



Using Field Experiments to Test Equivalence between Auction Formats: Magic on the Internet

David Lucking-Reiley

The American Economic Review, Vol. 89, No. 5. (Dec., 1999), pp. 1063-1080.

Stable URL:

<http://links.jstor.org/sici?sici=0002-8282%28199912%2989%3A5%3C1063%3AUFETTE%3E2.0.CO%3B2-U>

The American Economic Review is currently published by American Economic Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aea.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet

By DAVID LUCKING-REILEY*

William Vickrey's predicted equivalences between first-price sealed-bid and Dutch auctions, and between second-price sealed-bid and English auctions, are tested using field experiments that auctioned off collectible trading cards over the Internet. The results indicate that the Dutch auction produces 30-percent higher revenues than the first-price auction format, a violation of the theoretical prediction and a reversal of previous laboratory results, and that the English and second-price formats produce roughly equivalent revenues. (JEL C93, D44)

Perhaps the most fundamental result in the auction literature is the revenue equivalence theorem, which dates back to William Vickrey (1961). Comparing four different basic auction formats (English, Dutch, first-price, and second-price auctions), Vickrey showed that in a simple model of Nash equilibrium bidding behavior, the expected revenue to be collected by the auctioneer will be the same under all four auction mechanisms. In this paper I report the results of Internet auctions of *Magic* game cards designed to test Vickrey's predictions.

In the *Magic: the Gathering* game scenario, players assume the roles of dueling wizards, each with their own libraries of magic spells (represented by decks of cards) that may potentially be used against the player's opponent. Cards are sold in random assortments, just like baseball cards, at retail stores ranging from small game and hobby shops to large chains such as Toys "R" Us and Waldenbooks. Launched in August 1993, this product has already grossed hundreds of millions

of retail dollars, and now has over a million players worldwide. There are more than a thousand distinct cards which have been printed for use in this game, each of which has a slightly different role in game play.

The game has resulted in a thriving secondary exchange economy, in which each Magic card is a separate commodity. The Internet, with its convenient methods of transmitting messages, greatly facilitates trading among individuals. A Magic player can auction off a few hundred unwanted cards more easily than one could hold a garage sale for excess household goods. The Internet lowers transaction costs, enabling ordinary individuals to make trades without the assistance of retail or auctioning specialists. Transaction costs are particularly low for cards as opposed to, for example, computer hardware, because it is so easy to mail cards across the country.

In 1994 and 1995, I observed auctioneers in this market employing a variety of auction mechanisms, including English, first-price sealed-bid, and Dutch auctions, along with variants and hybrids of these three auction types. In September 1994, I began to participate in this market by running my own auctions, using a variety of different rules. This provided a unique opportunity to perform tests of the revenue equivalence theorem in auctions for real goods.

The paper is organized as follows. Section I reviews the four basic auction mechanisms: English, Dutch, first-price sealed-bid, and second-price sealed-bid, and briefly describes the findings

* Department of Economics, 415 Calhoun Hall, Vanderbilt University, Nashville, TN 37235 (e-mail: reiley@vanderbilt.edu). Although I financed the actual experiments myself while a graduate student, I very gratefully acknowledge the financial support of the National Science Foundation, Grant No. SBR-9811273, for subsequent work on this paper. I wish to thank Marius Hauser and Mary Lucking-Reiley for their research assistance, and Skaff Elias for product information about *Magic: the Gathering*. I am grateful to an anonymous referee, Ann Bell, Jim Cox, Rachel Croson, Ron Harstad, Elton Hinshaw, John List, Jennifer Reinganum, Bill Sethares, Vernon Smith, David Wildasin, participants at the Economic Science Association annual meetings, and especially Glenn Ellison, for their advice and constructive criticism.

of previous empirical investigations of revenue equivalence, most of which have taken place in laboratory settings. Section II describes my experimental design, while Section III reports data analysis. A brief Section IV concludes.

I. The Four Basic Auction Mechanisms

The four basic auction mechanisms outlined by Vickrey (1961) are the English, Dutch, first-price sealed-bid, and second-price sealed-bid auctions. Probably the most familiar auction format is the English auction, an ascending-price auction in which the last remaining bidder receives the good and pays the amount of his high bid. The Dutch auction involves decreasing prices: a public price clock starts out at some very high level, and the price falls until the point when the first participant finds the price low enough to submit a bid. The first bidder is declared the winner of the Dutch auction, and receives the good at the price at which he stopped the clock. These two auction types can be grouped together as “real-time” auctions, to distinguish them from auctions in which bids are sealed and there is no real-time bidding process involved.

First-price sealed-bid auctions are also common. Each bidder has the opportunity to submit a single bid by a particular deadline. After the deadline expires, the bids are examined and the highest bidder wins the good at the bid price. In a second-price sealed-bid auction, the winning bidder does not pay the amount of her own bid, but the amount of the second-highest bid. Somewhat rarer in practice than the other auction formats, this type of auction was proposed by Vickrey (1961) because of its desirable theoretical properties, which will be discussed in the next section. Stamp auctions by mail have traditionally employed this format,¹ as did the New Zealand government’s auction of communications spectrum rights in 1990.²

A. Theoretical Equivalence Between Auction Formats

If auctions A and B are *strategically* equivalent, then not only do they produce the same amount of expected revenue for the auctioneer, they also have the property that an identical bidder would follow the same strategy in auction A as in auction B. Vickrey demonstrated that the second-price sealed-bid auction format is strategically equivalent to the English auction format, because in either auction format, the dominant strategy for a bidder is the same: to bid one’s valuation (that is, submit a sealed bid equal to one’s valuation, or stay in the oral auction until one’s valuation has been reached).³ Similarly, Vickrey demonstrated that the Dutch and first-price auctions are strategically equivalent, because the information available before bidding is the same in each case (no bidder gets to learn anything about other bidders’ willingness to pay before submitting one’s own bid). Bidders in first-price or Dutch auctions will bid strictly less than their valuations, in order to leave themselves some surplus when they win.

The two pairs of strategically equivalent auctions are not strategically equivalent to each other: we expect higher bids in first-price than in second-price sealed-bid auctions, for example. However, Vickrey (1961) showed that all four auction types are still revenue equivalent. With independent private values (IPV) and risk-neutral bidders, *all four basic auction types yield the same expected revenue to the auctioneer*.⁴

Under different modeling assumptions, auction theorists have shown that Vickrey’s revenue equivalence theorem does not always hold. First, if bidders are risk averse, the first-price auction should yield higher revenues than the second-price auction (Riley and William F. Samuelson, 1981; Eric S. Maskin and Riley,

¹ Economists have credited Vickrey with inventing the second-price auction as a brand-new auction format [see, for example, Michael H. Rothkopf et al. (1990)], but its use in practice actually predates Vickrey’s paper. In a separate paper, Lucking-Reiley (2000), I document stamp auctions using the Vickrey rule as early as 1893.

² See John McMillan (1994 p. 148) for details.

³ This assumes a private-values model of bidder valuations. If bidder valuations are privately uncertain, then the strategic equivalence may be violated.

⁴ A particularly readable exposition of this topic is that of John G. Riley (1989). Other useful references appear in the surveys by R. Preston McAfee and John McMillan (1987) and by Robert Wilson (1992).

1984).⁵ Second, if bidders are risk neutral but values are affiliated⁶ rather than independent, then the first-price auction should yield lower revenues than the second-price auction (Milgrom and Weber, 1982). If the bidders also have private uncertainty about their own values, so that their affiliated estimates of value are not perfect until after the auction is over, then the second-price auction itself yields lower revenues than the English auction. The only two auction formats which remain revenue equivalent under all the standard theoretical models are the Dutch and first-price auctions, which are always strategically equivalent for bidders.

Table 1 summarizes the predictions of the various theoretical models. In the field experiments reported in this paper, I can observe neither bidder risk preferences nor the distribution of bidder valuations (although, I will argue, bidder valuations are likely to be private rather than having a common-value component). Therefore, my results focus on the two predicted strategic equivalences, which do not depend on risk preferences.

B. Previous Empirical Studies of Revenue Equivalence

It is difficult to obtain field data that allow testing of the equivalences between the basic auction types. Sotheby's does not run both a second-price auction and an English auction for the same piece of antique furniture to see whether both formats yield the same amount of revenue, for example, nor does the Treasury Department do so with its bond auctions. I am aware of only two sets of field data which allow revenue comparisons between auction formats. First is a set of data on English and first-price auctions run by the U.S. Forest Service for timber harvesting rights. Walter J. Mead (1966) and Ronald Johnson (1979) found a tendency

TABLE 1—THEORETICAL PREDICTIONS OF AUCTION MODELS

Model	Revenue rankings ^a
IPV ^b and risk neutrality (Vickrey)	D = F = S = E
IPV and risk aversion	D = F > S = E
Affiliated, privately known values and risk neutrality	D = F < S = E
Affiliated, privately unknown ^c values and risk neutrality	D = F < S < E

^a Abbreviations: D = Dutch auction revenue; F = first-price sealed-bid auction revenue; S = second-price sealed-bid auction revenue; E = English auction revenue.

^b Independent, privately known values.

^c Privately unknown values are those where a bidder is uncertain about his own valuation until the auction is over, although he does have some noisy signal of his value. "Common-value" auctions are a special case.

for the first-price auctions to raise more revenue than the English auctions. However, Robert G. Hansen (1985, 1986) pointed out a selection bias caused by the way the Forest Service chose which auction to use for each timber lot; after correcting for this bias, he found that the increased revenue in English auctions was no longer statistically significant. The second set of data is from sealed-bid, multiunit currency auctions in Zambia. Rafael Tenorio (1993) finds that in these auctions, a discriminatory (pay-your-bid) pricing rule raises significantly higher revenues than a uniform (lowest-accepted-bid) pricing rule, mainly due to higher bidder participation in the latter auction format. In general, despite the fact that revenue equivalence is one of the most fundamental issues in auction theory, there is very little evidence from field data on the four basic auction formats.⁷

By contrast, there have been a number of tests of revenue equivalence performed in laboratory experiments. [John H. Kagel (1995) provides a comprehensive review.] The Dutch auction has yielded lower revenues than the first-price auction (Vicki M. Copping et al., 1980; James C. Cox et al., 1982, 1983). Revenue in first-price sealed-bid auctions was significantly higher than the risk-neutral Nash

⁵ Strategic equivalence between first-price and Dutch auctions continues to hold. If values are privately known (i.e., bidders do not have any private uncertainty about the values they will have after the auction), the English and second-price auctions remain equivalent as well.

⁶ Paul R. Milgrom and Robert J. Weber (1982) introduce this technical term. Roughly, it means that if my information indicates my value is likely to be high, then it indicates that your value is also likely to be high.

⁷ For a review of the auction topics which have been studied empirically with field data, see the article by Kenneth Hendricks and Harry J. Paarsch (1995).

TABLE 2—PREVIOUS EXPERIMENTAL RESULTS ON STRATEGIC EQUIVALENCE^a

Experiment	Results
Coppinger et al. (1980)	<ul style="list-style-type: none"> • $F > D$ • English bids follow theoretical prediction
Cox et al. (1982, 1983)	<ul style="list-style-type: none"> • $F > D$
Kagel et al. (1987)	<ul style="list-style-type: none"> • $S > E$ • English bids follow theoretical prediction
Kagel and Levin (1993)	<ul style="list-style-type: none"> • Second-price bids exceed theoretical prediction

^a This table employs the same abbreviations as Table 1.

equilibrium (RNNE) prediction, while revenue in Dutch auctions was approximately equal to or slightly below the RNNE level. Similarly, the English auction has yielded lower revenues than the second-price auction, as bidders tend to bid their valuations in the English format (Coppinger et al., 1980; Kagel et al., 1987), but higher than their valuations in the second-price format (Kagel et al., 1987; Kagel and Dan Levin, 1993). Thus, in both cases of theoretical strategic equivalence, laboratory results show the dynamic auction raising less revenue than the equivalent sealed-bid auction.⁸ Table 2 summarizes the results of past laboratory experiments.

II. Experimental Procedure

For this study, I purchased over \$2,000 of Magic cards and resold them via auctions in the Internet marketplace. The basic procedure was to auction two copies of the same card via two different auction mechanisms, to make direct comparisons of the revenue earned in each one. Complications arise because of the real-world

⁸ The revenue equivalence between the two pairs of formats also fails in the laboratory. First-price auctions consistently earn higher revenues than do second-price auctions (Coppinger et al., 1980; Cox et al., 1982; Kagel and Levin, 1993). This finding is potentially consistent with bidder risk aversion. The result tends to disappear when the number of bidders becomes large, or when bidders' private values are affiliated rather than independent (Kagel and Levin, 1993).

nature of the auction market, and therefore the description of my experimental design is more lengthy than that of the typical laboratory experiment. The first subsection describes details of the Internet marketplace for Magic cards. The second subsection describes several ways in which my experiments differ from laboratory experiments in the literature. The final two subsections give details about the two different types of paired experiments: Dutch-first, and English-second, respectively.

A. Institutional Details of the Market

Soon after the introduction of *Magic: the Gathering*, Internet users formed a newsgroup, `<rec.games.deckmaster>`, devoted to the discussion of this new game. In addition to discussions about the rules of the game and strategies for constructing decks, many of the messages on this newsgroup were buy, sell, and trade offers for individual cards. Parties agreed to the terms of trade via electronic mail, and then carried out the transactions through postal mail. Messages devoted to economic trades soon overwhelmed the discussion group, so a new group, `<rec.games.deckmaster.marketplace>`,⁹ was devoted exclusively to the trading of cards.

Both the quantity and the variety of messages posted to this newsgroup were stunning. By the spring of 1995, nearly 6,000 messages were being posted each week, making `<rec.games.trading-cards.marketplace>` the highest-volume newsgroup on the Internet.¹⁰ Approximately 90 percent of the 26,000 messages per month were devoted to the trading of Magic cards, with the remaining 10 percent devoted to the trading of cards from other games. Interestingly, the messages utilized several different kinds of market mechanisms. Some people posted cards they were willing to trade, along with "wish lists" of cards they would be willing to accept in return, and solicited responses by private electronic mail. Others posted fixed prices at which they were willing to sell cards for cash. Many hope-

⁹ This newsgroup was later renamed `<rec.games.trading-cards.marketplace>`.

¹⁰ According to the April 1, 1995 edition of the monthly publication *Top 40 Newsgroups in Order by Traffic Volume*, this newsgroup edged out the second-place newsgroup, `misc.jobs.offered`, by 3,000 messages per month.

ful sellers conducted auctions of their unwanted cards, using a variety of auction mechanisms, including English, first-price sealed-bid, and even Dutch auctions, along with other variants and hybrids of these three auction types. The only basic mechanism from auction theory that I did not witness was the second-price sealed-bid auction. Thus, when I began to run my own auctions with a variety of different rules, I could expect bidders to treat me no differently than they would treat any other auctioneer. This provided a unique opportunity to perform tests of the revenue equivalence theorem in auctions for real goods.

Laboratory experiments, which to date have provided the vast majority of data on bidders' behavior in auctions, can be criticized on the grounds that subjects' behavior in an artificial laboratory environment may not be exactly the same as their behavior would be in the "real world." The Magic card market provides an opportunity to run controlled experimental auctions in the field, to check the robustness of laboratory results. Since in any given week there are dozens of auctioneers holding Magic auctions on the Internet, an experimenter can be a "small player" who does not significantly perturb the overall market.¹¹

One of the strengths of this study is its subject pool.¹² The bidders in these experiments have wide diversity in their geographic and professional backgrounds, but share an intense interest in Magic cards. Their keen interest in the auctioned goods should make them representative of people who bid in auctions in general, ideal for a test of auction theory.¹³

¹¹ Since it is unnecessary to pay subjects for their participation, this technique of running "field experiments" has a potential financial advantage over traditional laboratory techniques. Laboratory studies frequently cost over \$1,000 in subject payments. By contrast, for my set of experiments on Dutch and first-price auctions, I purchased cards for approximately \$1,600 and resold them for \$2,000, realizing a 25-percent return.

¹² For more detailed demographic information about the bidders, see my doctoral dissertation [Reiley (1996), available on request]. The bidders, from all over the United States and several foreign countries, included college students, professional engineers, government employees, and military personnel.

¹³ Some bidders clearly indicated that they were thinking carefully about the relatively unfamiliar auction formats. In a second-price auction, one bidder submitted a bid of \$15 on

B. Differences from Previous Auction Experiments

As noted above, my experiments differ from previous laboratory experiments on revenue equivalence by auctioning real goods in a real market. This leads to two other significant experimental design differences from previous experiments. First, auctions for different individual goods were simultaneous, rather than sequential. Second, the auctions had an uncertain and endogenously determined number of bidders, rather than a fixed number of bidders. In this subsection, I discuss these differences in more detail.

Previous laboratory experiments on revenue equivalence in auctions have involved sequential auctions: subjects participate in repeated "rounds" of auctions for single objects. By contrast, I conducted simultaneous auctions of 85 to 100 different Magic cards at a time. I chose to do this for two primary reasons. First, simultaneous auctions were the norm in this marketplace; practically all Magic auctions on the Internet were simultaneous auctions by a single auctioneer for dozens or even hundreds of different cards. Second, simultaneity enhanced data collection, as it enabled me to collect data on over 80 auctions at a time. A single English auction might take as much as a month to complete, which would make data collection via sequential auctions infeasible. My lack of repeated trials with the same subjects precludes my investigating learning effects, something frequently studied in the laboratory. However, because my experiments take place in a preexisting market with already experienced bidders, and because my bidders have plenty of time to make their decisions (days, as compared with minutes or seconds in laboratory experiments), I assume that learning effects are mostly irrelevant.

These experiments also differ from laboratory tests of the revenue equivalence theorem in their treatment of bidder participation. In lab experiments, as in Vickrey's original theoretical

a Shivan Dragon card, then added the following note: "I shudder to think what would happen if I put \$100 as an attempt to get the second highest bid and they put \$99 on the same token (shudder, shudder, shudder)."

model, the number of bidders is fixed and known to all. By contrast, my auctions involved an uncertain number of bidders each making their own participation decisions. I advertised each auction to a large number of people, but I did not know how many would enter. I decided on this design in order to make my auctions resemble the other real auctions in the market.

I used two different methods to invite bidders: posting advertisements on the newsgroup, and sending individual invitations via e-mail to people interested in Magic auctions. This is standard practice for auctioneers in this marketplace, as direct e-mail invitations produce high response rates, but advertisements allow one to reach a larger potential audience. I kept the number of invited bidders secret. As was standard practice by auctioneers in the marketplace, I posted to the newsgroup daily status updates for the English¹⁴ auctions, every day until the auction ended (typically about three weeks). By contrast, the sealed-bid auctions lasted only one week, and were each advertised only three times on the newsgroup.¹⁵ One might expect the number of bidders attracted by advertisements to be higher in the auctions with daily updates than in the sealed-bid auctions, and this might cause an upward bias in revenue for the English and Dutch auctions.

In an effort to control for this potential bias, I made sure that anyone who entered one auction via a newsgroup advertisement received an e-mail invitation to the next auction in the pair. For example, for my first English auction I sent out 90 direct e-mail invitations in addition to the newsgroup postings. It attracted a total of 40 bidders: 18 from the e-mail and 22 from the newsgroup. After this auction ended, I ran a paired second-price auction for which I sent direct e-mail invitations to the 22 "new" bidders

as well as to the original 90-person mailing list.¹⁶ In my analysis of the revenue results below, I will also consider the effects of excluding those observations which may have been affected by such advertising effects.

C. Dutch and First-Price Auctions

Table 3 gives an overview of the Dutch and first-price auction experiments, which consisted of two pairs of auctions. Paired experiment FD consisted of auction FD1, a first-price auction, followed by auction FD2, a Dutch auction, with matched pairs of cards for sale. Auction FD1 lasted one week, with 88 different cards for sale, while the matching Dutch auction FD2 lasted 12 days, until the last card was sold. Experiment DF, with auctions DF1 and DF2, reversed the time order of the auction formats, with a separate set of matched cards. Reversing the time order allowed me to control for a number of potential time-related biases.¹⁷

The announcement and rules for the first-price auction FD1 stated that bidders had one week to submit their bids.¹⁸ I placed constraints on the acceptable bid amounts,¹⁹ designed to make the set of possible bids equal to the set of prices possible in the Dutch decrement auction. At the end of the week, I compiled the results of

¹⁶ The second mailing list included only 85 rather than 112 people, because some people from the original list of 90 asked to be deleted from my mailing list.

¹⁷ One such bias is the asymmetric nature of my e-mail invitations. A second is demand saturation over time, as in Orley Ashenfelter's (1989) "revenue decline anomaly," with falling prices in sequential auctions for wine and art. These experimental auctions differ from the wine and art auctions in that my bidders did not necessarily know that another auction for the same cards would be coming. In fact, I actively tried to discourage such speculation in two ways: by including singleton cards along with the paired cards so that two paired auctions were not obviously identical, and by occasionally running other auctions consisting entirely of singleton cards (including some auctions designed for other experiments not included in this paper).

¹⁸ The complete text of the instructions for the four auction types is available on request from the author at <http://www.vanderbilt.edu/econ/reiley>.

¹⁹ For example, bids less than \$1 had to be in even multiples of a nickel (\$0.05), bids between \$1 and \$5 had to be in even multiples of a dime (\$0.10), bids between \$5 and \$10 had to be in even multiples of a quarter (\$0.25), bids between \$10 and \$20 had to be in even half-dollar amounts (\$0.50), and bids over \$20 had to be in whole-dollar amounts.

¹⁴ Although little precedent existed for Dutch auctions on the newsgroup, I also followed an identical strategy of daily updates in order to make the Dutch auctions resemble the standard English auctions as much as possible.

¹⁵ Three was the maximum number of times I felt I could post an identical message in a week without suffering a loss of goodwill from other market participants. Repeated posts of English auctions each provided new information on the current bid price, but repeated posts of sealed-bid auctions were completely identical, and "spamming" the newsgroup with many identical messages would violate norms in this market.

TABLE 3—OVERVIEW OF DUTCH-FIRST EXPERIMENTS

	Auction FD1	Auction FD2	Auction DF1	Auction DF2
Auction format	First-price	Dutch	Dutch	First-price
Card type	Black/Blue	Black/Blue	Red/Green	Red/Green
Start date	Fri., May 5	Wed., May 17	Wed., Sept. 27	Tues., Oct. 31
End date	Fri., May 12	Mon., May 29	Tues., Oct. 24	Tues., Nov. 7
Total number of cards auctioned	88	87 ^a	88	88 ^b
Total list value	550.42	533.67	308.24	308.24
Number of matched cards	87	87	86	86
Maximum selling price	27.00	26.00	22.00	25.00
Minimum selling price	0.10	0.15	0.20	0.10
List value of matched cards	533.67	533.67	303.25	303.25
Total revenue on matched cards	431.25	446.35	348.45	327.05
Number of participating bidders	32 ^c	63	88	42
from newsgroup announcements	3	7	7	3
from e-mail invitations	29	56	81	39
Number of e-mail invitations sent	403	379	586	472
Number of winners	20	22	22	22
Maximum number of cards to a winner	18	15	19	16
Maximum payment by a winner	92.75	80.50	57.25	92.75

^a I was unable to locate a second "Mana Drain" card in time for the second auction.

^b Two of these cards were only "near mint" instead of the "mint" condition of their counterparts in Auction DF1, and thus only 86 of the cards are strictly comparable as "matched" cards.

^c This equals to the number of people submitting bids in the first-price auctions, and to the number of people who asked to receive daily updates for the Dutch auctions.

the first-price auction and e-mailed the final prices to each of the participating bidders. Winners also received a congratulatory message listing the cards they had won, along with their total bill.²⁰ After receiving payment, I mailed out winners' cards.

I determined the starting prices for the Dutch auction FD2 by taking the winning bid amounts from the first-price auction FD1 and increasing them by five days' worth of Dutch clock decrements to get the starting prices for auction FD2, the Dutch auction. Thus, if each card sold after exactly five days, the auction revenues would have been identical. Prices fell by approximately 5 percent per day, with the exact decrement amounts corresponding to the acceptable bids in the first-price auction. In the instructions, bidders did not learn the exact decrement amounts, only that prices would fall by a "small amount" each day. The initial announcement of the Dutch auc-

tion stated the rules of the auction and the starting prices of each card. I mailed updates once per day to a list of interested bidders, indicating new prices on the unsold cards. To my surprise, 25 of the 87 cards in auction FD2 sold on the very first day of the Dutch auction, at prices which were all at least 15 percent higher than their corresponding prices in the first-price auction. Therefore, the revenues in auction FD2 represent a lower bound on the possible Dutch auction revenues.

The second pair of auctions described in Table 3 are labeled auction DF1 and auction DF2, which were, respectively, a Dutch auction followed by a first-price auction for the same cards. With no first-price auction benchmark to start with, I determined the starting prices by finding a list value²¹ for each card and marking it up by five days' worth of Dutch clock

²⁰ Although auctioneers on the Internet typically ask winning bidders to pay shipping costs, I included free shipping in the bid price, in order to encourage bidders to bid independently on each card.

²¹ For my "list" values, I used the Cloister price list (Jason Black, 1995), a standard reference list compiled using the thousands of trades in the Internet marketplace for Magic cards. Though there is lots of dispersion in the prices of trades for a given card, the Cloister value is a mean trading price, which remains rather constant over timescales of weeks.

TABLE 4—OVERVIEW OF ENGLISH-SECOND EXPERIMENTS

	Auction ES1	Auction ES2	Auction SE1	Auction SE2
Auction format	English	Second-price	Second-price	English
Card type	Legends	Legends	White/Gold	White/Gold
Start date	Thurs., Feb. 9	Thurs., Mar. 9	Sat., May 6	Sat., May 20
End date	Sat., Mar. 4	Thurs., Mar. 16	Sat., May 13	Wed., June 7
Number of cards auctioned	85	85	99	99
Total list value	214.64	210.14	828.63	828.63
Number of matched cards	66 ^a	66	98	98
Maximum selling price	12.00	10.00	23.00	21.00
Minimum selling price	0.10	0.25	0.05	0.05
List value of matched cards	107.94	107.94	804.33	804.33
Total revenue on matched cards	79.50	85.50	517.05	600.40
Number of participating bidders	40	27	43	38
from newsgroup announcements	22	13	3	12
from e-mail invitations	18	14	43	26
Number of e-mail invitations sent	90	85	385	372
Number of winners	26	15	27	17
Maximum number of cards to a winner	11	39	15	15
Maximum payment by a winner	26.50	69.20	154.00	136.50

^a Of the 85 cards, 13 were unmatched card types and 6 were unmatched in condition ("mint," "near mint," etc.).

decrements. None of the cards sold on the first day, but several sold on the second day. All but one card sold within two weeks; the final card remained unsold for an additional 11 days. One week after DF1 finished, I started DF2 for comparison, a first-price auction with the same set of 88 cards. As Table 3 shows, the number of participants increased somewhat relative to the first experiment, due to an increase in the size of my invitation mailing list.

D. English and Second-Price Auctions

Table 4 displays the four auctions designed to test the strategic equivalence between English and second-price auctions. Auctions ES1 and ES2 were an English auction followed by a second-price auction for one set of cards, while auctions SE1 and SE2 were a second-price auction followed by an English auction for another set of cards. The second-price auctions ES2 and SE1 used the same basic set of sealed-bid rules as the first-price auctions discussed above, each lasting one week. For the English auctions ES1 and SE1, I issued daily updates, displaying the current high bid on each card. Bidders received these updates via e-mail only if they had submitted a bid or asked to be added to the update

mailing list, but the updates were also posted daily to the newsgroup. I employed a "Going, Going, Gone!" technique for declaring a card sold: a card which had lasted a full day without a bid raise received an exclamation mark as a warning on the next update, a card which had lasted two full days received two exclamation marks as a second warning, and a card which had lasted three full days was declared "SOLD!" on the next update.

I again placed constraints on acceptable bid amounts. In auctions ES1 and ES2, the first two auctions I ran for this paper, all bids were constrained to be multiples of a nickel. After feedback from bidders, I used larger minimum bid increments for the higher-priced cards in auctions SE1 and SE2 (the same set of bid increments as in the first-price and Dutch auctions discussed above).

Table 4 presents summary statistics for the English and second-price auctions, analogous to those in Table 3 for the Dutch and first-price auctions. Experiment ES had a relatively large number of unmatched cards: only 66 of its 85 cards were perfect matches. However, the unmatched cards were chosen to be comparable in value across auctions, to keep the auctions as similar as possible overall. All subsequent data

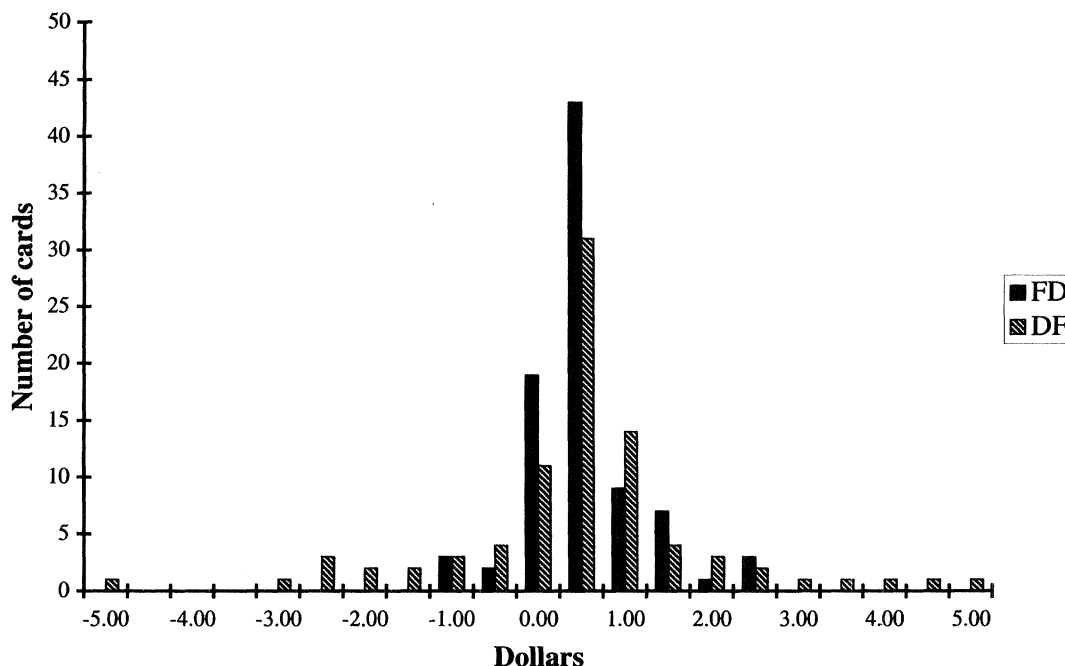


FIGURE 1. DUTCH REVENUE MINUS FIRST-PRICE REVENUE

analysis in this paper excludes observations that were not perfect matches. The number of e-mail invitations was much higher in the second pair (auctions SE1 and SE2) than in the first (auctions ES1 and ES2), but the total number of participating bidders was roughly comparable between the two pairs, with newsgroup advertisements making up the difference. In both experiments, newsgroup posts yielded more bidders in the English than in the second-price auctions, perhaps due to the greater frequency of posts in the English case. Finally, the time-order effect on revenues is the opposite of what one might have expected: the revenue on matched cards is greatest for the second auction in each pair.

III. Results

A. Dutch and First-Price Auctions

Figure 1 displays the distributions of revenue outcomes for the Dutch and first-price auctions. Specifically, it plots the difference between Dutch revenue and first-price revenue for each of the 87 matched pairs of cards in experiment

FD and the 96 matched pairs of cards in experiment DF. There is a clear tendency for the difference to be positive: Dutch auctions appear to raise more revenue than first-price auctions. This is a surprising result, conflicting with results from laboratory experiments, where Dutch auctions systematically raise *less* revenue than first-price auctions.

The top section of Table 5 displays statistical results for all matched pairs of cards. The revenue was higher in the Dutch auction for 63 of 87 matched pairs in experiment FD, and 59 of 86 pairs in experiment DF. The average difference between the revenues in the Dutch and first-price auctions, plotted in Figure 1, has a mean of 38 cents in experiment FD, and 25 cents in experiment DF. A two-tailed *t*-test at a 5-percent level of significance rejects the hypothesis of a zero mean difference for experiment FD, but not for experiment DF. The table also reports the mean log difference in revenue; this figure indicates that the Dutch auctions raised 30 percent more revenue than the first-price auctions on average. A two-tailed *t*-test on this variable results in a rejection of the revenue equivalence hypothesis in each sample, as does

TABLE 5—THE DIFFERENCE BETWEEN DUTCH AND FIRST-PRICE REVENUES

	Experiment FD	Experiment DF
All observations	87	86
Dutch revenue higher	63	59
Equal revenue	12	5
First-price revenue higher	12	22
Mean Dutch-first difference	\$ 0.38	\$ 0.25
(Standard deviation of mean)	(0.07)	(0.17)
Mean log difference	0.293	0.288
(Standard deviation of mean)	(0.038)	(0.059)
Wilcoxon statistic	5.66	2.78
Restricted sample ^a	82	86
Dutch revenue higher	62	59
Equal revenue	8	5
First-price revenue higher	12	22
Mean Dutch-first difference	\$ 0.40	\$ 0.25
(Standard deviation of mean)	(0.07)	(0.17)
Mean log difference	0.303	0.288
(Standard deviation of mean)	(0.039)	(0.059)
Wilcoxon statistic	6.15	2.78

^a The restricted sample excludes any observation where the winning bidder was attracted to the auction via newsgroup advertisement rather than via e-mail invitation.

a nonparametric Wilcoxon signed-rank test of revenue equivalence.²²

The bottom section of Table 5 contains results for a restricted sample of observations, in an attempt to correct for the advertising effects discussed in Section II, subsection B. The restricted sample excludes observations where the winner had been attracted to the auction via a newsgroup advertisement rather than an e-mail invitation. This turned out to exclude only five observations, all from the FD experiment. As can be seen in the table, the effect of this restriction on the sample was only to increase the statistical significance of the Dutch auction format's superior revenues.

In order to present pooled results from both auction pairs, I ran a regression of the Dutch-first revenue difference against a dummy variable for

the experiment (FD versus DF). The results are as follows (standard errors in parentheses):

$$(1) \quad REVDIFF = 0.25 + 0.13FD \\ (0.13) (0.18).$$

The estimated intercept indicates that Dutch auctions earn statistically significantly more revenues than first-price auctions, while the slope estimate indicates that the treatment order does not have a statistically significant effect on the revenue difference. I have presented results for the full sample ($N = 173$, $R^2 = 0.003$), although the restricted sample yields qualitatively similar results.

What are the possible explanations for this violation of revenue equivalence between Dutch and first-price auctions, opposite in direction to the violation observed in laboratory experiments? One possible explanation is endogenous bidder entry: my bidders decided whether or not to participate after reading the auction announcement, while the laboratory auctions had a fixed number of bidders. As can be seen in Table 3, the number of participating bidders was higher in the Dutch auctions than in their corresponding first-price auctions (63 versus 32 bidders, and 88 versus 42 bidders, respectively). These increased numbers of bidders may have been the reason for the increased revenues in the Dutch auctions, assuming that these additional bidders might occasionally bid fairly high.²³ Note that the higher frequency of advertisements in the Dutch auction does not fully explain the increased participation, as Table 1 shows that the majority of additional Dutch auction bidders come from the e-mail invitation. Perhaps the additional bidders were attracted merely by the novelty of a Dutch auction, an effect that would disappear with repeated trials, but as the vast majority of Magic

²² A tie (meaning revenue differences equal to zero for a given card) may be counted either as positive or negative differences in the computation of the Wilcoxon signed-rank statistic. This makes little difference in my data, since there are few ties. To be conservative, in computing my statistics I counted all ties as if they were negative differences, thus biasing the test against finding that Dutch revenues were higher than first-price revenues.

²³ On the other hand, it is worth noting that the measure of "participation" differs somewhat between the two formats. To be counted as a "participant" in the Dutch auction, one did not necessarily have to submit any bids, but merely had to ask to receive updates of the current Dutch prices. Thus, we might expect higher measured participation in the Dutch auction, even if there are the same number of serious bidders (which would provide the same revenues) in the two different auctions.

auctions on the Internet during this period were English auctions, it is hard to believe that Dutch auctions would seem significantly more novel than sealed-bid auctions.²⁴ In any case, if the difference in revenues were due merely to the higher bidder participation attracted by the Dutch auction, this would require no violations of the strategic equivalence of Dutch and first-price auctions by individual bidders.

To pursue the question of individual bidder behavior, I now examine disaggregated, bid-level data. There are two different ways a bidding anomaly (that is, a violation of strategic equivalence) can arise. First, someone might bid an amount X in the first-price auction, but pass up the opportunity to bid at a price less than or equal to X in the corresponding Dutch auction. Such an anomaly violates strategic equivalence by producing higher revenues in the first-price auction. Second, someone might bid in the Dutch auction at a price greater than the amount X he bid in the corresponding first-price auction, which would be an anomaly generating higher revenues in the Dutch auction. Even if such anomalies exist, they are difficult to observe, because of incomplete data generation in the Dutch auction: the Dutch auction generates bids only from the winning bidder, while the first-price auction generates bids from everyone. Despite the relative dearth of Dutch auction bid data, a few statistics could be generated, as follows.

The two Dutch auctions generated a total of 257 observed bids, 38 of which presented the opportunity to look for anomalies with higher Dutch than first-price bids. These 38 (from a total of 10 different bidders) each had a matching first-price bid observation from the same bidder. Of these 38 observations, 30 were higher in the Dutch auction, 4 were equal across auctions, and 4 were higher in the first-price auction. The 30 bid differences favoring the Dutch auction ranged in size from \$0.25 to \$16.00, with a mean of \$2.52, while the 4 differences favoring the first-price auction ranged only from \$0.10 to \$0.50. That the vast majority

of differences favored the Dutch auction indicates that bidders might systematically bid higher in Dutch auctions.²⁵

To look for bid anomalies favoring the first-price auction, I examine the bidders who participated in both auctions in a pair (that is, who submitted bids in a first-price auction and also asked to receive updates in the matching Dutch auction). Twenty such bidders submitted a total of 889 first-price auction bids. If a bidder passed up the opportunity to bid in the Dutch auction at a price less than or equal to the amount she bid in the corresponding first-price auction, then she would be violating strategic equivalence in favor of the first-price auction. A search of the bid data indicates that such violations are rare. Of the 889 first-price bid observations, only 24 were higher than, and 11 equal to, the corresponding winning bids in the Dutch auction. These 24 violations favoring the first-price auction ranged in size from \$0.10 to \$3.00, with a mean of \$0.71. With only 24 such violations from 889 possibilities, bidders seldom bid "too high" in the first-price auction.²⁶

In summary, the experimental data reveal a violation of revenue equivalence: the Dutch auctions raised significantly more revenue than the first-price auctions. The difference was approximately 30 cents per card, or 30 percent of card value, on average. Part of the explanation may be the endogenous entry of bidders, with the Dutch auction format generating more participation than the first-price auction format in this market. However, individual bidders also appear to violate the predicted strategic equivalence between Dutch and first-price auctions. I observed 30 cases of subjects bidding higher in the Dutch auction and 24 cases higher in the

²⁴ Dutch auctions were not entirely novel to this market, either. One bidder wrote me a message to explain that he did not want to participate in my Dutch auction, because, as he said, "I participated in one previous Dutch Auction on the Net, and I found it to be very very frustrating."

²⁵ The result is not as strong as it might first sound. Suppose bidders make random bidding errors that are equally likely to favor the Dutch auction as the first-price auction. Then, because individual Dutch bids are only observed when they win (or tie), one would be much more likely to observe Dutch bids which have randomly been made too high than those which have randomly been made too low. In such a model of bidding errors, the above sample of Dutch bids is biased towards finding bid differences that favor the Dutch auction.

²⁶ This statistic may be somewhat misleading, since the method would never be able detect violations favoring the first-price auction with a low-valuation bidder (for whom the Dutch bid data are heavily censored).

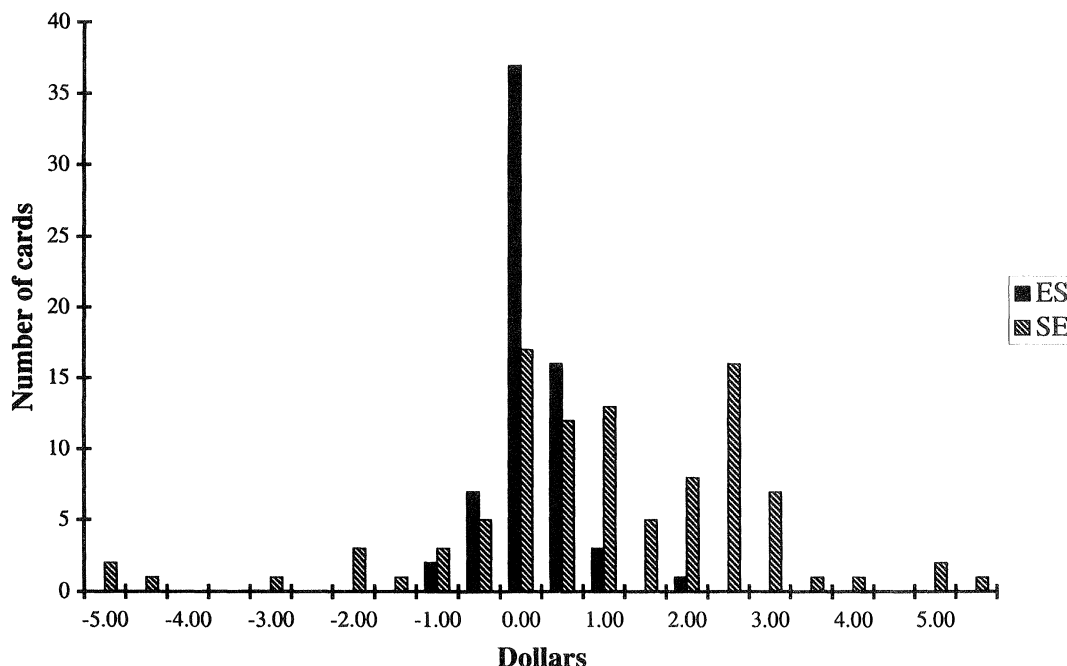


FIGURE 2. ENGLISH REVENUE MINUS SECOND-PRICE REVENUE

first-price auction format,²⁷ with the differences favoring Dutch auction revenues considerably larger than those favoring first-price revenues, though these data are likely biased by selection effects. A possible explanation is that in auctions for real goods, the Dutch clock may have a psychological anchoring effect, suggesting a high price to a person who may not have an explicit cash valuation in mind. Another is that bidders may be impatient over a timescale of days, and bidding early in a Dutch auction provides additional gratification to the winner by ending the auction several days earlier.

B. English and Second-Price Auctions

Table 4 showed that experiment ES included 66 matched pairs,²⁸ while experiment SE in-

cluded 98. Figure 2 shows the distribution of the revenue differences between the English and the second-price auctions for each card. There is a tendency for this difference to be positive in experiment SE, and a slight tendency for it to be negative in experiment ES. Table 6 displays numerical statistics for the two experiments. As before, the upper half of the table shows results for the full set of matched pairs of cards. A majority of observations (38 of 66) in experiment ES had higher English than second-price revenues, but a minority (24 of 98) had this feature in experiment SE. A *t*-test rejects the null hypothesis of revenue equivalence in both treatments, but in opposite directions. The difference between English and second-price revenues is positive and significant for experiment SE, but negative and significant for experiment ES. The Wilcoxon signed-rank test results are

²⁷ Nine different bidders generated the 30 bid violations favoring the Dutch auction, while 11 different bidders generated the 24 bid violations favoring the first-price auction. Two bidders were in both groups.

²⁸ Experiment ES consisted of two of the earliest auctions in my research program; the reason for the relatively low number of observations is that I included a number of

unmatched cards in this auctions in an effort to ensure that subjects would not come to expect every card I auctioned to have a matched pair in a later auction. Later, I realized I could achieve this same end by running occasional "decoy" auctions with no matches whatsoever, to offset the existence of pairs of auctions in which every card matched.

TABLE 6—THE DIFFERENCE BETWEEN ENGLISH AND SECOND-PRICE REVENUES

	Experiment ES	Experiment SE
All observations	66	98
English revenue higher	20	65
Equal revenue	8	9
Second-price revenue higher	38	24
Mean English-second difference	-\$0.09	\$0.85
(Standard deviation of mean)	(0.06)	(0.19)
Mean log difference	-0.408	0.436
(Standard deviation of mean)	(0.092)	(0.085)
Wilcoxon statistic	-2.44	4.78
Restricted sample ^a	22	73
English revenue higher	6	44
Equal revenue	2	7
Second-price revenue higher	14	22
Mean Dutch-first difference	-\$0.15	\$0.79
(Standard deviation of mean)	(0.08)	(0.22)
Mean log difference	-0.425	0.451
(Standard deviation of mean)	(0.180)	(0.105)
Wilcoxon statistic	-1.74	3.56

^a The restricted sample excludes any observation where the winning bidder was attracted to the auction via newsgroup advertisement rather than via e-mail invitation.

qualitatively the same. Since the results favor the second auction in each pair, and since my invitation policy involved sending a greater number of direct e-mail invitations to the second auction in each pair,²⁹ this argues for restricting attention to those auctions which were unaffected by the difference in invitation policies.

The bottom half of Table 6 includes only those cards for which both winners came from the original, smaller e-mail invitation list. Restricting the sample filters out many observations, but it does not change the qualitative results. The majority of cards in each sample have higher revenues in the second auction in each pair, and the mean difference in revenues remains opposite in sign between the two auctions. The *t*-test and Wilcoxon test continue to reject revenue equivalence of mean revenues for experiment SE, but no longer reject the null for

²⁹ Recall that the second auction's invitation list included everyone on the first auction's e-mail list, plus bidders who joined the first auction via a newsgroup advertisement.

experiment ES (whose sample size is reduced to only 22 by the restriction).

Pooled results for the two experiments come in the form of a regression of the difference between English and second-price revenues, on a dummy variable for the experiment treatment order (ES versus SE). The results are as follows (standard errors in parentheses):

$$(2) \quad REVDIFF = -0.09 + 0.94SE \\ (0.18) \quad (0.24).$$

The intercept indicates that the revenue difference is not statistically significantly different from zero in experiment ES, while the slope indicates that the second auction in the pair tends to have higher revenues (SE has a significantly larger difference between English and second-price revenues than does ES). I have presented results for the full sample ($N = 164$, $R^2 = 0.090$); the results for the restricted sample are qualitatively similar.

Next, I turn to data at the level of the individual bidder. Looking for cases where the same subject submitted both a second-price auction bid and an English auction bid on the same card,³⁰ I find 72 bid pairs submitted by a total of 8 subjects in experiment ES, and 159 bid pairs submitted by 10 subjects in experiment SE.³¹ The data for this comparison are more complete than for the Dutch-first comparison, as both the second-price and English auctions yield bid amounts on each card that interests a bidder. A violation of strategic equivalence in favor of the

³⁰ One might argue that I should also have considered instances where, despite participating in both auctions in a pair, a bidder submitted a bid for card X in one auction but not in the other auction. (There were, in fact, dozens of such instances.) I could have included such observations by taking the nonbid to be a bid equal to zero. However, I believe that in such cases, the decision not to bid on a card is probably due to changes in the bidder's demand for that card during the time between auctions (for example, the bidder obtains the card in some other way), rather than being caused by the choice of auction format. Therefore, I restrict my attention to those cases where the same bidder submits two bids on the same card in two different auctions. My assumption is that changes in bid levels may be due to the auction format, while discrete decisions whether or not to bid on a particular card are more likely to be due to demand shifts.

³¹ One bidder overlapped between these two sets.

English auction is relatively straightforward: the bidder's highest English auction bid exceeds her second-price auction bid on the same card.³² However, the opposite violation is more difficult to establish, as there are two other reasons why a bidder's highest English bid might be lower than her second-price bid. First, winning bidders have their bid data censored from above: a winning bidder would typically not reveal her full valuation but bid only a minimal amount more than the second-highest bidder. Second, a "jump bid" by an opponent might prevent a losing bidder from revealing her full valuation in the English auction.³³ With these potential pitfalls in mind, I attempt to see to what extent individual bidders may violate strategic equivalence.

To avoid sample-selection bias, I shall first define the "English dropout price" for a particular bidder and a particular card to be the smallest bid by an opponent which exceeds our bidder's final bid.³⁴ For example, if bidder X had the current high bid of \$1.50 on Tuesday, then bidder Y raised the price to \$1.60 on Wednesday, and bidder X never again bid on that card, then the "English dropout price" was \$1.60 for bidder X. An observed violation of strategic equivalence favoring the second-price auction occurs when the bidder's English bid is less than or equal to her English dropout price, which in turn is strictly less than her second-price bid.

Of 231 matched-bid observations, 39 were violations of strategic equivalence in favor of the second-price auction format, ranging from \$0.05 to \$11.00, with a mean of \$2.58. There were 81 violations in favor of the English format, ranging from \$0.20 to \$8.00, with a mean

of \$1.69. Of the remaining observations, nine had equal second-price and English bid strategies, while 102 were indeterminate.³⁵ Note that the categories used here have been designed to avoid introducing a sample-selection bias on the observations of bid violations. The violations favoring the English auction format outnumber the violations favoring the second-price format by a margin of more than two to one, although they are smaller in magnitude on average.³⁶ Figure 3 displays the distribution of the difference in bid level for the 129 nonindeterminate observations. The data indicate a failure of strategic equivalence for this sample of subjects, whose strategies tend to involve higher amounts in the English bid than in the second-price auction.³⁷

Breaking down the matched-bid observations by individual bidder, it becomes apparent that some bidders behaved differently than others. Table 7 displays statistics on the mean difference (normalized by list value) between the two bids submitted, with separate means computed for each bidder. Of the 17 different bidders with matched bid observations, nine tended to bid higher in the English auctions (that is, more of their observed bid differences were positive than were negative), while five tended to bid higher in the second-price auctions. Of these, only bidders 1, 6, 9, and 14 seem convincingly and consistently to have bid differently than strategic equivalence would suggest. (Each of these bidders yielded at least four different determinate observations, and for each the vast majority of observations favored one auction

³² This type of violation might sometimes be unobservable. If a bidder won the English auction at a price lower than her bid in the second-price auction, then her maximum English bid might not have been revealed.

³³ For example, suppose a bidder was willing to bid \$5.00 either in a second-price or an English auction. In the second-price auction, I would actually observe her bid of \$5.00. But if in the English auction her bid of \$4.50 was raised immediately to \$6.00 by someone else, then our bidder would never have the opportunity to submit a bid of \$5.00. The bid of \$4.50 would mistakenly suggest that she was willing to bid less in the English auction than in the second-price auction.

³⁴ If our bidder won the card in that English auction, then the dropout price is zero.

³⁵ Indeterminacy results when either a jump bid or a too-low English dropout price prevents me from observing the bidder's English bid strategy precisely enough to compare it to her second-price bid.

³⁶ Note that the category of "indeterminate" observations was developed to eliminate sample-selection bias. For example, suppose that a bidder would have bid higher in the English than in the second-price auction, but that the English bid price never rose as high as his second-price auction bid. Then, although it might at first appear that his English bid were lower than his second-price bid, it is only because the auction never gave him the opportunity to raise his English bid to be greater than or equal to his second-price bid. I therefore classify such an observation as indeterminate, because it is as likely to represent one type of violation as the other.

³⁷ Differences in logarithms produce qualitatively similar results.

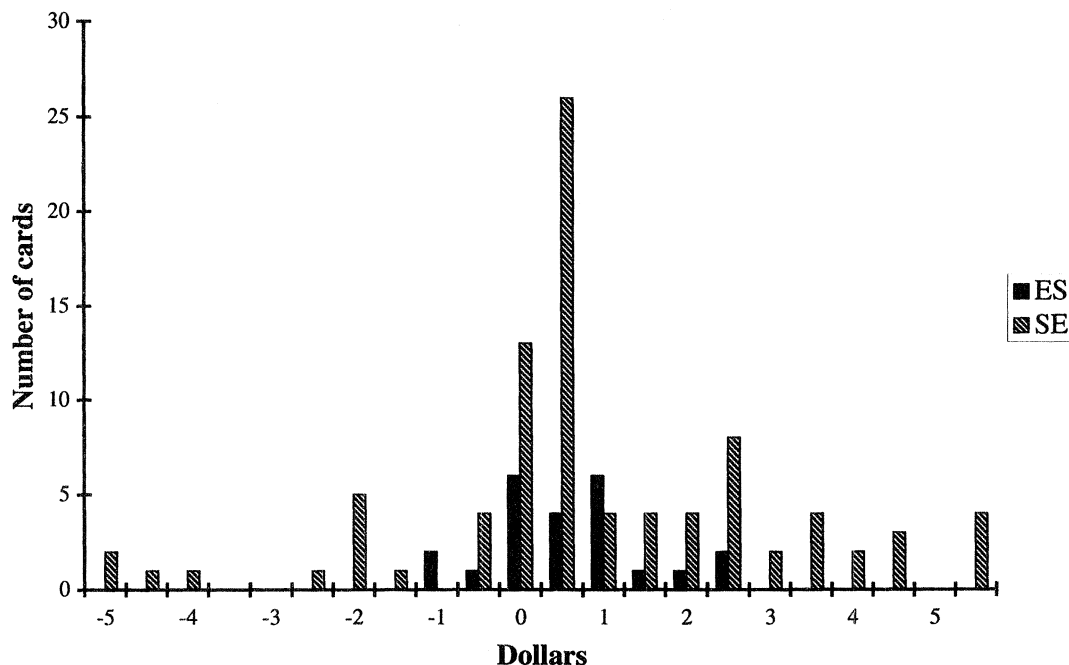


FIGURE 3. ENGLISH-SECOND DIFFERENCES OBSERVED IN MATCHED BIDS

over the other.) Three bidders consistently favored the English auction, while one consistently favored the second-price auction. Thus, although the bidders taken as a whole tend to bid higher in the English rather than the second-price auction format, there is some heterogeneity across individuals.

To summarize, the card-level evidence indicate that English and second-price auctions do exhibit revenue equivalence: the overall differences in mean revenues is small and statistically insignificant. The matched bid-level data show some tendency for subjects to bid higher in the English than in the second-price auctions, although this behavior is heterogeneous. Taking the average across the bidder data displayed in Table 7, bids were 3.0 percent higher in the English auction. This is somewhat surprising, as laboratory experiments with private values have typically observed overbidding in second-price auctions relative to English auctions. A possible explanation is that bidders' valuations for Magic cards could be privately uncertain and affiliated across bidders, causing English bids to be higher for reasons proposed by Milgrom and Weber (1982). I feel, however, that private un-

certainty is unlikely in this market. Bidders' valuations differ by their individually known uses for the cards (playing with a special deck, completing a collection) and their individually known opportunities to buy them outside the auction (friends, local retail stores with varying prices). Privately uncertain affiliated values would instead imply that bidders were speculating in the cards for some unknown future resale value. In my experiences talking with bidders, I found that most planned to use the cards for game play or to complete their personal collections. In any case, the observed English-second revenue differences are minimal compared to the Dutch-first differences, so the effects of affiliation on revenue are at best rather small.

IV. Concluding Remarks

The question of equivalence between auction formats has been extensively studied by auction theorists, and studies subsequent to Vickrey's pioneering article have shown reasons (such as affiliation of values) why which revenue equivalence might fail to hold. In this study, I use field experiments to rank the auction formats

TABLE 7—INDIVIDUAL BIDDERS' DIFFERENCES BETWEEN ENGLISH BID STRATEGY AND SECOND-PRICE BID STRATEGY

Bidder	Experiment	Number of observations				Difference ^a	
		Positive	Zero	Negative	Indeterminate ^b	Mean	Standard deviation
1	SE	14	0	0	1	3.30	1.89
2	ES	1	0	0	0	1.75	
3	ES	3	0	0	1	0.60	0.18
4	SE	0	0	1	0	-3.50	
5	SE	8	1	8	3	-0.06	2.78
6	ES	0	0	9	18	-1.28	1.12
7	SE	0	1	2	16	-0.12	0.10
8	SE	1	0	0	0	2.50	
9	Both	39	3	10	37	0.03	3.21
10	ES	0	0	1	0	-2.30	
11	SE	3	2	3	11	-0.47	1.94
12	ES	4	1	2	10	0.28	1.06
13	SE	4	0	0	2	0.75	0.54
14	ES	1	0	0	0	4.50	
15	SE	0	0	2	1	-3.88	1.24
16	ES	1	0	1	2	0.05	0.28
17	SE	2	1	0	0	1.17	1.81
Overall		162	18	78	204	0.28	2.77

^a The difference is defined as the English bid minus the second-price bid, in dollars.

^b An indeterminate observation is one for which data censoring in the English auction makes it impossible to observe the amount of the difference in bid strategy. Indeterminate observations are not included in the means and standard deviations reported in the table.

according to their generated revenues, in order to investigate which theoretical predictions are closest to the truth in a real-world auction market. I find, contrary to theoretical predictions and to previous laboratory results, that Dutch auctions earn approximately 30 percent more revenue than first-price auctions. I also find revenues to be roughly equivalent between English and second-price auctions, consistent with Vickrey's theoretical independent-private-values model, although bid-level data indicate some tendency for individual bidders to bid higher in the English than the second-price auction format.

Typical laboratory tests of economic theory are designed to make the assumptions of the theory true in the laboratory setting. They test whether behavior follows the predictions of the theory, conditional on the environmental assumptions being true. My field experiments conduct *unconditional* tests of a theory's predictions, because I do not observe whether or not the underlying assumptions of the theory are true. Both types of tests are useful. Laboratory tests provide specific feedback to theorists

about exactly where their theories fail. Field tests assess the practical predictive power of a theory, since most theoretical assumptions in economic models are intrinsically unobservable in practice.

These field experiments provide, for the first time, evidence from real-world auctions that allow revenue comparisons between the four basic auction formats. They indicate that Dutch auction revenues exceed first-price auction revenues, and English auction revenues are not significantly higher than second-price auction revenues. However, a number of questions about revenue rankings still remain. How do revenues in the two pairs of auctions compare with each other? Can these results be replicated in other auction markets? The findings in this paper also point out useful directions for laboratory research, such as adding the potential for endogenous bidder entry to laboratory studies of revenue equivalence. Other directions, given the recent advances in communications technology which have led both to card auctions on the Internet and simultaneous English-style auctions for FCC communications spectrum, would

involve explorations of the differences of simultaneous versus sequential auction formats, and of longer timescales in dynamic auctions (days and hours, versus minutes and seconds).

REFERENCES

- Ashenfelter, Orley. "How Auctions Work for Wine and Art." *Journal of Economic Perspectives*, Summer 1989, 3(3), pp. 23–36.
- Black, Jason. *Cloister's Magic card price list*. <<http://www.hhhh.org/cloister/pricelists/>>, various weekly issues, 1995.
- Coppinger, Vicki M.; Smith, Vernon L. and Titus, John A. "Incentives and Behavior in English, Dutch and Sealed-Bid Auctions." *Economic Inquiry*, January 1980, 18(1), pp. 1–22.
- Cox, James C.; Roberson, Bruce and Smith, Vernon L. "Theory and Behavior of Single Object Auctions," in Vernon L. Smith, ed., *Research in experimental economics*. Greenwich, CT: JAI Press, 1982, pp. 1–43.
- Cox, James C.; Smith, Vernon L. and Walker, James M. "A Test that Discriminates Between Two Models of the Dutch-First Non-Isomorphism." *Journal of Economic Behavior and Organization*, June–September 1983, 4(2–3), pp. 205–19.
- Hansen, Robert G. "Empirical Testing of Auction Theory." *American Economic Review*, May 1985 (*Papers and Proceedings*), 75(2), pp. 156–59.
- . "Sealed-Bid Versus Open Auctions: The Evidence." *Economic Inquiry*, January 1986, 24(1), pp. 125–42.
- Hendricks, Kenneth and Paarsch, Harry J. "A Survey of Recent Empirical Work Concerning Auctions." *Canadian Journal of Economics*, May 1995, 28(2), pp. 315–38.
- Johnson, Ronald. "Oral Auctions versus Sealed Bids: An Empirical Investigation." *Natural Resources Journal*, April 1979, 19(2), pp. 315–35.
- Kagel, John H. "Auctions: A Survey of Experimental Research," in John H. Kagel and Alvin E. Roth, eds., *Handbook of experimental economics*. Princeton: Princeton University Press, 1995, pp. 501–85.
- Kagel, John H.; Harstad, Ronald M. and Levin, Dan. "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." *Econometrica*, November 1987, 55(6), pp. 1275–304.
- Kagel, John H. and Levin, Dan. "Independent Private Value Auctions: Bidder Behaviour in First-, Second-, and Third-Price Auctions with Varying Numbers of Bidders." *Economic Journal*, July 1993, 103(419), pp. 868–79.
- Lucking-Reiley, David. "Vickrey Auctions Pre-date Vickrey." *Journal of Economic Perspectives*, 2000 (forthcoming).
- Magic: the Gathering*. <<http://www.wizards.com/magic>>.
- Maskin, Eric S. and Riley, John G. "Optimal Auctions with Risk Averse Buyers." *Econometrica*, November 1984, 52(6), pp. 1473–518.
- McAfee, R. Preston and McMillan, John. "Auctions and Bidding." *Journal of Economic Literature*, June 1987, 25(2), pp. 699–738.
- McMillan, John. "Selling Spectrum Rights." *Journal of Economic Perspectives*, Summer 1994, 8(3), pp. 145–62.
- Mead, Walter J. *Competition and oligopsony in the Douglas fir lumber industry*. Berkeley, CA: University of California Press, 1966.
- . "Natural Resource Disposal—Oral Auction Versus Sealed Bids." *Natural Resources Journal*, April 1967, 7(2), pp. 194–224.
- Milgrom, Paul R. and Weber, Robert J. "A Theory of Auctions and Competitive Bidding." *Econometrica*, September 1982, pp. 1089–122.
- Reiley, David. "Magic on the Internet: Experimental Tests of Auction Theory." Ph.D. dissertation, Massachusetts Institute of Technology, 1996.
- Riley, John G. "Expected Revenue from Open and Sealed Bid Auctions." *Journal of Economic Perspectives*, Summer 1989, pp. 41–50.
- Riley, John G. and Samuelson, William F. "Optimal Auctions." *American Economic Review*, June 1981, 71(3), pp. 381–92.
- Rothkopf, Michael H.; Teisberg, Thomas J. and Kahn, Edward P. "Why Are Vickrey Auctions Rare?" *Journal of Political Economy*, February 1990, 98(1), pp. 94–109.
- Tenorio, Rafael. "Revenue Equivalence and Bidding Behavior in a Multi-unit Auction Market: An Empirical Analysis." *Review of*

- Economics and Statistics*, May 1993, 75(2), pp. 302–14.
- Vickrey, William.** “Counterspeculation, Auctions, and Competitive Sealed Tenders.” *Journal of Finance*, May 1961, 16(1), pp. 8–37.
- Wilson, Robert.** “Strategic Analysis of Auctions,” in Robert J. Aumann and Sergiu Hart, eds., *Handbook of game theory with economic applications*, Vol. 1. Amsterdam: North-Holland, 1992, pp. 227–79.