations for which there exists no blocking coalition, such that no group of members of the system can jointly deviate to a different allocation which gives all those members a higher surplus. In the present example, the group consisting of bidder G and the auctioneer constitutes a blocking coalition: G could offer the auctioneer a payment of $\tilde{p}_G = p_{L1} + p_{L2} + \varepsilon < b_G$, with $\varepsilon > 0$. This increases the auctioneer's revenue, and gives G a non-zero profit - so the allocation that assigns the items to L1 and L2 is not a core allocation, and the price-triplet $(p_{L1}, p_{L2}, 0)$ does not lie in the core.[8]

When $b_{L1} + b_{L2} > b_G$, the set of core payments can be defined as:

$$(p_{L1}, p_{L2}) \in \{(x, y) \,|x + y \geq b_G, x \in [0, b_{L1}], y \in [0, b_{L2}]\}.$$

This is the set of payments such that neither L1 or L2 pays more than their bid, but the sum of their payments weakly exceeds the bid of G. This set, along with the bids and Vickrey payments are shown in Figure 1; the core is shaded in gray. The dotted diagonal line denotes the 'minimum revenue line', which contains all the points where the payments of L1 and L2 equal the payment of G exactly. The bold segment of this diagonal line depicts the 'minimum revenue core' (MRC),[9] which contains the points that are simultaneously in the core, and on the minimum-revenue line. The MRC depicts the combination of the lowest amounts that each of the L-types can bid, subject to them jointly out-bidding the G-type. From the seller's viewpoint, this is analogous to a 'second-price' in a single-unit auction: this is the highest observed bid after the actual winning bids have been removed..

[8]In the case when G wins the Vickrey payment is in the core, as then $b_j > b_{i1} + b_{i2}$.

[9]For a further detailed discussion of the MRC, see Day and Milgrom (2008).

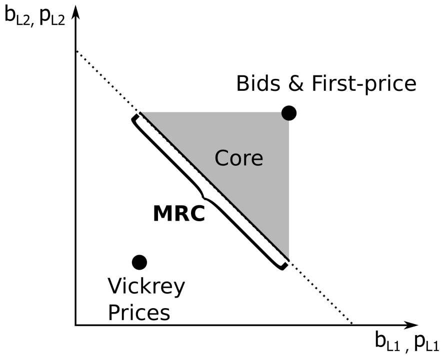


Figure 1: Vickrey prices, first-price payments and the MRC

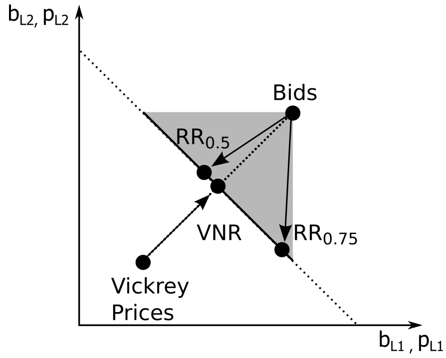


Figure 2: Vickrey Nearest, and Reference Rule with $\alpha = 0.5$ and $\alpha = 0.75$

The second weakness of the Vickrey auction is its susceptibility to collusion. We see from equation (1) that when L1 and L2 win, the payment of one is decreasing in the bid of the other. [10] If L1 and L2 behave cooperatively, they can both bid aggressively, which will reduce their joint payments. To collude perfectly L1 and L2 can both bid $b_{i1} = b_{i2} = 200$ - the highest possible value that G can have. Such bids makes sure that L1 and L2 always win, and both pay a price of 0. In less extreme cases, if both bidders overbid, they can still induce payments that are lower than their Vickrey prices under truthful bidding.

1.2 The First-Price Auction

The first-price auction, usually used for the sale of a single item, can be naturally extended to cover the case of package bidding. After the winner-determination problem has been solved, each winning bidder pays their bid in full. The payments in the first-price auction are:

$$P^{FP}\left(b_{L1}, b_{L2}, b_G\right) = \begin{cases} \left(b_{L1}, b_{L2}, 0\right) & if \quad b_{L1} + b_{L2} \geq b_G \\ \left(0, 0, b_G\right) & if \quad b_{L1} + b_{L2} < b_G \end{cases}.$$

Unlike the payments in the Vickrey auction, in the first-price auction the winners' payments are always in the core, as shown in Figure 1. In the case when L1 and L2 win, the first-price payments will also always lie (weakly) above the minimum-revenue line.

[10] Consequently the Vickrey auction revenue is not always monotonic in bids: it is possible that an auciton with higher (individual) bids can lead to lower revenue.

Despite its simplicity, the first-price auction with package bidding has been successfully used in practice, including the auctioning of bus routes in London (see Cantillon and Pesendorfer, 2006) and mobile telephony spectrum in Norway in 2013. [11]

1.3 The Vickrey Nearest Rule

The Vickrey Nearest Rule (VNR) is currently the most widely used of the core-selecting auction rules. One motivation behind these payment rules is to increase the revenue from Vickrey-type auctions while retaining most of their efficiency and truth-telling properties. Such a trade-off is achieved by making the winners' payments less dependent on their own bids, but still requiring that the payment vector lies in the core.[12] The VNR auction, as introduced by Day and Cramton (2012), first uses the submitted bids to calculate Vickrey prices, and then picks a price vector that minimises the Euclidian distance to the Vickrey payments subject to the prices being in the core.

In the case when bidder G wins, the Vickrey payment is in the core already, and VNR implements that payment. If L1 and L2 win, the VNR will select the point on the MRC which is closest to the Vickrey payment vector, as shown in Figure 2.

Mathematically, finding the point on the MRC that is closest to the Vickrey payments involves taking an orthogonal projection of the bid vector onto the MRC. I label the outcome of such a projection as the 'preliminary shares' of bidders L1 and L2, and denote them as s_{L1} and s_{L2}. The VNR payments then are:

[11] Information taken from the Norwegian Post and Telecommunications Authority document "800, 900 and 1800 MHz auction - Auction Rules" (2013).

[12] The intuiton is that if incentives to deviate from truth-telling are small, bidders will bid in a near-truthful way, which would mitigate efficiency losses due to misallocation.

$$P^{VNR}(b_{L1}, b_{L2}, b_G) = \begin{cases} (s_{L1}, s_{L2}, 0) & if \quad \begin{matrix} b_{L1} + b_{L2} \geq b_G, \ and \\ s_{i1}, s_{i2} > 0 \end{matrix} \\ (b_G, 0, 0) & if \quad b_{L1} \geq b_G + b_{L2} \\ (0, b_G, 0) & if \quad b_{L2} \geq b_G + b_{L1}, \\ (0, 0, b_{L1} + b_{L2}) & if \quad b_{L1} + b_{L2} < b_G \end{cases} \tag{2}$$

$$where \quad : \quad \begin{matrix} s_{L1} = \frac{1}{2}(b_{L1} + b_G - b_{L2}) \\ s_{L2} = \frac{1}{2}(b_{L2} + b_G - b_{L1}) \end{matrix} \tag{3}$$

The payments of local bidders in the VNR are broken down into three cases, depending on the asymmetry of the bids. If, say, $b_{L1} > b_G + b_{L2}$, so that L1 on his own out-bids G by a large margin, then $s_{L2} < 0$, which implies a negative price for L2. By the non-negativity constraint on prices, we then truncate $p_{L2} = 0$, and $p_{L1} = b_G$ to remain on the MRC. The converse case applies if $b_{L2} > b_G + b_{L1}$. When the asymmetry moderate and $s_{L1}, s_{L2} > 0$, both bidders pay their preliminary share. [13]

1.4 The Reference Rule Auction

The Reference Rule, introduced by Erdil and Klemperer (2010), is another payment rule for core-selecting package auctions. The motivation behind the rule is to make it more robust to small local deviation incentives than the VNR by further de-coupling local bidders' payments from bids. In the VNR, local bidders can influence their payment share by influencing the Vickrey prices, which depend on their own bid, as shown in equation (3). The innovation behind the Reference Rule is to define the bidders' payment shares in a way that further reduces the dependence on their own bids, while maintaining the core-selecting property. This is achieved defining a 'reference point' which is independent

[13]My interpretation of the VNR rule is slightly different from that of Ausubel and Baranov (2010). Under my reading, the Vickrey prices towards which VNR projects are not bounded by zero from below; in their interpretation this zero-bound is imposed, prior to calculating the projection. In Ausubel and Baranov's terminology, my reading of the VNR makes it equivalent to what they call a "nearest bid" rule (because the un-bounded Vicrkey prices are symmetric about the MRC, relative to submitted bids). When Vickrey prices are positive, both interpretations pick the same point.

of the L-types' bids, and then selecting the final payments that are closest in Euclidian distance to that point.

I define each local bidder's reference price based on the bid of the global bidder G and a sharing parameter α; the corresponding Reference Rule is RR(α). The reference price of bidder L1 is $r_{L1} = \alpha \cdot b_G$, and the reference price for bidder L2 is $r_{L2} = (1-\alpha) \cdot b_G$, with $\alpha \in [0,1]$. By varying α the reference point can be moved smoothly along the minimum-revenue line, with higher α setting the reference point closer L1's axis. The bidder payments in the Reference Rule then are:

$$P^{RR(\alpha)}\left(b_{L1}, b_{L2}, b_G\right) = \begin{cases} (r_{L1}, r_{L2}, 0) & if \quad \begin{array}{c} b_{L1} + b_{L2} \geq b_G,\ and \\ r_{L1} < b_{L1}, r_{L2} < b_{L2} \end{array} \\ (b_G - b_{L2}, b_{L2}, 0) & if \quad \begin{array}{c} b_{L1} + b_{L2} \geq b_G,\ and \\ r_{L1} < b_{L1}, r_{L2} > b_{L2} \end{array} \\ (b_{L1}, b_G - b_{L1}, 0) & if \quad \begin{array}{c} b_{L1} + b_{L2} \geq b_G,\ and \\ r_{L1} > b_{L1}, r_{L2} < b_{L2} \end{array} \\ (0, 0, b_{L1} + b_{L2}) & if \quad b_{L1} + b_{L2} < b_G \end{cases} \tag{4}$$

$$where \quad : \quad \begin{array}{l} r_{L1} = \alpha \cdot b_G \\ r_{L2} = (1-\alpha) \cdot b_G \end{array}$$

Since reference prices are only required to lie on the minimum-revenue line, and not on the MRC, it is possible that the reference point will lie outside the core. Then the point on the MRC that is closest to the reference point is a payment vector where one local bidder (say, L1) pays his bid in full, while the other local bidder's payment makes up the difference (between G's and L1's bid).

In the VNR, each local bidder's payment share always depends in part on his own bid. In the Reference Rule, so long as the *realised* reference point is on the MRC, the payment for each local bidder is completely *in*sensitive to their own bid. The *only* case in which a local bidder's payment depends on his bid is in the situation when the realised reference point is outside the MRC *and* he is the bidder that has to pay his bid in full.

This sensitivity occurs only under certain realisation of bids, and hence has limited impact on average. [14]

In general, as Figure 2 shows, the Reference Rule with $\alpha = 0.50$ generates payments different from VNR.[15] However, with $\alpha = 0.50$, the reference payments are the same as they would be in the Proxy Rule auction of Ausubel and Milgrom (2002). Hence to make the Reference Rule look significantly different from the VNR and Proxy Rule auctions, I use $\alpha = 0.75$ in the main experiment. Supplementary data for the Reference Rule with $\alpha = 0.50$ was obtained from an additional experiment, which is described in the Appendix.

1.5 Comparison of the four Auction Rules

To give a concrete comparison of the four auction rules, Figure 3 summarizes the outcome from applying each rule to the bid-triplet $(b_{L1}, b_{L2}, b_G) = (48, 40, 60)$. The L-types win, and the G-type pays zero in every auction. To show the influence of varying α on the behaviour of the Reference Rule, I calculate the payments for three values of α. For RR(0.25) the reference prices will be $r_{L1} = 15$ and $r_{L2} = 45$, which is outside the core, so the Reference Rule payments will be truncated to lie on the boundary of the MRC. This is not the case for $RR\,(0.75)$, and the payments in that case are not in the corner of the core.

	p_{L1}	p_{L2}	Revenue
First-price	48	40	88
Vickrey	20	12	32
VNR	34	26	60
RR(0.75)	45	15	60
RR(0.50)	30	30	60
RR(0.25)	20	40	60

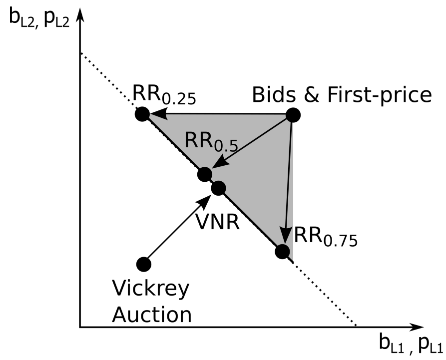


Figure 3: A numerical example of the four rules, with $(b_{L1}, b_{L2}, b_G) = (48, 40, 60)$

[14] Erdil and Klemperer (2010) show that under plausible conditions the Reference Rule has a lower sum of 'local deviation incentives' than VNR, while the sum of 'maximum deviation incentives' is unchanged. The proof proceeds by trading off the cases where bidders have zero incentives with those where incentives are maximal, and comparing these with the VNR, which has moderate incentives everywhere.

[15] The Reference Rule with $\alpha = 0.50$ generates reference payments on the mid-point of the minimum-revenue line, while the VNR selects payment shares at the mid-point of the MRC. Unless $b_{L1} = b_{L2}$, these two points will differ.

1.6 Bidding Restrictions and Collusion

None of the auctions I analyse require bidding above value in a competitive equilibrium, so in theory a restriction prohibiting such bids should have little bite. Investigating the impact of such restrictions is nonetheless worthwhile for two reasons. Firstly, even in simpler single-item auction contexts many experimental papers, such as Kagel and Levin (1995), find that overbidding is a frequent phenomenon. Bidders bid more than theory would predict, sometimes even above their value.[16] It is useful to gauge how such overbidding influences the performance of the rules examined here, and whether it is the driving force behind any revenue or efficiency findings.

The second reason for investigating bidding restrictions is that it allows me to look for collusion in the Vickrey auction. Here both individual profits as well as auction revenue are very sensitive to the presence of overbidding. For the other three auctions no collusive strategies have been found.[17] Running a set of sessions with the same instructions, with and without bidding restrictions, allows for a clean assessment of collusion.

2 Hypotheses

Testing competitive equilibrium bidding theory is the most direct application of auction experiments - thus I survey the relevant theory in Section 2.1. Yet even in simpler settings and when complementarities are absent, the experimental auction literature frequently rejects theoretical predictions.[18] In addition, the standard models do not consider collusion, an effect with potentially significant implications for practical auction performance. Hence I propose some additional intuitively plausible hypotheses in Section 2.2, which can also be tested on my data.

[16] For a good summary of this literature and further references, see Section 1.4 of Kagel & Levin (2008), and Section I.b2 in Kagel (1995).

[17] As of yet, there is no clear analysis as to the collusion incentives in VNR and the Reference Rule. The presumption is that being core-selecting auction rules, they should be robust to attempted collusion.

[18] Kagel (1996 and 2008) are a good overview of this literature.

2.1 Related Theory and Experimental Literature

Optimal bidding functions for the auctions I analyse, under an analogous valuation model, have been derived by Ausubel and Baranov (2010), Goeree and Lien (2009) and Sano (2010). I will refer to these bidding functions as the Bayesian-Nash equilibrium (BNE) biding functions. To obtain optimal bidding functions for the case of the first-price auction, Baranov (2010)uses numerical methods, since a solution cannot be found analytically; I do the same for the case of RR(0.75). Figure 4 shows that for local bidders, BNE bidding requires shading - bidding below value - in all auctions except Vickrey.

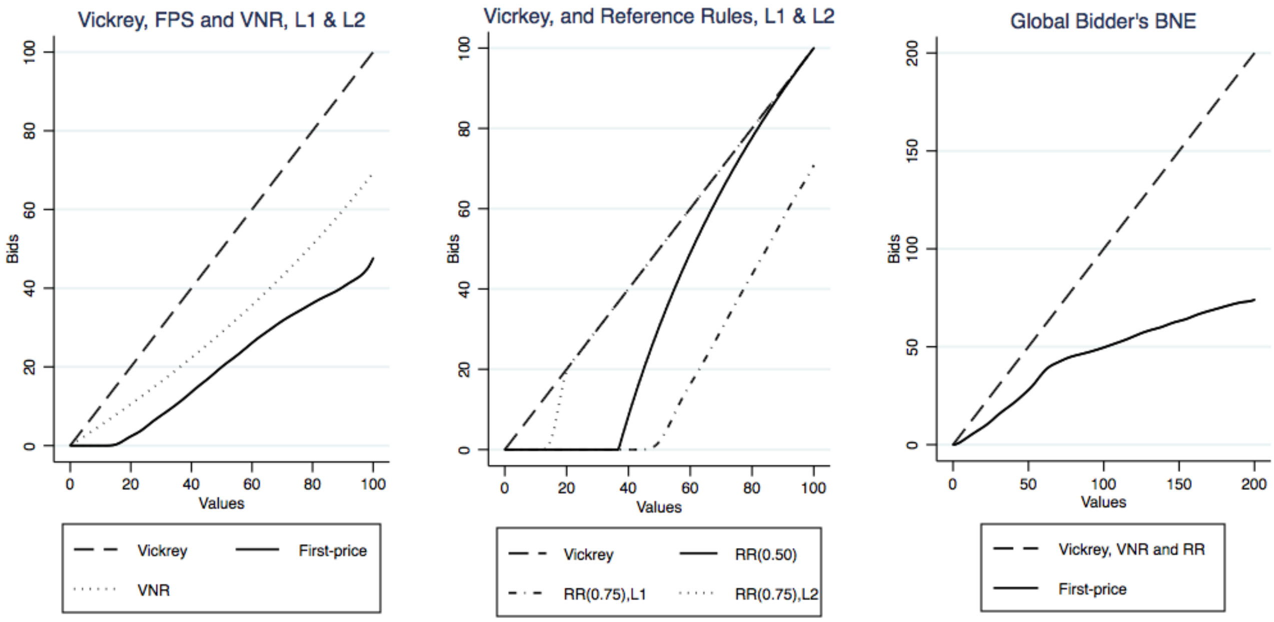


Figure 4: BNE Bidding Functions for local and global bidders. In all cases when the bidder's payment is above the Vickrey price, bidding below values occurs in equilibrium.

In both the first-price auction, and the both Reference Rules, local bidders with low values pool to bid precisely zero. In all these rules there is always a strictly positive marginal effect of the bid on the price, conditional on winning, when values are near zero. Thus a low-value local bidder has an incentive to free-ride on his co-bidder, and bid strictly zero. In VNR there is no such incentive for bidders with near-zero values because if a local bidder submits a very low bid, it is possible that his price conditional on winning is zero nonetheless.[19]

[19]This true in my interpretation of VNR, but not on the Ausubel and Baranov (2010) reading. Under

For the global bidder, the payment rule for all auctions except first-price is the same, and is equivalent to paying his Vickrey-price. Therefore, in the Vickrey auction, VNR and reference truthful bidding is a dominant strategy for the global bidder. In the first-price auction, the global bidder shades his bid below value considerably, as seen on Figure 4.

At the auction level, Ausubel and Baranov (2010) find that the Vickrey auction gives highest revenue, followed by the first-price auction, with VNR and Proxy Rule giving almost identical revenue, below the other two auctions. The efficiency ranking follows the same pattern as revenue.

Combining the findings of Ausubel and Baranov (2010) with the well-known prediction of truthful bidding in the Vickrey auction, I test the following set of theory-based hypotheses:

- Hypothesis HT: Bidders follow the competitive BNE bidding strategies.
- Hypothesis HR: The revenue ranking has Vickrey auction first, followed by first-price, with VNR and the RR(0.50) joint last.
- Hypothesis HE: The ranking for efficiency is the same as in HR.

The most relevant experimental work on package auctions, for my paper, are Kazumori (2010, 2014) and Kagel, Lien and Milgrom (2010, 2014). Kazumori (2010) investigates generalized Vickrey auctions, in addition to clock-proxy and simultaneous-ascending auctions. He finds that clock-proxy auctions out-perform the generalized Vickrey auction, and also outperform the simultaneous-ascending auction when the value structure mirrored exposure. Kagel et al. compare the performance of a combinatorial clock-auction with that of a simultaneous ascending auction for a variety of value and complementarity settings. Their interest is assessing how well the auctions perform when bidders bid only on a subset of profitable packages in each round, rather than bidding on all packages. They find that straightforward bidding - submitting bids a few most profitable packages only - leads to efficient outcomes (Kagel et al. 2010), though bidders sometimes diverge

their interpretation, the rule always projects towards (weakly) positive Vickrey prices, so there is always a positive marginal effect of bid on price; this re-establishes the pooling at zero result for low-value local bidders.

from such bidding patterns to push up prices for their competitors (Kagel et al. 2014). All these papers, however, have looked at dynamic auctions, with complicated value and complementarity structures, and their focus has been on efficiency and package-selection.

My work, in contrast, looks at static one-shot auctions, with a fixed package structure, and allows me to check whether in a simpler context the bidding will diverge from predictions once the package-selection aspect is removed.[20] In practice, in many high-value package auctions a hybrid design is used, where a clock phase is followed by a single supplementary bidding round which determines final prices and package allocation.[21] My research is thus a complement to, rather than a substitute for, the dynamic experimental auction literature.

2.2 Intuition-based Hypotheses

Even if bidders do not follow BNE strategies, they may still respond to auction incentives to some extent. It is thus worthwhile to assess the broader intuitions that could influence behaviour under the different rules.

In the Vickrey auction, every bidder's price conditional on winning is independent of their bid, while there is a partial dependence in the core-selecting rules. We should hence expect to see more aggressive bidding in the Vickrey than in the core-selecting auctions. In the first-price auction, conditional on winning the price equals the bid exactly, which we should expect to invite more cautious bidding. This ranking of incentives does not apply to the G-type bidders, who face the same payment rule under all auctions except first-price. Testing whether such bidders bid truthfully is contained in the hypothesis HT, but even if that hypothesis fails, it is possible that the G-types follow a similar non-truthful bidding pattern. I propose the following intuition-based hypotheses:

- Hypothesis HB: Local bidder types bid highest in the Vickrey auction, and submit lowest bids in the first-price auction. The Reference Rule and VNR rank interme-

[20] Kazumori (2014) has also conducted an experiment on one-shot package auctions, in a setting similar to mine, but his analysis only compares the Vickrey and Ausubel-Milgrom (2002) proxy auctions. He finds that proxy auctions revenue-superior, which is congruent with the results of this paper.

[21] The dynamic phase thus determines which packages are relevant, but does not necessarily fix the final allocation of packages to bidders.

diate.

- Hypothesis HG: Bidder G bids similarly in all auctions other than first-price.

In the discussions of Day and Cramton (2012) and Erdil and Klemperer (2010), part of the motivation for core-selecting auctions is that bidders may in fact not use full equilibrium strategies, but rather follow a rule-of-thumb. The VNR and the Reference Rule were developed to minimise incentives for deviation from truthful bidding. The intuition is that because payments are 'close to independent of own bids' then bidders could find it 'close to optimal' to bid truthfully. This intuition naturally generates another hypothesis:

- Hypothesis HA: Local bidders bid truthfully in the VNR and Reference Rule.

The final set of hypotheses I test relate to collusion in the Vickrey auction. Collusion can be defined as behaviour that deviates from an individually optimal competitive strategy towards one that aims to maximise joint profits of the colluding parties.[22] The general tendency in the collusion literature is to provide bidders in rich bidding contexts with many opportunities to collude, and look for periods of play when collusion is successfully sustained. Examples of this approach include Goswami, Noe and Rebello (1996) and Sade, Schnitzlein and Zender (2005), who look at collusion in discriminatory and uniform-price auctions with communication. Kwasnica and Sherstyuk (2007) similarly investigate Simultaneous Ascending Auctions with repeated play (within the same bidder group), but no communication. The survey of Kagel and Levin (2008) finds that repeated play with the same opponents, and communication, tend to facilitate collusion, though their survey does not cover any experiments on multi-unit Vickrey auctions.

In light of the above papers, the setup of my experiment is not inherently conducive to collusion: the matching is random across periods, and communication is prohibited. My experiment was the first auction study ever run at the laboratory I used, hence few of the participants are likely to have prior auction experience.[23] The valuation setup, however, is very simple and the Vickrey auction rules are straightforward, so the collusive strategies

[22] Playing a collusive strategy in itself is not necessarily non-equilibrium behaviour - in games where multiple equilibria exist, a 'collusive' outcome can be one of such equilibria.

[23] I cannot exclude the possibility that they would have participated in auction experiments elsewhere.

are easy to deduce: under perfect collusion, the L-types should bid exactly 200. Even if bidders do not notice this corner solution, it is possible that the L-types realise that they can mutually benefit by bidding significantly above value. None of the other auctions in the experiment give obvious incentives for bidding in excess of value, so I would not expect bidding behaviour to change much irrespective of whether a bidding restriction is in place or not. If we observe significant change of bidding patterns in the Vickrey auction across the two treatments, together with numerous bids in excess of value, these findings would be consistent with attempted collusion. I thus test the following hypotheses:

- Hypothesis HS: In auctions other than the Vickrey auction, the presence of bidding restrictions does not significantly affect bidding.
- Hypothesis HC: Removal of bidding restrictions in the Vickrey auction influences bidding behaviour. Without bidding restrictions the L-types bid more aggressively, and in excess of their value.

3 Experimental Design

The experiment was run over four sessions, and the participants were recruited from the population of Oxford graduate and undergraduate students via the mailing list at the Centre for Experimental Social Sciences (CESS) laboratory at the University of Oxford. Only students from science and social science subjects were included in the recruitment mailshot, and no participant was allowed to play in more than one session. The experiment itself was programmed using the zTree software of Fischbacher (2007), and run at the CESS laboratory. Sessions lasted up to two and a half hours, with average earnings of around £35 ($\approx$ \$55).[24]

During each session, the same group of participants played in each of the four auctions. The attendance was between 18 to 30 participants per session. After receiving the instructions for a given auction type, the participants were allowed to ask clarifying questions, and then were presented with an understanding test. Upon passing the test they

[24] A sample of the instructions is available in the Online Appendix.

participated in two payoff-irrelevant practice rounds, followed by the ten payoff-relevant rounds of the same auction rule. This design yielded 140 auction-round observations for each rule from the sessions without bidding restrictions and 160 auction-rounds with bidding restrictions present. The matching of participants to groups and bidder types was random each round, and communication was not permitted. Once the paying rounds of a given auction type were complete, the instruction sheets for that auction were collected, and the instructions for the next auction were distributed.[25]

A sample of the understanding test that the participants were required to complete is provided in the Online Appendix. The test was administered on paper, and there were few failures.[26]

To allow for an analysis of the importance of overbidding and possible collusion in the Vickrey auction, two of the four sessions were run with the bidding restrictions in place, prohibiting the bidders from bidding above value. In the other sessions the bidding restrictions were removed, and all three bidders were allowed to bid any number in $[0, 200]$.[27] The participants were paid for each auction rule based on their profits in two randomly selected rounds (out of the ten played); if the sum from these two rounds was negative, the payoff for that auction was truncated to zero. Final payments were calculated as the sum of payoffs from all four auction types, plus a show-up fee.

4 Verifying Data Quality

Since the experimental design is within subjects, I need to verify that bids are independent across auctions. To assess this degree of dependence, I ran a set of pairwise estimations of Kendall's τ correlation parameter and tested its significance.[28] None of the tests for L-type bidders reject a no-correlation null, with all p-values ¿ 0.15. The tests on the G-type

[25]The ordering of the auction rules was: [VCG,VNR,RR,FPS] in one set of sessions, and [VCG, FPS,RR,VNR] in another. These orderings were generated randomly, but for consistency the same pair of orderings was used in both restricted and unrestricted bidding sessions.

[26]On average, between one or two out of every thirty subjects failed the test.

[27]Bidders were made aware that under unrestricted bidding, though they would never pay more than their bid, they could end up with a negative payoff if they overbid and win at a price above their valuation.

[28]The purpose of this test is to check that the assumptions of the statistical test I use later are satisfied. While values are independent by design, I must check that the bidding process itself did not induce a strong pattern of dependece.

bidders also fail to reject the no-correlation null at the 95% level. These results suggest that there is little correlation between bidding pattern across auction types, and that the assumption of independence between treatments for testing purposes is acceptable.

In addition to the four sessions where bidders bid in all four auction rules, I also ran another set of experiments in an analogous setting, but focusing only on the effects of α in the Reference Rule; the details of these experiments are outlined in the Appendix.[29] Due to time-constraints (and participant fatigue), it was not feasible to run both $\alpha = 0.75$ and $\alpha = 0.50$ treatments in the main sessions. Since the data for RR(0.50) is available, I have included it in the comparisons for the present paper, though with the caveat that it is possible that participants' behaviour in RR(0.50) would be somehow influenced by their *not* playing in the other three auctions.

The supplementary experiment also contained a control treatment, where $\alpha = 0.75$. I can therefore compare the bidding patterns in the two experiments as a consistency check. Standard tests for differences between samples, however, do not reject a 'no difference' null, even at the 90% level.[30] These results suggest that the behaviour for the $\alpha = 0.75$ case is similar in both the main experiment as in the supplementary sessions, so the effects of presenting the Reference Rule in the two different settings are likely to be minor.

5 Auction-level Results

Revenue, surplus and efficiency are the three main parameters of interest for evaluating auction performance. Revenue is often of foremost importance to sellers, while bidders are primarily interested in their own surplus. From a welfare or policy point of view efficiency is also relevant, so that the items are allocated to the highest-value buyers.[31] One immediately visible characteristic of Table 1 is how distinct the first-price auction looks from the others under these criteria: the revenue is higher, surplus is lower, and

[29]The data collected in the supplementary experiment consisted of 140 auction-rounds for each rule - the same number as in the unrestricted bidding sample of the main experiment.

[30]The tests I used include the Mann-Whitney and Kolmogorov-Smirnov tests on the raw bid data, as well as direct tests of means and medians.

[31]Efficiecny here is calculated as: $100\% \cdot \frac{\text{sum of winning bidders' values}}{\text{sum of values under value-maximising allocation}}$

both variables have lower variance than in the other auctions.[32] Efficiency is high in all auctions except Vickrey, which is the only one with efficiency below 90%.

Table 1: Revenue, Efficiency and Surplus Summary. The first-price auction is revenue dominant, while the Vickrey auction is least efficient.

	Vickrey	First Price	VNR	RR(0.50)	RR(0.75)
revenue	67.6	91.5	68.2	77.0	71.1
	(56.9)	(37.1)	(41.2)	(42.3)	(46.3)
surplus	44.1	29.8	57.9	48.9	46.7
	(67.6)	(28.1)	(39.1)	(49.3)	(49.6)
efficiency (%)	88.9	97.5	97.7	94.9	95.1
	(22.2)	(8.4)	(9.1)	(13.8)	(12.8)

Means reported, standard deviation below. Revenue and surplus reported as points. The calculations are based on all 140 experimental auction rounds.

Results from the Vickrey auction, in Table 1 also show higher variability than corresponding figures for other auctions. This pattern is consistent with above-truthful bidding in the Vickrey auction: in this case, the local bidders may win, despite the global bidder having a higher value. Due to the Vickrey pricing formula (equation 1), if local bidders then win, prices and revenue will be low, and surplus high. If the local bidders overbid, but lose nonetheless, the price paid by the global bidder will be higher than in the truthful-bidding equilibrium, and surplus correspondingly lower. Though average surplus is not much lower in the Vickrey auction on average it is more variable, relative to other auctions.

The first-price auction revenue-dominates all other rules in pairwise median tests, as shown in Table 2. Pairwise comparisons between the Vickrey, VNR and Reference Rule cannot reject revenue equivalence. Though revenue in the Vickrey auction is lower than under VNR and Reference Rule, this difference is not statistically significant. I also cannot reject equivalence between the two kinds of Reference Rules with different values of α. This revenue ranking runs contrary to hypothesis HR, which I reject. The first-price auction performs better than predicted, and the Vickrey auction underperforms.[33]

[32] A parallel analysis for the restricted-bidding sample is conducted in the Online Appendix.

[33] Since values for each bidder and auction are drawn randomly, there is some variation in the average values across treatments. This variation is not the driving factor behind my results - in fact, the realised bidder values are on average lowest in the first-price auction (and highest in RR(0.50)). In pairwise median-difference tests, only this one pair rejects the no-difference null for values, at 95%. No other pairings reject in the median-difference test, even at the 90% level.

Table 2: Pairwise Auction Revenue and Surplus Comparisons. The first-price auction gives significantly higher revenue, and lower surplus, compared to every other rule. No other pairwise comparisons are statistically significant.

Revenue	Vickrey	VNR	RR(0.50)	RR(0.75)
FirstPrice	29.0***	24.0***	15.0**	23.0***
Vickrey		−3.0	−13.0	−7.0
VNR			−9.0	−1.0
RR(0.50)				8.0
Surplus	Vickrey	VNR	RR(0.50)	RR(0.75)
FirstPrice	−16.0**	−24.0***	−17.0***	−17.0***
Vickrey		−10.0	−2.0	−1.0
VNR			8.8	8.0
RR(0.50)				0.0

Reported values are for median-difference of (row - column), as points.
Rejections of zero-difference null at 90%/95%/99% level
indicated by */**/***; Bonferroni-Holm corrections applied.
Calculations based on all 140 experimental auction rounds.

Mirroring the results from the revenue figures above, the first-price auction generates less bidder surplus than any of the other three rules: all pairwise tests reject in this direction at a confidence level of 95% or stricter (see Table 2). All other pairings fail to reject the zero-difference null. Pairwise testing confirms the intuitive conclusion from Table 1: the first-price auction is different from the others, giving higher revenue and lower surplus.[34]

Assessing efficiency using a direct median-comparison test is unhelpful, because in all the treatments the median efficiency is 100%. A Kruskal-Wallis test nonetheless rejects with p-value ¡ 0.005, suggesting that efficiency is not homogenous across auctions. Hence I run a series of Mann-Whitney tests, pairwise for each combination of auctions; this allows me to check the distribution of efficiency in each pairing. All but one pairwise comparisons against the Vickrey auction reject at the 95% level or stricter, with Vickrey auction giving lower efficiency.[35] No other strict ranking pattern emerges. These findings provide evidence to reject hypothesis HE, according to which the Vickrey auction should

[34]The revenue and surplus conclusions of this section are precisely mirrored in the results from the restricted-bidding sample, and are included in the Online Appendix.

[35]The single auction that does not reject pairwise efficiency equivalence with the Vickrey auction is RR(0.50).

be most efficient.

All the statistical tests in this section have been median, or rank-based. As a robustness check, I have run a parallel analysis using standard cross-sectional econometric methods, and the results are reported in Appendix B. The analysis in Appendix B confirms the findings reported in this section - the first-price auction is still revenue superior, and the Virckrey auction least efficient.

6 Bidder-level Results

6.1 Bidding Constraints and Bidder Behaviour

I check the impact of bidding constraints by comparing the raw bid patterns across the two treatments, as summarised in Table 3.[36] Removing bidding constraints only significantly changes behaviour in the Vickrey auction. The bids are higher when restrictions are lifted, with a median difference of 30 for bidder L1, and 20 for L2. To put these numbers in perspective, recall that L-type values are uniform on [0,100] implying a median value of 50; the median increase in bids is at least 40% of this. The median-difference test accordingly rejects for all bidder types under the Vickrey auction at the 99% confidence level;[37] none of the other auctions register any rejections.

On this evidence, I cannot reject hypothesis HS: bidding constraints have no impact on first-price, VNR and Reference Rule auctions. In subsequent portions of the paper, I will conduct the analysis using data from the sessions with unrestricted bidding; a parallel analysis for the restricted-bidding sessions is available in the Online Appendix. The large difference registered in the Vickrey auction is consistent with hypothesis HC on collusion, and this finding will be further analysed in Section 6.5.

[36] The RR(0.50) auction is not included in this comparison, since none of the supplementary sessions were run with bidding restrictions.

[37] These are calculated using the Hodges-Lehmann method, implemented through the SomersD package in Stata (Newson, 2006).

Table 3: The influence of bidding restrictions on bids. Only the Vickrey auction shows a significant change in bidding across the two treatments.

Case		Vickrey	First-Price	VNR	RefRule(0.75)
Bidder L1	Medians	84.0 — 50.0	35.0 — 34.5	45.0 — 40.0	45.0 — 39.5
	Median Difference	30.0***	−2.0	3.0	5.0
Bidder L2	Medians	75.0 — 56.5	30.0 — 30.0	50.0 — 39.5	45.5 — 44.0
	Median Difference	20.0***	−2.0	5.0	4.0
Bidder G	Medians	136.0 — 90.0	65.0 — 79.5	100.0 — 90.0	106.5 — 91.0
	Median Difference	27.0***	−8.0	7.0	11.0

Medians reported as: Unrestricted — Restricted. Median difference tested via the Hodges-Lehmann method, using all 140 auction rounds. Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

6.2 Testing Bidder-level Intuitions

With the exception of the Reference Rule with $\alpha = 0.75$, all other auction settings analysed in this paper offer symmetric incentives for both L-type bidders, and the data from these two sub-cases can be pooled for analysis. This intuition is confirmed by the data: in the symmetric auctions, Mann-Whitney tests for the zero-difference null fail to reject on both the bid and shading variables (all p-values ¿ 0.15). For the purpose of further analysis in this section, the data for L1 and L2 types will thus be pooled in all auctions except RR(0.75), where I will consider both types separately.

To give an overview of local bidder's behaviour and assess hypothesis HB, Table 4 shows a set of pairwise median-difference tests across auctions for the bid variable. L-types bid the most in the Vickrey auction, and the least in first-price. The core-selecting auctions rank as intermediate, and show no significant difference from each other. The intuition of hypothesis HB cannot be rejected - the data shows that indeed Vickrey auction induces aggressive bidding, while first-price discourages it.

When assessing the validity of Hypothesis HG - that the G-type bidders bid similarly in all auctions except first-price - the Kruskal-Wallis tests for equality of populations rejects (p-value=0.005), suggesting that there are differences in bidding behaviour across auction types. On this evidence, I reject hypothesis HG.

Table 4: Pairwise comparison of L-types' bidding behaviour. Bidders bid most conservatively in the first-price auction, and most aggressively in the Vickrey auction.

Bids	Vickrey	VNR	RR(0.50)	RR(0.75)[L1]	RR(0.75)[L2]
FirstPrice	-44.0^{***}	-14.0^{***}	-16.0^{***}	-13.0^{***}	-13.5^{***}
Vickrey		30.0^{***}	26.0^{***}	30.0^{***}	27.0^{***}
VNR			-2.0	0.0	0.0
RR(0.50)				3.0	1.0

Reported values are for median-difference of ("row" - "column"), calculated as points from the raw bids, using all 140 auction rounds. Rejections of zero-difference null at the 90%/95%/99% level indicated by */**/***; Bonferroni-Holm corrections applied.

6.3 Bidder-level Tests of the Theory

The theory results being tested in this section base on the equilibrium bidding functions derived for the first-price, VNR, and RR(0.50) auctions by Ausubel and Baranov (2010). As no analytical results are available for RR(0.75) due to the asymmetry between L1 and L2, I obtained the equilibrium bidding functions numerically.[38] In first-price, and both Reference Rules, equilibrium bidding requires the L-types to bid exactly zero when their values are sufficiently low, and attempt to free-ride on the other L-type out-bidding the G-type on their own. In VNR, though such pooling at zero does not occur, theory still suggests bidding very cautiously in equilibrium. Table 5 shows that experimental results diverge significantly from theory.[39] Figure 4 provides an illustration of how experimental bidding functions for L-types compare to their theoretical counterparts; I have also included a set of "empirical best-response" curves, which are numerically calculated best-responses to bids actually submitted in the experiment.[40] Though the actual best response bids don't precisely coincide with Bayesian-Nash results from Ausubel and Baranov (2010), the two look more similar to each other than to the bidding functions observed in the experiment.

For L-types, the bidding variable rejects in all sub-cases, with the exception of the

[38] The method I use is similar to that of Baranov (2010).

[39] In Table 5, I use a permutation test for surplus. The surplus is calculated conditional on winning, which introduces a complex dependence pattern across the two samples: there are situations where an actual bid won in the experiment, but the corresponding theory-based bid would not have won (and vice versa). The samples are neither independent, nor matched-pairs. Thus I cannot use bootstrapping, and use permutation-based tests instead. For further discussion of permutation tests, see Good (1994).

[40] Analogous graphs for the G-types are provided in the Online Appendix.

Table 5: Bidder-level Tests of the Theory, calculated from all 140 auction-groups. In 9 of 11 bidder-type/auction pairings theory is rejected due to overbidding, at the 95% level. Surplus is lower than predicted by theory in 7 of the 11 cases, at the 95% level.

L-types	Vickrey	First-Price	VNR	RR(0.50)	RR(0.75), L1	RR(0.75), L2
Bid	80.0(48.0)***	31.5(18.3)***	48.5(30.8)***	45.0(2.9)***	50.0(32.7)***	45.5(48.5)
Win%	67.1(52.1)***	47.1(45.0)	47.9(35.0)***	39.3(32.9)**	52.9(35.7)***	52.9(35.7)***
Surplus	31.0(39.0)*	14.3(35.1)***	26.5(33.4)**	21.0(32.6)**	14.9(41.4)***	25.8(29.9)
G-type						
Bid	136.0(92.0)***	65.0(47.3)***	100.0(98.5)	122.5(112.0)**	106.5(94.5)**	
Win%	32.9(47.9)***	52.9(55.0)	52.1(65.0)***	60.7(67.1)	47.1(64.3)***	
Surplus	31.0(48.0)*	25.0(70.2)***	55.0(77.2)**	45.0(63.7)***	47.0(62.3)	

For bid and surplus, experimental medians reported; theory-based medians in parentheses. Calculations done using all 140 auction rounds. Sign-test used for testing bid and win% variables, median-based permutation test used for surplus. Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

L2-bidder in the RR(0.75) auction; the general pattern indicates that L-type bidders bid more than predicted by theory. Furthermore, the L-types bid exactly zero much too rarely: theory predicts a total of 100 bids at zero in my data, whereas only 38 are observed.[41] Beyond the misunderstanding of bidding incentives, it is likely that 'boundary effects' - the aversion to bid exactly at the boundary of the bidding support - may contribute to this finding.[42]

The G-types also overbid relative to theory in all auctions except VNR. However, in the core-selecting auctions and the Vickrey auction, the overbidding of the L-types dominates, which results in them winning more often than expected. Consequently the L-types also receive lower surplus, conditional on winning, in all cases except the L2-bidder in RR(0.75). The variable for winning probability does not reject in the first-price auction, suggesting that though both L- and G-types overbid considerably, this does not affect their relative winning chances. Conditional on winning, both types make less profit in the first-price auction than theory predicts.

The broad conclusions from Table 5 and Figure 5 suggest that in all auctions the

[41] Of the actually submitted zero-bids, only three occur when when BNE predicts they should; in the other 35 cases, BNE predicts strictly bids.

[42] A good analysis of this effect is Palfrey and Prisbrey (1997) in the context of public-goods contributions. In the present experiment, there is no way to test for this effect directly.

L-types overbid significantly relative to theory, therefore winning too often, but making lower profits than predicted. Correspondingly, in all auctions except first-price, the G-type wins too rarely, and when he does win he makes little profit. Jointly, these findings lead me to reject hypothesis HT - competitive BNE bidding theory is not supported by my data.

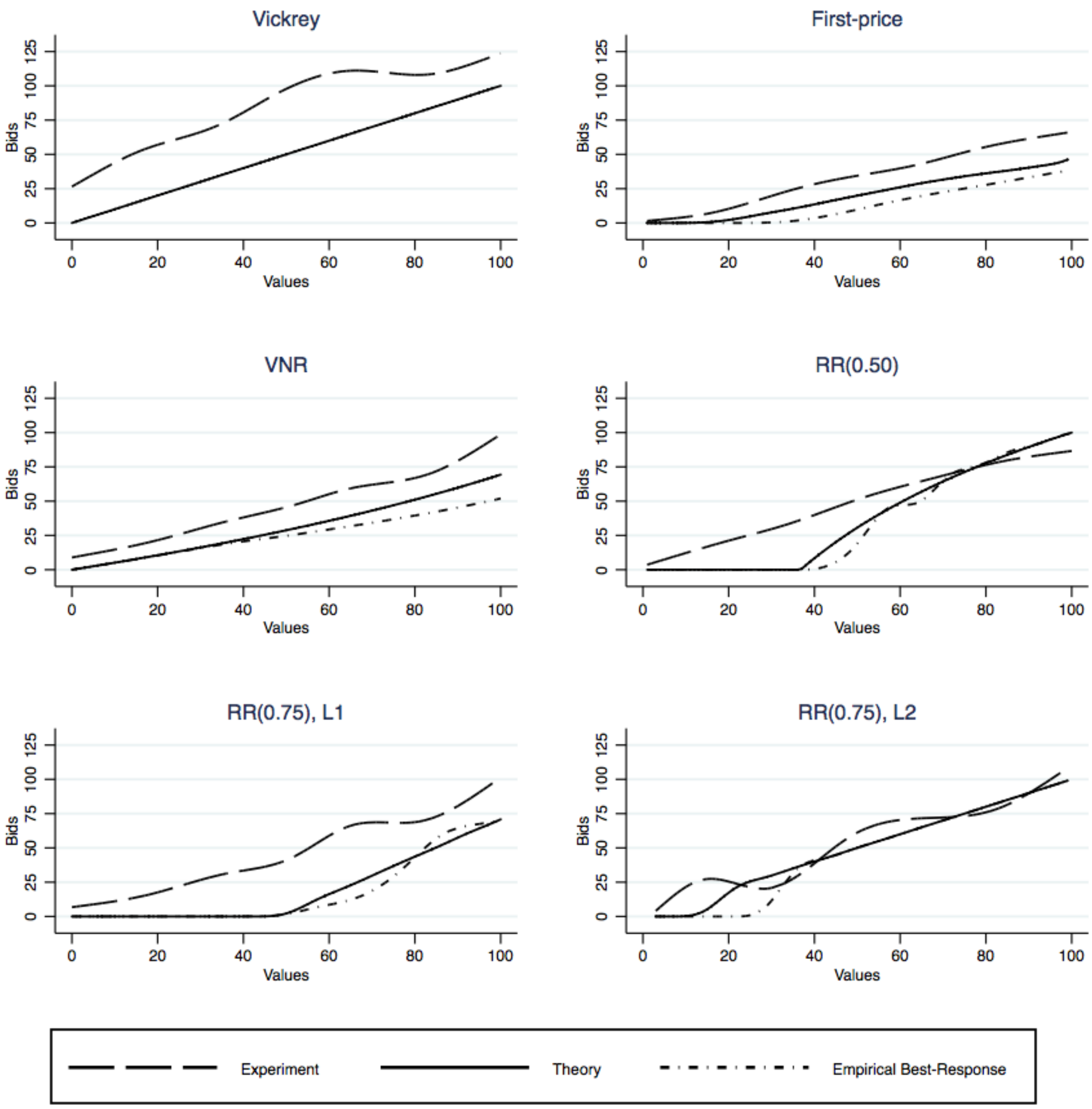


Figure 5: Spline fits for experimental local bidder's observed bids, theory-based best response functions, and numerically calculated best response functions to actual bidding in the experiment. With the exception of L2 in RR(0.75), the observed bidding functions diverge significantly from the best-response functions.

Hypothesis HA, on truthful bidding in core-selecting auctions, similarly finds no support in my experiment. A sign-test for the truthful-bidding null rejects for each bidder type at confidence level of 95%, or stricter. When deviating from theory, the bidders do not use a truth-telling rule-of-thumb. The intuition that core-selecting auctions induce a reversion to truthful bidding proves incorrect.

6.4 Evaluating Bidder Sophistication

The standard theoretical benchmark assumes that all bidders follow their Bayesian-Nash equilibrium strategies. But this benchmark may be inappropriate for experiments: perhaps bidders in the experiment *expect* that their opponents deviate from BNE-bidding. According to a 'sophisticated behaviour' hypothesis of Costa-Gomes, Crawford and Broseta (2001), the bidders may be trying to best-respond to the actual play of their opponents, rather than to theoretical predictions. If this is the case, then the fact that BNE-bidding is rejected should be unsurprising: such a strategy may not be a best response to actual play.

To assess whether sophisticated bidding could explain the divergence from theory, I calculate profits and winning probabilities for all bidder types under the additional scenario where each of the three bidder types unilaterally plays the BNE strategy, while the other two bidders play as they did in the experiment. If profits from actual bidding are higher than they would be if that bidder type unilaterally engaged in equilibrium play, then the observed bids may indeed be a best response to actual behaviour of the opponents. The results from this comparison are shown in Table 6, which finds little support for the sophistication hypothesis.

For local bidders, the winning probability and conditional profit variables reject the zero-difference null in all cases except for the L2-type in the RR(0.75) auction. In all these cases, the unilateral deviation towards BNE-bidding would lead to a (slightly, but significantly) lower winning probability, but a much higher surplus conditional on winning.[43] Since in Table 5 the L2-type's bidding in RR(0.75) was not significantly different from theory, it is unsurprising that a unilateral deviation towards theory does not lead

[43]If instead of 'surplus conditional on winning' I used 'unconditional surplus' instead, a sign-test on this variable rejects even more strongly. It would also reject in the additional case of the I2 bidder in RR(0.75).

Table 6: Testing for sophisticated bidding: surplus from actual bids vs. unilateral deviation to Bayesian Nash bidding. In 6 of 11 cases, a unilateral deviation gives a significantly higher surplus, at the 90% level.

L-types	Vickrey	FirstPrice	VNR	RR(0.50)	RR(0.75)-L1	RR(0.75)-L2
Win%	67.1(55.7)***	47.1(38.2)***	47.9(43.2)***	39.3(34.3)***	52.9(35.7)***	52.9(49.3)
Surplus	31.0(40.5)*	14.3(31.7)***	26.5(35.9)***	21.0(31.2)**	14.9(40.4)***	25.8(29.0)
G-type						
Win%	32.9(26.4)**	52.9(27.9)***	52.1(50.7)	60.7(55.7)	47.1(42.9)	
Surplus	31.0(39.0)	25.0(73.3)***	55.0(58.0)	45.0(48.5)	47.0(57.5)	

For surplus, experimental medians reported; 'unilateral deviation' medians in parentheses.
Sign-test used for testing the win% variable, median-based permutation test used for surplus.
Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

to higher conditional profit for this bidder. The results suggest, however, that the vast majority of L-type bidders are not engaging in sophisticated bidding.

The results for the G-type are more varied. In the first-price auction a unilateral deviation is profitable for the G-type for the same reason as it is for the L-types: the payment conditional on winning is then much lower. A similar deviation does not significantly improve profits in any of the other auctions, nor does it much affect winning probabilities in VNR and Reference Rule. In these auctions, the L-types' bids influence their payments in addition to the winning probability, but since G's payment depends only on L-types' bids, the foremost effect of equilibrium bidding is to reduce the probability of winning. The only way in which such a change in strategy would increase the profit, conditional on winning, is by excluding some of the cases where G-type wins after overbidding (and makes a negative profit). Table 6 shows that this effect is present, since benefits from deviation towards theory are positive, but not sufficiently to be significant.

Since the sophistication hypothesis is rejected in six of eleven sub-cases, it does not offer a plausible explanation for bidders' deviation from the theory. Following the BNE-bidding functions would leave each bidder type no worse off, even if their opponents did not follow suit.

The conclusions on the sophistication hypothesis don't change significantly if bidders were to unilaterally deviate towards the numerically-calculated best-response functions, instead of BNE.[44] The hypothesis still gets rejected in the same six out of eleven cases, though the expected profits from unilateral deviation are higher than in the present (BNE) case. This conclusion is unsurprising: the BNE-bidding functions assume that each player is best-responding the BNE-bidding by others, whereas the numerically calculated best-response functions take into account actual bidding in the experiment, and thus we should expect them to give higher expected profits.

6.5 Collusion in the Vickrey Auction

The most direct method for checking whether collusion is present is to look for instances of perfect collusion, where both L1 and L2 bid 200. This criterion is very stringent and of limited use if mis-coordination is frequent. Perfect collusion occurs in only 5 out of 140 rounds of play. In these 5 instances, the joint profit of the L-types is 110 - over twice average for the whole sample, which is 54. If successful, collusion is highly profitable.

To move beyond checking for perfect collusion, we need another plausible benchmark. Looking for overbidding in excess of value alone is insufficient because such bidding is frequently found even in single-item auctions where no collusive motive is present.[45] Furthermore, overbidding is sometimes attributed to a 'desire to win' effect: if bidders enjoy the phenomenon of winning in itself, they will bid more aggressively, even if this reduces their profit.[46] The significance of this effect is higher in rules where the influence of the bidder's own bid on their price is lower: the increased likelihood of winning looks evident, while the payoff-consequences are less obvious.

The experimental setup allows me to construct a benchmark that approximates the 'desire to win' effect, and use that to deflate the data from the Vickrey auction. The L-types' payments in VNR and RR(0.75) auctions are designed so as to mitigate the effect of own bids on the payment. While this isolation is not perfect, it does nonetheless provide the bidders with an opportunity to bid more aggressively without expecting large payoff-consequences. Looking at the differences in bids in these two auctions with, and

[44] Numerical results of this comparison are in the Online Appendix, Section 10.3.

[45] In second-price auctions, overbidding is found by Kagel and Levin (1995) and more recently Cooper and Fang (2008).

[46] For an overview, see Kagel and Levin (1995).

without, bidding restrictions allows me to construct a proxy for the 'desire to win' effect. I use this measure as my non-collusive benchmark.

To gauge the extent of the collusion attempts, I use the amount of overbidding (in excess of the benchmark) and the frequency with which such bids are submitted. If a significant portion of the data feature overbidding by a considerable amount, it is unlikely that such behaviour is purely accidental. Conversely, only moderate and occasional overbidding, makes collusion less plausible: such deviations could be attributed to miscalculation.

Table 7: Median Increase in Bids, after Removal of Bidding Restrictions

Auction	Vickrey	VNR	RefRule(0.75), L1	RefRule(0.75), L2
Median Decrease	13***	0	2**	1***

Median difference tested via the Hodges-Lehmann method.
Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

From Table 7, the largest median difference between restricted and unrestricted bidding treatments occurs in the Reference Rule for the L1-type. As expected, when bidding restrictions are lifted, this bidder type bids more aggressively, but only by 2 points.[47] A sign-test to check whether the shading by L-types in the Vickrey auction exceeds the 'desire to win' benchmark rejects with a one-sided p-value ≈ 0.008, and triggers suspicions of collusion.

Table 8: Numbers of Overbidding L-types, and conditional profit as points. Overbidding is most prevalent, and most profitable, in the Vickrey auction.

Overbid by more than:	Vickrey	First-price	VNR	RR(0.75)
0	166 (15.8)	7 (-6.4)	67 (12.5)	77 (4.3)
5	151 (13.7)	5 (-8.8)	52 (7.8)	59 (2.3)
10	136 (12.5)	4 (-11)	34 (2.3)	42 (-1.1)
20	116 (9.8)	1 (-26)	19 (-6.1)	23 (-8.5)
30	101 (6.7)	0 (NA)	12 (-15.0)	16 (-21.5)
50	79 (3.7)	0 (NA)	5 (-32.4)	6 (-53.7)
75	55 (-0.1)	0 (NA)	3 (-61.3)	5 (-67.2)

Mean surplus in brackets. Total number of L-type bids is 280 under all rules.

[47]This is the median increase in bids, and though the median amount of shading is still positive, 25% of the bids of this bidder type involve overbidding above value.

To further illustrate how the consequences of overbidding differ by auction, Table 8 shows the numbers of overbidding L-types, and their mean surplus. The number of overbidding L-types is highest in the Vickrey auction at all overbidding levels. As the ex-ante expected value of an L-type bidder is 50, overbidding by 30 is already 60% above the expected value, and over 40% of bids are in this group. Furthermore, almost 20% of all submitted bids are 75 points or more above value; this magnitude of overbidding is unlikely to be accidental, especially given how rarely similar deviations occur in the other auctions.

Bidders in the Vickrey auction still make more profit than they would by behaving similarly in any of the other auctions. By overbidding as much as 50 points, the L-types in the Vickrey auction still make a positive surplus (with a mean of 3.7), whereas in other auction types by this point the surplus is negative. Since overbidding is both most prevalent and most profitable in the Vickrey auction, it is likely that this pattern can be attributed to attempted collusion.[48]

Despite its prevalence, overbidding is not overall profitable for the bidders involved. The rejection of the 'sophisticated bidding' hypothesis showed that L-types in the Vickrey auction would do better by unilaterally deviating towards truthful bidding. The data describes a pattern where L-type bidders attempt to collude, despite frequent miscoordination. As a result, the Vickrey auction underperforms doubly: even though in Section 5 it gave low revenue to the seller, at the bidder level this has not translated into higher surplus. Both the seller and the bidders end up significantly worse off than theory predicts.

7 Discussion

Table 9 summarizes the outcomes of the hypotheses tested in this paper. At the auction level, the theory-based hypothesis HR, on revenue, is rejected due to the superior performance of the first-price auction, and the equally poor outcomes form the Vickrey

[48]The findings of Table 8 would not significantly change if I looked at the amount of 'bidding in excess of equilibrium prediction' rather than looking at overbidding relative to true values.

auction. The data do not support the hypothesis of full efficiency in the Vickrey auction either: instead, it ranks as least efficient. No significant differences among the other rules emerge, so overall hypothesis HE is also rejected.

Table 9: Outcome of the hypothesis tests

Hypothesis	Outcome
HR: The revenue ranking is Vickrey>First-price> VNR≈RR(0.50)	Rejected
HE: The efficiency ranking is the same as in HR	Rejected
HB: Bidding is most aggressive in the Vickrey auction, least in first-price	Accepted
HT: Bidders follow competitive equilibriums strategies	Rejected
HA: L-types bid truthfully in VNR and Reference Rule	Rejected
'Sophistication hypothesis'	Rejected
HG: G-types bid similarly in all auctions except first-price	Rejected
HS: Bidding constraints have no effect in first-price, VNR and RR	Accepted
HC: Bidding behaviour in Vickrey Auction is consistent with collusion	Accepted

The acceptance of hypothesis HB shows that bidders were broadly responding to auction incentives in the ways we would intuitively expect. However, the data rejects more precise hypotheses on bidding behaviour. For the first-price auction, this finding is similar to results on overbidding in single-unit experiments. In the core-selecting auctions, the VNR and Reference Rule, the picture is more complex. Participants with low values do not submit zero bids often enough, and all types bid more than predicted. This leads to the rejection of hypothesis HT. Furthermore, the participants do not bid truthfully in any of the core-selecting auctions, whereby I reject hypothesis HA. Neither theory, nor rule-of-thumb behaviour offer a satisfactory explanation of the experimental results.

The rejection of the 'sophistication hypothesis' showed that unilateral deviations towards equilibrium bidding would be profitable for L-type bidders in five out of six cases, which suggests that participants were also not best-responding to each other's actual bidding behaviour. The current experimental design cannot explain the cause of such a pattern. Future work in this area will look at the influence of expectations to evaluate whether the divergence from theory is due to incorrect expectation formation, or sub-optimal bidding in response to correct expectations.

The behaviour of L-type bidders in the Vickrey auction is consistent with attempted collusion, even if full collusion rarely manifests. In all other auctions the presence of bidding constraints has no impact, as shown by the acceptance of hypothesis HS. In the Vickrey auction extensive overbidding is observed when constraints are removed. The extent of the overbidding was above what I could attribute to a 'desire to win' effect, and the number of extremely high bids is higher than in all other auctions.

A natural interpretation of finding collusion in the setting of my paper is to relate it to practical one-shot auctions, in contrast to the collusion literature which looks at repeated play. An example of this would be a one-off sale of government assets with a pure efficiency objective, and no concern for revenue. My results suggest that even if revenue in itself is unimportant, the potential for collusive bidding in a Vickrey auction is high, and that is sufficient to undermine its efficiency properties. A policy with a pure efficiency objective could be counterproductive.

8 Conclusions

My main finding is the surprisingly good performance of the first-price auction: it generates most revenue, without any corresponding efficiency loss. Conversely, the performance of the Vickrey auction is unexpectedly poor: contrary to the expectation of full efficiency, it ranks last on this criterion. Given that efficiency concerns are frequently used to argue against the use of first-price mechanisms in high value auctions, my experimental results provide evidence to allay such worries. The core-selecting auctions tie with the first-price auction on efficiency, and are revenue-equivalent with the Vickrey auction; they are not "the best of both worlds", but also never rank last, contrary to theoretical predictions.

At the individual level, I find that bidding diverges significantly from Bayesian Nash equilibrium predictions. Bidders frequently bid in excess of the theoretical benchmark, and occasionally even above their valuation. Overbidding can not be attributed to sophistication, as the observed bids never resulted in higher profits compared to a unilateral deviation towards Nash equilibrium bidding. In the core-selecting auctions, bidders also do not use a truth-telling rule-of-thumb: I find no evidence to support the intuition that payments close-to-independent of own bids induce close-to-truthful bidding. The behaviour I observe in the Vickrey auction is consistent with attempts at playing collusively, even though such attempts are rarely successful. The Vickrey auction generates neither high revenue, nor high bidder surplus. My results suggest that even with complementarities, simple first-price rules are unlikely to fail as badly as feared, and opportunity-cost based pricing rules may not realise the benefits that we intuitively expect.

9 Appendix A: The Variable-α Experiment

In the proofs that Erdil and Klemperer (2010) use to analyse the incentive properties of the Reference Rule, the reference point itself does not change the deviation incentives on aggregate. However, it affects the relative amount that each bidder pays, conditional on winning, and this may have non-trivial behavioural implications. Numerical calculations have shown that as α changes, so do the optimal bids, resulting in extremely disparate optimal bidding functions for the two types as α tends to either 0 or 1.[49] This additional experiment set out to examine whether such variation would also emerge in the laboratory.

Let K denote the upper end of the support of the value distribution of the L1-type. Then asymmetries in the valuations of the two L-types can be modelled as follows: set $v_{L1} \sim U[0, K]$ and $v_{L2} \sim U[0, 200 - K]$. This keeps the sum of supports (and hence the expected total value) of the two L-type bidders the same as that of the G-type bidder, but when $K \neq 100$ the L-types are no longer symmetric. The nature of asymmetry in my experiment can then be summarised by two parameters: α and K. I consider four cases:

- Setting 1: $\alpha = 0.50$ and K=100 (i.e. $v_{L1}, v_{L2} \sim U[0, 100]$)
- Setting 2: $\alpha = 0.75$ and K=150 (i.e. $v_{L1}, \sim U[0, 150], v_{L2} \sim U[0, 50]$)
- Setting 3: $\alpha = 0.75$ and K=100 (i.e. $v_{L1}, v_{L2} \sim U[0, 100]$)
- Setting 4: $\alpha = 0.50$ and K=150 (i.e. $v_{L1}, \sim U[0, 150], v_{L2} \sim U[0, 50]$)

This particular combination of α and K allows me to investigate two main questions. Firstly, I can check whether it is the asymmetry of the α parameter itself that influences

[49] In the limit, as $\alpha \to 0$ or $\alpha \to 1$ an analytical solution is possible. The solution entails the L-type bidder with the infinitesimal 'reference share' bidding truthfully, while the other L-type shades by a large amount.

behaviour; for this comparison, I look at the cases where the support of the two bidders' valuations stays constant, and α varies. Secondly, I can assess whether it is the magnitude of α *relative to* the 'expected valuation' of the bidders that matters; here I compare the cases where the ratio of $\frac{E(v_{L1})}{E(v_{L2})} = \frac{\alpha}{1-\alpha}$, to those where it is not.

The experimental setup of these session was analogous to the main experiment in this paper, with the exception that here only one set of instructions was given out at the beginning of the experiment. These instructions outlined how variations in the α parameter influenced reference payments in the Reference Rule.[50] The participants were allowed to ask questions whereafter they proceeded to complete an understanding test.[51] Upon successful completion of the test, the participants were informed which α parameter and which valuation model would apply in the given section of the experiment. They subsequently played two practice rounds, followed by ten payment-relevant rounds in each setting.[52] The duration of the sessions in the Alpha-experiments was two hours on average, generating mean earnings of £27 (∼\$43).

9.1 Results of the Variable-α Experiment

Comparing bidder-level results in the asymmetry experiment poses complications that are not present in the main experiment. Direct tests of bidding variables cannot be conducted across settings where K varies, because these tests will reject by default due to the bidding support being different across the compared cases.

This problem does not arise, however, when performing tests while holding K fixed. When I test for the effects of varying α only, holding K fixed, none of the four test-pairing for the L-type bidders reject a zero-difference null even at the 90% level. Hence α on its own does not significantly influence individual bidding.

An alternative to using direct bid data is to look at bid ratios,[53] but this approach will artificially inflate differences in the cases where $K \neq 100$. Here the two L-types have

[50] The instructions are available from the author on request.

[51] The rate of failures was three out of 45 participants in this phase of the experiment.

[52] The order of the Cases in the experimental sessions was from 1 to 4 in the first session. The ordering was reversed for the other session.

[53] These are calculated as the ratios of bid relative to the value of the bidder.

a different value support, and the L2-bidder with a narrower support is more likely to exhibit large variation in the bid ratios. The tests are hence likely to over-reject a zero-difference null, though using non-parametric tests reduces the likelihood of this mistake. However, when I run a battery of median-difference tests for both L-types on their bid ratios, only one statistically significant difference emerges. The L2-type's bid-ratios in Setting 4 ($\alpha = 0.50,\ K = 150$) test as significantly lower than in all other cases. This is an intuitive finding, as in this case the L2-type can be seen to be in a particularly weak position: they have a bidding support of only [0,50], but their 'preliminary share' of the payments is a disproportionately higher 50%. As a result, in this setting the L2 type bids more cautiously. No other ranking emerges from the pairwise tests.

A final hypothesis that I test on the individual bidder data is to check whether setting the α proportionately to the ratio of expected values of the two L-types affects bidding. It is, for example, possible that bidders would have a preference for equality or some notion of fairness, as found by Battalio, Van Huyck and Gillette (1992) in the context of two-person coordination games. To test for this effect, I pool the data from settings 1 and 2, where α is set 'proportionately', and test it against the pooled data from settings 3 and 4. Median-difference tests for both L1's and L2's bidding ratios fail to reject the zero-difference null (p-values ¿0.22 in both cases). Thus I cannot find any influence of proportionality on bidding at the individual level.

From the G-types' perspective, all four settings are identical, so we should expect them to bid similarly in all four cases. A Kruskal-Wallis test for this hypothesis marginally rejects with a p-value=0.046, indicating that the G-types do not bid the same way across the four settings. In pairwise tests for bidding and shading, various individual pairings reject, but no coherent pattern emerges. It appears that the G-type bidders are trying to best respond differently to the G-types' actual bidding across the different settings, ignoring the prediction that truthful bidding should be optimal every time.

At the auction level, the main variables of interest are again revenue, surplus and efficiency. A summary of these parameters across the four settings is shown in Table 10. Setting 1 immediately stands out: revenue is almost 10 points higher than in the other three settings, while surplus is lower by a similar amount. Efficiency is high in all four settings, and the differences are small.

Table 10: Revenue, Surplus and Efficiency Summary from alpha experiment

	K=100—α=0.50	K=150—α=0.75	K=100—α=0.75	K=150—α=0.50
revenue	77.0 (42.3)	65.5 (41.0)	62.6 (38.4)	64.2 (40.9)
surplus	48.9 (49.3)	61.1 (51.4)	58.2 (44.1)	63.8 (49.1)
efficiency	94.9 (13.8)	95.3 (15.0)	96.9 (12.0)	96.0 (15.1)

Means reported, standard deviations below.

A series of pairwise median-difference tests for revenue is summarised in Table 11. The results hence confirm that the symmetric setting with K=100 and $\alpha = 0.50$ is revenue-superior to the other three cases, with the tests rejecting the zero-difference null with 90% confidence or stricter. No significant revenue differences emerge amongst the other pairings. Correspondingly, Setting 1 also yields significantly lower surplus than Setting 4 (p-value=0.009). Finally, a Mann-Whitney test for differences in efficiency fails to reject between Settings 1 and 2, but it does reject the zero-difference null between Setting 1 and Settings 3 and 4 with p-value=0.015 and p-value=0.002; after applying the Bonferroni-Holm corrections, these rejections remain significant at the 90% and 95% levels, respectively. This implies that Setting 1 is less efficient, but no other pairings yield a rejection of the zero-difference null. Using the RR(0.50), or the Proxy Rule, in a symmetric setting yields superior revenue, but lower efficiency.

Table 11: Pairwise Revenue-difference Tests for variable-α experiment

	K=150—α=0.75	K=100—α=0.75	K=150—α=50
K=100—α=0.50	12.5*	14.0**	13.0*
K=150—α=0.75		2.0	0.0
K=100—α=0.75			-1.0

Reported values are for median-difference of (row - column).
Rejections of zero-difference null at 90%/95%/99% level
indicated by $^{\star}/^{\star\star}/^{\star\star\star}$; Bonferroni-Holm corrections applied.

The final test of interest at the auction level checks whether revenue and efficiency are sensitive to setting the α proportionately to the bidders' expected values. If the proportional cases where $\frac{E(v_{L1})}{E(v_{L2})} = \frac{\alpha}{1-\alpha}$ perform significantly better, this would be supporting evidence in favour of the flexibility inherent in the Reference Rule. A median-difference test for revenue rejects with a p-value=0.037; the median-difference is 7 points in favour of the proportional settings. A corresponding Mann-Whitney test for efficiency rejects with a p-value ¡ 0.001. In practice the differences in efficiency are low - on average around 1.3 points - so the statistical significance here has limited economic importance. This pair of findings gives some support to the view that selecting a reference point appropriately in relation to the relative values of the assets for sale may yield superior revenue results.

Overall, the findings of the sessions on asymmetries do not offer conclusive answers as to the influence of α. Though I find some significant auction-level results in favour of setting α appropriately, the bidder-level data show little sensitivity to α.

References

Ausubel, L. and Baranov, O.: 2010, Core-Selecting Auctions with Incomplete Information, *University of Maryland Working Paper* .

Ausubel, L. and Cramton, P.: 2004, Auctioning many divisible goods, *Journal of the European Economic Association* **2**(2-3), 480–493.

Ausubel, L., Cramton, P. and Milgrom, P.: 2006, The clock-proxy auction: A practical combinatorial auction design, *Combinatorial Auctions* pp. 115–138.

Baranov, O.: 2010, Exposure vs. Free-Riding in Auctions with Incomplete Information, *Working Paper* .

Cantillon, E. and Pesendorfer, M.: 2006, Auctioning bus routes: the london experience, *in* C. P. Ausubel L. and S. R. (eds), *Combinatorial Auctions*, MIT Press, Cambridge, Mass., USA.

Cooper, D. and Fang, H.: 2008, Understanding overbidding in second price auctions: An experimental study, *The Economic Journal* **118**(532), 1572–1595.

Costa-Gomes, M., Crawford, V. and Broseta, B.: 2001, Cognition and Behavior in Normal-Form Games: An Experimental Study, *Econometrica* **69**(5), 1193–1235.

Day, R. and Milgrom, P.: 2008, Core-selecting package auctions, *International Journal of Game Theory* **36**(3), 393–407.

Day, R. W. and Cramton, P.: 2012, Quadratic core-selecting payment rules for combinatorial auctions, *Operations Research* **60**(3), 588–603.

Erdil, A. and Klemperer, P.: 2010, A new payment rule for core-selecting package auctions, *Journal of the European Economic Association* **8**(2-3), 537–547.

Fischbacher, U.: 2007, z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* **10**(2), 171–178.

Goeree, J. and Lien, Y.: 2009, On the impossibility of core-selecting auctions, *Institute for Empirical Research in Economics University of Zurich Working Paper No.* .

Good, P.: 1994, *Permutation Tests: a practical guide to resampling methods for testing hypotheses*, Springer, New York.

Goswami, G., Noe, T. and Rebello, M.: 1996, Collusion in uniform-price auctions: experimental evidence and implication for treasury auctions, *Review of Financial Studies* **9**(3), 757.

Kagel, J.: 1995, Auctions: A Survey of Experimental Research, *The Handbook of Experimental Economics* pp. 501–585.

Kagel, J. and Levin, D.: 2008, Auctions: A survey of experimental research, 1995-2008, *Handbook of Experimental Economics Vol 2.* **2**.

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

Kazumori, E.: 2010, Auctions with package bidding: An experimental study, *Working Paper* .

Kazumori, E.: 2014, Vickrey Auctions Versus Proxy Auctions: An Experimental Study, *working paper* .

Kwasnica, A. and Sherstyuk, K.: 2007, Collusion and Equilibrium Selection in Auctions, *The Economic Journal* **117**(516), 120–145.

Newson, R.: 2006, Confidence intervals for rank statistics: Somers' d and extensions, *Stata Journal* **6**(3), 309.

Palfrey, T. and Prisbrey, J.: 1997, Anomalous behavior in public goods experiments: How much and why?, *The American Economic Review* **87**(5), 829–846.

Sano, R.: 2010, An Equilibrium Analysis of a Package Auction with Single-Minded Bidders, *Working Paper* .

Sherstyuk, K.: 2002, Collusion in private value ascending price auctions, *Journal of economic behavior & organization* **48**(2), 177–195.

Sun, N. and Yang, Z.: 2006, Equilibria and indivisibilities: gross substitutes and complements, *Econometrica* **74**(5), 1385–1402.

Sun, N. and Yang, Z.: 2009, A double-track adjustment process for discrete markets with substitutes and complements, *Econometrica* **77**(3), 933–952.

ComReg: 2012, Multi-band spectrum release: Information memorandum, **URL:** `http://www.comreg.ie/_fileupload/publications/ComReg1252.pdf`.

Danish Business Authority: 2012, 800 mhz auction: Information memorandum, **URL:** `http://erhvervsstyrelsen.dk/file/251159/information-memorandum-800mhz-auction.pdf`.

Federal Aviation Administration: 2008, Congestion management rule for laguardia airport; final rule, *Federal Register* **73**(198), 60573–60601.

Norwegian Post and Telecommunications Authority: 2013, 800, 900 and 1800 mhz auction - auction rules, **URL:** `http://eng.npt.no/portal/page/portal/PG_NPT_NO_EN/PAG_NPT_EN_HOME/PAG_RESOURCE_TEXT?p_d_i=-121&p_d_c=&p_d_v=128156`.

Ofcom: 2012, The award of 800 mhz and 2.6 ghz spectrum information memorandum, **URL:** `http://stakeholders.ofcom.org.uk/binaries/consultations/award-800mhz/statement/IM.pdf`.

Teytelboym, A.: 2013, Essays on networks and market design, *D.Phil Thesis, University of Oxford* .

10 APPENDIX C - ONLINE APPENDIX

10.1 Bidding Functions for G-type bidders

Figure A1 shows the G-types' bidding curves under the five auction rules tested in my experiment. In all cases except first-price, the theory predicts that truthful bidding should be optimal. In Table 5, I showed that tests for equilibrium bidding for the G-types were rejected in all auctions except VNR. Looking at Figure A1, that conclusion is consistent with the graphs: in VNR the bidding function is closest to the truthful-bidding prediction.

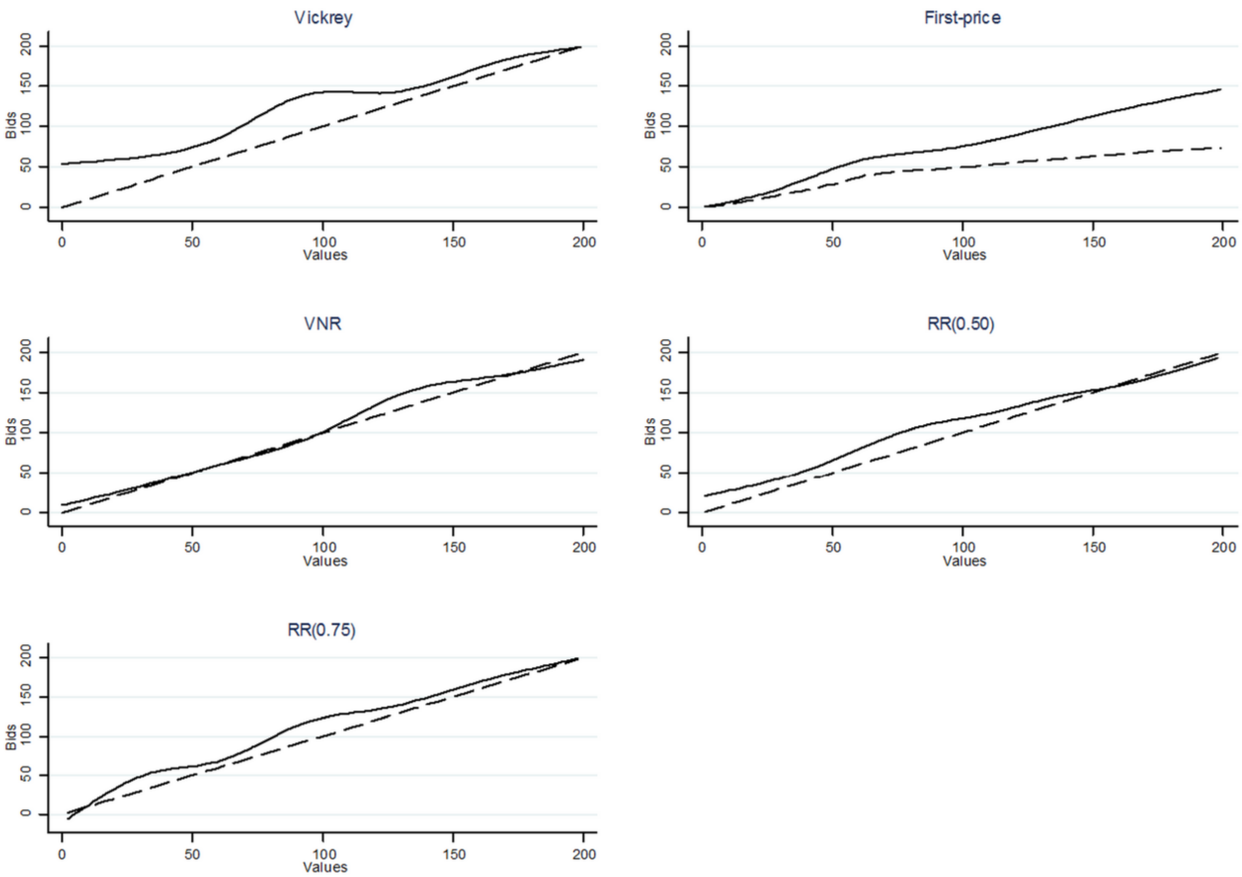


Figure 6: Empirical bidding functions (solid) and Nash-Equilibrium (dashed) bidding functions for G-type bidders, when bidding is unrestricted. With the exception of VNR, actual bid functions considerably diverge from theoretical predictions.

10.2 Analysis of the Restricted-bidding Sessions

Analogously to Section 5, here I conduct a bidder- and auction-level evaluation of the four auction rules using the restricted bidding data. As Table 12 indicates, the evidence does not conform to theory.

Table 12: Comparison of actual v.s. theoretical bidding under bidding restrictions

L-types	Vickrey	FirstPrice	VNR	RR(0.75)-L1	RR(0.75)-L2
Bid	50.4(52.6)***	35.1(20.9)***	42.7(31.8)***	40.5(16.8)***	43.1(46.5)***
Win%	51.9(51.2)	48.1(48.1)	46.3(31.3)***	43.8(35.0)***	43.8(35.0)***
Surplus	44.8(45.6)	16.6(36.1)***	30.0(38.5)***	29.5(43.1)***	39.2(34.2)
G-type					
Bid	95.6(99.1)***	77.3(45.8)***	94.7(97.2)***	93.6(96.7)***	
Win%	48.1(48.8)	51.9(51.9)	53.8(68.8)***	56.3(65.0)***	
Surplus	57.6(54.3)	29.4(74.6)***	60.7(68.5)	63.2(74.5)	

For bid and surplus, experimental medians reported; theory-based medians in parentheses. Sign-test used for testing bid, shading and win% variables, median-based permutation test used for surplus. Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

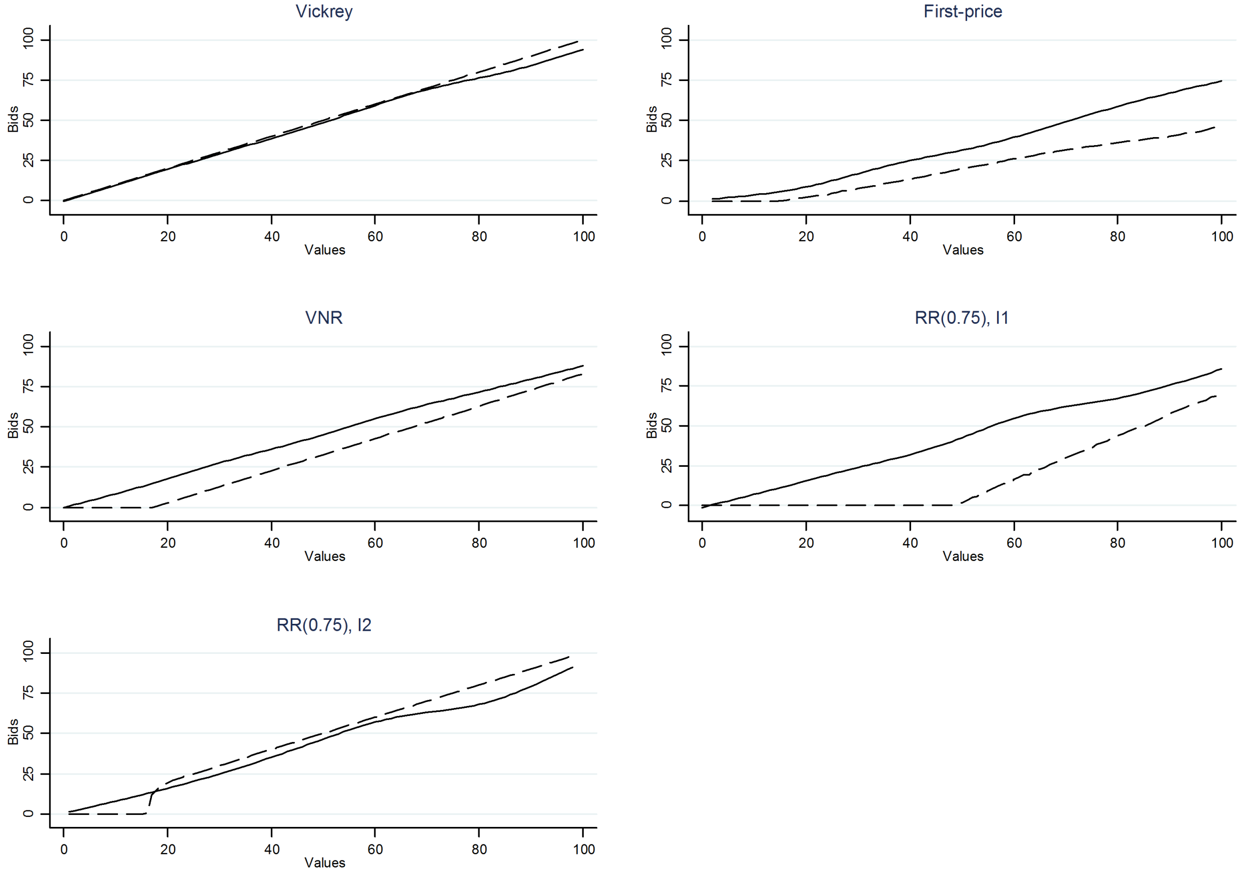
The theoretically predicted bidding functions are rejected at the 99% level, similarly to the results from the unrestricted bidding sample. There is, however, a statistical complication in testing the theory in those sub-cases where 'truthful bidding' is the equilibrium strategy and bidding is restricted.[54] If we use rank-based robust statistics (as I have done elsewhere), all the median differences will have the same sign by necessity, since a bidder can never overbid his value under bidding restrictions. Rank-based statistics will over-reject in all these cases, and the sign-test that I have used in Table 12 is particularly affected. This would bias the findings for all bidders in the Vickrey auction, as well as the G-types in VNR and Reference Rule.

Looking further at Figures 6 and 7, the bidding function in all the affected cases appear to be very close to the truthful-bidding line. It is likely that the rejection of theory in these cases is a statistical artifact. To verify this, I re-run the comparisons in these four cases using a mean-based permutation test, as in Section 6.3. This test did not reject

[54]This applies to all bidders in the Vickrey auction, and also Bidder J in VNR and the Reference Rule.

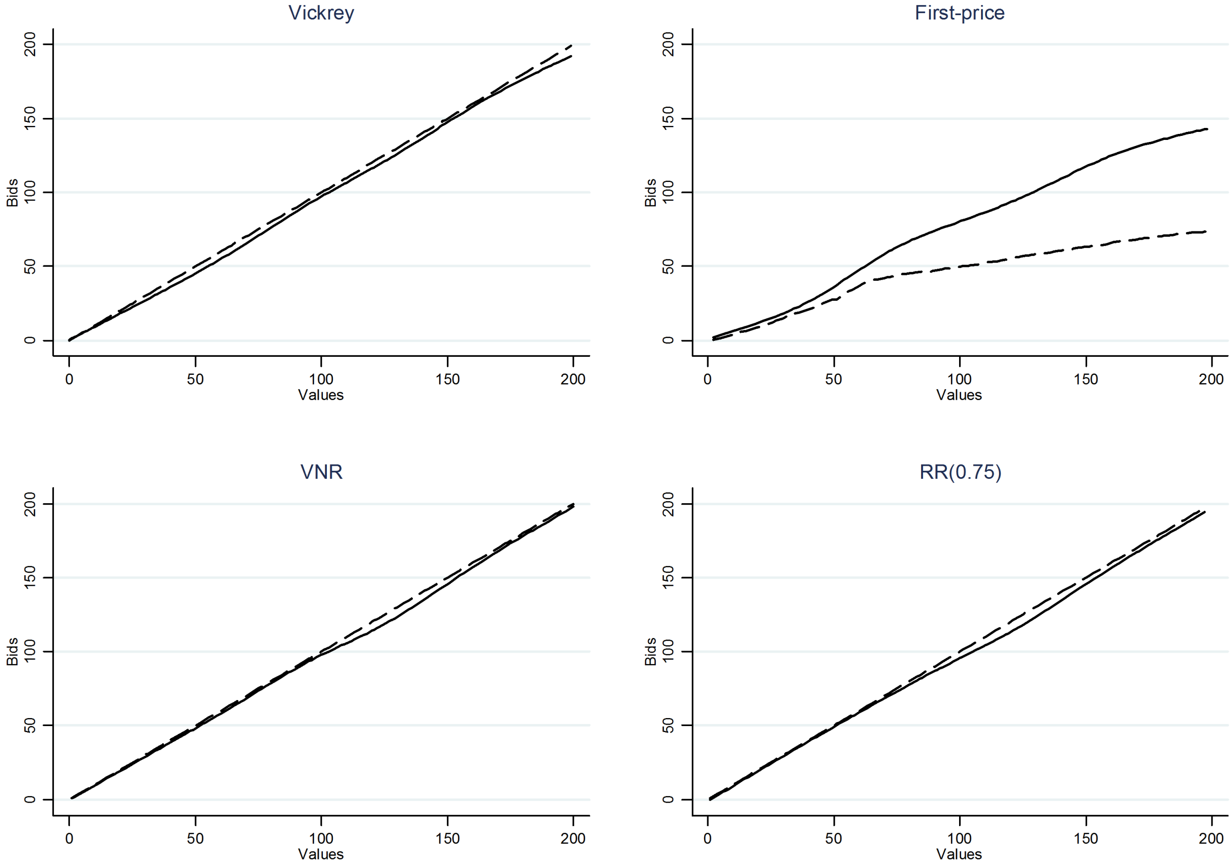
in any of the four cases where truthful bidding was the equilibrium strategy (all four p-values > 0.33), suggesting that when bidding restrictions are in place, bidders follow the equilibrium strategy closely.[55] The fact that I cannot reject truthful bidding for the G-types also means that I cannot reject hypothesis HC on the restricted bidding sample. A Kruskal-Wallis test also fails to reject the null that bids in these four cases come from the same population (p-value=0.96), whereby I also cannot reject hypothesis HS.

Figure 7: Empirical bidding functions (solid) and Nash-Equilibrium (dashed) bidding functions for L-type bidders, when bidding is restricted. In the Vickrey auction the truthful-bidding equilibrium is confirmed by data, but actual bidding diverges from Bayesian-Nash predictions significantly in the other four cases.



[55]To check for consistency, I also ran this same test for those cases where truthful bidding was not the equilibrium strategy; consistently with the sign-test results in Table 12, the permutation test also rejected the null of bid equivalence. Thus in the cases where the theory benchmark did not include truthful bidding, the sign-test and permutation test outcomes overlap.

Figure 8: Empirical bidding functions (solid) and Nash-Equilibrium (dashed) bidding functions for G-type bidders, when bidding is restricted. Bidding restrictions have prevented overbidding, resulting in near-truthful bidding in auctions other than first-price. In the first-price auction Bayesian-Nash equilibrium predictions are not confirmed.



There are three likely explanations for the discrepancy between the results here and those of Section 6.3, where all the truthful-bidding equilibrium bidding hypotheses get rejected. Firstly, it is possible that bidders simply understood the rules of the auction better in these two sessions, and understood how to pick an equilibrium strategy. Secondly, putting a cap on bids could have created a 'focal-point' in auctions where bidders notice their bids don't strongly affect their payments. The bid-functions in for VNR, as well as Reference Rule in Figure A2 lend some support to this view: the bid functions are very close to truthful bidding, more so than in the case when bidding is unrestricted, even though in both cases the observed behaviour is far from the equilibrium prediction. Finally, the bid-cap may simply be imposing a bid-ceiling in all those cases where bidders would wish to overbid relative to their value, and making these bids observationally equivalent to Nash equilibrium behaviour.

Pairwise comparisons of bidding patterns, in Table 13, show that bids are again lowest in the first-price auction. Similarly to Section 6, I also find that bidding in the Vickrey auction is significantly higher than in the other three auctions.

Table 13: Pairwise comparisons of bidding under bidding restrictions

Bids	Vickrey	VNR	RR(0.75)[L1]	RR(0.75)[L2]
FirstPrice	$-16.0^{\star\star\star}$	$-7.0^{\star\star\star}$	-5.0	$-8.0^{\star\star}$
Vickrey		$8.0^{\star\star\star}$	$10.0^{\star\star\star}$	$8.0^{\star\star}$
VNR			2.0	0.0

Reported values are for median-difference of (row - column).
Rejections of zero-difference null at 90%/95%/99% level indicated by $\star/\star\star/\star\star\star$.
Bonferroni-Holm corrections applied.

An auction-level summary of revenue, surplus and efficiency is presented in Table 14. The revenue results in Section 5 were in part driven by overbidding, and with bidding restrictions in place the revenue is lower in all four auctions. Bidder surplus has correspondingly increased, and the efficiency of all auctions is very high.

Table 14: Auction-level summary of revenue, surplus and efficiency under bidding restrictions

	Vickrey [N=160]	FirstPrice [N=160]	VNR [N=160]	RR(0.75) [N=160]
revenue	56.6 (47.7)	97.9 (31.7)	62.7 (35.3)	59.6 (37.1)
surplus	74.2 (46.1)	31.2 (18.9)	60.3 (40.3)	65.6 (40.4)
efficiency	99.2 (5.3)	98.6 (5.6)	99.5 (3.0)	99.5 (2.9)

Means reported, standard deviation below.

Pairwise tests of revenue and surplus are shown in Table 15. As in the unrestricted bidding case, the first-price auction is revenue-dominant over the other three rules at the 99% level, while no other tests reject revenue equivalence. Surplus in the first-price auction is accordingly lower than under the other rules. In addition, the pairwise test between the Vickrey auction and VNR rejects at the 95% level, with surplus being lower under VNR. A test between the Vickrey auction and the Reference-Rule does not reject a zero-difference null, hence I cannot obtain a full ranking.

Table 15: Pairwise comparison of revenue and surplus under restricted bidding

Revenue	Vickrey	VNR	RR(0.75)
FirstPrice	47.0***	37.0***	41.0***
Vickrey		−10.0	−6.0
VNR			4.0
Surplus	Vickrey	VNR	RR(0.75)
FirstPrice	−38.0***	−25.0***	−30.0***
Vickrey		13.0**	8.0
VNR			−5.0

Reported values are for median-difference of (row - column). Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***; Bonferroni-Holm corrections applied.

When bidding is restricted, the efficiency properties of the four rules are very similar. A Kruskal-Wallis test fails to reject the null that the efficiency draws for all four auction come from the same population. In Section 5 I found the Vickrey auction to be least efficient due to prevalent overbidding above value, but here bidding restrictions prevent such behaviour.

In sum, the auction-level findings from the experiments with restricted bidding are close to the findings in Section 5. The first-price auction is still revenue-dominant, and no less efficient than any of the other rules analysed. The Vickrey auction does not perform as poorly under bidding restrictions as it does under unrestricted bidding: removing the possibility for overbidding, and with it the opportunity for collusion, eliminates the cases in which the Vickrey auction underperforms most acutely.

10.3 Further results from testing for sophistication

As shown in Table 6, the sophistication hypothesis gets rejected because local bidders have a profitable unilateral deviation towards BNE-bidding. However, the BNE bidding functions from Ausubel and Baranov (2010) are not the best best response to actual bids in the experiment: as seen in Figure 5, the numerically calculated best-responses don't exactly overlap with the Bayesian-Nash results. Intuitively, the unilateral deviation towards this calculated best-response bidding function should be even more profitable than a deviation towards Bayesian-Nash bidding, so the sophistication hypothesis should be even more soundly rejected. Table **??** shows that the the hypothesis is rejected in exactly the same cases as when deviating towards Bayesian-Nash bidding, but the expected profit is now higher. For the global bidder, the only difference emerges in the first-price auction; in the other auctions the Bayesian-Nash and numerically calculated best responses coincide. The number of cases in which sophisticated bidding thus remains the same, though the profit difference is now larger.

Table 16: Testing for sophisticated bidding: surplus from actual bids vs. unilateral deviation to numerically-calculated best response. In 6 of 11 cases, a unilateral deviation gives a significantly higher surplus, at the 90% level or stricter. This

L-types	Vickrey	FirstPrice	VNR	RR(0.50)	RR(0.75)-L1	RR(0.75)-L2
Win%	67.1(55.7)***	47.1(33.9)***	47.9(37.9)***	39.3(33.6)***	52.9(35.7)***	52.9(47.1)*
Surplus	31.0(40.5)*	14.3(42.0)***	26.5(40.7)***	21.0(33.5)**	14.9(44.3)***	25.8(32.2)
G-type						
Win%	32.9(26.4)**	52.9(39.3)***	52.1(50.7)	60.7(55.7)	47.1(42.9)	
Surplus	31.0(39.0)	25.0(61.0)***	55.0(58.0)	45.0(48.5)	47.0(57.5)	

For surplus, experimental medians reported; 'unilateral deviation' medians in parentheses.
Sign-test used for testing the win% variable, median-based permutation test used for surplus.
Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

10.4 Sample Instructions for the Experiment