Auctions For Complements - An Experimental Analysis*

Daniel Marszalec†

This draft: October 2014

Abstract: I evaluate the performance of four static sealed-bid package auctions in an experimental setting with complementarities. The underlying valuation model comprises two items, and three bidders; two 'individual' bidders demand one item only, while the third bidder wants to win both. The rules being compared include the Vickrey and first-price auctions, as well as the Vickrey Nearest Rule and the Reference Rule. Auction-level tests find the first-price auction revenue dominant over the other three rules, while the Vickrey auction performs worst, with the other two rules ranking intermediate. Bidder-level tests of the experimental data reject the competitive equilibrium bidding functions that can be derived in this setting, and I find that overbidding is widespread in all four auctions, as is aversion to submitting boundary bids. I also observe behaviour consistent with collusive bidding by participants in the Vickrey auction. Contrary to theoretical predictions, the Vickrey auction performs worst on efficiency, primarily for this reason.

In its currents state, the theory of auctions for single items and substitutes is well-developed, providing equilibrium bidding predictions and revenue rankings under a broad variety of assumptions. Yet many economic context fail to conform to those fundamental assumptions: bidders often demand multiple items and exhibit complementarities across units. Practical examples of such demand patterns include bidding on mobile telephony spectrum, contracts for serving bus routes or airport take-off and landing slots.[3] Until recently, there were no theoretical results as to the behaviour of auction rules even in simple settings of this type. Now that a few basic theoretical results are available, I aim to bridge the gap between theory and practice, by experimentally evaluating the extent to which bidder behaviour conforms to theory and our broad expectations of what optimal bidding in auctions should look like when complementarities are present.

I focus on four static sealed-bid auction rules: the Vickrey and first-price auctions, and the Vickrey-Nearest Rule and the Reference Rule. The motivation for picking these

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

*I would like to thank my supervisors Vince Crawford and Paul Klemperer, as well as Ian Jewitt,Thees Spreckelsen and Alex Teytelboym for their comments and criticisms. Financial support from the ESRC and RES is also gratefully acknowledged.

[3]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).

auction pairings is that the Vickrey auction embodies the well-known theoretical property of truthful optimal bidding which should induce a fully efficient allocation of items, while the first-price auction gives strong-incentives for bidding below the true valuation and may result in inefficient allocations. These two auctions make a good yardstick for assessing the pair of newer rules, both of which possess the core-selecting property, and were designed to mitigate the shortcomings of simpler mechanisms.

Recent experimental auction literature has focused primarily on dynamic auctions, such as the combinatorial-clock auction, and simultaneous ascending auctions.[4] This strand of research has been primarily concerned about investigating efficiency properties of those auctions, and to bidders' selection of packages in settings with complex valuation patterns. However, many practical implementations of these dynamic designs feature a one-shot static auction at the end: 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.[5] 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 was no prior experimental work in this area.

The primary aim of this paper is comparing the bidder and auction-level performance of the four aforementioned package auctions in a simple context with complementarities. I set up an experimental model with two items and three bidders, of whom two bid individually on one item each, while the third bidder bids for the bundle of both items together. A practical example of a similar setting would be an auctioneer selling a jacket and a pair of trousers: some buyers may only want the jacket, others may only need the trousers, but there may also be buyers who wish to buy both together to form a complete suit. Within this kind of context, I assess how bidders behave under each of the four auction rules, how much revenue is generated for the seller, and how close we get to a fully efficient allocation. Additionally, the resultant data also allow me to test theoretical predictions of equilibrium bidding, as derived by Ausubel and Baranov (2010).[6,7]

Experimental results obtained in this paper run contrary to theoretical expectations. Despite truthful bidding being theoretically optimal in the Vickrey auction, such behaviour does not emerge in my dataset, and this causes the Vickrey auction to be less efficient than predicted; in fact, it turns out to be the least efficient of the auctions considered. In terms of revenue, the first-price auction emerges dominant over its rivals, while the dataset cannot reject revenue-equivalence among the other three rules.

At the bidder level, I find that the data reject the theory-based equilibrium bidding functions for all bidder types. Furthermore, I observe that even a unilateral deviation towards equilibrium bidding would benefit the individual bidders, whereby the actual bidding patterns cannot be considered as a better response to other bidders' actual behaviour relative to the theoretical predictions.

The setup of my experiment also allows for investigating potentially collusive behaviour in the Vickrey auction. In practice, its susceptibility to collusion is frequently quoted as one of the major drawbacks the Vickrey auction, but the phenomenon has not been previously

[4] Kagle, Lien and Milgrom (2010 and 2014), and Kazumori (2010) are good examples of this.

[5] See ComReg (2012), Danish Business Authority (2012), and Ofcom (2012).

[6] Secondly, the experimental dataset created for this paper is intended as a benchmark for further, more complex, analysis of core-selecting auctions. Two series of follow-up experiments, investigating more complex competition and information patterns, as well as exposure, have already been conducted jointly with Alex Teytelboym, with results reported in Teytelboym (2013).

[7] Sun and Yang (2006, 2009) have also proved that in this setting there exists a dynamic incentive-compatible mechanism which finds the competitive equilibrium. In the present experiment, we only considered one-shot sealed-bid auctions, thus we could not include this mechanism in our comparison.

experimentally investigated in a multi-item sealed-bid context. In the present dataset, I find frequent overbidding behaviour consistent with attempts at collusive behaviour. However, perfect collusion is rarely achieved due to random matching and prohibition of communication among bidders. Nonetheless, the attempts at collusion are in themselves sufficient to undermine the revenue and efficiency properties of the Vickrey auction.

The rest of the paper is structured as follows. The four auction rules that I analyse are introduced in Section 1, and the precise formulation of hypotheses which are tested in this paper are discussed in Section 2. The experimental setup is presented in Section 3, and Section 4 performs a quality check of the data. Experimental results and tests of my hypotheses are performed in Section 5 and their interpretation is discussed in Section 6, while Section 7 concludes.

1. Auction Setup and Descriptions

In this paper I use a setting with three bidders and two items to model the context of bidding on goods with complementarities. I label the two items as '1' and '2', and assume that two of the bidders have a positive valuation on one item only. I call these the 'individual' bidders, and label them as I1 and I2, corresponding to whether they value item 1 or item 2 positively. The third bidder, J - the 'joint' bidder - has a positive value only on the bundle of items 1 and 2 together, and zero value on items 1 and 2 individually. Each bidder is permitted to bid only on the bundle they value positively, so the auctioneer will always receive three bids.

To model complementarities, I assume that the individual bidders' values are drawn from a uniform distribution on [0,100], while the joint bidder's value is drawn from a uniform distribution on [0,200]. I will use b_{i1} to denote the bid of bidder I1, b_{i2} for the bid of bidder I2, and b_j for the bid of joint bidder J. To denote the auction rule itself, I will use $P(b_{i1}, b_{i2}, b_j)$ to denote the payment vector conditional on the bid-triplet (b_{i1}, b_{i2}, b_j). Correspondingly, the individual payments assigned by an auction mechanism to the three bidder types will be labelled as p_{i1}, p_{i2} and p_j respectively, such that $P(b_{i1}, b_{i2}, b_j) = (p_{i1}, p_{i2}, p_j)$.

Prior to calculating the bidders' payments, the auctioneer solves a winner-determination problem where he picks a feasible bid-maximising allocation where each item gets allocated to at most one bidder. Given that in the present setting there are only two sensible allocations,[8] the winner-determination problem only entails checking whether the sum of the individual bidders' bids exceeds that of the joint bidder. If the sum of individual bids is higher, the I-types win one item each, whereas in the converse case the J-type wins both. This 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 whether it is the individual or joint bidders that win, in the Vickrey auction the price paid by the each winning bidder is determined solely by the bids of the other bidders. The price is calculated such that each bidder receives a payoff equal to the incremental surplus that he brings to the auction.

For a numerical example, consider $(b_{i1}, b_{i2}, b_j) = (48, 40, 60)$. Bidders I1 and I2 win the item, as the sum of their bids exceeds J's bid. The surplus that bidder I1 brings to the

[8] 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.

system is 28: without I1's bid, the auctioneer only faces the bids of $b_j = 60$ and $b_{i2} = 40$, whereby J would win both items, and the surplus - here evaluated at the bidders' bids - would be 60. With I1's bid of 48, however, I1 and I2 win, and the total surplus is 88 - an increase of 28. To give I1 a surplus of 28, the payment that the Vickrey auction charges I1 must thus solve $b_{i1} - p_{i1} = 28 \implies p_{i1} = 48 - 28 = 20$. By similar calculations, I2's payment is $b_{i2} = 12$.

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

$$P^{VA}(b_{i1}, b_{i2}, b_j) = \begin{cases} (VP_{i1}, VP_{i2}, 0) & if \quad b_{i1} + b_{i2} \geq b_j \\ (0, 0, b_{i1} + b_{i2}) & if \quad b_{i1} + b_{i2} < b_j \end{cases} \tag{1}$$
$$where : \begin{array}{l} VP_{i1} = \max[(b_j - b_{i2}), 0)] \\ VP_{i2} = \max[(b_j - b_{i1}), 0)] \end{array}$$

Despite its theoretically appealing properties of truth-telling and efficiency, there are a few well-known concerns about the Vickrey auction which limit its practical usefulness. Two of these reasons are low revenue generating potential, and susceptibility to collusion. From equation (1) we see that in the case when $b_{i1} + b_{i2} > b_j$ with $0 < b_{i1} < b_j$ and $0 < b_{i2} < b_j$,[9] the Vickrey auction will essentially 'leave money on the table', in the sense that $p_{i1} + p_{i2} < b_j$ - so 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.

More generally, the core is defined as a set of allocations for which there exists no 'blocking coalition', in the sense that no (sub)group of members of the system can jointly deviate to a different allocation which gives all members of that group a higher surplus. In the present example, the group consisting of bidder J and the auctioneer together constitutes a blocking coalition: J could offer the auctioneer a payment of $\tilde{p}_j = p_{i1} + p_{i2} + \varepsilon < b_j$, with $\varepsilon > 0$. This increases the seller's revenue, and gives bidder J a non-zero profit - so the allocation that assigns the items to I1 and I2 is not a core allocation, and correspondingly the price-triplet $(p_{i1}, p_{i2}, 0)$ does not lie in the core.[10]

When $b_{i1} + b_{i2} > b_j$, the set of core payments can be defined as:

$$(p_{i1}, p_{i2}) \in \{(x, y) \,|x + y \geq b_j, x \in [0, b_{i1}], y \in [0, b_{i2}]\}$$

This is the set of payments such that neither I1 or I2 pays more than their bid, but the sum of their payments (weakly) exceeds the bid of J. This set, along with the bids and Vickrey payments are shown in Figure 1, with the area corresponding to the core shaded in gray. The dotted diagonal line denotes the 'minimum revenue line', which contains all the points where the payments of I1 and I2 equal the payment of J exactly. The bold segment of this diagonal line depicts the 'minimum revenue core' (MRC),[11] which contains all the points that are simultaneously in the core, and on the minimum-revenue line. The notion that the MRC is capturing is to depict the combination of the least amount that each of the I-types can bid, subject to them jointly out-bidding the joint bidder. From the seller's point of view, 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.

[9] This case corresponds to the situation where I1 and I2 together out-bid J, but neither of the individual bids, on their own, would be sufficient to out-bid the joint bidder.

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

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

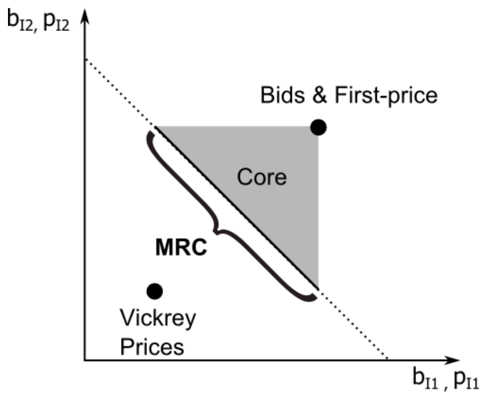

Figure 1. Vickrey prices, First–price payments and the MRC

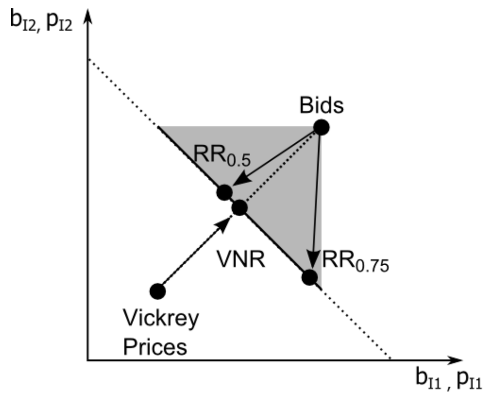

Figure 2. Vickrey Nearest, and Reference Rule with $\alpha = 0.5$ and $\alpha = 0.75$

The second concern with the Vickrey auction is its susceptibility to collusion. We see from equation (1) that in the case where I1 and I2 win, the payment of each of them is decreasing in the bid of the other.[12] If I1 and I2 decide - either explicitly or tacitly - to bid cooperatively, they can bid very aggressively, which will reduce their (joint) payments. To collude perfectly, in this setting, I1 and I2 can both bid $b_{i1} = b_{i2} = 200$, which is the highest possible value that J can have. Such bids makes sure that I1 and I2 always win, and both pay a price of 0. In less extreme cases, so long as both bidders overbid, they can still induce payments lower their Vickrey prices under truthful bidding. If only one of the two individual bidders attempts to collude and his co-bidder does not reciprocate, there is a possibility for the 'colluding' bidder to make significantly negative profits.

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, as in the present experiment. After the winner-determination problem has been solved, each winning bidder pays their bid in full for the item that they get allocated. The first-price auction can be thus be summarised as:

$$P^{FP}(b_{i1}, b_{i2}, b_j) = \begin{cases} (b_{i1}, b_{i2}, 0) & if \quad b_{i1} + b_{i2} \geq b_j \\ (0, 0, b_j) & if \quad b_{i1} + b_{i2} < b_j \end{cases}$$

Unlike the payments in the Vickrey auction, the first-price auction the winners' payments are always in the core, as shown in Figure 1. Indeed, in the case when I1 and I2 win, the first-price payments will also always lie (weakly) above the minimum-revenue line. Despite its simplicity, the first-price auction with package bidding has been used in practice numerous times, including the auctioning of bus routes in London (see Cantillon and Pesendorfer, 2006) and mobile telephony spectrum in Norway in 2013.[13]

1.3. The Vickrey Nearest Rule

The Vickrey Nearest Rule (VNR) is one of many recent implementations of core-selecting auctions. 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 can achieved by making the winners' payments less dependent on

[12] 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.

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

their own bids, but still require that the payment vector lie in the core.[14] The VNR auction, as introduced by Day and Cramton (2012), first uses the submitted bids to calculate Vickrey payments, 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 J wins, the Vickrey payment is in the core already, and the VNR implements that payment. However, if I1 and I2 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 'payments' using such a projection as the 'preliminary shares' of bidders I1 and I2, and denote them as s_{i1} and s_{i2} respectively. With this notation introduced, the VNR payments can be summarised as:

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

$$where : \begin{array}{l} s_{i1} = \frac{1}{2}(b_{i1} + b_j - b_{i2}) \\ s_{i2} = \frac{1}{2}(b_{i2} + b_j - b_{i1}) \end{array} \tag{3}$$

The payments of individual bidders in the VNR are broken down to consider three cases, depending on the relative asymmetry of the bids. If (say) $b_{i1} > b_j + b_{i2}$, so that I1 on his own out-bids J by a large margin, then $s_{i2} < 0$, which would suggest a 'negative' price for I2. By the non-negativity constraint on prices, we must then truncate $p_{i2} = 0$, and $p_{i1} = b_j$ to remain on the MRC. The converse case applies if $b_{i2} > b_j + b_{i1}$, and when the asymmetry moderate, so that $s_{i1}, s_{i2} > 0$, then both bidders pay a positive amount (which adds up to b_j).

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 individual payments from bids. In the VNR, individual bidders can influence their payment share by influencing the Vickrey prices, which depend (in part) on their own bid, as shown in equation 3. The innovation behind the Reference Rule is to define the bidder's payment shares in a way that further reduces the dependence on their own bids, while maintaining the core-selecting property by. This is achieved defining a 'reference point' which is independent of the I-types' bids, and then selecting the final payments that are closest in Euclidian distance to that point.

I will define each individual bidder as having a reference price which depends on the bid of the joint bidder J and a sharing parameter α. The reference price bidder I1 is $r_{i1} = \alpha \cdot b_j$, and the reference price for bidder I2 is $r_{i2} = (1 - \alpha) \cdot b_j$, with $\alpha \in [0, 1]$. The Reference Rule, with a particular value of α will be accordingly labelled as RR(α). Using this parametrisation, by varying α the reference point can be moved smoothly along the minimum-revenue line, with higher α putting the reference point closer I1's axis. With the reference point thus selected, the bidder payments in the Reference Rule can be defined as follows:

[14] 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.

$$P^{RR(\alpha)}\left(b_{i1},b_{i2},b_j\right) = \begin{cases} (r_{i1},r_{i2},0) & if \quad \begin{matrix} b_{i1}+b_{i2}\geq b_j,\ and \\ r_{i1}<b_{i1}, r_{i2}<b_{i2} \end{matrix} \\ (b_j-b_{i2},b_{i2},0) & if \quad \begin{matrix} b_{i1}+b_{i2}\geq b_j,\ and \\ r_{i1}<b_{i1}, r_{i2}>b_{i2} \end{matrix} \\ (b_{i1},b_j-b_{i1},0) & if \quad \begin{matrix} b_{i1}+b_{i2}\geq b_j,\ and \\ r_{i1}>b_{i1}, r_{i2}<b_{i2} \end{matrix} \\ (0,0,b_{i1}+b_{i2}) & if \quad b_{i1}+b_{i2}<b_j \end{cases} \tag{4}$$

$$where \quad : \quad \begin{matrix} r_{i1}=\alpha\cdot b_j \\ r_{i2}=(1-\alpha)\cdot b_j \end{matrix}$$

Since the reference prices are only required to lie on the minimum-revenue line, and not on the MRC, it is possible that the realised reference point will in fact lie outside the core. In such a case, the point on the MRC that is closest to the reference point is a payment vector that requires one individual bidder (say, I1) to pay his bid in full, while the other individual bidder's payment makes up the difference (between J's and I1's bid) such that the sum of payments ends up on the MRC.

We can now see how the deviation incentives under the reference rule differ from those in the VNR mechanism. In the VNR, each individual bidder's payment share always depends, to some extent, on his own bid. In the Reference Rule, so long as the *realised* reference point is on the MRC, the payment for each individual bidder is completely *in*sensitive to his own bid. The *only* case in which an individual 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. Indeed, in this sub-case, the relevant bidder's payment is as sensitive in the Reference Rule as it is in the first-price auction. However, this sensitivity occurs only under certain realisation of the bidder's values (and bids), and hence has limited impact on average.[15]

From Figure 2 it is also evident that even the reference rule with $\alpha = 0.50$ is not equivalent to the VNR payment in general.[16] 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 chose to use $\alpha = 0.75$ throughout the main sessions of our experiment; supplementary data for the Reference Rule with $\alpha = 0.50$ was obtained from an additional experiment, detailed in the Appendix.

1.5. Comparison of the four Auction Rules

A numerical and diagrammatic examples is natural way of illustrating the differences between the four mechanisms discussed above. Results from applying each of the auctions to the numerical example with $(b_{i1}, b_{i2}, b_j) = (48, 40, 60)$ are presented in Figure 3, below. Here the I-types win, and the J-type pays zero in each auction. To show the influence of varying α on the behaviour of the Reference Rule, I have calculated the payments for three differing values of α, denoted by RR(α) according to which α is picked. Note that for RR(0.25) the reference prices will be $r_{i1} = 15$ and $r_{i2} = 45$, which is outside the core, so the Reference Rule payments will be truncated to lie on the boundary of the MRC.

[15] 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.

[16] 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 $b_{i1} = b_{i2}$, these two points will differ.

This is not the case for $RR(0.75)$, and the payments in that case are not in the corner of the core.

	p_{I1}	p_{I2}	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

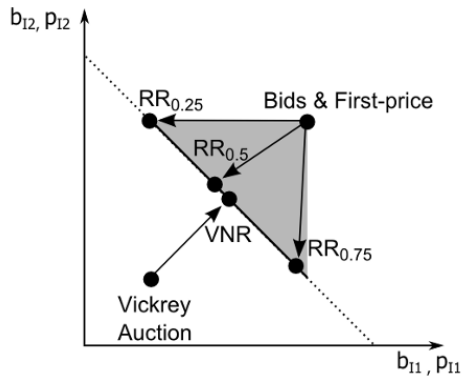

Figure 3. A numerical example of the four auction rules, with $(b_{i1}, b_{i2}, b_j) = (48, 40, 60)$

1.6. Bidding Restrictions and Investigating Collusion

None of the auctions discussed present a competitive equilibrium bidding solution that would require bids in excess of individual valuations. Hence a restriction of requiring bids to lie below the bidder's value 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. Not only do bidders bid more than theory would predict, but they also bid above their value, which can lead to negative payoffs.[17] It will thus be useful to gauge the extent to which such overbidding influences the performance of the four rules examined here, and whether it could in fact be the driving force behind any revenue or efficiency rankings.

A second reason to consider bidding restrictions is that it permits an evaluation of cooperative (or collusive) bidding in the Vickrey auction. Here both individual profits as well as auction revenue can be very sensitive to the presence of overbidding, as highlighted in Section 1.1. For the other three auctions that I compare, no obvious collusive strategies have been found.[18] Thus running a set of sessions with the same auction rules and instructions, but with adding or removing bidding restrictions, allows for a clean and direct assessment of this particular effect.

2. Hypotheses

The most direct application of an experiment such as mine is to test existing theory on the underlying auctions - this is surveyed in Section 2.1. Yet even in simpler settings and when complementarities are absent, the experimental auction literature frequently finds that theoretical predictions are not confirmed.[19] In addition, the existing theoretical results in my setting do not take into account the possibility of collusion, which may be a significant factor affecting the practical performance of that auction rule. Thus I propose

[17] 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).

[18] 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.

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

a few additional intuitively plausible hypotheses in Section 2.2, which can also be tested in the current experiment.

2.1. Related Literature and Theory

For my paper, the most relevant experimental work on auctions is Kagel, Lien and Milgrom (2010, 2014) and Kazumori (2010). 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 particular interest is in assesshing how well the auctions perform if bidders bid only on a subset of profitable packages in each round, rather than bidding on all packages, in each round. They find that straightforward bidding - submitting bids on the most profitable package only - leads to efficient outcomes (Kagel et al. 2010), bidders sometimes diverge from such bidding patterns to push up prices for their competitors (Kagel et al 2014). Kazumori investigates generalised Vickrey auctions, in addition to clock-proxy and simultatneous-ascending auctions. He finds that clock-proxy auctions out-performed the generalised Vickrey auction, and also outperformed the simultaneous-ascending auction when the value structure mirrored exposure. However, both these papers have looked at dynamic auctions, with complicated value and complementarity structures, and their focus has been the efficiency and package-selection questions.

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

The papers of as Ausubel and Baranov (2010), Goeree and Lien (2009) and Sano (2010) have provided theoretical foundations for optimal bidding in three of the four auctions I analyse. The work of Ausubel and Baranov (2010) provides theoretical results for optimal bidding in first-price package auctions, the VNR, and a the Proxy Rule auction of Ausubel and Milgrom (2002) under a valuation setting analogous to the one used in this paper.[21] The authors find that in the VNR the optimal strategy involves individual bidders with values below a certain cutoff to all pool into submitting a bid of zero, while all bidders with values above this cutoff should shade their bid by a constant amount. The optimal strategy for bidder type J is to bid his value. 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.

In the equilibrium of the Proxy Rule auction, which is equivalent to RR(0.50), low-value individual bidders will also pool in bidding zero, and higher value bidders will shade by a nonzero amount also. The degree of shading decreases with v, so for relatively low value of v, individual bidders will shade more in the Proxy Rule auction than in the VNR auction, though the ordering is reversed at higher values. At the extreme, when $v_i = 100$, there is no shading in the Proxy Rule auction. At the auction level, the results of 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 revenues, below the other two auctions. The efficiency ranking follows the same pattern as revenue.

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

[21] The optimal bidding functions for the VNR auction were independently found by Goeree & Llien (2009), and similarly the results for the Proxy Auction were independely found by Sano (2010).

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

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

2.2. Intuition-based Hypotheses

Even if bidders do not follow equilibrium strategies accurately, if we assume that participants are in fact responding to the auction incentives at all, we can try to predict their relative behaviour under the four rules. In the Vickrey auction, a bidder's price conditional on winning is independent of his bid, while there is a partial dependence in the core-selecting rules. I would thus 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 I would expect to invite more cautious bidding. This ranking of incentives does not apply to the J-type bidders, who face the same payment rule under all auctions except first-price. Testing whether these bidders all bid truthfully is contained in the hypothesis HT, but even if that hypothesis fails, it is possible that the J-types follow a similar non-truthful bidding pattern. I thus propose the following intuition-based hypotheses:

- Hypothesis HB: Individual bidder types will bid most aggressively (shade the least) in the Vickrey auction, and shade most in the First Price auction. The Reference Rule and VNR rank as intermediate.
- Hypothesis HJ: Bidder J 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 'rules of thumb' which appear suitable for the auction in question. The VNR and the Reference Rule were thus proposed as two mechanisms which aim to minimise (two different kinds of) incentives for deviation from truthful bidding. The intuition here is that because payments are 'close to independent of own bids' then bidders could find it 'close to optimal' to bid truthfully. This gives me another hypothesis that can be directly tested on my data:

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

The final set of hypotheses I set out to test in this experiment relate to collusion in the Vickrey auction. Collusion in games can be defined as behaviour that significantly 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 relatively 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

[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.

(2005), who look at collusion in discriminatory and uniform-price auctions with communication, and Kwasnica and Sherstyuk (2007), who in turn investigate Simultaneous Ascending Auctions with repeated play (within the same bidder group) but no communication.j The survey of Kagel and Levin (2008) finds that repeated play with the same opponents, and communication, tend to facilitate collusion, though this survey doesn't 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. Furthermore, the experiment was the first auction study run at the Oxford CESS lab, 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, whence the collusive strategies are easy to deduce: under perfect collusion, the I-types should bid exactly 200. Even if bidders don't fully notice this extreme solution, it is possible that the I-types realise that they can mutually benefit each other by bidding significantly above value. None of the other auctions in our 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 I do observe significant change of bidding patterns in the Vickrey auction across these two treatments, together with numerous bids in excess of value, these findings would be consistent with attempted collusion. I will thus test the following two hypotheses:

- Hypothesis HS: In auctions other than the Vickrey auction, the presence of bidding restrictions does not significantly affect bidder behaviour.

- Hypothesis HC: Removal of bidding restrictions in the Vickrey auction influences bidding behaviour, resulting in the I-types' bidding significantly more aggressively and in excess of value.

3. Experimental Design

The experiment was run over four sessions in 2010, 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 ($\sim$ \$55).[24]

During each session, the same group of participants played in each of the four kinds of auction. 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 participated in two payoff-irrelevant practice rounds, followed by the ten payoff-relevant rounds. 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 type were collected, and the instructions for the next auction type were distributed.

The understanding test that the participants were required to complete prior to proceeding to bid in a given auction type was carried out on paper. The test specified the

[23] I cannot exclude the possibility that they would have participated in auction experiments elsewhere.
[24] A sample of the instructions is available in the Online Appendix.

values and bids of all three bidders in the given type of auction, and required the participants to calculate who the winners were, what were their payments and profits. The participants were allowed to refer back to the instruction sheets while completing the tests, and everyone was provided with a calculator. Barring slight algebraic errors in calculator use, there were few problems with participants completing the tests and there were few failures.[25]

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 $[0, 200]$. The bidders were made aware that when bidding was unrestricted, even though they would never pay more than their bid, they could nonetheless end up with a negative payoff if they overbid and win at a price that exceeds their valuation. 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 tha auction was truncated to zero. Final payments were calculated as the sum of (possibly truncated) payoffs from all four auction types, plus a show-up fee.

4. A Comment on Data

Given that the experimental design is within subjects, I need to verify that bids are independent across auctions. To assess this degree of dependence, I ran series of pairwise estimations of Kendall's τ correlation parameter and tested its significance at 95% level. None of the tests for I-type bidders rejected, with all p-values>0.15. The tests on the J-type bidders also fail to reject the hypothesis that the bidding patterns are uncorrelated. 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 not too onerous.[26]

In addition to the four sessions wherein 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. Due to time-constraints (and participant fatigue), it would not have been feasible to run both $\alpha = 0.75$ and $\alpha = 0.50$ treatments within the main sessions. Since the data for RR(0.50) was 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 under other rules than RR. To allay concerns of such an effect being present, the supplementary experiments also contained a treatment wherein $\alpha = 0.75$, and this permitted for a consistency check between the two sets of data.

When I tested for differences in these two datasets for both the bidding and shading variables in RR(0.75), the two-sided Mann-Whitney and Kolmogorov-Smirnov tests both fail to reject with p-values >0.1, as do the tests for means and medians.[27] 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 setting are likely to be minor.

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

[26] The purpose of this test is to check that the assumptions of the statistical test I used, when evaluating bids, are satisfied.

[27] Even the t-test, which would be most prone to 'over-rejecting' the null of equal means, fails to reject at the 95% level.

5. Results

The discussion of the experimental results will start by evaluating auction-level performance in Section 5.1. I find that the first-price auction performs better than predicted, while the Vickrey auction underperforms relative to expectations. I subsequently look at bidder-level data in Sections 5.2 and 5.3 to analyse the possible reasons behind this aggregate result.

5.1. Auction-level Results

Revenue, surplus and efficiency are the three main parameters of interest for evaluating auction performance, each looking at the auction from a different perspective. Revenue is frequently of foremost importance to sellers, while bidders are primarily interested in their surplus; from a welfare or policy point of view efficiency is also relevant. An overview of these three parameters in the unrestricted bidding sample are provided in Table 1.[28]

Table 1: Revenue, Efficiency and Surplus Summary

	Vickrey [N=140]	FirstPrice [N=140]	VNR [N=140]	RR(0.50) [N=140]	RR(0.75) [N=140]
revenue	67.6 (56.9)	91.5 (37.1)	68.2 (41.2)	77.0 (42.3)	71.1 (46.3)
surplus	44.1 (67.6)	29.8 (28.1)	57.9 (39.1)	48.9 (49.3)	46.7 (49.6)
efficiency	88.9 (22.2)	97.5 (8.4)	97.7 (9.1)	94.9 (13.8)	95.1 (12.8)

Means reported, standard deviation below.

One immediately striking characteristic of Table 1 is how distinct the first-price auction looks from the other three rules: the revenue is higher, surplus is lower, and both variables have lower variance than their counterparts in other auctions.

Results from the pairwise tests and comparisons on revenue equivalence are shown in Table 2. The first-price auction revenue-dominates all other three rules, while pairwise comparisons between the Vickrey, VNR and Reference Rule cannot reject revenue equivalence. Though revenue in the Vickrey is lower than under the other rules, 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 thus reject; the first-price auction performs better than predicted, while the Vickrey auction underperforms.

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). None of the other tests for the null of zero median difference in surplus reject. The more rigorous pairwise testing thus confirms the intuitive conclusion from Table 1 - the first-price auction is very different from the others, giving higher revenue and lower surplus.[29]

Assessing efficiency using a direct median-comparison test, as above, is unhelpful, because in all the treatments the median efficiency is 100%.[30] A Kruskal-Wallis test nonethe-

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

[29] 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.

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

Table 2: Pairwise Auction Revenue and Surplus Comparisons

Revenue	Vickrey [N=140]	VNR [N=140]	RR(0.50) [N=140]	RR(0.75) [N=140]
FirstPrice [N=140]	29.0***	24.0***	15.0**	23.0***
Vickrey [N=140]		−3.0	−13.0	−7.0
VNR [N=140]			−9.0	−1.0
RR(0.50) [N=140]				8.0

Surplus	Vickrey [N=140]	VNR [N=140]	RR(0.50) [N=140]	RR(0.75) [N=140]
FirstPrice [N=140]	−16.0**	−24.0***	−17.0***	−17.0***
Vickrey [N=140]		−10.0	−2.0	−1.0
VNR [N=140]			8.8	8.0
RR(0.50) [N=140]				0.0

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.

less rejects with p-value<0.005, suggesting that efficiency is not homogenous across auctions. Hence I ran a series of Mann-Whitney tests, pairwise for each combination of auctions, to check whether the distribution of efficiency results is the same across the five rules, or whether a clear dominance pattern emerges. All but one pairwise comparisons against the Vickrey auction rejected at the 95% level or stricter, with Vickrey auction giving lower efficiency. The single auction that does not reject pairwise efficiency equivalence with the Vickrey auction is RR(0.50). No other strict ranking pattern emerges from the pairwise tests. These findings provide evidence to reject hypothesis HE, which would require the Vickrey auction to be most efficient, and the core-selecting auctions least.

5.2. Influence of Bidding Constraints

The impact of bidding constraints can be directly assessed by comparing the raw bid patterns across the two treatments and checking for differences. This comparison is done for each bidder type in Table 3.[31]

Only the Vickrey auction indicates that bidding restrictions have an effect, with bids significantly higher under unrestricted bidding. To put these numbers in perspective, recall that I-type values are uniform on [0,100] implying a median value of 50 (and median value of 100 for type J). The median-difference test accordingly rejects for all bidder types under the Vickrey auction at the 99% confidence level,[32] but none of the other

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

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

The test for median-differences used here is analogous to the Mann-Whitney test, but does not assume that the two compared samples have the same shaped distribution (which they do not, in our case, since 'bidding resrictions' truncate the strategy space).

Table 3: Testing the effects of bidding restrictions on Raw Bids

Case		Vickrey	First-Price	VNR	RefRule(0.75)
Bidder I1	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 I2	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 J	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 implemented via Hodges-Lehmann method.

Rejections of zero-difference null at 90%/95%/99% level indicated by */**/***.

auctions register any rejections. I thus accept hypothesis HS - bidding restrictions are only significant in the Vickrey auction. In subsequent portions of the paper, the analysis will be carried out using data from the sessions where bidding was unrestricted, though 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 5.3.3.

5.3. Bidder-level Results

With the exception of the Reference Rule with $\alpha = 0.75$, all other auction settings analysed in this paper offer symmetric incentives for both I-type bidders, and the data from these two sub-cases could thus be pooled for analysis. This intuition is confirmed by the data: in the Vickrey auction, as well as first-price and VNR and the symmetric RR(0.50), all Mann-Whitney tests for the null of ‘no difference’ between the I1 and I2 types fail to reject on both the bid and shading variables (all p-values >0.15). In the case of RR(0.75), the test rejects on the shading variable (p-value = 0.03). For the purpose of further analysis in this section, thus, the data for I1 and I2 types will be pooled in all auctions except RR(0.75), where I will consider both types separately.

To give an overview of individual bidding and assess hypothesis HB, Table 4 shows a set of pairwise median-difference tests across auctions for the bid and shading variables. The conclusions from these two variables are congruent: I-types bid the most in the Vickrey auction, and the least in first-price, while the three core-selecting auctions are intermediate, and show no significant difference from each other. Analogously, the amount of shading is greatest in the first-price auction, and lowest in the Vickrey auction, though I also find that the I2-type in the Reference Rule shades more than bidders in VNR and RR(0.50). The intuition of hypothesis HB cannot be rejected - the data shows that indeed Vickrey auction induces aggressive bidding, while first-price discourages it.[33]

When assessing the validity of Hypothesis HJ - that the J-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

[33] It does not, however, follow that bidders bid ‘closest to their value’ in the Vickrey auction: the aggressiveness in this auction results in over-bidding, which takes these bids *futher* away from the ‘true value’. If I test for the absolute deviation of bids from values, using the same methods as in Table 4, I find that bidders actually bid *closest* to their value in the VNR and Referene Rule.

Table 4: Pairwise Comparison of I-types' Bidding Behaviour

Bids	Vickrey [N=280]	VNR [N=280]	RR(0.50) [N=280]	RR(0.75)[I1] [N=140]	RR(0.75)[I2] [N=140]
FirstPrice [N=280]	-44.0^{***}	-14.0^{***}	-16.0^{***}	-13.0^{***}	-13.5^{***}
Vickrey [N=280]		30.0^{***}	26.0^{***}	30.0^{***}	27.0^{***}
VNR [N=280]			-2.0	0.0	0.0
RR(0.50) [N=280]				3.0	1.0

Shading	Vickrey [N=280]	VNR [N=280]	RR(0.50) [N=280]	RR(0.75)[I1] [N=140]	RR(0.75)[I2] [N=140]
FirstPrice [N=280]	31.0^{***}	10.0^{***}	10.0^{***}	10.0^{***}	13.0^{***}
Vickrey [N=280]		-20.0^{***}	-15.0^{***}	-19.0^{***}	-14.0^{***}
VNR [N=280]			0.0	0.0	3.0^{**}
RR(0.50) [N=280]				0.0	2.0^{***}

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.

auction types. In the same setting, a Mann-Whitney test comparing shading in the VNR and the RR auctions fails to reject (with p-value=0.12 for RR(0.75) and p-value=0.32 for RR(0.50)), which singles out the Vickrey auction as the anomalous one. On this evidence, the data reject hypothesis HJ.

5.3.1. 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 Baranov and Ausubel (2010). While no analytical results are available for RR(0.75) due to the asymmetry between I1 and I2 type bidders, equilibrium bidding functions can be obtained numerically.[34] In all of the core-selecting auctions, equilibrium bidding requires the I-types to bid exactly zero when their values are (sufficiently) low, and attempt to free-ride on the other I-type out-bidding the J-type on their own. Table 5 shows that experimental results diverge significantly from theory, while Figure 4 provides an illustration of how experimental bidding functions for I-types compare to their theoretical counterparts.[35,36]

For I-types, the bidding variable rejects in all sub-cases, with the exception of the I2-bidder in the RR(0.75) auction; the general pattern indicates that I-type bidders bid more

[34] The method used is similar to that which Baranov (2010) uses to obtain the equilibrium bid functions in the first-price auction.

[35] Analogous graphs for the J-types are provided in the Online Appendix.

[36] In Table 5 I use standard non-parametric tests for all variables except the 'surplus'. The surplus is calculated 'conditional on winning' which introduces a complex pattern of dependence across the 'experimental' and 'theoretical' samples: there are situations where an actual bid won in the experiment, whereas the corresponding theoretically predicted bid would not have won (and vice versa). Thus the samples are neither independent, nor matched-pairs. Given this dependence, I cannot use bootstrapping and use permutation-based tests instead. For further discussion of permutation tests, see Good (1994).

Table 5: Bidder-level Tests of the Theory

I-types	Vickrey [N=140]	FirstPrice [N=140]	VNR [N=140]	RR(0.50) [N=140]	RR(0.75) [I1, N=140]	RR(0.75) [I2, N=140]
Bid	80.0(48.0)***	31.5(18.2)***	48.5(35.8)***	45.0(3.1)***	50.0(32.7)***	45.5(48.5)
Win%	67.1(52.1)***	47.1(45.0)	47.9(36.4)***	39.3(32.9)**	52.9(35.7)***	52.9(35.7)***
Surplus	31.0(39.0)*	14.3(35.3)***	26.5(30.5)*	21.0(32.6)**	14.9(41.0)***	25.8(29.8)
J-type						
Bid	136.0(92.0)***	65.0(47.1)***	100.0(98.5)	122.5(112.0)**	106.5(94.5)**	
Win%	32.9(47.9)***	52.9(55.0)	52.1(63.6)***	60.7(67.1)	47.1(64.3)***	
Surplus	31.0(48.0)*	25.0(70.3)***	55.0(70.3)	45.0(63.7)***	47.0(61.7)	

For bid and surplus, experimental medians reported; theory-based medians in parentheses.
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 */**/***.

(and correspondingly shade less) than predicted by theory. Furthermore, the I-types bid exactly 'zero' much too rarely: theory would predict a total of 531 bids at zero in our data, whereas in practice only 83 (or 16%) of these materialise. 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 in part contribute to the scarcity of zero-bids, though I have not found a means of testing for this effect in the present set-up.[37]

The J-types also over-bid relative to theory in all auctions except VNR. However, in the core-selecting auctions and the Vickrey auction, the overbidding of the I-types dominates, which results in them winning more often than expected. Consequently the I-types also receive significantly lower surplus, conditional on winning, in all cases except the I2-bidder in RR(0.75). The variable for winning probability does not reject in the first-price auction, suggesting that though both I and J types overbid considerably, this does not affect their relative winning chances. Conditional on winning, both types make less profit in the

[37] A good analysis of this effect is Palfrey and Prisbrey (1997) in the context of public-goods contributions.

first-price auction than theory predicts.

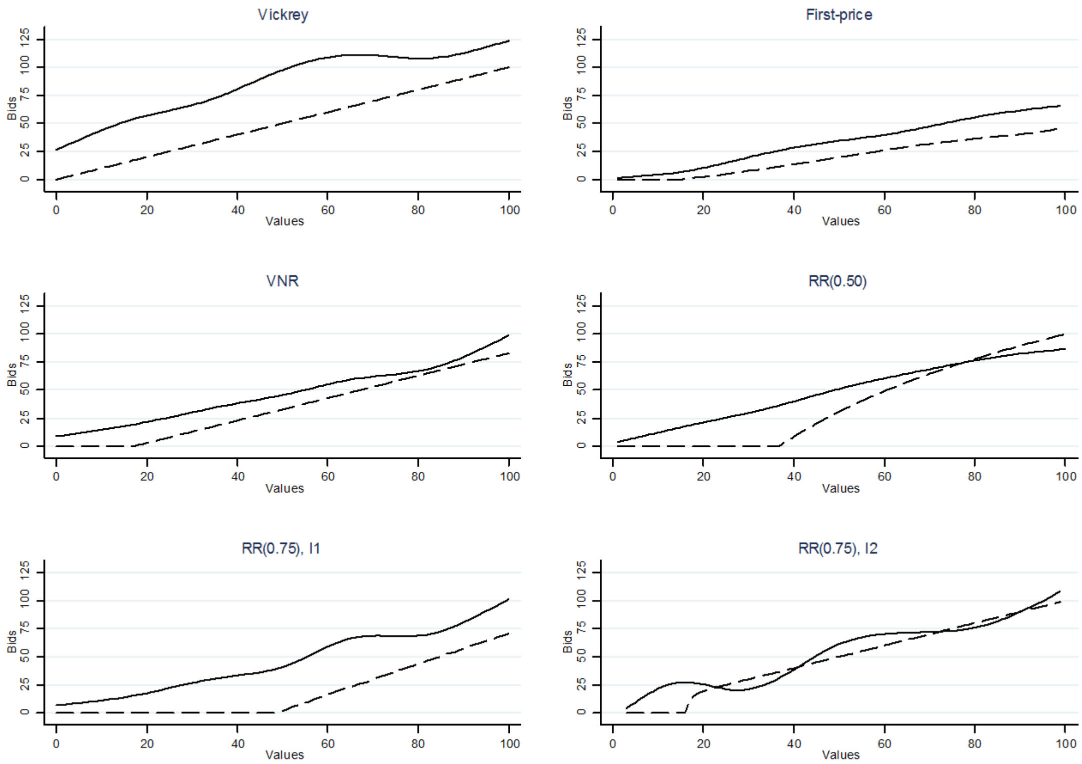

Figure 4. Bidding functions for I-type bidders: experimental (solid) and theoretical (dashed). Curves fitted are cubic splines with 7 knots and no monotonicity constraints imposed.

The broad conclusions from Table 5 and Figure 4 suggest that in all auctions the I-types over-bid significantly, relative to theory, thus winning too often, but making lower profits than predicted. Correspondingly, in all auctions except first-price, the J-type wins too rarely, and when he does win he makes little profit. Jointly, these findings lead me to reject hypothesis HT - competitive equilibrium bidding theory is not supported by my data.

Hypothesis HA, on truthful bidding in core-selecting auctions, similarly finds no support in my data. The highest p-value generated by a sign-test for truthful bidding is for the I2-type in the Reference Rule, and here p-value $=0.039$, which is still a rejection at the 95% level. For all other cases, the sign-test generates p-values <0.001. Thus neither equilibrium theory, nor intuitive arguments for truthful bidding are a good description of bidder behaviour in the experiment.

5.3.2. Testing for Sophistication

An objection that can be levelled against my rejection of theoretical equilibrium bidding is that I am testing against an invalid benchmark. The counterfactual that was used for comparison with experimental results was the scenario where all bidders follow the postulated equilibrium strategies, and under this counterfactual (with one exception) the bidders made less profit than theory would predict. In practice, however, since equilibrium bidding was so convincingly rejected, perhaps bidders in the experiment *know* that their opponents are not following equilibrium strategies, and hence they themselves engage in 'sophisticated bidding' rather than behaving in accordance with equilibrium theory.[38] The 'sophistication hypothesis' would suggest that the bidders' behaviour may in fact be a best-response to the actual (rather than theoretically predicted) behaviour of their rivals.

To assess whether sophisticated bidding could be a reason for rejection of the theory, I calculated profits and winning probabilities for all bidder types under the additional scenario where each of the three bidder types unilaterally plays the equilibrium 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 behaviour may indeed be a best (or a better) response to actual opponent behaviour. The results from this comparison are shown in Table 6.

Table 6: Testing for 'Sophisticated Bidding'

I-types	Vickrey [N=140]	FirstPrice [N=140]	VNR [N=140]	RR(0.50) [N=140]	RR(0.75) [I1, N=140]	RR(0.75) [I2, N=140]
Win%	67.1(55.7)***	47.1(38.6)***	47.9(44.6)**	39.3(34.3)***	52.9(35.7)***	52.9(49.3)
Surplus	31.0(40.5)*	14.3(31.8)***	26.5(31.6)**	21.0(31.3)**	14.9(40.3)***	25.8(29.0)
J-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; 'sophisticated bidding' medians in parentheses.
Sign-test used for testing the win% variable, median-based permutationtest used for surplus.
Rejections of zero-difference null at 90%/95%/99% level indicated by $^{*}/^{**}/^{***}$.

For individual bidders, the winning probability and conditional profit variables reject the zero-difference null in all cases except for the I2-type in the RR(0.75) auction. In all these cases, the unilateral deviation towards equilibrium bidding would lead to a (slightly, but significantly) lower winning probability, but a much higher surplus conditional on winning.[39] Since in Table 5 the I2-type's bidding in RR(0.75) was not significantly different from theory, it is consequently unsurprising that unilateral deviation towards theory here does not lead to higher conditional profit. Overall, however, the results suggest that the vast majority of I-type bidders are not engaging in sophisticated bidding.

The results for the J-type are more varied. In the first-price auction a unilateral deviation is profitable for the J-type for the same reason as it is for the I-types - the payment conditional on winning is now much lower. A similar deviation, however, does not significantly improve profits in any of the other auctions, nor does it much affect winning probabilities in VNR and Reference Rule. In all these auctions, the I-types' bids influence their payment in addition to the winning probability, but since J's payment depends only on I-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 J-type wins after over-bidding (and making a negative profit). Table 6 does reflect that this effect is

[38] The notion of 'sophisticated bidding' that I use here is analogous to 'sophisticated behaviour' in Costa-Gomes, Crawford and Broseta (2001).

[39] If instead of 'surplus conditional on winning' I were to look at the 'unconditional surplus', a sign-test on this variable rejects even more strongly than the permutation test of 'conditional surplus'. It would also reject in the additional case of the I2 bidder in RR(0.75), which does not reject under the present test.

present, since benefits from deviation towards theory are positive, but not sufficiently to be significant.

In informal post-experiment interviews, numerous participants noted that when playing as a J-type, they would bid 'more aggressively than felt sensible' in an attempt to counter the effects of overbidding of the I-types. Perhaps a part of this eagerness to overbid is due to the non-negativity of participant payments, or the lack of adequate training.[40] As shown by Table 6, such behaviour was unprofitable in Vickrey and first-price auctions, but made no significant difference in the core-selecting auctions.

The sophistication hypothesis thus gets rejected in seven of eleven sub-cases, and hence it is an implausible explanation for the rejection of theory, especially in the case of I-type bidders. Actual bidding never tested as giving significantly higher profits than an unilateral deviation towards theory.

5.3.3. Collusion in the Vickrey Auction

In Section 5.2 I found that without bidding restrictions, the I-types bid significantly more aggressively in the Vickrey auction, which is consistent with the collusive hypothesis HC. The next step is to evaluate whether such a bidding pattern can plausibly be attributed to collusion, or whether other explanations are more plausible.

The most direct method for checking whether collusion is present is to look for instances of 'perfect collusion', where both I1 and I2 bid 200. This criterion is very stringent and of limited use if mis-coordination occurs. While perfect collusion does occur in my data, this happens in only 5 out of 140 rounds of play. In these five instances, the joint profit of the I-types is 110, while the average for the whole sample is 54 when bidding is unrestricted.[41]

To move beyond checking only for perfect collusion, I must address two issues: firstly, picking a relevant non-collusive benchmark, and secondly, defining a criterion by which the degree of attempted collusion is measured. The most obvious benchmark for a given bidder type is to compare the restricted and unrestricted bidding patters - but we already know from Section 5.2 that bidders bid significantly more aggressively under no restrictions; in Section 5.3.1 I also showed that the I-types bid much higher than predicted by theory. This benchmark alone could, however, be misleading, since aggressive overbidding is frequently found even in simpler single-item auctions.[42]

A typical explanation of observed overbidding in auctions is that bidders 'like to win' and hence will bid more aggressively to win an item even if this reduces their profit conditional on winning. Such behaviour can look particularly attractive in auction rules where the payments lie below the submitted bid, and depend on bids of other players also: the increased likelihood of winning looks evident, while the payoff-consequences look less obvious.

The experimental setup allows me to construct a benchmark that measures this 'bid to win' effect, and use that to deflate the data from the Vickrey auction. The I-type 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, as it is in the Vickrey auction, it

[40] I did not find an obvious 'learning' pattern in the amount of overbidding by J-type bidders.

[41] Curiously, the average joint profit of the I-types conditional on winning when bidding is restricted is 89. The difference between this average, and the lower average surplus when bidding is unrestricted, comes from cases where collusion is 'attempted but not successful'. When only one I-type bids collusively and the other one does not, if they win, one of them is likely to pay in excess of his value. Another reason for this result is that with unrestricted bidding, the J-type also overbids, which inflates prices paid by I-types when they win.

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

does nonetheless provide the bidders with an opportunity to bid more aggressively without expecting large payoff-consequences. Thus looking at the differences in shading in these two auctions with, and without, bidding restrictions will allow me to construct a proxy for the 'bid to win' effect. This will be my 'non-collusive benchmark' - if the change in bidder behaviour in the Vickrey auction is significantly greater than this, then collusion may be present.

To gauge the 'extent' of the collusion attempts, I will use the amount of overbidding (in excess of the 'bid to win' amount) and the frequency with which such bids are submitted. If I only observe moderate and occasional overbidding, collusion is not very plausible, and such deviations could easily be attributed to miscalculation. On the other hand, if a significant portion of the data feature overbidding by a considerable amount, it is unlikely that such behaviour is purely accidental.

From Table 3, we saw that the largest median difference between restricted and unrestricted bidding treatments occurs in the Reference Rule for the I1-type, and the difference is -2. Thus when bidding restrictions are lifted, this bidder type does indeed bid more aggressively.[43] Using a sign-test to check whether the shading for I-types in the Vickrey auction is significantly more than this rejects equality with p-value =0.012; the change in bidding behaviour is thus sufficiently great to trigger suspicions of collusion. A summary of the numbers of overbidding I-types, as well as their median surplus, is show in Table 7, below.

Table 7: Numbers of Overbidding I-types

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)

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

The data shows that the number of overbidding I-types is much higher in the Vickrey auction at all overbidding levels than in any other auction. Recall that the average expected value of an I-type bidder is 50 - with this in mind, 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 above value; such magnitudes of overbidding are unlikely to be accidental, especially given how rarely similar deviations occur in the other auctions.

Even conditional on this pattern of overbidding, the bidders make more profit in the Vickrey auction than they would by bidding similarly in any of the other auction types. By overbidding as much as 50 points, the I-types in the Vickrey auction still make a positive surplus (mean of 3.7, median 3.0), whereas in other auction types by this point

[43] Note that this is the median decrease in shading, and though the median amount of shading is still positive, 25% of the bids of this bidder type involve overbidding relative to value.

the surplus is negative by a large margin. Since overbidding is both most prevalent and more profitable in Vickrey auction than under the other bidding rules, it is likely that the behaviour emerging from our dataset embodies aspects of collusive bidding.[44] It does not follow, however, that this type of bidding improves profits overall: in the rejection of the 'sophisticated bidding' hypothesis I showed that I-types in the Vickrey auction would do better by unilaterally deviating towards truthful bidding. What the data propose is a scenario where I-type bidders attempt to collude, despite frequent mis-coordination. As a result, the Vickrey auction underperforms doubly: even though in Section 5.1 it gave low revenue to the seller, at the individual level this has not translated into higher bidder surplus. Both the seller, and the bidders, end up significantly worse off than theory predicts.

6. Discussion

A summary of the assessment of the hypotheses of this paper is in Table 8. At the auction level, the theory-based hypothesis HR on revenue got rejected due to the superior revenue performance of the first-price auction, and the equally poor outcomes form the Vickrey auction. The expectation of full efficiency in the Vickrey auction was also not supported by the data, and indeed this auction ranked as least efficient. No significant differences among the other rules emerged, whence overall hypothesis HE was also rejected.

Table 8: Outome 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: I-types bid truthfully in VNR and Reference Rule	Rejected
'Sophistication hypothesis'	Rejected
HJ: J-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 I would intuitively expect. However, more precise hypotheses on bidding behaviour got rejected by the data. In the case of the first-price auction, this finding is similar to results on overbidding in single-unit contexts. For the core-selecting auctions, the VNR and Reference Rule, the picture is more complex. Relative to theory, experiment participants did not submit 'zero bids' nearly often enough when their valuations were low, and more generally they bid more than predicted. This lead to the rejection of hypothesis HT. Furthermore, the participants also did not bid truthfully in any of the core-selecting auctions, as I found in rejecting hypothesis HA. Thus neither theory, nor 'rule of thumb' behaviour is a good explanation of bidding in my experiment. Despite this, the rejection of the 'sophistication hypothesis' showed that unilateral deviations towards equilibrium bidding would be profitable for I-type bidders in five of

[44] The findings of Table 7 would not significantly change if I looked at the amoung of 'bidding in excess of equilibrium prediction' rather than looking at overbidding relative to true values. In paraticular, at high levels of overbidding, the Vickrey auction would still stand out.

six cases, which suggests that experiment participants were not best-responding to each other's actual bidding behaviour.

My interpretation of the behaviour of I-type bidders in the Vickrey auction proposes that they are attempting to collude, albeit usually with limited success. In all other auctions the presence of bidding constraints had no impact, as shown by the acceptance of hypothesis HS, while in the Vickrey auction extensive overbidding was observed when constraints were removed. The extent of the overbidding was far above what I could attribute to a 'bid to win' effect, and the number of extremely high bids was much higher than in any of the 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 such a situation where a Vickrey auction might be considered suitable option would be a one-off sale of, say, government assets or licences with a pure efficiency objective, and no concern for revenue. The results of this paper would suggest that even if revenue in itself is unimportant, the potential for attempts at collusive bidding in this auction is high, and that would be sufficient to undermine its efficiency properties.

7. Conclusions

This paper has taken a few first steps in the comparative experimental analysis of four package auction rules in the presence of complementarities. In terms of practical implications for auction design, the main finding of interest is the surprisingly good performance of the first-price auction: it yielded superior revenue compared to the other auctions, but with no significant efficiency loss.[45] The Vickrey auction performed worst on both of these criteria, while the core-selecting auctions ranked as intermediate. These conclusions run counter to the expectation that Vickrey auctions should yield 100% efficient outcomes, and the theoretical prediction of Ausubel and Baranov (2010) which would also rank it as revenue-dominant. Given that efficiency concerns are frequently used to argue against the use of first-price mechanisms in high value auctions, these experimental results provide evidence to allay such concerns.

At the individual level, I found that actual bidding diverged significantly from theoretical predictions, invariably due to bidders bidding in excess of the theoretical benchmark, and occasionally even above their own valuation. This behaviour could not be attributed to 'sophistication' of the bidders, as actual bidding never resulted in significantly higher individual profits compared to a unilateral deviation towards equilibrium bidding. The behaviour I observed in the Vickrey auction was consistent with attempts at playing collusively, even though such attempts were rarely successful. The Vickrey auction thus generated neither high revenue, nor high bidder surplus, confirming the widespread intuition that despite its theoretically attractive properties, auction designers are wise to avoid its use in practice even in one-shot setting.

8. Appendix A: The Variable-α Experiment

In the proofs and arguments that Erdil and Klemperer (2010) use to analyse the incentive properties of the Reference Rule, the actual reference point itself does not change the relevant deviation incentives on aggregate. However, it can significantly affect the

[45] This finding has subsequently turned out to be robust to changes in the complexity of the bidding setup, such as increase in the number of bidders, changes in amount of available information, and changes in the levels of exposure of the joint bidders. These follow-up experiments were carried out jointly with Alex Teytelboym, and the findings are reported in Chapter 5 of Teytelboym (2013).

relative amount that each bidder will have to pay, conditional on winning, and this may have non-trivial behavioural implications. Numerical calculations have shown that as α changes so do the optimal bidding functions, resulting in extremely disparate bidding by the two types as α tends to either 0 or 1.[46] This additional experiment set out to examine whether such variation would also emerge in a laboratory.

When investigating whether asymmetries matter in practice, it is useful to introduce some additional notation. Let K denote the upper end of the support of the value distribution of the I1 bidder. Then asymmetries in the valuations of the two I-types can be conveniently modelled as follows: set $v_{i1} \sim U[0, K]$ and $v_{i2} \sim U[0, 200 - K]$. This still keeps the sum of supports (and hence the expected total value) of the two I-type bidders the same as that of the J-type bidder, but by picking $K \neq 100$, the I-types are no longer symmetric. The nature of asymmetry in my experiment is thus summarised by the values of the two parameters, α and K. I considered four cases:

- Setting 1: $\alpha = 0.50$ and K=100 (i.e. $v_{i1}, v_{i2} \sim U[0, 100]$)
- Setting 2: $\alpha = 0.75$ and K=150 (i.e. $v_{i1}, \sim U[0, 150], v_{i2} \sim U[0, 50]$)
- Setting 3: $\alpha = 0.75$ and K=100 (i.e. $v_{i1}, v_{i2} \sim U[0, 100]$)
- Setting 4: $\alpha = 0.50$ and K=150 (i.e. $v_{i1}, \sim U[0, 150], v_{i2} \sim U[0, 50]$)

This particular combination of $\alpha' s$ and supports allows me to investigate two main issues. Firstly, I can check whether it is the asymmetry of the α parameter itself that influences behaviour: for this comparison, I look at the cases where the support of the two bidders' valuations stays constant, and the α varies.[47] Secondly, I can assess whether it is actually the magnitude of α *relative to* the 'expected valuation' of the bidders that matters: here I will compare the cases where the ratio of $\frac{E(v_{i1})}{E(v_{i2})} = \frac{\alpha}{1-\alpha}$, to those where it is not.[48]

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.[49] Again, the participants were allowed to ask questions whereafter they proceeded to complete an understanding test. The test was similar in format as in the main experiment, with the additional complication that the α–parameter varied from one question to another.[50] Upon successful completion of the test, the participants were informed which α parameters and which valuation model would apply in the given section of the experiment, and they subsequently proceeded to play two practice rounds, followed by ten payment-relevant rounds. The duration of the sessions in the Alpha-experiments was two hours on average, generating mean earnings of £27 (~$43).

8.1. 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 and shading variables

---

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

[47] Here the relevant comparisons are: Setting 1 v.s. Setting 3 , and Setting 2 v.s. Setting 4.

[48] Here the relevant comparisons are: Setting 1 v.s. Setting 4 , and Setting 2 v.s. Setting 3.

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

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

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. Thus when I test for the effects of varying α only, holding K fixed, none of the four test-pairing for the I-type bidders reject a zero-difference null even at a 90%. Hence α on its own does not significantly influence individual bidding.

An alternative to using direct bid and shading data is to look at bid and shading ratios,[51] but this approach will artificially inflate differences in the cases where $K \neq 100$. Here the two I-types have a different value support, and the I2-bidder with a narrower support is more likely to exhibit large variation in the bid ratios than I1, since the value, which is in the denominator, has a much lower expectation for I2 than it does for I1. The tests are hence likely to over-reject a zero-difference null. With that caveat in mind, I ran battery of median-difference tests for both I-types on bid as well as shading ratios, and still found only one statistically significant difference. The I2-type's bid-ratios in Setting 4 ($\alpha = 0.50,\ K = 150$) test as significantly lower (and shading ratios as correspondingly higher) than in all other cases. This is an intuitive finding, as in this case the I2-type can be seen to be in a particularly 'weak' position: they have a bidding support of only [0,50] (compared to I1's support of [0,150]), but their 'preliminary share' of the payments is a disproportionately higher 50%. As a result, in this setting the I2 type bids more cautiously. No other ranking beyond this, however, emerges from the pairwise tests.

A final hypothesis that can be tested on the individual bidder data is to check whether setting the α proportionately to the (relative) expected values of the two I-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. This could affect participants' bidding depending on where the α parameter is set, relative to their expected valuations. To test for such effects I pooled the data from settings 1 and 2, where α is set 'proportionately', and tested it against the pooled data from settings 3 and 4. Median-difference tests for both I1's and I2's bidding ratios failed to reject the zero-difference null (p-values>0.22 in both cases), and a similar pattern was observed for the shading ratio. Thus I could not find any influence of 'proportionality' on bidding at the individual level.

From the J-types' perspective, all four settings are identical, thus 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 J-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 J-type bidders are trying to 'best respond' differently to the I-types' actual bidding across the different settings, ignoring the prediction that truthful bidding should be optimal.

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 9. 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.

A series of pairwise median-difference tests for revenue is summarised in Table 10. The results hence confirm that the symmetric setting with K=100, $\alpha = 0.50$ 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); while the tests against Settings 2 and 3 yield p-values <0.05, these are

[51] These are calculated as the ratios of bid and shading relative to the value of the bidder.

Table 9: Revenue, Surplus and Efficiency Summary from alpha experiment

	K=100|α=0.50 [N=140]	K=150|α=0.75 [N=140]	K=100|α=0.75 [N=140]	K=150|α=0.50 [N=140]
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.

not significant due to the Bonferroni-Holm correction even at the 90% level. 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. In the case of these rejections, Setting 1 is less efficient, and no other pairings yield a rejection of the zero-difference null. It thus appears that using the RR(0.50), or the Proxy Rule in a symmetric setting yields superior revenue, but lower efficiency.

Table 10: Pairwise Revenue-difference Tests for variable-α experiment

	K=150|α=0.75 [N=140]	K=100|α=0.75 [N=140]	K=150|α=50 [N=140]
K=100|α=0.50 [N=140]	12.5*	14.0**	13.0*
K=150|α=0.75 [N=140]		2.0	0.0
K=100|α=0.75 [N=140]			−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.

A final test of interest at the auction level would be to assess whether revenue and efficiency are sensitive to setting the α proportionately to the bidders' expected values. This comparison is particularly significant for its policy implications relating to the relevance of 'reference points' - if the proportional cases where $\frac{E(v_{i1})}{E(v_{i2})} = \frac{\alpha}{1-\alpha}$ 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, with the median-difference being 7 points in favour of the 'proportional' settings. However, a corresponding Mann-Whitney test for efficiency rejects with a p-value<0.001, and indicates that there is only a 46% chance of drawing a higher efficiency value from the 'proportional' sub-sample. In practice the differences in efficiency are low (as seen also in Table 9), on average around 1.3 points, so the statistical significance here may not have much 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. This conclusion, however, is heavily influenced by the revenue-superior performance of the $RR\,(0.50)$ auction in the symmetric case.

Overall, the findings of the sessions on asymmetries did not offer many conclusive answers as to the influence of α. While I found some significant auction-level results in favour of setting α appropriately, the bidder-level data showed little sensitivity to α. Further exploration of different degrees and nature of asymmetry in the Reference Rule would be useful to illuminate the cause for this divergence, and to assess its policy implications more fully.

REFERENCES

Ausubel, L.M. & Baranov, O.V., 'Core-Selecting Auctions with Incomplete Information', *University of Maryland Working Paper* (2010).

Ausubel, L.M. & Cramton, P., 'Auctioning many divisible goods', *Journal of the European Economic Association*, vol. 2, no. 2-3, 480–493 (2004).

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

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

Cantillon, E. & Pesendorfer, M., 'Auctioning bus routes: the London experience', in Ausubel L., Cramton P. & R., Steinberg eds., *Combinatorial Auctions* (Cambridge, Mass., USA: MIT Press, 2006).

Cooper, D.J. & Fang, H., 'Understanding overbidding in second price auctions: An experimental study', *The Economic Journal*, vol. 118, no. 532, 1572–1595 (2008).

Costa-Gomes, M., Crawford, V.P. & Broseta, B., 'Cognition and Behavior in Normal-Form Games: An Experimental Study', *Econometrica*, vol. 69, no. 5, 1193–1235 (2001).

Day, R. & Milgrom, P., 'Core-selecting package auctions', *International Journal of Game Theory*, vol. 36, no. 3, 393–407 (2008).

Day, Robert W & Cramton, Peter, 'Quadratic core-selecting payment rules for combinatorial auctions', *Operations Research*, vol. 60, no. 3, 588–603 (2012).

Erdil, A. & Klemperer, P., 'A new payment rule for core-selecting package auctions', *Journal of the European Economic Association*, vol. 8, no. 2-3, 537–547 (2010).

Fischbacher, U., 'z-Tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics*, vol. 10, no. 2, 171–178 (2007).

Goeree, J.K. & Lien, Y., 'On the impossibility of core-selecting auctions', *Institute for Empirical Research in Economics University of Zurich Working Paper No.* (2009).

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

Goswami, G., Noe, T.H. & Rebello, M.J., 'Collusion in uniform-price auctions: experimental evidence and implication for treasury auctions', *Review of Financial Studies*, vol. 9, no. 3, 757 (1996).

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

Kagel, J.H. & Levin, D., 'Auctions: A survey of experimental research, 1995-2008', *Handbook of Experimental Economics Vol 2.*, vol. 2 (2008).

Kagel, J.H., Lien, Y. & Milgrom, P., 'Ascending Prices and Package Bidding: A Theoretical and Experimental Analysis', *American Economic Journal: Microeconomics*, vol. 2, no. 3, 160–185 (2010).

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

Kwasnica, A.M. & Sherstyuk, K., 'Collusion and Equilibrium Selection in Auctions', *The Economic Journal*, vol. 117, no. 516, 120–145 (2007).

Newson, Roger, 'Confidence intervals for rank statistics: Somers' D and extensions', *Stata Journal*, vol. 6, no. 3, 309 (2006).

Palfrey, T.R. & Prisbrey, J.E., 'Anomalous behavior in public goods experiments: How much and why?' *The American Economic Review*, vol. 87, no. 5, 829–846 (1997).

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

Sherstyuk, K., 'Collusion in private value ascending price auctions', *Journal of economic behavior & organization*, vol. 48, no. 2, 177–195 (2002).

Sun, N. & Yang, Z., 'Equilibria and indivisibilities: gross substitutes and complements', *Econometrica*, vol. 74, no. 5, 1385–1402 (2006).

Sun, N. & Yang, Z., 'A Double-Track Adjustment Process for Discrete Markets With Substitutes and Complements', *Econometrica*, vol. 77, no. 3, 933–952 (2009).

ComReg, 'Multi-band Spectrum Release: Information Memorandum', **URL:** `http://www.comreg.ie/_fileupload/publications/ComReg1252.pdf` (2012).

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

Federal Aviation Administration, 'Congestion Management Rule for LaGuardia Airport; Final Rule', *Federal Register*, vol. 73, no. 198, 60573–60601 (2008).

Norwegian Post and Telecommunications Authority, '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` (2013).

Ofcom, '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` (2012).

Teytelboym, A., 'Essays on Networks and Market Design', *D.Phil Thesis, University of Oxford* (2013).

9. APPENDIX B - ONLINE APPENDIX

9.1. Bidding Functions for J-type bidders

Figure A1, below, shows the J-types' biding curves under the five auction rules tested in our 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 J-types were rejected in all auctions except VNR. Looking at Figure A1. this conclusion is consistent with the presented graphs: in VNR the bidding function is indeed closest to the truthful-bidding prediction.

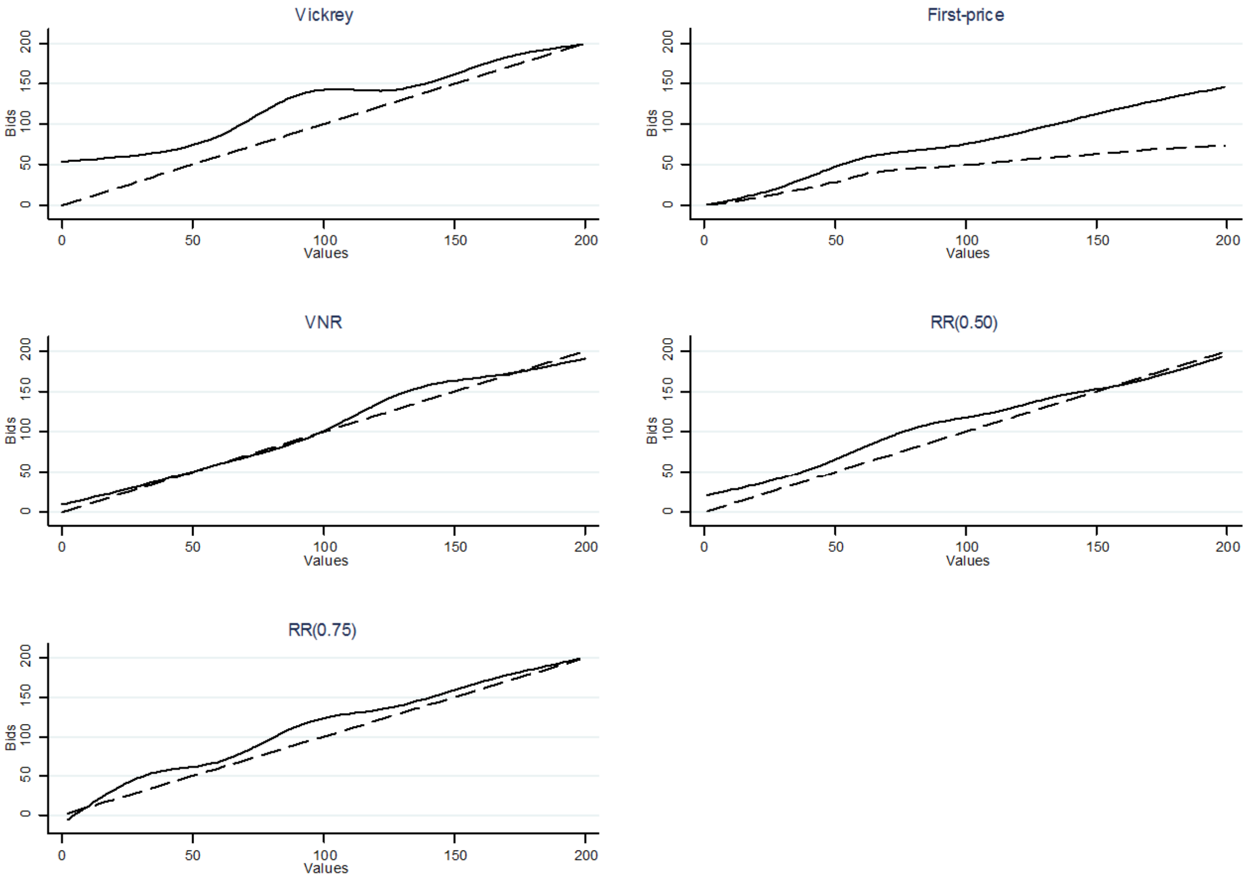

Figure 1: Figure A1. J-types' bidding functions: experimental (solid) and theoretical (dashed).

9.2. Further 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. Results of testing the extent to which actual bidding follows the theory are shown in Table 11, below.

Table 11: Comparison of actual v.s. theoretical bidding under bidding restrictions

I-types	Vickrey [N=160]	FirstPrice [N=160]	VNR [N=160]	RR(0.75) [I1, N=160]	RR(0.75) [I2, N=160]
Bid	50.4(52.6)***	35.1(20.9)***	42.7(31.8)***	40.5(16.8)***	43.1(46.5)***
Shade	2.2(0.0)***	16.7(30.9)***	4.8(15.7)***	7.3(31.0)***	5.6(2.2)***
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)
J-type					
Bid	95.6(99.1)***	77.3(45.8)***	94.7(97.2)***	93.6(96.7)***	
Shade	3.4(0.0)***	22.9(54.4)***	2.6(0.0)***	3.1(0.0)***	
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 shading, bid and surplus, experimental medians reported; theory-based medians in parentheses.
Sign-test used for testhing 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 get rejected at the 99% level, similarly to what I found on 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. 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 over-bid his value under bidding restrictions. Thus rank-based statistics will over-reject in all these cases, and the signed-test that I have used in Table 11 is particularly sensitive to this. This would affect the findings for all bidders in the Vickrey auction, as well as the J-types in VNR and Reference Rule.

Looking further at Figures A2 and A3, the bidding function in all the affected cases appear to be very close to the truthful-bidding line. It is likely, thus, that the rejection of theory in these cases is a statistical artifact. To verify this, I re-ran the comparisons in these four cases using a mean-based permutation test, the same kind I use to evaluate surplus in Section 5.3.1. This test did not reject 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.[52] The fact that I cannot reject truthful bidding for the J-types also means that I cannot reject hypothesis HC on the restricted bidding sample. Furthermore, a Kruskal-Wallis test 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.

[52] 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 11, 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 overlapped.

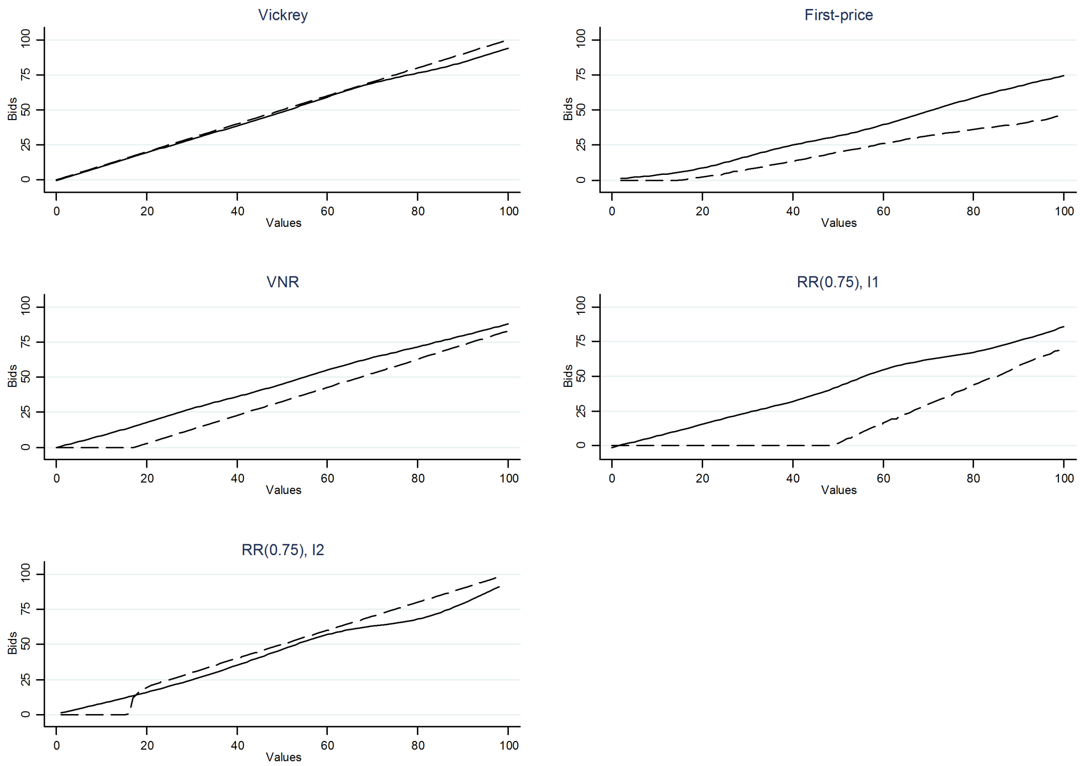

Figure A2. Bidding functions for I-type bidders under restricted bidding: experimental (solid) and theoretical (dashed).

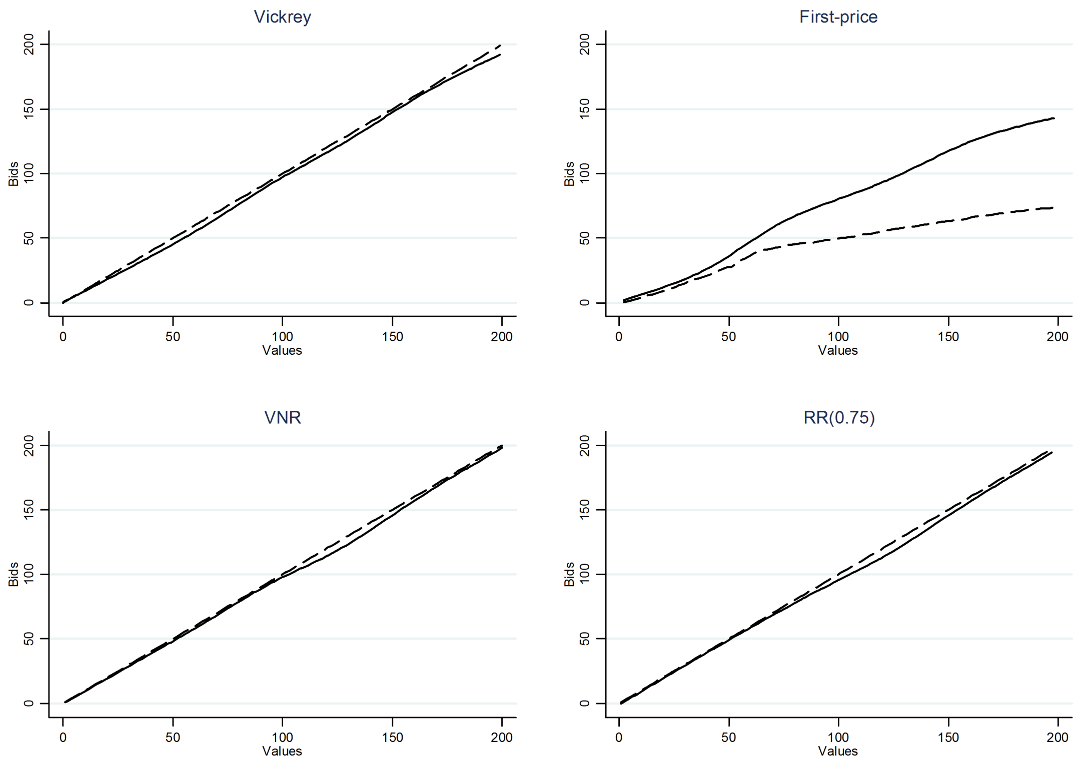

Figure A3. Bidding functions for J-type bidders under restricted bidding: experimental (solid) and theoretical (dashed).

There are three likely explanations for the discrepancy between the results here and those of Section 5.3.1, 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 create 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 'equilibrium behaviour'.

Pairwise comparisons of bidding and shading patters, in Table 12, show that bidders bids are lowest, and shading is highest, in the first-price auction. Similarly to Section 5.3, I also find that bidding in the Vickrey auction is significantly higher than in the other three auctions; though the shading variable also rejects, the median-differences are much smaller than with unrestricted bidding.

Table 12: Pairwise comparisons of bidding and shading under bidding restrictions

Bids	Vickrey [N=320]	VNR [N=320]	RR(0.75)[I1] [N=160]	RR(0.75)[I2] [N=160]
FirstPrice [N=320]	$-16.0^{\star\star\star}$	$-7.0^{\star\star\star}$	-5.0	$-8.0^{\star\star}$
Vickrey [N=320]		$8.0^{\star\star\star}$	$10.0^{\star\star\star}$	$8.0^{\star\star}$
VNR [N=320]			2.0	0.0

Shading	Vickrey [N=320]	VNR [N=320]	RR(0.75)[I1] [N=160]	RR(0.75)[I2] [N=160]
FirstPrice [N=320]	$13.0^{\star\star\star}$	$12.0^{\star\star\star}$	$9.0^{\star\star\star}$	$10.0^{\star\star\star}$
Vickrey [N=320]		$0.0^{\star\star\star}$	$-3.0^{\star\star\star}$	$0.0^{\star\star\star}$
VNR [N=320]			$-1.0^{\star\star\star}$	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 13. The revenue results in Section 5 were in part driven by over-bidding, 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 13: 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 14, below. As in the unrestricted bidding case, the first-price auction is revenue-dominant over the other three rules at the 99% confidence level, while no other pairwise tests reject revenue equivalence. Surplus in the first-price auction is correspondingly lower than under the other three rules. In addition, the pairwise test between the Vickrey auction and VNR rejects at the 95% level, with surplus being lower under VNR. The pairwise test between the Vickrey auction and the reference-rule does not reject, hence I cannot obtain a fuller unambiguous ranking.

Table 14: Pairwise comparison of revenue and surplus under restricted bidding

Revenue	Vickrey [N=160]	VNR [N=160]	RR(0.75) [N=160]
FirstPrice [N=160]	47.0***	37.0***	41.0***
Vickrey [N=160]		−10.0	−6.0
VNR [N=160]			4.0

Surplus	Vickrey [N=160]	VNR [N=160]	RR(0.75) [N=160]
FirstPrice [N=160]	−38.0***	−25.0***	−30.0***
Vickrey [N=160]		13.0**	8.0
VNR [N=160]			−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 did under unrestricted bidding, since removing the possibility for overbidding eliminates most of the cases in which the Vickrey auction fails.

9.3. Sample of Instructions for the Experiment

A sample of the instructions handed out to bidders during the experiment are attached below, followed by the understanding test that was administered prior to participation in the experiment itself. The first two pages were the same for all auctions, while the subsequent pages were auction-specific. The attached set of instructions is for the VNR auction, though the names of the auction rules were not revealed during the experiment.