How can we evaluate conservation auctions? Three possible methods

Steven G.M. Schilizzi

School of Agricultural & Resource Economics
The University of Western Australia
Steven.Schilizzi@uwa.edu.au

Contributed paper prepared for presentation at the 56th AARES annual conference, Fremantle, Western Australia, February7-10, 2012

How can we evaluate conservation auctions? Three possible methods

Steven G.M. Schilizzi

School of Agricultural & Resource Economics
The University of Western Australia
Steven.Schilizzi@uwa.edu.au

Abstract

By design, tenders are used when costs are unknown. But if costs are unknown, how can we evaluate the tenders, when their evaluation involves measuring their cost-effectiveness? We identify three approaches: theoretical, empirical and experimental. We first use experimental data to compare the efficiency of each approach, then apply the most efficient one to field data from the Scottish fishing vessel decommissioning program. We estimate the potential errors one would make in using a less efficient approach. In this case, we demonstrate a novel use of controlled lab experiments for interpreting field data and evaluating policy effectiveness.

Keywords

Auctions; conservation; experiments; evaluation; measurement; market-based policy instruments

Introduction

Conservation tenders are policy tools typically used by government agencies to buy environmental services from private landholders. By now, in 2012, they have acquired a rather rich history, with the Conservation Reserve Program in the United States reaching back to 1985 (Reichelderfer and Boggess, 1988). Other countries that have made significant use of them include Australia, the United Kingdom and Germany, with, recently, some developing countries also entering the fray, such as Indonesia and Malawi (Jack *et al*., 2009; Jack, 2011). It seems natural, therefore, that they should be submitted to some sort of evaluation, to see whether their use achieves the desired goals and whether, should they do so, other policy instruments might not do better.

Surprisingly, the art – or perhaps the science – of auction evaluation in this context is not well developed. In 2003, Stoneham et al. produced an evaluation of the BushTender program in south-eastern Australia that led to some controversy in academic circles. They asserted that this program had been seven times more efficient than an equivalent fixed price scheme where all landholders would get paid the same price per unit of environmental service. Such a statement was based on a specific method, or lack thereof, for evaluating the performance of this program. As this paper will show, this issue of auction evaluation is not trivial but several approaches to solving the problem are available.

Auction evaluation is a non-trivial problem only for auctions based on the buy-as-you-bid principle, or first-price auctions. This is because, as shown for example by Cox et al. (1982;

1984), this auction format does not provide incentives for bidders to bid their true costs, at least to the extent these costs are known to them; instead, in procurement settings, where bidders are sellers rather than buyers, they tend to bid higher. Such is not the case for so-called second-price auctions, where, typically, all successful bidders are paid the same price, set by the bid of the first rejected bidder (Vickrey, 1961). This auction format does provide the incentives for bidders to bid their true costs. As a result, the present analysis applies only to those settings where first-price auctions are involved. In the context of conservation tenders, this has to date nearly always been the case.

In addition, the problem addressed here concerns ex-post evaluations, that is, one wishes to evaluate an auction that has already been held; it does not apply to the ex-ante problem where one wishes to know whether to go ahead or not with an auction. The key difference between the two approaches lies in the fact that in the first case, one can use the bids that were submitted, whereas in the second case such bids are not forthcoming.

In this paper, we will describe, test, evaluate and compare three different approaches to the ex-post, first-price auction evaluation problem, then apply the one which obtains the best test results to the evaluation of the Scottish fishing vessel decommissioning auctions carried out in 2001 and 2003. Accordingly, the next section of the paper describes then tests and evaluates each of the three approaches; following that, the best performing method is used to evaluate two fishing vessel decommissioning auctions carried out in Scotland in 2001 and 2003. A final section concludes.

Methods

In this section we shall define then test and evaluate three different methods relevant to the ex-post, first-price auction evaluation problem. The three methods are theoretical, empirical and experimental. The latter two are innovations. The empirical approach, which is still work in progress, was proposed only a couple of years ago, and the experimental approach is, to our knowledge, presented for the first time. Only the theoretical approach is not new.

To date, only methods based on theoretical models have been used for auction evaluation in this context. Rosse (1970) seems to have been the first to have devised estimation procedures to recover cost functions from data on market-clearing prices and quantities. He was followed by such analysts as Porter (1983), Bresnahan (1987), Berry et al. (1995), and Goldberg (1995). More recently, Wolak (2002) recovers bidder costs from electricity auctions, Cantillon and Pesendorfer (2005: Chapter 22, section 5) from auctioning bus routes in London, and Krasnokutskaya and Seim (2008) from highway procurement auctions. All follow essentially the same principle and a similar procedure. The principle is that all cost estimates combine an assumed functional form for the aggregate demand of the products under consideration, and the assumed model of competition among individuals or firms. The procedure typically involves four steps: 1) choice of the game and model describing the auction data; 2) computation of optimal bids assuming rational behaviour based on maximising expected profits, followed by an estimation of the equilibrium distribution of bids; 3) computation of the probabilities of winning; 4) estimation of bidder costs. This approach typically leads to an econometric exercise, which explains why so many of the early papers were published in *Econometrica*. Fundamentally, the theoretical approach involves deducing bidder costs assuming bidders bid optimally and maximise their expected benefits. The approach is theoretical as there is no guarantee that real bidders actually behave in this way, e.g. perfect rationality.

However, in the special case of conservation tenders, a purpose-built model was proposed in 1997 by Latacz-Lohmann and Van der Hamsvoort, which, although following the same rationale, is specifically suited to this context.

This section first presents and describes the three approaches, then proceeds to test each one using data from controlled lab experiments.

Section 1 : Defining the three evaluation methods

Theoretical method – using optimal bids

Latacz-Lohmann and Van der Hamsvoort (1997), hereafter LH, were the first to propose a theoretical model for sealed-bid, discriminatory-price, budget-constrained procurement auctions, or tenders, applied to conservation programs. Their model differs from the target-constrained model in that it is a best-response rather than a Nash equilibrium model. It introduces an external parameter β, which represents the bid cap, or highest bid acceptable to the tendering agency. Bidders to not know the bid cap, so they form expectations about β, around which a distribution can be assumed. In the LH model, it was assumed for any individual bidder i to be uniform between a lower and upper bound, $[\underline{\beta}_i, \overline{\beta}_i]$. Bidders will not put in bids higher than $\overline{\beta}$ given that this would contradict their expectations. The resulting model of optimal bidding behaviour, assuming risk-neutral bidders, was shown to have the following form:

$$b_i^* = \max\left(\frac{c + \overline{\beta}_i}{2}, \underline{\beta}_i\right) \tag{1}$$

where the optimal bid b_i^* is predicted to be both greater than the lowest estimate of the bid cap ($\underline{\beta}_i$) and a function of participation costs, c_i. However, optimal bids will not reveal bidder costs perfectly given that they also depend on the (highest) bid cap expectation, $\overline{\beta}_i$. It is well known from auction theory that sealed-bid first-price, pay-as-you-bid auction mechanisms involve an optimal degree of bid shading and therefore provide a built-in rent for bidders. By solving the model for c_i, one can express costs as a function of bids and expected bid caps (we drop the subscript i for simplicity):

$$\text{if } \frac{1}{2}(c + \overline{\beta}) < \underline{\beta} \quad \text{then} \quad b^* = \underline{\beta} \quad \text{in which case } c \text{ cannot be defined} \tag{2a}$$

$$\text{if } \frac{1}{2}(c + \overline{\beta}) > \underline{\beta} \quad \text{then} \quad b^* = \frac{1}{2}(c + \overline{\beta}) \tag{2b}$$

$$\text{in which case } c = 2b^* - \overline{\beta} \tag{3}$$

Empirical method – using bids from the field

In the absence of data including information on bidders' costs or, as in the LH model, on bid cap expectations, an approach based on theoretically optimal bids may not be practical to implement. One way to circumvent this difficulty is to adopt a direct approach based on a statistical analysis of the empirical bids generated in the field, plus some necessary assumptions. This is the approach adopted by White and Burton (2010), hereafter WB. They start with the assumption that a bid can be considered as the sum of the bidder's costs plus an element of rent (recall the bid shading property of first-price or discriminatory-price auctions):

$$b_i = c_i + r_i \tag{5a}$$

They then estimate the bid function together with a random iid residual term:

$$b_i = \beta_0 + \beta_i c_i + u_i \tag{5b}$$

where β_0 is the constant term, β_i is the cost coefficient and u_i is the iid residual.

BW then consider the algebraic deviations of each bid relative to the estimated bid and make the following assumption: that any deviation larger than the one with the largest negative value (that is, the smallest absolute residual) represents a rent. By implication, the largest negative value is taken as the benchmark for the minimum rent, which is set equal to zero:

$$\hat{r}_i = u_i - \min_i\{u_i\} \tag{6}$$

where $\hat{r}_i$ is the estimated rent and $\min_i\{u_i\}$ is the largest value of the residual below the fitted curve. Costs can then be deduced as

$$c_i = b_i - \hat{r}_i \ . \qquad (7)$$

Experimental method – using bids generated in the lab

The idea here is to design experiments that are isomorphic to the mechanism that generated the field data. In this case, one needs to design an auction that shares the same key characteristics as the field auction. The choice of these characteristics depends on what parameter one wishes to extrapolate. If, as is the case here, we believe that bidders across isomorphic settings will, on average, show the same bidding behaviour and, in particular, the same degree of bid shading, then the experiment must make sure it controls for all the parameters likely to influence bidders' bid shading. One parameter that is not easy to control for is the degree of bidder risk aversion. However, if one is careful to take a measurement before the auction, one should be able to adjust for this parameter, provided one knows something about the average risk aversion of the population in the field.

In this case, the key controls were the degree of competition intensity (the budget relative to the number of bidders) and the expected selection ratio: although unknown beforehand in a budget-constrained auction, it is possible to calibrate a lab experiment to obtain a fairly good ex-ante approximation of the selection ratio to reproduce that of the field auction, in this case two-thirds of bidders. Another important parameter is the distribution of costs, both in terms of its range and in terms of its shape, uniformly distributed (linear function) or otherwise. If the experiments are run once the bids are in from the field, one can choose the cost distribution to reflect the bid distribution. Experiments have not yet been run to test for the importance of the shape of the cost distribution on the degree of bid shading, so we cannot tell for sure how one will affect the other.

The solution chosen for this study was to extrapolate from the experiments the degree of bid shading as a function of the level of cost, which, as predicted by first-price auction theory, shows that bid shading is a decreasing function of cost (Cox *et al*., 1982; 1984).

Provided all these aspects are controlled for in the experiments, one can formulate the working hypothesis that bid shading will be an invariant across isomorphic auction settings. The rationale for this hypothesis is that the incentive structures between two isomorphic settings are identical; on average, therefore, bidder behaviour should also be identical.

The logic is as follows, with *b*, *c* and *r* denoting bidders' bids, costs and rents, respectively, and indexes *x* and *f* denoting experimental and field data, respectively:

$$r_f\,(b_f,\, c_f) = r_x\,(b_x,\, c_x) = r \qquad (8a)$$

It is of course not possible to know c_f nor r_f ; we therefore have no *direct* way of knowing whether our hypothesis above is a reasonable one to make or not: that is, we have no direct way of controlling for external validity. We have however an *indirect* approach at our disposal. The two experiments K and P were carried out in two different countries with two different populations; we can therefore make one hold the role of *f* while the other takes on the role of *x*, and vice versa. In other words, we can consider, say, the K experiment to represent our field data set, for which we assume costs and rents in K, c_k and r_k, unknown, and use relationship (8a) above to compute them using our knowledge of $r_p(b_p, c_p)$, where the indexes *k* and *p* take on the roles of *f* and *x* respectively ($f \rightarrow k$ and $x \rightarrow p$). Thus, our working assumption now becomes:

$$r_k\,(b_k,\, c_k) = r_p\,(b_p,\, c_p) = r \qquad (8b)$$

whence we deduce

$$c_k = c_k\,(b_k,\, r) \qquad (9)$$

and more specifically, since *r* is generically defined as $r = (b - c) / b$

$$c_k = b_k \times (1 - r) \qquad (10)$$

which holds for each individual bidder.

We now have the means to rigorously test whether our working hypothesis is tenable or not: we compare the predicted costs c_k, estimated using P data, with the known experimental costs of the K experiment. The transfer or extrapolation of each bidder-specific r_i is straightforward if the source and the destination data sets have the same number of bidders; if not, as is usually the case, the number of r_i must be adapted to meet the number of bidders in the destination data set. In addition, because bid shading in discriminatory-price auctions is a monotonically decreasing function of individual cost, bidders must be ordered. In this way, the bid-shading-to-cost relationship is preserved across the two data sets.

One can then consider two ways to adapt the number of r_i from one data set to the other: one is by statistically fitting a function $r = r(b)$, and the other is by mechanically shrinking or stretching the number of r_i. This will be illustrated further below.

Section 2: Testing the three methods on experimental data

Testing the theoretical method

The LH model cannot be directly applied to field data, given that neither costs (c_i) nor bid cap expectations [$\underline{\beta}_i$, $\bar{\beta}_i$] are observed. The model can however be applied to experimental data if the experiments are designed so that both these data are observed. This is what was done with the K and the P experiments, where bids, costs and expected bid caps were observed (see Appendix I). In addition, the experiments had costs constrained to $c_i \geq 0$, so equation (XXX) above becomes:

$$c_i = \max(0, 2b_i^* - \bar{\beta}_i) \qquad (4)$$

Before using the inverse bid function (4) to estimate costs, it was necessary to check, using the direct relationships (2a) and (2b), how well the optimal bidding model predicted the observed bids. For the K data, the model underestimates real bids by about 7% (Figure 1). This can be understood when one takes into account the fact that the model optimizes bids assuming bidder risk neutrality. A separate measurement of bidders' risk aversion prior to the experiments proper evaluated their average certainty equivalent ratio at 108%, slightly risk-loving relative to a risk-neutral ratio of 100%. The two results are in close correspondence, with the expected sign. Similar results were obtained with the P data, except that in that case bidders were measured as being risk-averse, with, as one should expect, the estimated bid line situated a little above, rather than below, the experimental bid line.

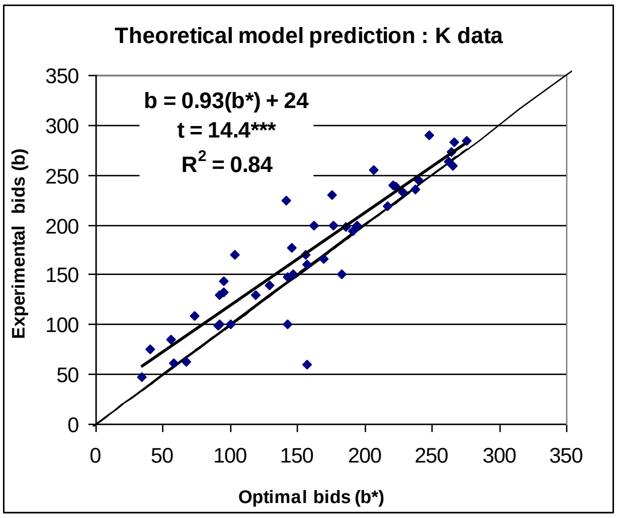


Figure 1

One would expect that, logically, the model should perform equally well in predicting costs as in predicting bids; but as the comparison with the K and P data revealed, this is not necessarily the case. Whereas the model predicted K bids better than P bids, the reverse was true for costs. This is because the max operator in equation (4) does not allow for complete symmetry between the two computations.

The error in cost prediction can be measured by regressing the estimated costs against the actual experimental costs (Figure 2). A perfect match would have the regression coefficient equal to 1.00 and a constant term equal to 0. In the K experiment the estimated cost coefficient is off by –13%, while in the P experiment it is off by +6% (both with confidence levels of better than 99%). The constant term is less accurate, being off by 19% in terms of the min–max cost range of 260, while its estimate is not statistically significant with the P experiment (smaller number of data, n = 27, compared to n = 44 in K).

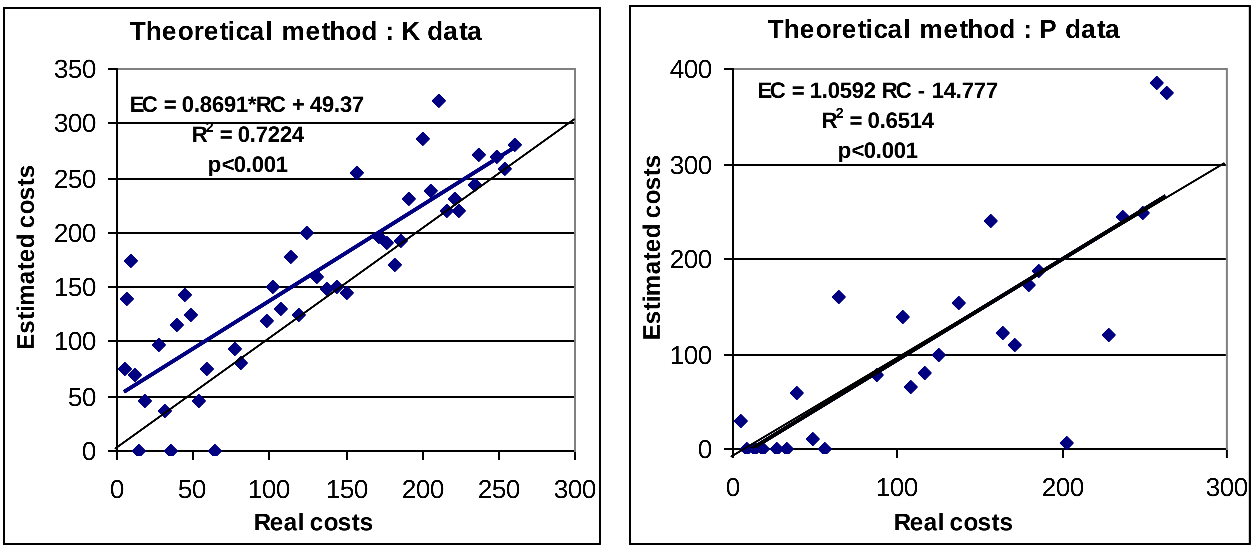


Figure 2

Legend: EC = Estimated Costs
RC = Real Costs = experimental data

In terms of the auction's cost-efficiency (CE), defined as total payment divided by total cost, one only includes in the calculation those bidders who were selected. In this case, the error made in estimating CE was –26% for the K data and +8% for the P data. These results did not improve when potential outliers were excluded. Note that the opposite signs for the K and P data reflect, as explained above, the fact that bidders were measured as risk-takers in the first case but risk-averse in the latter, the model itself assuming risk neutrality.

Testing the empirical method

To test the accuracy of this procedure, it was applied to the K and P experimental data, the result of which is given in Figures 3 and 4. It is clear from the comparison of the two data sets in Figure 3 that the efficiency of this method depends on the degree of variability of the bids around the fitted curve (in this case, a straight line, since costs were uniformly distributed). Bid variability was greater in the P data than in the K data (small triangles around the top-most line); as a result, this method yields better results for P than for K. In the K experiment, this method appears to perform quite poorly. This method is particularly sensitive to extreme bid values or outliers, a property which it shares with data envelopment analysis, a method with which it is closely related.

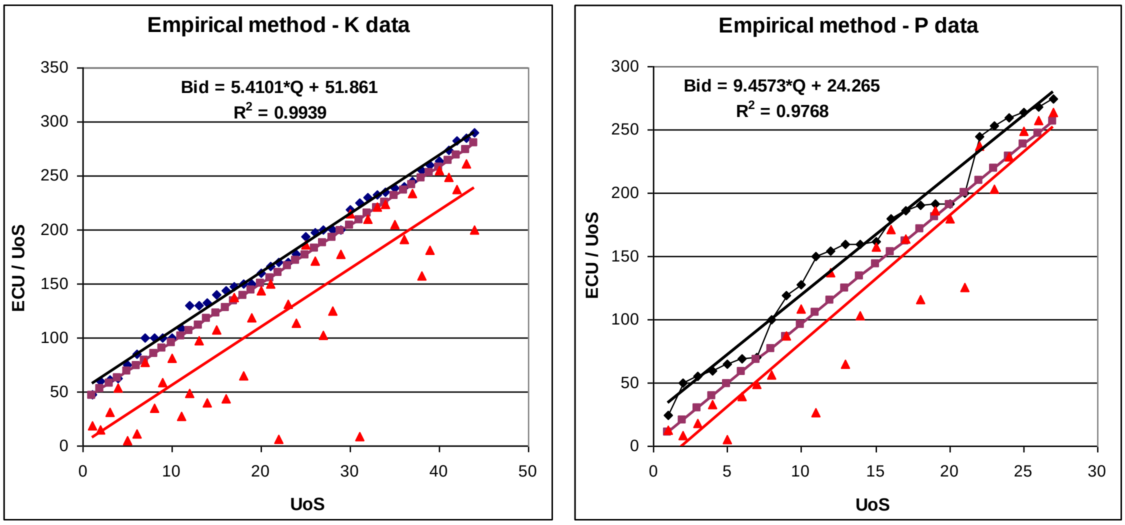


Figure 3

Legend: Bids (diamonds)
Estimated Costs (squares)
Real Costs, experimental (triangles)
ECU = Experimental Currency Units
Q = quantity, in UoS = Units of Service

The error in the linear regression cost coefficient, defined as a deviation from the perfect value of 1.00, was –27% and –18% for the K and P data respectively (Figure 4). The constant term was off by 24% and 8% in terms of the min–max cost range of 260: that is, 63/260 and 32/260 respectively. As was the case for the theoretical approach above, costs are better estimated for the P than for the K data. As highlighted above, with this statistical approach bids varied more widely around the trend in the P data than in the K data, where they were tightly grouped around the fitted trend. This is largely an artefact of this approach.

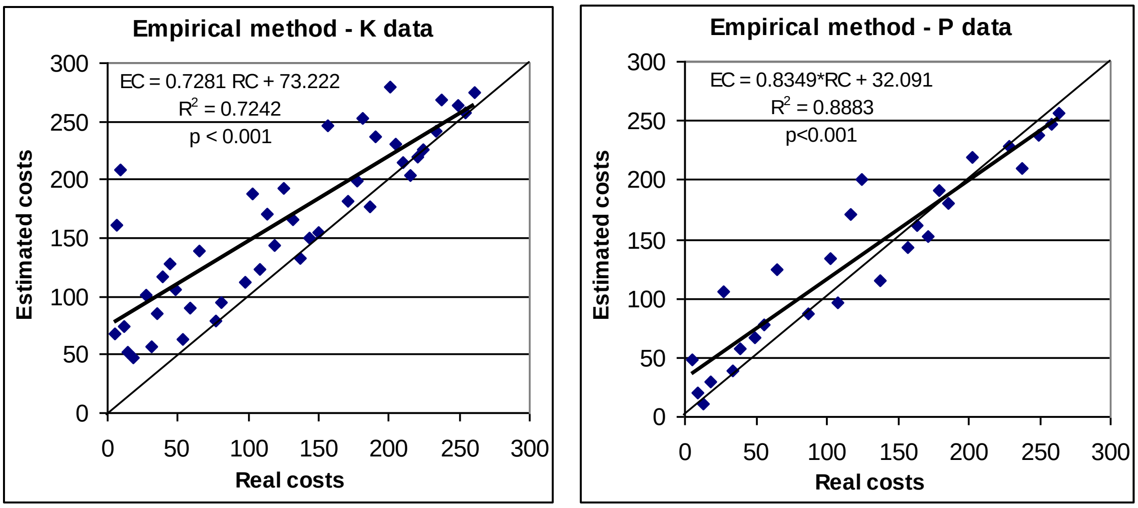


Figure 4

Legend: EC = Estimated Costs
RC = Real Costs = experimental data

In terms of predicting not the costs themselves, but the cost-efficiency of the auction (total payments / total cost), we need to take into account the selection ratio, as imposed by the budget constraint. Prediction errors on the tail end of bidders, those who are not selected, do not matter in this assessment. In this case, the error in cost-effectiveness was –33% for the K experiment and

–15% for the P experiment. If outliers are removed from the data, the error hardly diminishes for the K experiment (–29%) but falls by about one half for the P experiment (–8%).

Testing the experimental transfer method

Neither of the two previous methods – the one using a theoretical model and the one using a direct statistical inference – appears to be very accurate, erring by up to 33% in estimating auction performance. There is however the third possibility, the experimental transfer method, to which we now turn. As highlighted in the methodology section, the transfer can be made using a statistical best-fit or a mechanical shrink or stretch.

Best-fit method

Considering the transfer or extrapolation of bid shading ratios *r* from the source data set P to the destination data set K, one first regresses and fits a function using the source data: $r_p = r_p(b_p)$. We explored a number of different functional forms, searching for the best fit. It consistently turned out to be a (slightly) quadratic function, in this case

$$r_p = 10^{-5}\, b_p^2 - 0.0058\, b_p + 0.8558 \quad \text{(t-value = – 4.65; p < 0.001)}$$

We then used this function on the destination data set (K) to compute individual bid shading ratios r_k as a function of own bids b_k, that is:

$$r_k = 10^{-5}\, b_k^2 - 0.0058\, b_k + 0.8558$$

from which one can directly compute for each bidder the underlying costs as

$$c_k = b_k \times (1 - r_k)$$

The reverse prediction was also done, by swapping the roles of the source and destination data sets, now predicting from K to P; the best-fit function was also a quadratic:

$$r_k = 9\times10^{-6}\, b_k^2 - 0.0053\, b_k + 0.9217 \quad \text{(t-value = – 4.91; p < 0.001)}$$

The error in the regression cost coefficient, defined as a deviation from the perfect value of 1.00, was –15% in the P-to-K prediction and –17% in the K-to-P prediction. The constant term in the P-to-K prediction was off by +18% (t-value = 3.65) in terms of the min–max cost range of 260, while its estimated value was not statistically significant in the K-to-P prediction.

Prediction of the auction's performance in terms of cost-efficiency (total payment / total cost) was off by –20% in the P-to-K but only by +5% in the K-to-P. With outliers excluded (two in both cases), the first estimate improved to –15% but the second deteriorated to +16%. These results are all summarised in Table 1 below.

This best-fit transfer method is indifferent to whether the source data set is smaller or larger than the destination data set. This is not true of the second method, to which we now turn.

Shrink or stretch method

An alternative method to transfer the individual bid-shading ratios (r_i) is, after having ordered bids as previously, to shrink or stretch the number of such ratios keeping constant the lowest and the highest of them; that is, the highest ratio corresponding to the lowest bid, and vice versa. In the P-to-K transfer, one needs to stretch the number of ratios from 27 to 44, while in the reverse case one needs to shrink their number from 44 to 27. This is quite straightforward, provided one keeps intervals between each ratio equally spaced throughout, between the upper and lower bound. In the P-to-K prediction, destination costs can then be directly computed using the formula $c_k = b_k \times (1 - r_p)$, where the r_p are the ratios imported from the P data.

The error in the regression cost coefficient, defined as a deviation from the perfect value of 1.00, was −10% in the P-to-K prediction and −2% in the K-to-P prediction (see Table 1). The constant term in the P-to-K prediction was off by +12% (t-value = 3.91) in terms of the min–max cost range of 260, while, as in the best-fit approach, its estimated value was not statistically significant in the K-to-P prediction.

Prediction of the auction's performance in terms of cost-efficiency (total payment / total cost) was off by –9% in the P-to-K and by +5% in the K-to-P. With outliers excluded (two in both cases), the first estimate improved to –3% but the second deteriorated (as in the best-fit estimate) to +16%. These results are all summarised in Table 1 below.

Table 1

	Error in cost estimates		Error in Cost-Efficiency** (payment / cost)	
Cost estimation method	**Cost coefficient (relative to 1)**	**Constant term (relative to 0)**	**Outliers included**	**Outliers excluded**
Theoretical model (L&H)				
K (n = 44)	**– 13%** (– 5%) *	**+ 19%**	**–26%**	**–26%**
P (n = 27)	**+ 6%** (+ 8%)	**(NS)**	**+8%**	**+17%**
Statistical residuals (W&B)				
K (n = 44)	**– 27%** (– 18%)	**+ 28%**	**–33%**	**–29%**
P (n = 27)	**– 17%** (– 12%)	**+ 12%**	**–15%**	**–8%**
Experimental transfer - Best fit				
P to K fit	**– 15%** (– 4%)	**+ 18%**	**–20%**	**–15%**
K to P fit	**– 17%** (– 12%)	**(NS)**	**+5%**	**+16%**
Experimental transfer - Direct				
P to K (stretch)	**– 10%** (+ 2%)	**+ 12%**	**–9%**	**–3%**
K to P (shrink)	**– 2%** (+ 4%)	**(NS)**	**+5%**	**+16%**
	All p < 0.001	All sig. p < 0.05		

* Numbers between ( ) indicate results that exclude from the data clear outliers.

** Results show errors in predicting auction's cost-efficiency (total payment / total cost), defined as the difference between CE with estimated costs and CE with actual experimental costs.

Table 1 summarizes and compares the results obtained using the three cost estimation methods, for both the underlying cost functions and for the auction's cost-efficiency performance. On both accounts, it appears that the experimental transfer method performs best, in the sense that the errors committed are the smallest. This does not imply, of course, that this approach should always perform best: more studies of this kind will need to be carried out before one can make such a statement. However, given the information currently available, the experimental transfer approach appears to be at least as acceptable as any other. We shall therefore choose this method for evaluating the Scottish government's (SEERAD) fishing vessel decommissioning auctions of 2001 and 2003.

In addition, we can be more specific by noting that the P to K transfer is more appropriate for the field data investigation, given that the SEERAD auctions had a greater number of bidders than either of the two experiments; therefore, the number of bid-shading ratios needs to be 'stretched' rather than 'shrunk', which is what the P(n=27) to K(n=44) transfer does. The corresponding results, shown by the shaded values, compare well relatively to the other methods.

Results: Evaluating the Scottish fishing vessel decommissioning auctions

The previous section aimed at identifying and testing three methods for auction evaluation. Given the data at our disposal, it turned out that the direct experimental transfer method performed best when put to the test. We shall therefore choose this method to evaluate two real auctions that were carried out in Scotland for the purposes of decommissioning fishing vessels and reduce pressure on fisheries. These two auctions involved, in 2001 198 bidders and in 2003 159 bidders. They were not identical in that the first targeted fishing capacity in terms of vessel volume and power, whereas the second targeted fishing activity, especially cod landings. Both auctions were of the sealed-bid, discriminatory price format, with no announced reserve price. Although there might have been some transmission of information from the first to the second, the two auctions can be treated as separate, rather than as a repeat. As it happened, the then SEERAD agency of the Scottish government was interested in evaluating the performance of these auctions, but was unsure of to how to do it. Preliminary analysis by the author suggested that the two auctions showed contrasting performance outcomes and this led to a deeper analysis, from this paper emerged.

Evaluation using best method

Given the results of the previous section, we apply the direct experimental transfer approach to the field data of the two Scottish SEERAD auctions. It must be kept in mind that doing so means transferring the bid-shading ratios from the experimental setting to the field setting. This in turn assumes that the experimental behaviour of University students can reliably be extrapolated to the field behaviour of professional fishermen. Such external validity of experimental data is by no means guaranteed; however, we can take it as an acceptable working hypothesis given the identity of incentive structures between the two settings mentioned earlier and the performance of this approach in the previous section.

The 71 experimental data points were evenly 'stretched' to 198 in the 2001 SEERAD auction and to 159 in the 2003 SEERAD auction. Then, knowing the bids and the rent/bid ratios, defined as [B–C]/B, one can directly compute the underlying costs. The resulting cost curves are given in the two panels of Figure 5. In spite of appearances, the rent/bid ratios imported from the experiments were the same in both the 2001 and 2003 auctions: they averaged 31% over all bidders.

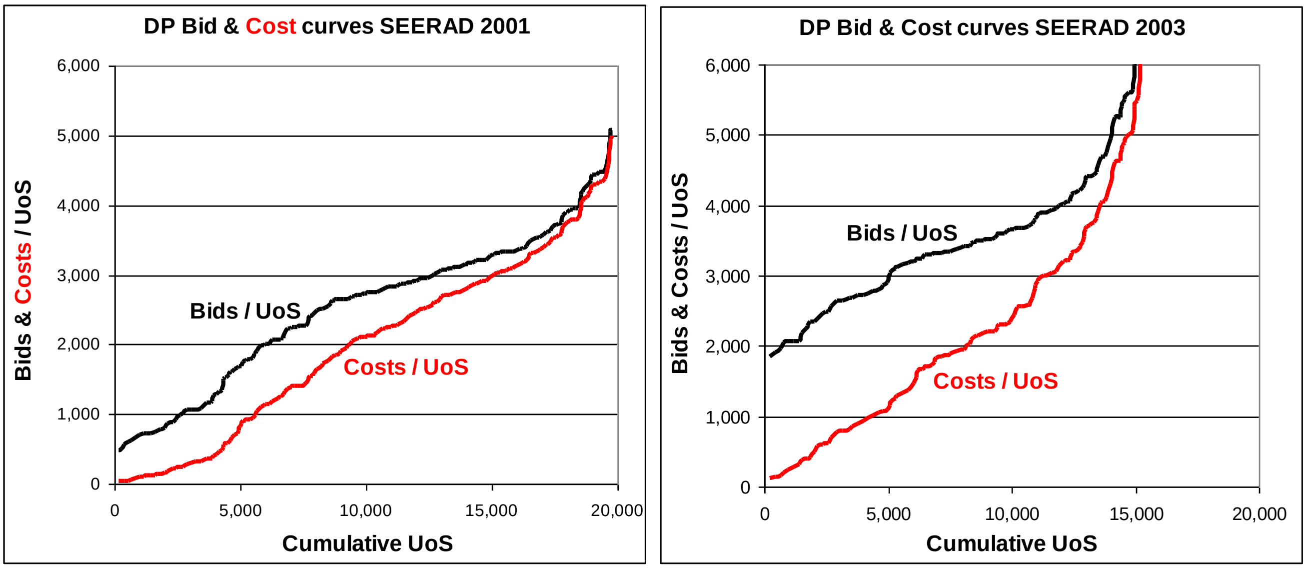


Figure 5

UoS = Units of Service = a measure of the quantity of fishing capacity put up for decommissioning

As detailed in Schilizzi and Latacz-Lohmann (2007), evaluating an auction's performance requires clear *evaluation criteria*. These typically include:

- the ratio of government payments to fishing capacity bought out (cost-effectiveness);
- the ratio of total bidders' rents over government payments (rent ratio); and
- the ratio of government payments to total bidders' costs, which we can call 'cost-efficiency'.

These ratios hold only over those bidders who have been selected into the scheme. In 2001, 129 out of 198 bidders, or 65% were selected (i.e. where awarded a vessel decommissioning contract), and in 2003, 69 out of 159 bidders, or 43%, were selected. However, the quantity of interest to the Scottish government was not the number of bidders per se, but the quantity of fishing capacity that was decommissioned. Measured in these terms, 65% of the capacity on offer was bought out in 2001 and 66% in 2003. The discrepancy in 2003 reflects the fact that a smaller number but larger-sized boats were decommissioned relative to 2001.

Evaluating the performance of a field auction also requires, as argued in Schilizzi and Latacz-Lohmann (2007), an *evaluation benchmark*. Basically, one is interested in whether it performs better or worse than an alternative policy instrument. Both 2001 and 2003 auctions were sealed-bid, discriminatory price, or pay-as-you-bid, auctions. We shall study their performance relative to two benchmarks, an equivalent uniform price auction and an equivalent fixed price scheme, in terms of their cost-efficiency as defined above.

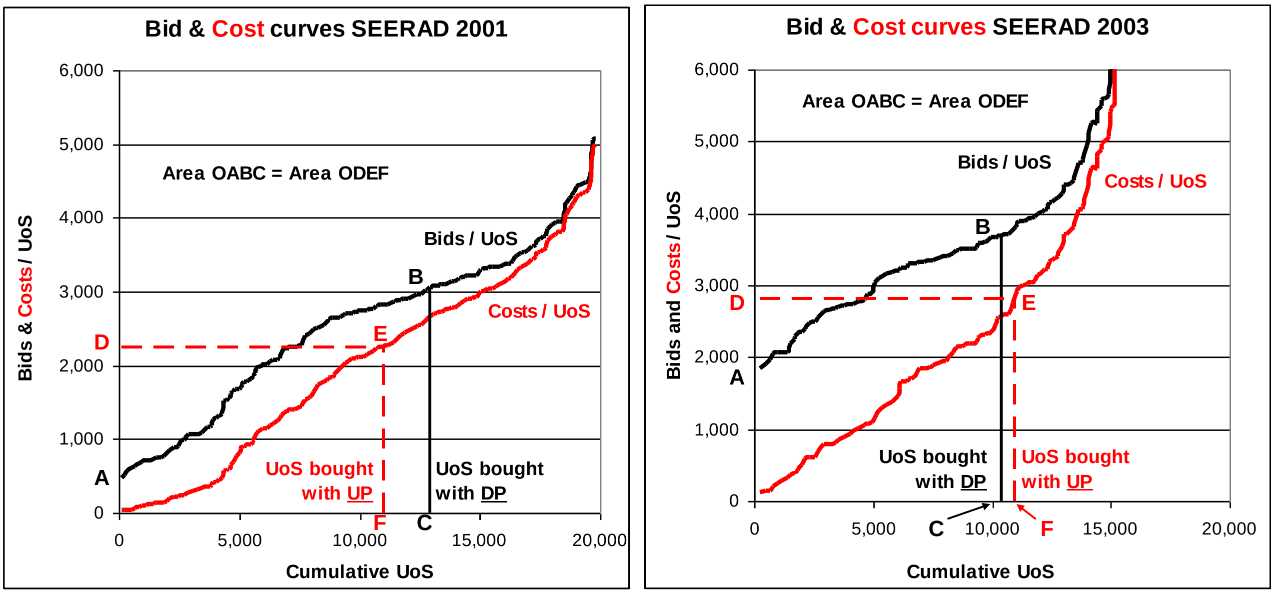
A uniform price (UP) auction is the multi-unit extension of the second-price Vickrey auction (Vickrey, 1961). It differs from the discriminatory price (DP) auction by the way the procurement authority pays the selected bidders. If a bidder is selected in a DP auction, he gets paid what he bid for: it is the 'pay-as-you-bid" rule. In a UP auction, he gets paid the amount bid by the first rejected bidder, which, in a procurement auction, corresponds to an amount always greater than his own bid. Although bidders in a UP auction get paid more than their bid, it was shown by Vickrey that this auction format is incentive-compatible, in that bidders have an incentive to bid their true cost. By contrast, in a DP auction, bidders have an incentive to bid over and above their true cost (Cox *et al.*, 1982 and 1984). These two incentives act in a countervailing way, making the overall performance of each auction format theoretically indeterminate and subject to empirical specifics.

A UP auction *equivalent* to a DP auction is a UP auction which, in our case, is run with an equal budget to that of the DP. (In target-constrained auctions, it would be run with an identical target). The DP auction will perform better than the equivalent UP if, using the cost-effectiveness criterion, it can buy out with the same budget more fishing capacity than the DP.

Similarly, a fixed price scheme (FPS) is *equivalent* to a DP auction if it corresponds to the minimum uniform payment rate that would have resulted in the same budgetary expenditure as the DP auction. (In a target constrained auction, it would be the minimum uniform payment rate that would have been needed to achieve the same outcome, in this case, buy out the same quantity of fishing capacity, as the DP auction.) In actual practice, this specific FPS is not feasible given that it would require perfect knowledge of individual bidder costs. A more realistic implementation of this idea is to use the estimated average cost of eligible participants as a base for choosing the fixed payment rate. When using a FPS, the participants are selected not through an auction mechanism, but using some automatic selection rule; typically, a random first-come, first served rule where, *a priori*, each participant has an equal probability of obtaining a contract.

Evaluating performance using an alternative **auction** *as benchmark*

Figure 6 shows the performance of the 2001 and 2003 DP auctions relative to their equivalent UP auction, based on our estimates of the cost curves derived in the previous section. The result is quite interesting in that the two auctions appear to perform in an opposite manner, when using this benchmark. The 2001 DP auction performs better in terms of cost-effectiveness than its UP equivalent in that, for the same budget, the former buys out more fishing capacity than the latter by an estimated amount of 17% (about 13,000 capacity units bought out instead of 11,000, given a budget of £25 million). In Figure 6, point C is to the right of point F.



UoS = Units of Service (a measure of the fishing capacity offered up by bidders for decommissioning)

Figure 6

The auction also performs better than its equivalent UP in terms of the other two criteria: its rent ratio is 37% instead of UP's 55% and its cost-efficiency (total payment / total cost) is 1.57 instead of 2.23. By contrast, the 2003 auction performs worse than its equivalent UP, albeit only slightly. It buys out 4% less fishing capacity (this time point C is to the left of point F), its rent ratio is 56% instead of 52% (8% worse), and its cost-efficiency ratio is 2.27 instead of 2.07 (9% worse).

These results are summarised in column 3 of Tables 2 a and b below.

Table 2-a

Performance criteria	2001 (DP) auction	Equivalent 2001 UP auction	Equivalent 2001 FPS
% FC decommis.	65%	56%	80%
£ / unit of FC	1,928	2,253 (+17%)	1,572 (–18%)
Rent ratio (%)	37%	55%	40%
Payt/costs ratio	1.57	2.23	1.00*

FC = Fishing Capacity
* by construction

Table 2-b

Performance criteria	2003 (DP) auction	Equivalent 2003 UP auction	Equivalent 2003 FPS
% FC decommis.	66%	69%	77%
£ / unit of FC	2,935	2,815 (–4%)	2,019 (–31%)
Rent ratio (%)	56%	52%	72%
Payt/costs ratio	2.27	2.07	1.00*

FC = Fishing Capacity
* by construction

Evaluating performance using an alternative policy instrument *as benchmark*

In comparing the SEERAD auctions to their fixed-price equivalents, one also needs to use the estimated cost curves derived previously. The average estimated cost for the 2001 auction was £197,701 per vessel (corresponding to £2,025 per unit of fishing capacity or UoS). If all eligible vessels were paid this amount, only 50% of their owners would volunteer to give up their vessel, which corresponds to a self-selection rate of 50%. The other 50% have costs that exceed this average payment. However, this fixed payment, given the acceptance rate of 50%, would only use up 33% of the available budget (£8.34m). By assumption, the policy maker has decided to spend the whole budget of £25m. To spend it all, an average fixed payment of £326,207 per vessel would need to be paid out (£3,341 per unit of fishing capacity), and 82% of eligible vessels (163 out of the 198) or 80% of fishing capacity would then be offered up for decommissioning.

With this fixed payment, the overall cost-effectiveness of the scheme would appear as total budget outlay per unit decommissioned: £25m for 163 vessels, that is £153,174 per vessel, or £1,572 per unit of fishing capacity decommissioned. This compares favourably to the £1,928 of the (DP) auction, being 18% cheaper. It also compares favourably in terms of percentage of eligible capacity decommissioned: 80% bought out instead of 65% for the auction. In terms of the proportion of rent in the total payment, however, it performs slightly worse than the auction: 40% compared to 37%.

The 2003 auction's FPS equivalent leads to a slightly greater percentage of vessels being decommissioned, namely 86%, or 77% of fishing capacity, which compares to the auction's 66%. The FPS scheme's overall cost-effectiveness runs at £2,019 per unit of fishing capacity decommissioned, which is about 30% cheaper than the auction's £2,935 per unit. The rent component of the government's layout appears as a much worse 72% compared to the auction's 56%.

These FPS results are summarized in column 4 of Tables 2 a and b above.

These results would not have been possible without the prior estimation of the cost curves underlying the auction bids. Without knowledge of these costs in discriminatory-price auctions, one would be reduced to making heroic assumptions similar to those made in Stoneham *et al*.

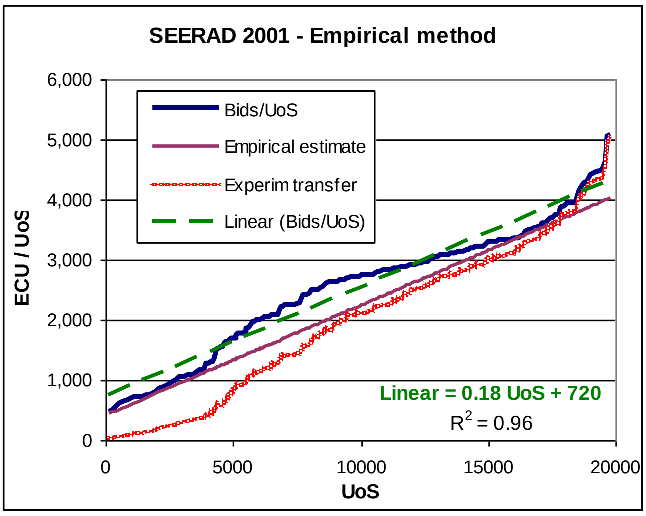
(2003). These authors simply assumed the cost curves were represented by the bid curves. The results obtained above demonstrate the extent of the errors that can be introduced by making this assumption.

Speaking of errors, it may be of some interest to see how large an error would be committed if a method other than the best one identified above was used instead. This is the object of the next section.

Evaluation error when using methods other than the best

Given the nature of the LH model described earlier, one cannot readily apply this approach to the SEERAD auctions: neither the underlying costs nor the bid caps are known. This leaves us with the empirical method.

The idea here is to apply the empirical method used by White & Burton (2010) to the two SEERAD auctions and compare the results obtained with the experimental transfer method, which the testing on experimental data has identified as the best benchmark at our disposal in this case. In Figure 7, the resulting cost curves for the two SEERAD auctions are shown as the continuous straight line ("empirical estimate") under the varying bid curve. Just by eye-balling Figure 7, it is clear that the error committed is much greater in the 2003 auction than in the 2001: the overall distance between the experimental and empirical cost curves is much greater in the latter. More specifically, in 2001, costs estimated using the empirical method totalled £20,693,483 while, using the benchmark method they totalled £15,815,256; an error of +31%. In the 2003 auction, the corresponding figures are £28,540,996 and £13,490,492, leading to an error of 112% – clearly not a negligible one.



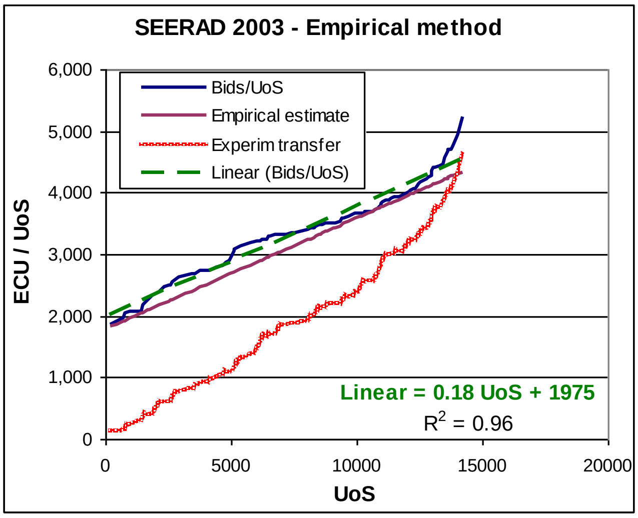


Figure 7

This translates into the following errors when evaluating each auction's performance in terms of cost-efficiency (total payment / total cost): in 2001, the CE estimated empirically is 1.27 while the benchmark estimate is 1.57, reflecting an error of –24%; in 2003, the corresponding figures are 1.07 and 2.27, reflecting an error of –53%. With such errors, the auction's evaluation could be very misleading for decision-makers.

The above results require an obvious caveat: the errors computed are not absolute, but relative. They depend on the choice of the benchmark evaluation method, which must be chosen as the one having generated the smallest errors when tested with controlled experimental data. In this case, it was the direct experimental transfer. But in other types of situations, it could be another method that performs best. It is too early now to make any general statement as to which method should most likely perform best in general.

Conclusions

The purpose of this paper was to compare three possible methods for evaluating the performance of conservation tenders and, more generally, for procurement auctions. To be able to do this when using mechanisms based on the 'pay-as-you-bid' rule, which involve a certain degree of bid-shading, knowledge of the underlying bidder costs is essential. Of course, direct knowledge is not only infeasible but also, in a sense, necessary; otherwise there would be no reason for carrying out an auction. If, however, the auction is based on the uniform-pricing principle, no such knowledge of underlying costs is, in principle, needed, since such auctions have been shown to be incentive-compatible. Given the necessity of estimating cost curves for the pay-as-you-bid auctions, this paper set out to investigate how this was possible and to what effect.

We identified three main methods: a theoretical method based on optimal bids computed with an auction model; an empirical method based on a direct use of bid data as generated by a field auction; and an experimental method based on transferring bid-shading behaviour from lab data to field data, so as to evaluate a field auction. The third approach is quite innovative and rests on the principle of an isomorphic experimental setup: the incentive structure in the lab must be identical to that in the field for average lab behaviour to be extrapolated to field behaviour. Even so, because the set of real or perceived constraints may differ, such external validity is by no means guaranteed.

We were mindful of this limitation when comparing the performance of each of the three methods. We submitted each one to a test using controlled lab data and, in particular, used different experiments in different countries for testing, to some extent, the method's external validity. In the present case, the experimental method based on a direct transfer of bid-shading ratios performed best when tested with different experimental data. It was therefore chosen to estimate the costs of two Scottish fishing vessel decommissioning auctions.

We were able to estimate their underlying cost curves and evaluate each auction's performance. In addition, we investigated how large an error would be made if another, less performing method was used. If the empirical method had been used, the error in auction performance exceeded 50%, signalling the potential for seriously misleading evaluations. Obvious policy consequences follow.

To sum, this paper has tried to highlight the importance of three aspects of auction performance evaluation and comparison:

- the use of a reliable method for estimating bidder costs
- the use of a reliable strategy for evaluating and comparing cost-estimation methods
- the notion of an isomorphic experimental setup
- a strategy for evaluating, to some extent, the external validity of experimental data

The problem investigated in this paper is one of evaluation ex-post: one wishes to evaluate the performance of an auction that has already taken place and for which the bids are already in. The purpose may be to find out whether an alternative policy instrument might have yielded better outcomes. A radically different approach is needed if the aim is to evaluate the performance of an auction ex-ante, before it has been held, for example to decide whether to go ahead and run it or not. This is the topic of a forthcoming paper.

References

Berry, S.T., Levinsohn, J.L. and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*: 63(4), 841-890.

Bresnahan, T. (1987). Competition and Collusion in the American Automobile Market: The 1955 Price War. *Journal of Industrial Economics*: 45(4): 457-482.

Cantillon, E. and Pesendorfer, M. (2005). Auctioning Bus Routes: The London Experience. Chapter 22, Section 5 (A method to infer costs from auction data) in: Ausubel, Cramton and Steinberg: *Combinatorial Auctions*, MIT Press: Boston, MA.

Cox, J., B. Roberson, and J. Walker (1982). Theory and Behavior of Single Object Auctions. In V.L. Smith (ed.)*Research in Experimental Economics*. Greenwich, Conn.: JAI Press.

Cox, J., V. Smith, and J. Walker (1984). Theory and Behavior of Multiple Unit Discriminative Auctions. *Journal of Finance* 39(4): 983–1010.

Goldberg, P.K. (1995). Product Differentiation & Oligopoly in International Markets: The Case of the U.S. Automobile Industry. *Econometrica*: 63(4): 891-952.

Jack, B.K, Leimona, B. and Ferraro, P.J. 2009. "Revealed Preference Approach to Estimating Supply Curves for Ecosystem Services: Use of Auctions to Set Payments for Soil Erosion Control in Indonesia." *Conservation Biology* 23(2): 359-367.

Jack, B. K. (2011). Designing Markets for Carbon Offsets: A Field Experiment in Malawi (January 30, 2011). Available at SSRN: http://ssrn.com/abstract=1638563

Krasnokutskaya, E. and Seim, K. (2008). Bid Preference Programs and Participation in Highway Procurement Auctions. Mimeo: Department of Economics and Wharton School, University of Pennsylvania, Philadelphia, PA 19104. November 2008, 44 p.

Latacz-Lohmann, U. and van der Hamsvoort, C. (1997). Auctioning conservation contracts: a theoretical analysis and an application. *American Journal of Agricultural Economics*, 79: 407-418.

Porter, R. H. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880-1886. *Bell Journal of Economics*, 14(2): 301-314.

Reichelderfer, K. and Boggess, W. G. (1988). Government decision-making and program performance - the case of the conservation reserve program. *American Journal of Agricultural Economics* 70, 1-11.

Rosse, J. N. (1970). Estimating Cost Function Parameters without Using Cost Data: Illustrated Methodology. *Econometrica*, 38(2): 256-275.

Schilizzi, S. and Latacz-Lohmann, U. (2007). Assessing the Performance of Conservation Auctions: An Experimental Study. *Land Economics*, 83 (4): 497–515.

Stoneham G., Chaudri, V., Ha, A. and Strappazon, L. 2003. Auctions for conservation contracts: an empirical examination of Victoria's BushTender trial. *Australian Journal of Agricultural and Resource Economics* 47 (4): 477-500.

Vickrey, W. (1961). Counterspeculation, Auctions and Competitive Sealed Tenders. *Journal of Finance* 16(1): 8–37.

White, B. and Burton, M. (2010). Measuring the Cost-effectiveness of Conservation Auctions Relative to Alternate Policy Mechanisms. Working Paper 1006, School of Agricultural and Resource Economics, University of Western Australia, Crawley, Australia. 19 p.

Wolak, F.A. (2002). Identification and Estimation of Cost Functions Using Observed Bid Data: An Application to Electricity Markets. Mimeo: Department of Economics, Stanford University, Stanford, CA 94305-6072. January 2002, 62p.