

Auctions for Government Securities: A Laboratory Comparison of Uniform, Discriminatory and Spanish Designs

by KLAUS ABBINK, JORDI BRANDTS, and PAUL PEZANIS-CHRISTOU

revised version, December 2003

Abstract

The Bank of Spain uses a unique auction format to sell government bonds, which can be seen as a hybrid of a uniform and a discriminatory auction. For winning bids above the average winning bid, buyers are charged the average winning bid, otherwise they pay their respective bids. We report on an experiment that compares this auction format to the discriminatory format, used in most other countries, and to the uniform format. Our design is based on a common value model with multi-unit supply and two-unit demand. The results show significantly higher revenue with the Spanish and the uniform formats than with the discriminatory one, while volatility of prices over time is significantly lower in the discriminatory format than in the Spanish and uniform cases. Actual price dispersion is significantly larger in the discriminatory than in the Spanish. Our data also exhibit the use of bid-spreading strategies in all three designs.

Keywords

Treasury, Spanish auctions, discriminatory auctions, uniform auctions, multi-unit demand, common values, experimental economics

JEL Classification Codes

C92, D44, E42, E58, H63

Acknowledgements

We thank seminar participants at the Economic Science Association 2000 Conference (Barcelona), the 2000 Econometric Society European Meetings (Lausanne), the LEA Workshop on Auction Institutions and Applications (Bruxelles, 2001), the CEEL Workshop on Experimental Economics (Siena, 2001) and the annual meeting of the Verein für Socialpolitik (Magdeburg, 2001) for helpful comments. We are also grateful to Cristina Mazón and Gonzalo Rubio for helpful discussions and advice on the organization of the Spanish auction for securities and to David Rodríguez for valuable research assistance. Financial support by the European Union from a TMR-ENDEAR network grant (FMRX-CT98-0238), by the Spanish *Ministerio de Educación y Cultura* (PB98-0465), from the Barcelona Economics program of CREA, and by the Nuffield Foundation is gratefully acknowledged. Remaining errors are ours.

Affiliations

Klaus Abbink

School of Economics
The University of Nottingham
University Park
Nottingham, NG7 2RD
United Kingdom
Phone +44-115-9514768
Fax +44-115-9514591
klaus.abbink@nottingham.ac.uk

Jordi Brandts and Paul Pezanis-Christou

Institut d'Anàlisi Econòmica CSIC
Campus UAB
08193 Bellaterra
Spain
Phone +34-93-5806612
Fax +34-93-5801452
Jordi.Brandts@uab.es
ppc@iae.csic.es

1. Introduction

In the last ten years, the issue of the appropriate design for treasury auctions has received special attention in at least two instances. In 1992, the US Treasury Department launched the so-called “Treasury experiment”. Since then, securities with a maturity of 2 and 5 years have been sold using a uniform-price instead of the discriminatory-price auction used before. In 2000, the European Central Bank (ECB) decided to move from a fixed rate tender allocation procedure to an auction mechanism. After some internal discussions, the ECB eventually opted for a discriminatory-price mechanism.

In treasury auctions, buyers typically submit bids that specify a quantity and a price (or a yield) at which they wish to purchase the quantity demanded.¹ Once submitted, these bids are ranked from the highest to the lowest price (or from the lowest to the highest yield) and the quantity for sale is awarded to the best bids (i.e., highest prices or lowest yields). In a discriminatory-price auction, the winning bidders pay the price they have bid (or the price according to the yield they have asked for), whereas in a uniform-price auction, they are all charged a price that is equal to the *cut-off* price, the highest market-clearing price. Bartolini and Cottarelli (1997) survey the security trading practices of 77 countries and report that 39 of the 42 countries that auction off securities use the discriminatory-price format.

An alternative to these two mechanisms that has been used by the Bank of Spain since January 1987 is the so-called “Spanish auction”, which is a hybrid between the uniform- and the discriminatory-pricing mechanisms. On the one hand, as in the uniform-price auction, winning bids that are *above* the weighted average winning bid pay the same price; this price is equal to the weighted average winning bid. On the other hand, as in the discriminatory-price auction, winning bids that are *below* the weighted average winning bid are fully paid. The Bank of Spain uses this auction format for all marketable securities issued by the government: Treasury bills with a maturity of up to 18 months (*letras del tesoro*), notes with a maturity of up to 10 years (*bonos del estado*), and bonds with more than 10 years maturity (*obligaciones*).²

In this paper, we use laboratory experiments to compare the uniform, the Spanish and the discriminatory auction formats with respect to their revenue generating properties and to some other performance features. The volumes sold in treasury auctions are typically very large figures so that even small revenue advantages of one format over another can substantially reduce the government’s refinancing costs and are thus worth investigating.³

¹ Bids for smaller quantities can often be submitted without specifying a price. These *non-competitive* bids are always awarded at a price resulting from the auction.

² Official statements from the Bank of Spain explaining why they use this particular format are, to our knowledge, not available.

³ Taking the example of Spain, in 2001, the gross volume of securities sold in the Spanish treasury auctions summed up to 65.2 billion euros (Banco de España (2002)).

In this paper we use experiments to study some issues for which the theoretical approach does not yield any directly applicable results and field data are scarce. In such a context experiments can of course neither replace theory nor field studies, but they can be used to obtain distinct answers to the questions at hand. From an empirical standpoint, as the three auction formats do not co-exist in reality, it is difficult to carry out proper field studies to compare their properties.⁴

Laboratory experiments have recently been used to study the performance of various market institutions in a variety of contexts. Plott (1997), Abbink, Irlenbusch, Pezanis-Christou, Rockenbach, Sadrieh and Selten (forthcoming), Goeree and Offerman (2002) and Banks, Olson, Porter, Rassenti and Smith (forthcoming) assess and compare different spectrum auction designs in the US, the UK and the Netherlands. Rassenti, Smith and Wilson (2003) and Abbink, Brandts and McDaniel (2003) present experimental comparisons of different market mechanisms for the electric power industry.

There are a few experimental studies that specifically deal with treasury auction issues. In a model with perfect information, Goswami, Noe and Rebello (1996) first investigated the effect of allowing buyers to communicate with each other before the auction takes place. Such pre-play communication is found to depress prices in the uniform-price but not in the discriminatory-price auction, an indication that collusive behavior is more likely to emerge in the former than in the latter format.⁵ Ehrhart (2000) considers the ECB repo auctions and reports a comparison of discriminatory- to uniform-price auctions with multi-unit demands.⁶ Using linear demand functions, he finds that although bids are higher in the uniform auction, there is no significant difference in revenues. Ehrhart (2001) studies the fixed rate tender mechanism (previously used by the ECB) and finds that bidders massively overstate their demands, as observed in practice. A comparison of these two studies, which used the same experimental set-up, therefore suggests that the allocation by means of auctions helps to alleviate the demand overstatement problem which is inherent to the fixed rate tender mechanism.

We designed our experiment so as to capture the main strategic features of treasury auctions while keeping an experimentally tractable bidding environment. To this end, and in contrast to the above-mentioned experimental studies, we assume a framework in which the asset to be sold has a common value. This is a standard assumption in the theoretical literature on these markets. We believe this framework to be the simplest representation of the situation faced by the major institutional bidders in a treasury auction. These bidders act mainly as resellers;

⁴ Tenorio (1993) and Malvey, Archibald and Flynn (1997) compare the performance of discriminatory to uniform auctions with field data. Other empirical studies of treasury auctions include Simon (1994), Nyborg and Sundaresan (1996), Gordy (1999) and Nyborg, Rydqvist and Sundaresan (2002).

⁵ An early theoretical comparison of the stability of collusion across two single-unit auction formats is given in Robinson (1985). He finds that as long as cartel members share the same information, cartels are – both for independent private values and common values - stable in open ascending bid auctions, but not in sealed high-bid auctions.

⁶ In repo auctions, the central bank auctions off the money to the banks who need it to refinance their obligations. These auctions follow the same procedures as treasury auctions. Since they are used mainly as instruments for monetary policy, revenue maximisation may not be the seller's primary goal.

they purchase government securities in the auction and resell them to their customers. As government securities are traded in an organized secondary market, their resale price can be considered the same for all bidders in the primary market. However, prior to the auction, this resale price is uncertain, and individual bidders' expectations about this price may well differ.⁷

Two other features of our design are worth highlighting here. First, in contrast to most of the literature on common value auction experiments, we consider a multi-unit demand framework. Second, we run long auction series of 75 rounds with the same set of participants. There are several reasons for this design choice. The main reason is that, effectively, treasury auctions are held on a regular basis with the same set of bidders meeting repeatedly. This opens the door to the appearance of supergame effects. However, since such effects are possibly present in natural environments and given that our motivation is fundamentally applied we think that it is better that our experiments allow for repeated interactions. A further reason for this design choice is that it gives subjects substantial experience with the game and, at the same time, is a good way to obtain a sufficient number of statistically independent observations. The alternative procedure based on random matchings of subjects makes it very hard to reconcile these two objectives.

2. Theoretical Background

In this section we briefly review the relevant theoretical background for our experiment. Treasury auctions are typically modeled as auctions of a perfectly divisible common value asset, for which bidders receive private independent signals. In these markets, bidders are assumed to compete with each other for the purchase of asset shares in terms of demand functions that specify bid-quantity pairs. The treasury then allocates the available supply of assets by fulfilling the demands of the highest bidders.

Wilson (1979) first showed by means of examples that the uniform-price version of these auctions is unfavorable to the seller in terms of revenues, whether bidders are risk neutral or constant absolute risk averse. Back and Zender (1993) elaborate on this model to study the rationale for the US "treasury experience." They assume the existence of a reserve price and of an uncertain supply of assets, which are important features of the US and other treasury auctions, and analyze discriminatory and uniform-price auctions with risk neutral bidders.⁸ Their main result shows that both auction institutions yield multiple equilibria and that the uniform-price auction permits self-enforcing collusive bidding strategies, which unravel in discriminatory auctions. Wang and Zender (2002) further show that if bidders submit identical (piecewise continuously differentiable) demand functions, then the equilibrium demand

⁷ For simplicity we do not model the "When Issued" market, which in Spain is called "grey market" (*mercado gris*).

⁸ For a theoretical analysis of the effect of supply uncertainty on the outcomes of treasury auctions, see Back and Zender (2001) and Mazón and Núñez (1998).

functions in both auction institutions are downward sloping if bidders are risk averse (because risk aversion implies decreasing marginal valuations) and flat if bidders are risk neutral.

To allow for a revenue ranking of the uniform and discriminatory auction formats, the above analyses need to assume a symmetric information framework where all bidders receive the same signal instead of independent private signals that are drawn from a given probability distribution.⁹ Within this symmetric framework, and assuming risk neutral bidders, Back and Zender (1993) and Wang and Zender (2002) show that the discriminatory auction yields a unique equilibrium with a greater expected revenues than the uniform auction. However, Wang and Zender (2002) also find that the opposite revenue ranking will be observed when bidders are risk averse and the number of bidders is large compared to the available supply.

As for a theoretical analysis of the Spanish auction format, the combination of its price setting rule with the multi-unit demand feature introduces the additional difficulty that the prices to be paid by the winning bidders do not only depend on the ranking of buyers' bid-quantity pairs but also on a function that weighs the winning bids by the quantities demanded (i.e., the weighted average of winning bids). To this extent, the quantities chosen affect the determination of all prices, along with the winning bids. Nevertheless, Alvarez, Cerdá and Mazón (1999) show, in a model *à la* Wang and Zender (2002) with symmetric information but where (constant absolute) risk averse bidders can only submit linear demand functions, that there exists a unique equilibrium. In this equilibrium, the Spanish auction yields greater expected revenues than the other two formats if the number of bidders is large in terms of the expected supply of assets.

Gordy (1996) uses numerical methods to analyze discriminatory-price auctions in which bidders have (constant absolute) risk averse preferences and two-unit demands. His focus is on the use of bid spreading strategies, which are straightforward to test for in a two-unit demand framework.¹⁰ According to his conjecture, bidders submit multiple different bids for different quantities to hedge against the winner's curse or to express their risk aversion, as shown in Wang and Zender (2002). Gordy's intuition for this conjecture is that by submitting downward sloping demands, a bidder reduces the winner's curse by aligning his/her outcome more closely to the aggregate outcome.

Using similar numerical methods to those used by Gordy (1996), Alvarez and Mazón (1999) compare the revenue performance of the Spanish format to those of the discriminatory format when there are at maximum four buyers who each bid for two units of a common value asset.

⁹ Ausubel and Cramton (1998) provide a framework that allows for a general comparison of the equilibrium revenue properties of various formats when bidders receive private information and are risk neutral. In particular, they show that if bidders are allowed to bid for the total supply of assets then the multiple equilibrium revenues generated by the uniform- and the discriminatory-price formats are bounded from above by the revenues generated by second- and first-price auctions for a single unit, respectively. To this extent, and according to the predictions of Milgrom and Weber (1982) for single unit auctions with affiliated signals, uniform-price auctions would generate higher maximum expected revenues than discriminatory auctions.

¹⁰ A two-unit demand framework can also be seen as being close to the average number of bids per bidder submitted in the real treasury auctions of Portugal (i.e., 2.8, see Gordy (1999)) and Spain (2-3, see Martinez Mendez (1996)).

Interestingly, the Spanish auction mechanism is found to yield higher revenues to the seller for a range of different parameter values (i.e., the number of buyers, the degree of buyers' (constant absolute) risk aversion, the parameter of the common value's distribution and the total number of units sold). Their analysis also suggests that, in equilibrium, the Spanish auction would induce bidders to increase the spread of their multiple bids (i.e., bid-spreading) more than the discriminatory format.

One can see that the theoretical revenue comparisons across various auction formats are model-dependent and that, in particular, there are no results available for the discrete multi-unit environment. In such a case, experiments can provide an independent perspective on the issues at hand.¹¹ In the next section we describe our experiments.

3. Experimental Design and Procedure

3.1. Design choices

We use an extension of a specific design frequently used in experimental auction studies (Kagel and Levin (1986, 1999)). Consider a unit of an asset put up for sale in a treasury auction. We assume that the unit has a "true value" v , which is drawn (with replacement) from a uniform distribution defined on the interval $[a; b]$. As mentioned above, this value can be interpreted as the unit's resale price in the secondary market. We treat this value as exogenous and thereby abstract from possible interactions between the primary and the secondary market.¹²

Bidders are not informed about the realization of v , but each bidder i receives a private estimate, s_i , which is independently drawn from a uniform distribution defined on $[v-e; v+e]$. The parameter e is common information to all bidders. The signal stands for an individual bidder's assessment of the resale price in the secondary market, which may be optimistic (this corresponds to a high signal) or pessimistic (this corresponds to a low signal). Each bidder knows his signal, but not those of the other bidders.

We chose the support of the common value's distribution to be $[1000, 5000]$ and the signal's distribution to be $[v-200, v+200]$. To simplify experimentation, we use only integer numbers. By choosing a support for the common value's distribution that is relatively wider than the one of the signal's distribution, we ensure that the probability that subjects receive private signals that are in the interval $[1000; 5000]$ is close to one.

¹¹ While theoretical revenue comparisons are model-dependent, experimental revenue comparisons are potentially design-dependent. We think that our experimental design is rather natural, but of course it is not the only one.

¹² This assumption is in line with previous empirical work on treasury auctions, see e.g. Nyborg and Sundaresan (1996), Malvey, Archibald, and Flynn (1997). The interaction between primary and secondary markets has been analyzed by Bikhchandani and Huang (1989).

Subjects were allowed to bid for up to two units of a homogeneous (and fictitious) asset. Compared to the alternative of asking subjects to submit price-quantity pairs, the submission of unit bids has the advantage of making it easier for subjects to predict what the weighted average winning bid could be. With unit bids, the weights of the winning bids are all equal to one. Further, the restriction to two bids allows us to check for the use of bid-spread strategies in a straightforward manner, without having to account for the quantity dimension.

We conducted auctions with eight bidders, which is in the mid-range of the actual number of major buyers in treasury auctions. Since each bidder can submit up to two bids¹³ (i.e., no bid at all, one bid or two bids), there was a maximum possible demand for 16 units. In each auction, seven units were put up for sale.¹⁴

Bids were expressed in *talers*, the fictitious currency used in the experiment. They had to be no lower than 500 *talers*. We imposed this minimum bid in order to avoid ambiguity in the interpretation of zero bids. Without a positive minimum bid, zero bids could have been interpreted both as a refusal to bid and as an attempt to receive an item for free. The maximum possible bid a bidder could submit in the Spanish and the uniform-price auction treatments was set to 6000 *talers*. In the discriminatory auction treatment, subjects were not allowed to bid more than the highest possible value of the item given their signal. We introduced such a ceiling on bids to prevent subjects from submitting by error ruinously high bids.¹⁵ However, these bid restrictions did not turn out to be a constraint for subjects' actual bidding behavior.

Subjects submitted their sealed bids simultaneously through interconnected computer terminals. Once all bids were collected, the software ranked them in decreasing order and awarded a unit to the seven highest bids. If two or more bids were equal, the order in which they were considered was randomly determined; this is important for the case where the seventh highest bid corresponds to more than one unit. The software computed the average winning bid and specified the prices to be paid. In discriminatory auctions, winning bids were fully paid. In Spanish auctions, winning bidders were charged the average winning bid for all bids above this average, while all winning bids that fell below this average were fully paid. In uniform auctions bidders paid the seventh highest bid for all winning bids. Upon winning a unit, a bidder made a profit equal to the difference between the value of the item and the price he paid.

¹³ Most experimental common-value auctions involve single-unit demand. An exception is Manelli, Sefton and Wilner (2000), who report an experiment that considers a multi-unit demand framework, where valuations either have a common value component or are private information. They compare the revenue and efficiency properties of uniform-price auctions to those of a dynamic English auction and look at the effect of introducing a common value component in the buyers' valuations. They find revenue to be lower in the dynamic English auction in both the common value component and the private value cases.

¹⁴ In the Spanish auction for securities, a little less than half of total demand is usually satisfied (see the tables in Martínez Méndez (1996)).

¹⁵ In the discriminatory auction, a bid that is higher than the highest possible value will *always* lead to losses. In both the Spanish and uniform auction formats, it is not straightforward that bidding more than the highest possible value is always disadvantageous so that we did not impose a bid ceiling for these formats.

After each auction, subjects were informed about the realization of the common value. In addition, all subjects received information on the average winning bid, all selling prices, and the number of units sold at each selling price. Subjects were not informed about bids and signals. All the information provided to subjects was appended to a history window that could be retrieved at any time during the experiment.

3.2. The conduct of the experiment

The experiment was conducted at the *Universitat Pompeu Fabra* in Barcelona. We conducted one treatment with the Spanish auction design, one with the discriminatory-price design and one with the uniform-price design (henceforth SPAN, DISC and UNIF). The experiment was computerized with software developed using RatImage (Abbink and Sadrieh (1995)). Subjects were recruited by public advertisement on campus and were mostly economics students, but typically not extensively trained in auction theory in particular.

We conducted a total of twelve sessions with 192 subjects and each subject was allowed to participate in only one experimental session. Hence, sixteen subjects participated in each of the sessions. They were randomly assigned to one of two separate markets of eight participants. Each market consisted of 75 consecutive repetitions (hereafter, *rounds*), which was known to the subjects. By letting subjects play that many rounds, we made sure that they had the opportunity to gain experience; this seems especially important to us because the participants in treasury auctions are typically experienced professionals. As already mentioned, the same eight subjects played in the same market throughout the session to reflect the repeated interaction in treasury auctions. We will refer to such a set of subjects as a *market*. Subjects were not told with whom of the other session participants they were in the same market.

We used eight samples of the common values and the private signals. Each sample was generated in the following way. The common value was independently drawn for each round and, conditional on the realization of this common value, the private signals were also independently drawn. To control for sampling errors in the across-treatment analysis of the data, each sample was used for one group in SPAN, one group in DISC and one in UNIF, i.e. the realizations of the common value and of the signals were the same for one group of SPAN, one of DISC and one of UNIF.

In each session, two samples were needed because two markets were conducted in parallel. We report the results of twenty-four markets, eight with each treatment.¹⁶ Thus, we gathered eight independent observations per treatment because in each session, subjects bidding in one market did not interact with those bidding in the other market.

¹⁶ Due to a technical problem, two additional sessions were held with the Spanish treatment using the same samples of random variables as in one previous session. These sessions are not reported because we do not have data from additional sessions from the other two treatments to match them with. However, none of our conclusions changes if we replace the original data by these additional sessions. We consider this to be reassuring for our findings.

At the outset of each session, subjects were granted capital balances of 4000 talers (€12.04 or \$11.00 at the time) to which gains were added and from which losses were subtracted. The total earnings of a subject from participating in this experiment were equal to his capital balance plus the sum of all the profits he made during the experiment minus the sum of his losses. A session lasted for about 2 to 3 hours (this includes the time spent to read the instructions; an English translation of which is reproduced in Appendix I). At the end of the experiment, subjects were paid their total earnings anonymously in cash, at a conversion rate of one Spanish peseta for two talers. Subjects earned on average 3951 Spanish pesetas (€3.79, \$21.73), which is considerably higher than the students' regular wage. One hundred pesetas are equivalent to 0.602 euros. At the time of the experiment, the exchange rate to the US dollar was approximately \$0.55 for 100 pesetas.

4. Results

We are particularly interested in the revenue outcomes and subjects' bidding behavior towards the end of the experiment. Therefore, we focus on the last third of the experiment (rounds 51-75). Before the first session was held, we decided to concentrate on these late rounds to make sure that subjects would have gained sufficient experience. However, for comparative purposes, we will also report results for other rounds.

Recall that groups constitute the independent observations in our experiment; bidders in the same market interact with each other and are very likely to influence each other's behavior. Therefore, statistical tests that require independence of observations can only be conducted at the group level. We test for treatment differences in the following manner. First, we compute the respective test statistic (e.g. the revenue level) for each group in each treatment. Then, we compute pair-wise treatment differences for groups that used the same set of realizations of values and private signals. As will be seen below, this test allows us to characterize significant treatment differences in behavior. On the basis of eight observations, we then test the significance of these differences by applying the non-parametric binomial test on the frequencies of positive and negative differences.

4.1. The Seller's Revenue

Figure 1 plots the average revenue raised (the sum of the seven selling prices) in the last 25 rounds for each of the eight samples of each treatment. Each triple of bars refers to the groups of SPAN, DISC and UNIF that used the same sample of common values and private signals. Notice that seven out of eight times, the seller's revenue is higher in SPAN and in UNIF than in DISC. According to the binomial test, these differences in the revenues are significant at a one-sided $p = 0.035$. In contrast, the difference between SPAN and UNIF are not significant.

To verify whether the data shown in Figure 1 are the result of stable behavior we need to look at a representation of revenue over time. However, for that purpose we need to transform the

data. Given the high round-to-round variation in the realizations of the common value v , the seller's revenue varies substantially across rounds. To allow for a meaningful comparison of the revenues generated by the different treatments across rounds, we define the *standardized* revenues, SR , as the difference between the actual revenue, i.e. the sum of selling prices, and the total value of the seven items for sale, that is:

$$SR = \sum_{i=1}^7 p_i - 7v.$$

This measure indicates how much of the total surplus is extracted from buyers. Figure 2 plots the evolution of this measure aggregated over the eight markets of each treatment. We show moving averages over five rounds to smoothen the plots (Appendix II shows the information separately for each triple of markets).¹⁷

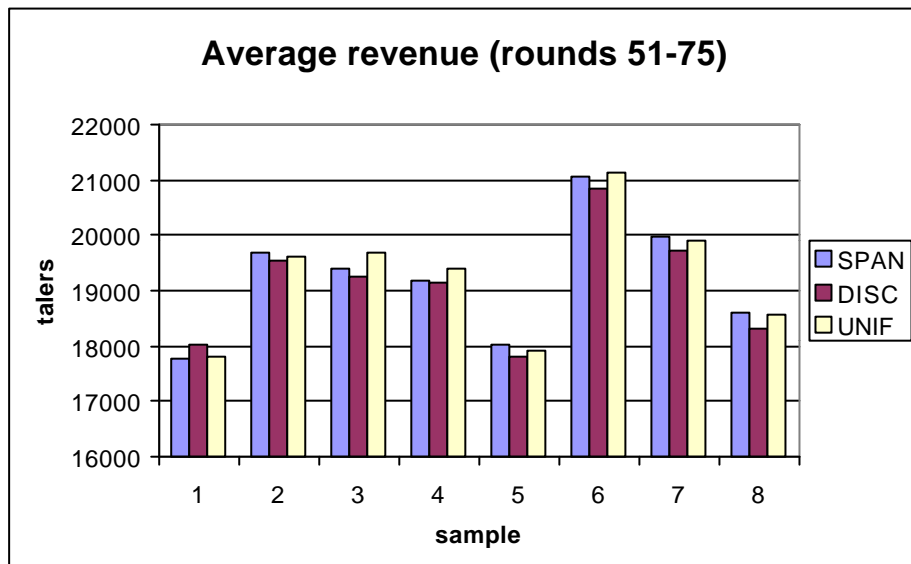


Figure 1

In a common value environment like the one we study, subjects often incur in what is known as *the winner's curse*, as documented in Kagel (1995).¹⁸ Notice that a buyer's signal is symmetrically distributed around the true value so that it is an unbiased estimator for the asset's true value. However, buyers may make losses if they incautiously base their bids on their estimates of the items. The bidder with the highest (over-)estimate then wins the auction and is likely to pay more than the asset's true value; on average they will incur losses. This phenomenon is especially pervasive with inexperienced bidders in the early rounds.

¹⁷ Notice that despite the standardization, some parallelism between the graphs prevails because the signals relative to the true value is the same across treatments, but varies over time.

¹⁸ This phenomenon was first empirically observed by Capen, Clapp and Campbell (1971) in the context of auctions for mineral drainage rights.

Perhaps surprisingly, Figure 2 shows that the aggregate standardized revenue is strongly negative for the first 20 rounds, indicating that the split of the total surplus is strongly in favor of buyers. This is due to subjects' very low bidding in the early rounds of several sessions.¹⁹ This low bidding may be due to a very cautious approach to the task, or possibly an attempt to establish collusive strategies in the early rounds.

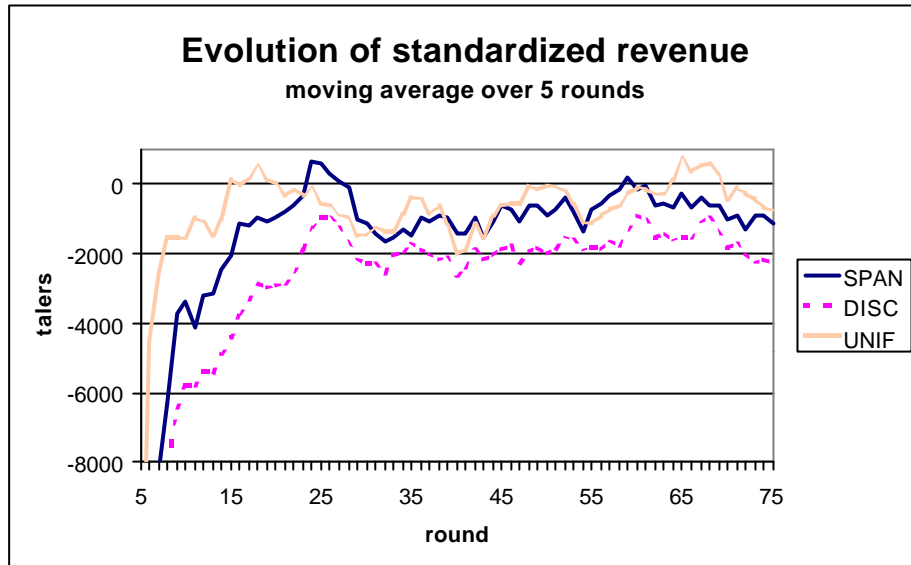


Figure 2

However, behavior in the early rounds is not representative of what happens later. As the experiment proceeds, competition starts having its effect and buyers' earnings decrease rapidly. After about round 30, we observe high and almost constant standardized revenues. We do not observe a statistically significant predominance of either positive or negative trends in the last 25 rounds. We compute, for each session separately, non-parametric Spearman rank correlation coefficients between the standardized revenue and auction rounds. Using these as summary statistics, the binomial test cannot reject the null hypothesis that positive and negative correlation coefficients are equally likely, so that behavior in the last 25 rounds can be seen as stable.²⁰

The adjustment in early rounds does not lead to a total erosion of buyers' profits. In all three treatments, the plots in Figure 2 mostly lie below the zero line. This indicates that on average, buyers earn positive profits so that overall they do not fall into the winner's curse. However, not all purchases of the last 25 rounds are profitable to buyers. In UNIF, an average of 2.23 bidders incurs a loss in a round, whereas this figure drops to 1.91 in SPAN and to 1.40 in

¹⁹ This is in contrast to other auction experiments with common values in which subjects usually incurred severe losses in the early rounds and often earned overall negative profits (see Kagel (1995)). A possible explanation of this discrepancy is that, with 7 units supplied and 16 units demanded, our treasury auctions are comparatively less competitive than those in which the winner's curse is usually observed.

²⁰ We observe positive coefficients in 2 of 8 sessions of SPAN and 3 of 8 sessions of DISC and UNIF. In the remaining sessions, the coefficients are negative.

DISC. The differences between DISC and both SPAN and UNIF are significant at $p = 0.035$ (one-sided) according to the binomial test. However, SPAN and UNIF are not statistically different from each other. Hence, the higher revenue generated by SPAN and UNIF is related to a greater propensity of buyers to overpay.

The plots in Figure 2 show that the revenue generated from inexperienced subjects is quite different from the one obtained from experienced subjects. This suggests that in rather complex environments, conclusions based on behavior observed in the early rounds of an experiment may be misleading. In addition, and contrary to what Figure 2 suggests, we find no significant difference in the revenues raised by the three auction institutions.

Hence, our data are in line with the numerical predictions of Alvarez and Mazón (1999) since we also find that the Spanish auction generates greater average revenues than the discriminatory format, and they are *partially* in line with the analysis of Alvarez, Cerdá and Mazón (1999), since we do not find the predicted revenue superiority of the Spanish format over the uniform format. However, these comparisons should be kept in perspective since the fundamentals of our experiments are very different from those of these numerical and theoretical studies.

4.2. Inter-Auction Dispersion of Prices

Besides average revenue, the choice of an auction procedure may also affect the *volatility* of prices *over time*. Malvey, Archibald and Flynn (1997) report that auction-to-auction volatility increased significantly after the U.S. Treasury used a uniform pricing rule for some types of securities. Nyborg, Rydqvist and Sundaresan (2002) and Gordy (1999) analyze bills and bond data from the Swedish and Portuguese treasury auctions, which are both organized as a discriminatory auction, and report that volatility increases with the duration of assets and various measures of market uncertainty, respectively. In our experiment, we measure volatility by the standard deviation of the standardized revenue over the last 25 rounds. Figure 3 reports the standard deviations for each market in each treatment and indicates that they are higher in SPAN and UNIF than in DISC. According to the binomial test, the difference is significant at $p = 0.004$ (one-sided) so that SPAN and UNIF produce indeed a higher inter-auction dispersion of prices. To some extent, this would be in line with the empirical finding of Malvey *et al.* (1997) of an increased volatility as the “treasury experiment” has been implemented in the US. By contrast, the difference between SPAN and UNIF is not significant.

4.3. Intra-Auction Dispersion of Bids and Prices

Our multi-unit design allows a straightforward test of the intra-auction dispersion of winning bids and prices. This issue is more relevant for auctions of multiple indivisible and homogenous goods, like spectrum licenses, which are also related to our set-up.²¹

²¹ One possible goal for the seller of multiple homogenous goods is to adopt an auction mechanism that achieves some “fairness” in the prices paid by winning bidders. Van Damme (2002) reports that it apparently was unde-

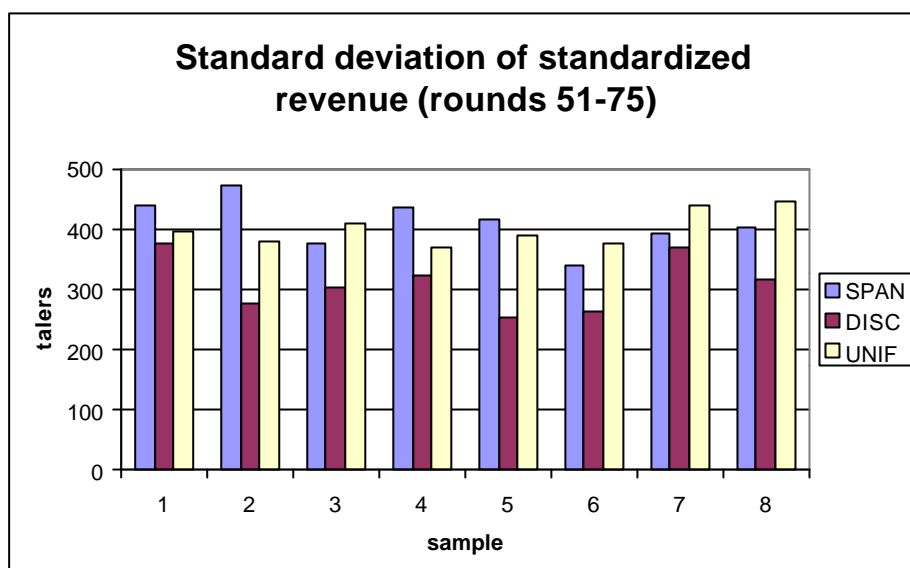


Figure 3

By definition, the uniform-price auction achieves the lowest possible intra-auction dispersion of prices (i.e., 0), whereas the other two formats will typically generate a positive one. The ranking of the discriminatory and the Spanish auction formats remains unclear since their respective pricing rules are likely to involve very different bidding strategies. We therefore compare the intra-auction dispersion of winning bids and of prices across treatments. We measure this dispersion of winning bids (prices) by the difference between the highest and lowest winning bids (prices) in a given round.

Figure 4 plots the average intra-auction dispersions of winning bids and prices for each sample in the three treatments and for the last 25 rounds. Although the high bars for sample 1 and 3 in UNIF are striking, they reflect well the buyers' incentives to bid in the uniform format. As bidders know that they will most likely not pay their bid, some of them decide to submit extremely high bids, far above their signals. Such bidders apparently wanted to ensure to win a unit, taking the risk of paying virtually any price and discounting the fact that other bidders may overbid as well. This, however, was observed for only two bidders (one in each sample). Such an overbidding has also been reported in (single unit) second-price auctions and has been attributed to the lack of negative feedback from winning an auction by overbidding. Typically, in these auctions the probability of making a loss upon winning an item by overbidding is very small when compared to the probability of making a positive profit upon winning by overbidding (Kagel (1995)). Clearly, this explanation assumes that bidders did not follow the dominant strategy logic but that they learn to bid from past experiences.

Despite the two exceptionally high bars in figure 4, we cannot even reject the null hypothesis of equal intra-auction dispersion of winning bids in UNIF than in SPAN, since dispersion was

sirable for the Danish regulator that winners of spectrum licenses would have to pay different prices for identical licenses.

higher in SPAN two out of eight times. On the other hand, Figure 4 reveals that bid dispersion is higher in SPAN and UNIF than in DISC seven out of eight times. According to the binomial test, this difference is significant at $p = 0.035$ (one-sided). In the aggregate, and for the last 25 rounds, the average dispersion of winning bids is 246.1 in SPAN, 184.7 in DISC, and 999.4 in UNIF (279.8 if we do not count the two extreme outliers).

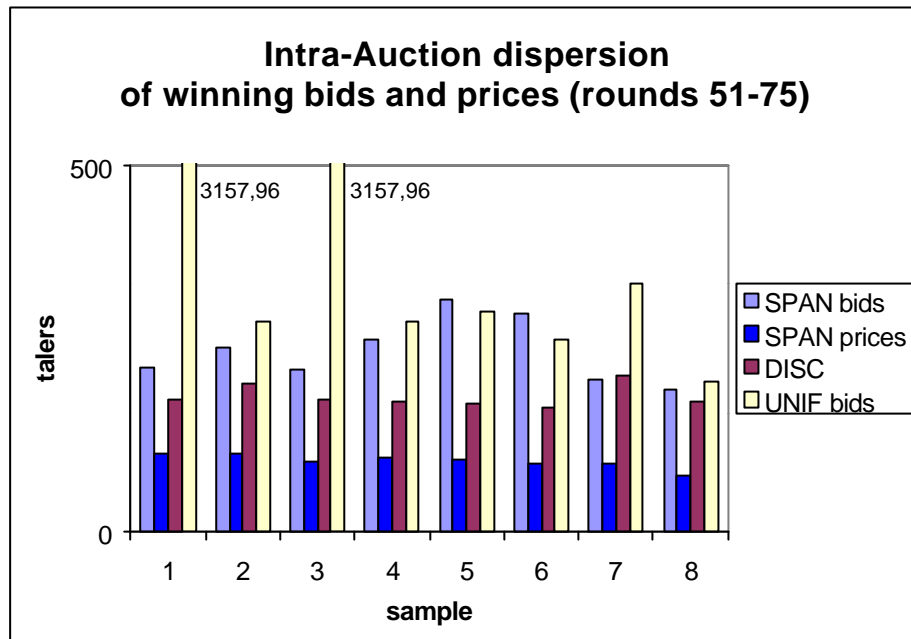


Figure 4

The higher dispersion of winning bids in SPAN is not strong enough to generate a greater dispersion of prices than in DISC. Actually, we observe a *smaller* intra-auction dispersion of prices in SPAN than in DISC (8 out of 8 times). The difference is significant at $p = 0.004$ (one-sided) according to the binomial test. In the aggregate, for the last 25 rounds, the average dispersion of prices is 96.0 in SPAN and 184.7 in DISC. To this extent, the above analyses of revenues and prices suggests that the Spanish auction format has the interesting properties of generating higher revenues with less intra-auction dispersion of prices than the discriminatory format, but it generates higher inter-auction volatility of prices.

Figure 5 illustrates the dispersion of all submitted bids, not restricted to the winning bids only. The figure shows the average level of the first, second, third, etc. highest bid in a market. To make them comparable, they are standardized to the signal they correspond to.²² For the higher bids, the line is steepest for UNIF and well above the ones for the other treatments. This reflects the greater dispersion of winning bids as found earlier. For the lower bids, however, the differences between treatments diminish.

²² The plot for the uniform treatment shows a slightly rising average level of bids as we move from the 15th to the 16th highest bid. This is due to the fact that not always all bidders submit two bids. In the rounds in which all 16 bids were submitted, the average of the 16th highest bid happened to be slightly above the average of all 15th highest bids, which contain also rounds in which only 15 bids are submitted.

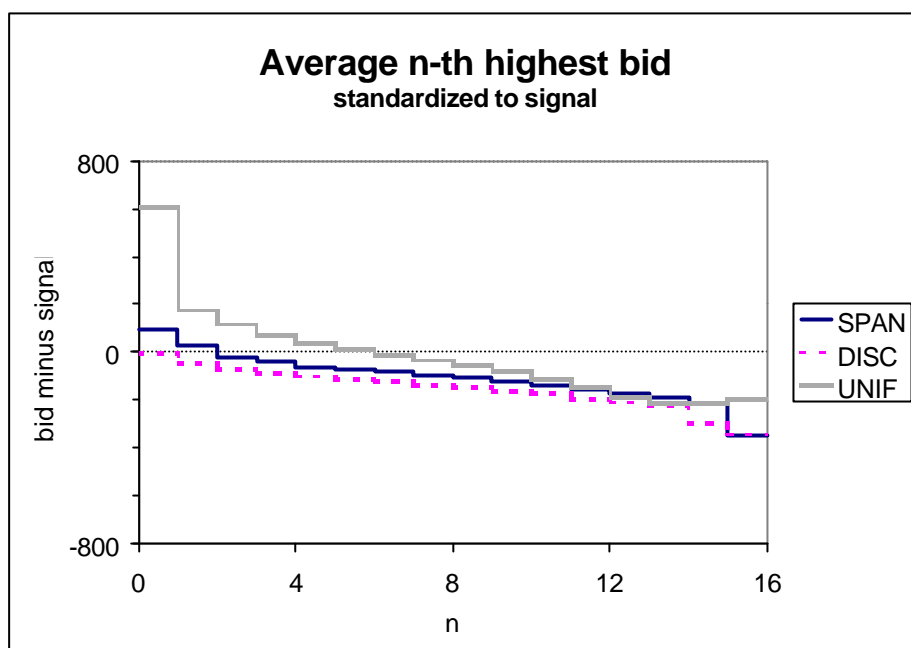


Figure 5

4.4. Individual Bids Related to Signals

We now compare the relation between bids and signals in the three formats. Recall that, in contrast to the discriminatory auction, not all bids are fully paid in the Spanish auction format and only marginal ones are in the uniform format. We may expect that this feature induces a ranking in the level of bidding in the different auctions. Figure 6 shows that this expectation is clearly confirmed. The figure shows the aggregate cumulative distribution of the differences between bids and the private signals.²³

In DISC, 95.4% of all bids are below the signal, compared to 89.0% in SPAN and only 60.3% in UNIF. The median bid is 146 talers below the signal in DISC, 101 talers in SPAN, and only 41 in UNIF. In all eight markets, we observe highest average bids in UNIF, lowest in DISC, with the average bids in SPAN in between. These differences are significant. The binomial test rejects the hypothesis of equally high bids for all pairwise comparisons.

Bidding in the Spanish auction format is lower than in the uniform auction. Recall that in the Spanish auction, all bids below the average winning bid are fully paid. Further, increasing the own bid increases the price also if the own bid is above the average winning bid, as it raises the average. Therefore, bidders need to be more cautious in the Spanish auction than in the uniform format, but they do not shade their bids as much as under the discriminatory scheme.

²³ This figure does not distinguish between high or low signals on the range of possible common values. We also conducted non-parametric (spline) regressions of observed bids on signals for the aggregate data. In all treatments, the estimated bid functions displayed a very strong linear relation between bids and signals that was almost parallel to the diagonal bid = signal.

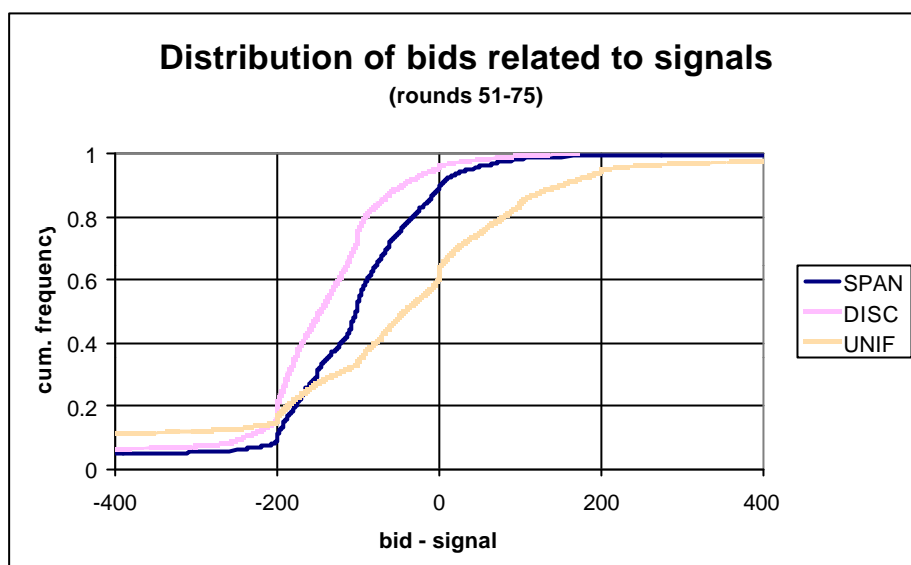


Figure 6

4.5. Bid-spread strategies

In this section, we investigate whether subjects used the opportunity to submit different bids for the two units, as suggested by Gordy (1996).²⁴ His conjecture is that bidders spread their bids more as the potential for a winner's curse increases (i.e., as the expected number of well-informed bidders increases or the volatility of market rates increases). The submission of downward sloping demand functions has been reported in empirical studies of the treasury auctions of Portugal (Gordy (1999)) and Sweden (Nyborg, Rydqvist and Sundaresan (2002)). In both cases, bid spreading increases with volatility of market rates.

The numerical simulations of Alvarez and Mazón (1999) indicate that in equilibrium, bidders increase the spread of their bids more in the Spanish than in the discriminatory auction. The intuition is that a bidder can increase his bid for the first unit at a lower expected cost in the Spanish format than in the discriminatory one. On the other hand, he also has an incentive to lower his bid for the second unit because if he wins two units, his bid for the second unit reduces the price he has to pay for the first unit (i.e., the average winning bid).

Our data provide evidence that subjects used bid-spreading strategies. On average, the absolute difference between the two bids is 76.5 talers in SPAN, 64.6 talers in DISC, and 156.6 in UNIF (see figure 7). Of course, this applies only to bidders who submit two bids in a given round. In the last 25 rounds, subjects submitted only one bid 7% of times and treatment differences could not be detected.

²⁴ Engelbrecht-Wiggans and Kahn (1998a and 1998b) study bid spreading in the context of independent private value auctions.

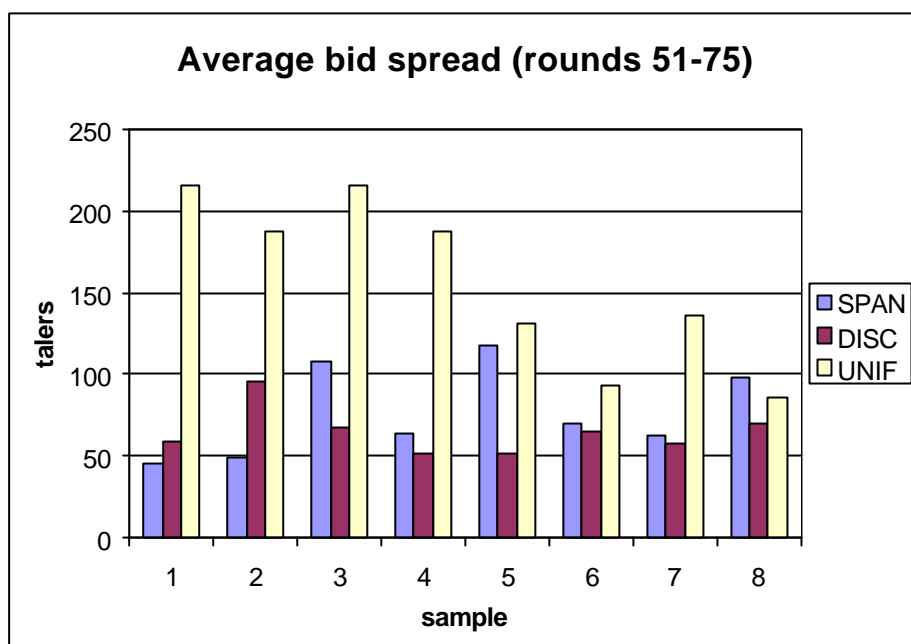


Figure 7

The highest bid spreads are not observed in the Spanish auction treatment, but in the uniform auction one. In UNIF bid spreads are significantly higher than in the other formats (one-tail $p = 0.032$, according to the binomial test, for each comparison). Average bid spreads are higher in SPAN than in DISC six out of eight times, but we cannot conclude that bid-spreading is significantly stronger in anyone of the two treatments (one-tail $p = 0.145$, according to the binomial test). Therefore, we cannot reject the null hypothesis of equal bid-spreading in the SPAN and the DISC treatments at a sufficient significance level.

5. Summary and Concluding Remarks

We report on an experiment that analyses buyers' bidding behavior and the revenue performance of three auction institutions that are being used for the sale of government securities. We designed our experiment so as to capture some of the main features of treasury auctions.

Our main result is that the uniform and the Spanish auction both raise significantly more revenue than the discriminatory auction. The fact that in the discriminatory auction bids are fully paid leads to lower bidding than in the other two cases. Bids are somewhat higher in the uniform than in the Spanish auction, but not enough to induce a significant revenue difference. We conjecture that this is due to the fact that, in the Spanish auction, precisely the higher winning bids are not fully paid.

With respect to the comparison of the uniform and the discriminatory auctions, the revenue ranking that we find is in line with the findings of Milgrom and Weber (1982) for single-unit demands and the one suggested by Ausubel and Cramton (1998) in the context of share auc-

tions. Our results only partly coincide with theoretical analysis of Alvarez, Cerdá and Mazón (1999), since the Spanish auction generates more revenue than the discriminatory but not more than the uniform auction. However, the two institutions that raise more revenue do not extract all buyers' surplus; on average buyers do not fall into the winner's curse in any of the three auction institutions we study. As indicated, this may be attributed to the fact that the ratio between the number of demand units and that of supply units is smaller than in other common value auction experiments, in an attempt to stay close to what is more common in treasury auctions.

Hence, the performance of the Spanish and the uniform auctions can be considered to be satisfactory, since they lead to comparatively high government revenue, but do not leave buyers empty-handed. In contrast, we observe that auction-to-auction volatility is higher in the Spanish and the uniform auctions than in the discriminatory auction and that price dispersion is significantly lower, about half, in the Spanish than in the discriminatory auction.

Finally, we observe that subjects use bid-spread strategies (or equivalently, they submit downward sloping demand functions) to a considerable extent. Moreover, in accordance with Gordy (1999) and Nyborg, Rydqvist and Sundaresan (2002) we find that bid-spreading is linked to volatility. We do not find significantly more bid-spreading in the Spanish auction format than in the discriminatory format, as reported by Alvarez and Mazón (1999) in the context of numerical simulations and attributed to the particular pricing rule. With respect to the interpretation of the use of downward sloping demand functions in uniform auctions in terms of self-enforcing collusive strategies (not possible under the discriminatory or Spanish formats), the significantly higher revenues observed in the uniform and the Spanish auctions suggest that such collusive bidding cannot explain the observed behavior.

Our results thus suggest that the auction format used by the Bank of Spain and the more standard uniform auction are interesting alternatives to the widely used discriminatory-price auction format, recently adopted by the ECB, since it could well reduce the governments' financing costs.

Of course, our results are not the last word on that matter. Although our model captures the essential features of a treasury auction, some features have not been included to keep the experiment simple. For example, we did not investigate the effects of buyers' asymmetric preferences on the revenue ranking of these auction mechanisms. Further, we abstracted from interactions between the auction for government securities and the secondary market where these securities are traded. These aspects should be dealt with in future work.

References

- Abbink K., Brandts J. and McDaniel T. (2003), "Asymmetric demand information in uniform and discriminatory call auctions: an experimental analysis motivated by electricity markets", *Journal of Regulatory Economics*, 23, 125-144.
- Abbink K., Irlenbusch B., Pezanis-Christou P., Rockenbach B., Sadrieh A., and Selten R. (forthcoming), "An Experimental Test of Design Alternatives for the British 3G/ UMTS auction", *European Economic Review*.
- Abbink K. and Sadrieh A. (1995), "RatImage – Research Assistance Toolbox for Computer-Aided Human Behavior Experiments", *SFB Discussion Paper B-325*, University of Bonn.
- Alvarez F., Cerdá E., and Mazón C. (1999), "Treasury bill auctions in Spain: A linear approach", *mimeo*, Universidad Complutense, Spain.
- Alvarez F. and Mazón C. (1999), "Multiple bids in a multiple-unit common value auction: simulations for the Spanish auction", *mimeo*, Universidad Complutense, Spain.
- Ausubel L.M. and Cramton P.C. (1998), "Auctioning Securities", *mimeo*, University of Maryland.
- Back K. and Zender J.P. (1993), "Auctions for Divisible Goods: On the Rationale for Treasury Experiment", *The Review of Financial Studies*, 6:4, 733-764.
- Back K. and Zender J.P. (2001), "Auctions of Divisible Goods with Endogeneous Supply", *Economics Letters*, 73, 29-34.
- Banco de España (2002), "Boletín Estadístico", Octubre 2000.
- Banks J., Olson M., Porter D., Rassenti S., Smith V.L. (forthcoming), "Theory, experiment and the Federal Communications Commission Spectrum Auctions", *Journal of Economic Behavior and Organization*.
- Bartolini L. and Cottarelli C. (1997), "Designing Effective Auctions for Treasury Securities", *Current Issues in Economics and Finance*, 3:9, 1-6, Federal Reserve Bank of New York.
- Bikhchandani S. and Huang C.-H. (1989), "Auctions with Resale Markets: An Exploratory Model of Treasury Bill Markets", *The Review of Financial Studies*, 2:3, 311-339.
- Capen, E.C., Clapp R.V., and Campbell W. M. (1971), "Competitive Bidding in High-Risk Situations", *Journal of Petroleum Technology*, 23, 641-653.
- Ehrhart K.M. (2000): "Repo Auctions: Experimental Investigation of Variable Rate Tender", *Mimeo*, University of Karlsruhe.
- Ehrhart K.M. (2001): "European Central Bank Operations: Experimental Investigation of the Fixed Rate Tender", *Journal of International Money and Finance*, 20, 871-893.
- Engelbrecht-Wiggans, R. and Kahn C.M. (1998a), "Multi-Unit Auctions with Uniform Prices", *Economic Theory*, 12:2, 227-258.
- Engelbrecht-Wiggans, R. and Kahn C.M. (1998b), "Multi-Unit Pay-Your-Bid Auctions with Variable Awards", *Games and Economic Behavior*, 23: 1, 25-42.
- Goeree J.K. and Offerman T. (2002), "Using First-Price Auctions to Sell Heterogeneous Licenses", *mimeo*, University of Amsterdam.
- Gordy M.B. (1996), "Multiple bids in multiple-unit common value auction", *mimeo*, Board of Governors of the Federal Reserve System.
- Gordy M.B. (1999), "Hedging winner's curse with multiple bids: Evidence from the Portuguese Treasury bill auction", *The Review of Economics and Statistics*, 81:3, 448-465.
- Goswami G., Noe T.H. and Rebello M.J. (1996), "Collusion in Uniform-Price auctions: Experimental evidence and implications for Treasury auctions", *The Review of Financial Studies*, 9:3, 757-785.
- Kagel J.H. (1995), "Auctions: A Survey of Experimental Research", *The Handbook of Experimental Economics*, eds. J. Kagel and A. Roth, Princeton University Press.
- Kagel J.H. and Levin D. (1986), "The Winner's Curse and Public Information in Common Value Auctions", *American Economic Review*, 76:5, 894-920
- Kagel J.H. and Levin D. (1999), "Common Value Auctions with Insider Information", *Econometrica*, 67:5, 1219-1238.
- Malvey P.F., Archibald C.M., and Flynn S.T. (1997), "Uniform-Price Auctions: Evaluation of the Treasury Experiment", Office Of Market Finance, U.S. Treasury, Washington, D.C.
- Manelli A., Sefton M. and Wilner B. (2000), "Multi-unit auctions: A Comparison of Static and Dynamic Mechanisms", *mimeo*, Arizona State University.
- Martínez Méndez P.M. (1996), "The Spanish market for government securities", *mimeo*, Bank of Spain.
- Mazón C. and Núñez S. (1998), "On the optimality of Treasury Bill auctions: The Spanish case", *mimeo*, Universidad Complutense and Bank of Spain.
- Milgrom P.R. and Weber R. (1982), "A Theory of Auctions and Competitive Bidding, Part 2", *mimeo*, Northwestern University.

- Nyborg K.G., Rydqvist K. and Sundaresan S. (2002), "Bidder Behavior in Multi-Unit Auctions: Evidence from Swedish Treasury Auctions", *Journal of Political Economy*, 110(2), 394-424.
- Nyborg K.G. and Sundaresan S. (1996), "Discriminatory versus Uniform Treasury Auctions: Evidence from When-Issued Transactions", *Journal of Financial Economics*, 42:1, 63-104.
- Plott C.R. (1997), "Laboratory Experimental Testbeds: Application to the PCS Auction", *Journal of Economics and Management Strategy*, 6:3, 605-638.
- Rassenti S., Smith V. and Wilson B. (2003), "Discriminatory price auctions in electricity markets: Low volatility at the expense of high price levels", *Journal of Regulatory Economics*, 23, 109-123.
- Robinson, M. S. (1985), "Collusion and the choice of Auction", *RAND Journal of Economics*, 16, 141-145.
- Simon D.P. (1994), "The Treasury's Experiment with Single-Price Auctions in the Mid-1970s: Winner's or Taxpayer's Curse?", *The Review of Economics and Statistics*, 76:4, 754-760
- Tenorio R. (1993), "Revenue Equivalence and Bidding Behavior in a Multi-Unit Auction Market: An Empirical Analysis", *The Review of Economics and Statistics*; 75:2, 302-314.
- Van Damme E. (2002), "The European UMTS-Auctions", *European Economic Review*, 46,4-5, 846-869.
- Wang J.J.D. and Zender J.P. (2002), "Auctioning divisible goods", *Economic Theory*, 19(4), 673-705.
- Wilson R. (1979), "Auctions of Shares", *Quarterly Journal of Economics*, 93(4), 675-689.

Appendix I: Instructions

(original text in Spanish)

General Information: You are about to participate in an experiment that consists of 75 rounds. In each of these rounds, you will participate in an auction in which 7 identical units of a fictitious good will be sold. In each auction, you will be in a group of 8 potential bidders: yourself and 7 other participants. Groups of bidders remain the same for the 75 rounds. In each auction, you can bid for a maximum of 2 units. All of you received the same instruction sheet.

Values and Estimates: At the outset of the each auction, the value V for one of these units will be randomly determined. V is the same for the seven units and is equally likely to be any integer number in the interval $[1000;5000]$. It is drawn with replacement that is to say, in each round, V does not depend on draws for the previous rounds.

Once V is determined, it will not be revealed to anyone of you.

At the outset of each experiment, each of you will receive private estimate of V which will be randomly drawn (with replacement) from the interval $[V-200;V+200]$. We will call this estimate E . E is equally likely to be any integer number in the interval $[V-200;V+200]$. Together with E , you will be given a maximum possible value and minimum possible value, given E . That is to say, $E-200$ and $E+200$.

Each of you will know her of his own E , but not the one of the other participants.

Bids: You are allowed to submit one bid, two bids or no bid at all. A bid is the amount of money that you are ready to pay for one unit if you win the auction. If you submit two bids, these bids can be equal or different. Bids can be any integer number from the interval $[500;6000]$.

Allocation and Prices: The seven highest bids will be called the *winning bids*. The seven highest bid will be called the *cut-off price*. If there are tied bids for the cut-off price, those which are assumed to be winning bids are randomly chosen among those tied bids. If you have one or two winning bids, you will be awarded one or two units, respectively. If you have no winning bids, you will not be awarded a unit. If in a given round there are less than 7 bids then the number of units sold will be equal to the number of bids submitted.

The computer will calculate the average winning bid which will be called the *average price*. The price that you will pay for the units that you have been awarded is determined as follows:

If your winning bid is above the average price, you will pay the average price.

If your winning bid is below the average price, you will pay your bid.

Gains and Losses: With each of your winning bids you can make either a gain or a loss. For each unit that you have been awarded, your gain or your loss is equal to V minus the price that you have to pay. Hence if the price you have to pay is below V , you will make a gain and if it is above V , you will make a loss.

Information Feedback: At the end of each round, you will be given information about the average price; V ; the total number of units that were sold; the number of units that you have been awarded for a price equal to the average price and the number of units that you have been awarded for a price between the cut-off price and the average price.

At anytime during the experiment, you will be able to open a "History window" which provides you with information about the outcomes of all the previous round in which you participated.

Rewards: The currency used in this experiment are talers so that all values, estimates, bids, prices, gains and losses are in talers. At the outset of the experiment, each of you will receive a capital balance of 4000 talers. Your total gain from participating in this experiment is equal to the sum of all your gains and your capital balance minus your losses. If ever your total gain is negative during the experiment, you will not be allowed to participate in the experiment anymore. If ever there are more then three participants with negative total gains, the experiment will end.

At any time during the experiment, you will be able to check your total gain in talers on the monitor. At the end of experiment, your total gain will be converted in Pesetas at the exchange rate of 0.5 Pesetas for each taler.

Appendix II: Revenue Generated in the Eight Markets

The following plots depict the standardized revenue over the 75 rounds of the experiment. To smoothen the graphs, moving averages over five rounds are shown. The three groups with the same samples of values and private signals are shown in one figure, where the Spanish auction group is represented by dark solid lines, the discriminatory-price auction group by dotted lines, and the uniform auction group by lighter solid lines.

