Empirical methods for determining a reserve price in conservation auctions

Mark Eigenraam, Joselíto J. Chua, Claire Edwards

Department of Sustainability and Environment
PO Box 500, East Melbourne, Victoria, 3002

Abstract

Conservation auctions are increasingly being used to procure public environmental goods on private land. In the absence of demand-side price information, the majority of conservation auctions in Australia have been designed without a reserve price. In these instances bids have been accepted in order of cost-effectiveness until the budget constraint binds. It is widely recognised that in situations where auctions are run repeatedly a reserve price strategy could allow for a more efficient allocation of funds across multiple rounds, both spatially and temporally.

This paper provides a brief overview of methods for determining a reserve price for application in conservation auctions. It is concluded that information deficiencies and the high transaction costs involved in the application of these methods to conservation auctions often render them unsuitable for application to real-world auctions.

This paper presents an empirical approach to determining a reserve price using data obtained during an auction - the supply curve. The approach stems from the C4.5 algorithm, developed in the field of data mining to construct decision trees from training data using the concept of information entropy. The algorithm establishes a reserve price by determining the cut-off price that results in the "best fit" of two normal distributions to the frequency distribution of bid-price per unit environmental benefit.

Empirical data from conservation auctions in Victoria is used to demonstrate the algorithm and compare auction results obtained using the algorithm and traditional "budget" methods. The paper presents a discussion on the situations where the algorithm could be appropriately used, and advantages and limitations of the approach are identified. The paper concludes that the use of the algorithm can result in efficiency gains over the traditional budget method in situations where alternative reserve price strategies are impractical.

1 Introduction

It is widely recognised that there is a need to invest in the maintenance and improvement of ecosystem health in Australia. As many priority ecosystems exist solely, or predominantly, on land that is privately owned (Stoneham et al, 2000; DSE, 2008), government agencies are interested in investing in the maintenance and improvement of ecosystems on private land.

Auctions for conservation are a favoured policy instrument for targeting environmental improvements on private land. Their argued benefits include cost revelation incentives provided by competitive bidding (Stoneham et al, 2003; Conner et al, 2008) and the use of a scientifically based scoring system to identify high value projects (Connor et al, 2008).

Historically, most conservation auctions have been implemented in a one shot setting without a reserve price (Stoneham et al, 2002; Connor et al, 2008; Windle et al, 2009). Typically both a budget and the auction area are determined *ex anti.* Bids are accepted in decreasing order of cost effectiveness until the budget is exhausted.

Stoneham et al, 2003, argued that while a reserve price is less important in a budget constrained single shot auction, it would become more important in a repeated auction setting where funds can be allocated across multiple rounds. This notion can be extended to the allocation of funds across multiple tenders in different locations (within the same time period), as well as allocating funds between auctions and alternative policy mechanisms (e.g. fixed price grants schemes).

Another issue arises in the absence of a reserve price when there is inadequate competition (bids) to reach the *ex ante* budget allocation. If an auction area is chosen that does not result in sufficient bids to expend the whole budget should all bids be accepted? It is likely that some bidders would have attempted to game the auction and bid high in the hope of extracting greater rents. Without a binding budget constraint all the high-cost bids will be successful impacting negatively on cost effectiveness. Alternatively, there may be sufficient bids however it is evident from the bid data that some of the bids accepted are very high unit cost when compared to others selected. In these instances bid selection panels are often required to make a subjective decision on which bids to fund. There is no information (apart from intuition) to aid bid selection panels with this decision.

This paper presents an empirical method to estimate an *ex post* reserve price using the unit bid data[1] from the auction. The approach is based on the theoretical framework applied in the C4.5 algorithm (Quinlan, 1993) which is used widely in data mining. In C4.5, the concept of information gain is used as the partitioning criterion for constructing decision trees. The algorithm applies the same concept to find a threshold to partition the unit bid data into two sets. The algorithm utilises information gain as the criterion for partitioning the unit bid data set into high cost and low cost suppliers of environmental benefit. The threshold is used as the reserve price. In this paper the application of the bid

[1] Unit bid data refers to the bid price per unit of environmental benefit.

cut-off algorithm to conservation auctions is demonstrated using empirical data from past EcoTender auctions run in Victoria, Australia.

2 Conservation auctions on private land

Currently 63 percent of Victorian land is privately owned. When private landholders make land management decisions, they face strong economic incentives to generate private benefits. However, it is often the case that landholders are not provided with sufficient incentives to produce public environmental benefits. This market failure results in under-investment in environmental outcomes on private land (Rolfe, 2002).

A range of policy mechanisms have been implemented to provide incentives for the production of public environmental benefit on private land. Historically, fixed price grants schemes have been the primary mechanism used for purchasing environmental outcomes from private landholders (Latacz-Lohmann & Hodge, 2003). In a fixed price grant scheme, a landholder is paid a set price based on input units (for example, $20.00 per kilometre of fence) for the management actions they undertake. Latacz-Lohmann and Hamsvoort (1997) argue that an information asymmetry exists between the landholder and the governing agency. The landholder has more accurate information about their costs, whilst the agency has better information about the environmental benefit obtained. While grant mechanisms provide incentives for landholders to engage in the production of public environmental outcomes, fixed price input based schemes do not address problems arising from asymmetric information.

More recently, there has been a transition towards market based approaches, such as conservation auctions, in an effort to procure environmental outcomes cost-effectively (Windle & Rolfe, 2007). In a conservation auction, landholders competitively bid to be paid to produce environmental outcomes on their land. Generally a scoring system, or metric, is used to put a value on the predicted environmental outcomes from different bids. Bids can then be ranked from best to worst value for money and selected down the list until the budget is exhausted, or an alternative cut-off point is reached.

It is argued in (Connor, Ward, & Bryan, 2008) that the competitive bidding process addresses asymmetric information by providing incentives for bidders to reveal information about their costs. Moreover, the use of a scientific metric to measure the value of environmental outcomes across competing projects allows for the identification of high value projects (Connor et al., 2008). Cost revelation incentives provided by competitive bidding, coupled with the use of scientific metrics to value environmental outcomes, allows for the identification of low cost suppliers of environmental outcomes.

The Victorian Governments BushTender auction for terrestrial biodiversity outcomes (Stoneham, Chaudhri, Ha, & Strappazzon, 2003) and EcoTender auction for multiple environmental outcomes (Eigenraam, Strappazzon, Lansdell, Beverly, & Stoneham, 2007) are examples of conservation auctions in Victoria. Other conservation auctions that have been implemented in Australia include the Onkaparinga Catchment Care Auction (Connor et al., 2008) and the

Southern Desert Uplands Landscape Linkage Auction (Windle & Rolfe, 2007). A well known Conservation Auction outside of Australia is the US Conservation Reserve where private land-holders competitively tender for payments to remove their land from agricultural production (Vukina, Zheng, Marra, & Levy, 2008).

3 Supply and demand in conservation auctions

In conservation auctions information about land-holders costs is elicited through a competitive bidding process. The output of the auction process is a supply curve depicting the marginal cost to the agency of acquiring environmental benefits, defined by an environmental metric. The nature of the supply curve will be influenced by the auction design as well as the distribution of land-holder bids (costs). In a discriminatory price auction, the optimal bidding strategy involves seeking rents, where the quantity of rents sought is a decreasing function of both the landholders ranking in the distribution of costs (ascending order), and the number of bidders. In a uniform price conservation auction where a landholder can only submit a single bid, the optimal bidding strategy for landholders is to bid their cost. An example of a conservation auction supply curve is given in Figure 1 below.

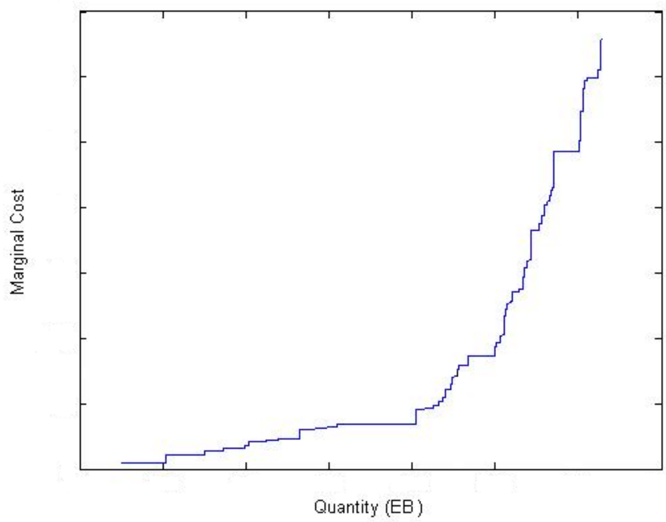

Figure 1: Typical supply curve from conservation auction.

While competitive auctions facilitate the revelation of supply side information, demand side information is much harder to obtain. Ideally, when a government agency runs a conservation auction, it would have a demand schedule articulating the quantity of environmental benefits they wish to purchase for any given price. Assuming that the agencies demand schedule is monotonically decreasing in cost, the optimal quantity and marginal price will be given by the intersection of the supply curve (above) and the demand curve of the agency. Figure 2 below shows the optimal quantity and marginal price given the availability of an agency demand schedule.

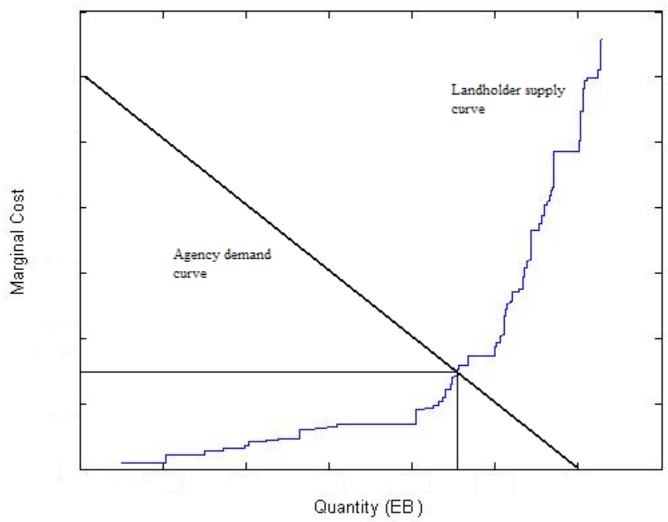

Figure 2: Optimal quantity of EBI with known demand schedule.

However, information on the agencies demand schedule is often unknown. An agencies demand for a marginal unit of environmental benefit is dependent on the opportunity cost of investing in that unit. That is, the benefit that is forgone by not investing in the highest value alternative.

For simplicity, let us assume that the government agency has only two options; invest in a marginal unit of environmental benefit in the current tender, or save the funds to invest in a future tender. Consider an agency that has a priori information on the expected shape of the supply curve in a future tender. Assume that the agency is indifferent between allocating funds between the present and future tender. In this case, the agency could allocate funds between the two tenders such that the marginal cost of purchasing the final environmental benefit unit in the current tender is equal to the expected marginal cost of procuring the final unit of environmental benefit in the future tender as shown in Figure 3 below.

In many instances it is difficult to accurately predict the cost of environmental benefits in future tenders. Further, the costs may be different if the future tender is held in a different area. In the new area landholder costs may be higher or lower depending on the opportunity cost of the land. For example, land in a sheep grazing area generally has a lower opportunity cost than dairy production land. If the tender is to be held in the same area, costs may be higher in the second tender due to low cost suppliers being funded in the prior tender.

In addition, communities may have preferences that place a higher value on environmental benefits obtained geographically closer to them; it can be argued that government has a responsibility to allocate funds in a manner that accounts for these preferences. This notion is supported in (Hajkowicz, 2007)

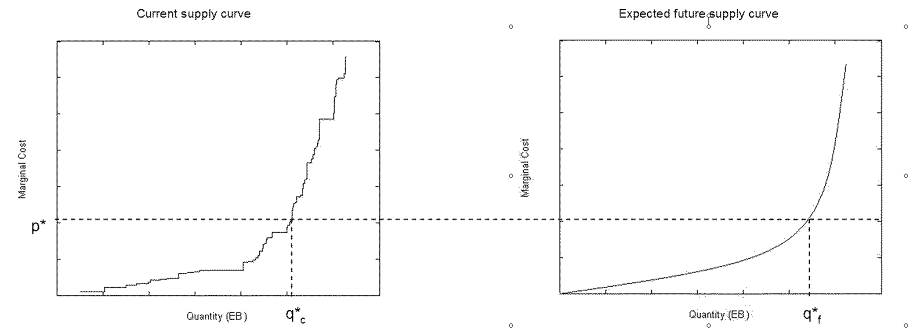

Figure 3: Optimal quantity of EBI purchased with known current and expected future supply curves.

where it is found that fairness is a strong policy objective in the allocation of agency funds to regions for environmental management. Consequently, it can be contended that the role of government in allocating funds to manage natural resources extends beyond allocating funds to areas where environmental goods and services can be procured at least cost.

4 Reserve prices and conservation auctions

In theory, a conservation auction can be either budget or target constrained. In the former, the environmental benefit is maximised within a maximum budget constraint. In the latter, the budget is minimised within a minimum environmental benefit constraint. Both of these optimisation problems have the potential to result in the selection of some bids with a high unit price. This can be inefficient if the marginal cost of purchasing these bids is greater than the marginal cost of purchasing the same level of benefit through an alternative means. The use of an appropriately designed reserve price can improve the efficiency of both budget and target constrained tenders by placing an upper bound on the marginal cost of environmental benefit.

The majority of conservation auctions in Australia have been implemented with no reserve price (Stoneham et al., 2003; Windle & Rolfe, 2007; Connor et al., 2008). The use of a reserve price was discussed in (Stoneham et al, 2003) where it was identified that a reserve price strategy would become more important in a repeated auction setting to optimally allocate funds across multiple auctions both temporally and spatially.

The US Conservation Reserve Program (CRP) is an example of a conservation auction where a reserve price strategy has been implemented. In the CRP, the reserve price was set to the average land rental price (by county and soil type) scaled by area and contract duration (Vukina et al., 2008). In another example, in the Victorian Stormwater Tender [2] , a reserve price was determined

[2]In Stormater Tender, landholders competitively tendered to undertake actions to reduce stormwater runoff from their property

ex post by equating the expected marginal costs of public and private supply of environmental benefit (Nemes et al., 2010). The public supply curve was obtained from data on the costs and environmental benefits obtained from public programs that had been collected previously (Nemes et al., 2010). This method relied upon the existence of information on costs and environmental benefit which may not be available in many instances. Further, most non market-based programs do not employ a metric when evaluating alternative investment options. Therefore it is not possible to calculate a unit cost for use as a reserve price.

This concept of equating marginal costs across programs can be applied if both programs have a comparable metric. For example if a tender and traditional grants scheme are run simultaneously, a budget could be optimally allocated across the two programs by equating marginal costs. This method has the potential to yield efficiency gains over running the two programs separately. However, it is likely to be the case that when the administrative costs of implementing both programs are considered, it is more efficient to allocate the total budget to one program.

5 Motivation for an empirical estimate

In 2007 the Victorian government committed $4.5m to three EcoTenders without specifying the breakdown of funding between the tenders. The development of the bid-threshold algorithm was motivated by the desire to cost effectively allocate funds across the three tenders. In particular, the development of the algorithm was motivated by the need to develop an alternative to funding all projects in a tender that is under-subscribed.

It was also recognised that it was not possible to use results from one tender to assist in setting a reserve price in subsequent tenders due to the heterogeneity in opportunity costs across regions. In addition, over the three years of the EcoTender trials, scientific metrics would be updated to reflect improved measurement and modelling capabilities. Therefore the environmental benefits would not be directly comparable across the tenders, adding to the difficulty of using information from prior tenders in setting a reserve price.

In response to these issues it was identified that an *ex post* reserve price strategy was required with the following properties. The strategy:
1. relies solely on data obtained within the tender;
2. is replicable and transparent; and
3. selects a cut-off that has merit from a theoretical point of view.

An overview of landholder bid signalling and the application of information gain is provided in Section 6 below. Section 6.1 provides a brief background on the information-theoretic principle behind the algorithm. Section 6.2 describes how the algorithm is applied to situations where there is no sufficient prior knowledge about the reserve price. We shall also describe how the algorithm would behave in exceptional circumstances where more complex models do not provide greater information than the simpler model. Section 6.3 describes how the algorithm can be extended to “soft” thresholds that take uncertainty into account, and to situations where there is prior information on the reserve price from previous tenders.

6 Information Gain

In this section we describe an algorithm for estimating an *ex post* reserve price in a conservation tender. The unit bid prices received are interpreted as signals about the unit payment landholders require in order to provide the EBIs – the landholders' "willingness to accept". Without a reserve price it is not possible to identify which of the signals (unit bids) come from low-cost and high-cost sources.

The aim of the algorithm is to estimate a threshold from the signal using a concept known as "information gain". The threshold is then used as the *ex post* reserve price and all bids below the threshold are accepted. The bid prices are a sample of the signals from a population of landholders.

In order to apply the information gain concept, a model that best reflects the signals must be chosen. In this case, we assume that the unit cost of providing EBIs is a random variable with a normal probability distribution. We can use that model to describe the sample. For instance, we can use the mean to partition the sample into low-cost and high-cost bids. However, the mean is not necessarily the best statistic to partition the sample. For example, outliers can dramatically influence the mean value. Rather than using an arbitrary statistic (such as the mean) we can use to the concept of information gain to make better use of the information contained in the unit bids.

The "source" of the signal is an important concept used in information theory. In this case, the source is a "typical" landholder which is modelled as a random variable with a probability distribution. The signal is the unit bid, which is the realisation of the random variable. The challenge is to model the source (i.e. model a "typical" landholder). A simple model is to assume that the random is normally distributed. In this case, the expected value of the random variable is the mean of the normal distribution.

Information gain requires an alternative model for the source. Consider a model where there are two types of "typical"landholders. For example, assume that landholders may be categorised as sheep farmers or cattle farmers, and that cattle farmers always have higher opportunity costs than sheep farmers. If an agency knows that a bid is coming from a specific type of farmer, then they can more accurately estimate the bid price. Knowing the type of farmer leads to an information gain about the unit bid price. In reality, there could be several types of landholders, and the distributions might overlap between different types. For the purpose of determining a reserve price, the only relevant types of sources are typical low-cost and typical high-cost bidders. The low-cost and high-cost bidders are separated by a threshold into two mutually-exclusive groups.

A model with two sources is considered to be more complex than a model with a single source. Typically, we would expect to gain more information about the signals from more complex models than from simpler models. The threshold where the alternative model leads to the greatest amount of information gain (compared to the simple model) forms the reserve price. This threshold separates the bids into two groups such that knowing which group a bid belongs to provides the greatest information of a bid's location in the modelled distribution of bids.

The use of information gain to determine a reserve price is motivated by the need to find an empirical model that separates the low-cost and high-cost bid prices using only the bid data. If information on opportunity costs is available, traditional economic approaches to setting a reserve price are likely to be more appropriate.

6.1 Information-theoretic principles

Given continuous random variables X and Y, the concept of *information gain* provides a quantitative measure of the amount of information one variable provides about the other. In information and communications theory, the amount of information received from observing the value of a random variable is a function of the probability distribution. If $p(x)$ is the probability density function (p.d.f.) of X, the *Shannon entropy* (Shannon, 2001; Wallace, 2005) is the expected amount of information[3] from any observed value of X:

$$H(X) = -\int_{-\infty}^{+\infty} p(x) \log p(x) dx \tag{1}$$

If X and Y are not independent, then an observed value of Y provides some information about X. The entropy of X conditional on Y is:

$$H(X|Y) = -\int_{-\infty}^{+\infty} p(y) \left(\int_{-\infty}^{+\infty} p(x|y) \log p(x|y) dx \right) dy \tag{2}$$

where $p(y)$ is the probability density function of Y, and $p(x|y)$ is the conditional probability density function of X given Y. The amount of information "gained" about X from observing Y is the difference between the entropy of X, and its conditional entropy given Y:

$$\mathcal{I}(X;Y) = H(X) - H(X|Y) \tag{3}$$

This is also known as the *Kullback-Leibler divergence* (Bishop, 2007; Duda, Hart, & Stork, 2001) of the univariate probability distribution of X from the conditional probability distribution of X given Y. If X and Y are independent, then $\mathcal{I}(X;Y) = 0$, which means Y does not provide any gain in information about X. If X and Y are identical, then $\mathcal{I}(X;Y) = H(X) = H(Y)$, which means Y provides the same amount of information as observing X.

Given observed values $X = x$ and $Y = y$, the amounts of empirical information received and gained are functions of the estimated probabilities

[3]The logarithm usually has base 2, with gives a measure of information in *bits*.

of the values:

$$h(x) = -\log \Pr(X = x) \tag{4}$$

$$= -\log \int_{x-\delta}^{x+\delta} p(\xi)d\xi \tag{5}$$

$$h(x|y) = -\log \Pr(X = x|Y = y) \tag{6}$$

$$= -\log \int_{y-\delta}^{y+\delta} \int_{x-\delta}^{x+\delta} p(\xi|\nu)d\xi d\nu \tag{7}$$

$$I(x;y) = h(x) - h(x|y) \tag{8}$$

where δ is a constant quantisation (discretisation) parameter which determines numerical accuracy in the estimated probabilities of observed values (Christofides et al., 1999). Note that $H(X) = \mathbb{E}[h(x)]$, $H(Y) = \mathbb{E}[h(y)]$ and $\mathcal{I}(X;Y) = \mathbb{E}[I(x;y)]$, where $\mathbb{E}[\cdot]$ is the expectation operator.

6.2 Bimodal Price Threshold (BPT) Algorithm

In a conservation tender, the bid prices from the landholders are regarded as the observed values of random variables. Suppose there are n bids (unit prices), denoted by the set $S = \{x_1, x_2, \ldots, x_n\}$. Assuming the landholders are acting independently, S can be sorted in ascending order (i.e., $x_i \leq x_{i+1}$) without loss of generality.

The observed values are regarded as signals from the landholders indicating the minimum price at which they are willing to sell the environmental goods and services they produce. Without any prior information about the population of landholders, a simple model is to assume that the signals are coming from n independent identically-distributed (i.i.d.) sources, denoted by the random variables $X = \{X_1, X_2, \ldots, X_n\}$, each with a Gaussian probability distribution whose parameters can be estimated from S. That is, the simple model assumes that the sources are a group of "typical" landholders.

Since the X_i's are independent, the amount of information from receiving the signals in S can be calculated as follows:

$$h(S) = -\log \Pr(X = S) \tag{9}$$

$$= -\log \Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) \tag{10}$$

$$= -\log \prod_{i=1}^{n} \Pr(X_i = x_i) \tag{11}$$

$$= -\sum_{i=1}^{n} \log \Pr(X_i = x_i) \tag{12}$$

$$= -\sum_{i=1}^{n} \log \int_{x_i-\delta}^{x_i+\delta} f(\xi;\mu,\sigma^2)d\xi \tag{13}$$

where δ is a constant quantisation parameter, and:

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \tag{14}$$

is the normal probability density function with parameters μ and σ estimated from S.

The reserve price, on the other hand, provides information for the buyer side. It indicates the maximum price that a purchasing agent is willing to pay for environmental goods and services. The reserve price can be represented as a threshold, $t \notin S$, for assigning a binary label, y_i, to each bid price in S according to the following rule:

$$y_i = \begin{cases} \text{“low”} & x_i < t \\ \text{“high”} & t < x_i \end{cases} \tag{15}$$

The labeling of the bid prices with respect to the threshold (reserve price) is effectively an alternative, slightly more complex model of the sources of the signals. The alternative model suggests that the sources comprise two categories of landholders: those who typically bid below the threshold, and those who typically bid above the threshold. That is, the model partitions the set X into two subsets:

$$X_{\text{low}} = \{X_i \text{ where } y_i = \text{“low”}\} \tag{16}$$

$$X_{\text{high}} = \{X_i \text{ where } y_i = \text{“high”}\} \tag{17}$$

where the random variables in X_{low} are assumed to be i.i.d., each with a Gaussian probability distribution whose parameters μ_{low} and σ_{low} are estimated from:

$$S_{\text{low}} = \{x_i \in S \mid x_i < t\} \tag{18}$$

Similarly, the random variables in X_{high} are assumed to be i.i.d., each with a Gaussian probability distribution whose parameters μ_{high} and σ_{high} are estimated from:

$$S_{\text{high}} = \{x_i \in S \mid t < x_i\} \tag{19}$$

Let $\theta = \{y_1, y_2, \ldots, y_n\}$ denote the labels of the bid prices in S for a given threshold. The amount of information gained about S given θ can be approximated empirically as follows:

$$I(S;\theta) = h(S) - h(S|\theta) \tag{20}$$

$$= h(S) - (h(S_{\text{low}}|\theta) + h(S_{\text{high}}|\theta)) \tag{21}$$

where $h(S)$ is as defined in Equation 13, and:

$$h(S_{\text{low}}|\theta) = -\sum_{x_i \in S_{\text{low}}} \log \int_{x_i-\delta}^{x_i+\delta} f(\xi; \mu_{\text{low}}, \sigma^2_{\text{low}}) d\xi \tag{22}$$

$$h(S_{\text{high}}|\theta) = -\sum_{x_i \in S_{\text{high}}} \log \int_{x_i-\delta}^{x_i+\delta} f(\xi; \mu_{\text{high}}, \sigma^2_{\text{high}}) d\xi \tag{23}$$

In a conservation tender, the reserve price can be determined by finding a threshold that maximises the empirical information gain in Equation 21. Algorithm 1 on page 13 describes the procedure. The algorithm is based on the following assumptions:

1. There is no prior knowledge about the distribution of the reserve price. Assuming a normal distribution with parameters μ and σ estimated from S essentially biases the reserve price towards μ. In fact, if the distribution is symmetric around μ (e.g. a normal distribution or a uniform distribution), then the information gain is likely to be maximised with a threshold value near μ. In which case, the reserve price partitions S into two equally-sized subsets. (Section 6.3 discusses how to extend the algorithm to situations where there is prior knowledge on the probability distribution of the reserve price).

2. The quantity of supply available at each bid price is not taken into account. This constraint prevents high-quantity suppliers from dominating the reserve price, unless they are willing to spread out the quantity of their supply across several bids. (Section 6.3 discusses how to extend the algorithm to weight the bid prices by the available quantity).

3. The set S is likely to contain outliers, requiring robust estimates of the parameters of the normal distribution. The algorithm considers three methods for estimating μ and σ from S: "mean" which uses the arithmetic mean and standard deviation; "median" which uses the median and half the range of values in S; and, "midpoint" which uses the mean of the maximum and minimum values in S. We are assuming that these methods are sufficient to address the effects of outliers.

4. Preference is for a threshold (reserve price) at an interval with a wide gap between consecutive values in S. The algorithm incorporates this constraint by using the width of the gaps to vary the quantisation parameter, δ, at each iteration. This parameter determines the numerical accuracy of the estimated probabilities. A δ which is either too narrow or too wide would overestimate or underestimate the probabilities of the observed values (Figure 4). Either case would result in low information gain.

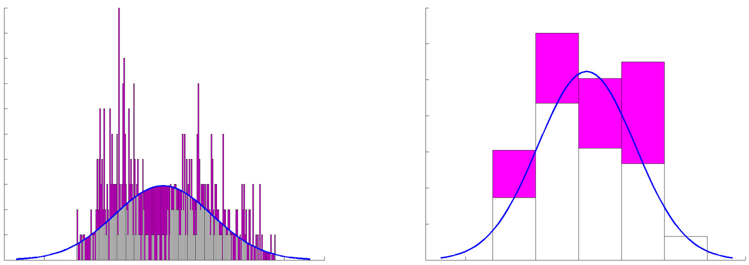

Figure 4: Overestimation or underestimation of probabilities when δ is too narrow or too wide.

5. If there are two thresholds (reserve prices) that lead to the same amount of information gain, preference is for the threshold in an interval where the margin, δ, is largest. In the unlikely event that there are two thresholds with the same margins and same information gains, then preference is

for the higher threshold—that is, the reserve price that procures more environmental benefits. If a conservative lower reserve price is preferred, then the conditional expression $(\hat{\delta} \leq \delta)$ in Line 16 in Algorithm 1 should be replaced with $(\hat{\delta} < \delta)$ instead.

Algorithm 1: `BimodalPriceThreshold`

Data: Bid prices $S = \{x_1, x_2, \ldots, x_n\}$, sorted such that $x_i \leq x_{i+1}$
Result: Reserve price $\hat{t}$

1 **begin**
2   initialise $\hat{t} \leftarrow$ undefined ;
3   initialise $\hat{I} \leftarrow -\infty$;
4   initialise $\hat{\delta} \leftarrow -\infty$;
5   **for** $i \leftarrow 1$ **to** $(n-1)$ **do**
6   **if** $x_i \neq x_{i+1}$ **then**
7             $t \leftarrow 0.5\,(x_i + x_{i+1})$;
8             $\delta \leftarrow 0.5\,(x_{i+1} - x_i)$;
9             $S_{\text{low}} \leftarrow \{x \in S \mid x < t\}$;
10             $S_{\text{high}} \leftarrow \{x \in S \mid t < x\}$;
11   **for** *method* $\leftarrow$ *"mean", "median", "midpoint"* **do**
12   use *method* to estimate μ and σ from S ;
13   use *method* to estimate μ_{low} and σ_{low} from S_{low} ;
14   use *method* to estimate μ_{high} and σ_{high} from S_{high} ;
15                 $I \leftarrow h(S) - (h(S_{\text{low}}|\theta) + h(S_{\text{high}}|\theta))$;
16   **if** $(\hat{I} < I)$ *or* $((\hat{I} = I)$ *and* $(\hat{\delta} \leq \delta))$ **then**
17                     $\hat{I} \leftarrow I$;
18                     $\hat{t} \leftarrow t$;
19                     $\hat{\delta} \leftarrow \delta$;
20   **end**
21   **end**
22   **end**
23   **end**
24 **end**

6.3 BPT extensions

An important question in determining a reserve price using Algorithm 1 is the degree of uncertainty that could be associated with the threshold value. In a conservation tender, for example, it can be argued that the algorithm should take into account uncertainties in the unit bid price, particularly in relation to the precision of the metric for EBI, and the impact of outliers.

One approach is to estimate uncertainty through a stochastic simulation of "trials" involving samples of the unit bids, as shown in Algorithm 2. In each trial, a subset of S is selected randomly, and a threshold is calculated using that subset. The trial is repeated several times, resulting in several possible thresholds. The probability distribution of the thresholds over several trials can be used to estimate some degree of uncertainty. For example, the mean, μ_t, and standard deviation, σ_t, of the thresholds over several trials can be calculated, and the value $t = \mu_t \pm \sigma_t$ can be regarded as a "soft" reserve price. This means that we can choose a threshold anywhere between $\mu_t - \sigma_t$ and $\mu_t + \sigma_t$. However, due to random sampling, repeated applications of Algorithm 2 may produce different values of μ_t and σ_t for the same S. In situations where a deterministic soft threshold is required, or if the size of S is small, a leave-one-out method (Algorithm 3) could be used instead of random sampling.

Algorithm 2: `BPTSoft`

Data: Bid prices $S = \{x_1, x_2, \ldots, x_n\}$, sorted such that $x_i \leq x_{i+1}$
Parameter: M = number of trials, and R = proportion of n per trial
Result: Reserve price mean μ_t, and standard deviation σ_t

begin
    for $m \leftarrow 1$ **to** M **do**
        $S_m \leftarrow$ set of $\lfloor nR \rfloor$ randomly-selected bids from S ;
        $t_m \leftarrow$ `BimodalPriceThreshold(`S_m`)` ;
    end
    $\mu_t \leftarrow$ arithmetic mean of $\{t_1, \ldots, t_M\}$;
    $\sigma_t \leftarrow$ standard deviation of $\{t_1, \ldots, t_M\}$;
end

Algorithm 3: `BPTLeaveOneOut`

Data: Bid prices $S = \{x_1, x_2, \ldots, x_n\}$, sorted such that $x_i \leq x_{i+1}$
Result: Reserve price mean μ_t, and standard deviation σ_t

begin
    for $m \leftarrow 1$ **to** n **do**
        $S_m \leftarrow S \setminus \{x_m\}$;
        $t_m \leftarrow$ `BimodalPriceThreshold(`S_m`)` ;
    end
    $\mu_t \leftarrow$ arithmetic mean of $\{t_1, \ldots, t_n\}$;
    $\sigma_t \leftarrow$ standard deviation of $\{t_1, \ldots, t_n\}$;
end

The algorithm can also be extended to incorporate arbitrary constraints on the reserve price. For example, if there was a policy requirement that at least half of the bids have to be accepted, then the first "for-loop" (Line 5) in Algorithm 1 could be modified so that i starts from $\lceil n/2 \rceil$ instead of 1.

Note that the algorithm can only select a threshold from a set of candidate thresholds. Information gain is the selection criterion, but it does not produce the list of candidate thresholds. In Algorithm 1, the candidate thresholds are constrained to a finite set consisting of mid-points between the unit bid prices. Thus, by default, the algorithm ensures that there will be at least one unit bid on either side of the threshold. The set of candidate thresholds can be adjusted to incorporate external information on the reserve price. For example, the set of candidate thresholds in Algorithm 1 can be extended to include reserve prices from past tenders, including those that are outside the range of values in S, effectively allowing for an (unlikely but possible) blanket acceptance or rejection of all the bids.

The algorithm can also be modified to incorporate prior models of the distribution of the reserve price. For example, suppose T is a random variable for the reserve price, and its probability distribution is known a priori (e.g. estimated from past tenders). It is possible to use this information in two ways:

1. If the probability distribution of T is known to all bidders, it would be reasonable to assume that they are going to quote bids centered around $\mathbb{E}[T]$, with variance $\mathrm{Var}[T]$. In which case, rather than estimating μ and σ from S in Algorithm 1, we can impose a presumption that μ and σ are equal to $\mathbb{E}[T]$ and $\sqrt{\mathrm{Var}[T]}$, respectively.
2. We can estimate the posterior probability, $\Pr(\theta|S)$, of the labelling given the unit bids. Let $Y = \{Y_1, \ldots, Y_n\}$ denote the random variables for the labeling of the unit bids in S. In Algorithm 1, each candidate threshold, $T = t$, has a corresponding unique labelling, $Y = \theta$, so that:
$$h(\theta) = -\log \Pr(Y = \theta) \propto \ -\log \Pr(T = t) \tag{24}$$
Thus we can maximise $\Pr(\theta|S)$ by minimising:
$$\begin{aligned} h(\theta|S) = -\log \Pr(\theta|S) &= -\log \left(\frac{\Pr(S|\theta)\Pr(\theta)}{\Pr(S)} \right) && (25) \\ &= -\log \Pr(\theta) - (-\log \Pr(S) + \log \Pr(S|\theta)) && (26) \\ &= h(\theta) - (h(S) - h(S|\theta)) && (27) \\ &= h(\theta) - I(S;\theta) && (28) \end{aligned}$$
Note that $h(\theta|S)$ can be interpreted as the amount of additional information needed to know θ given S. Minimising $h(\theta|S)$ is equivalent to choosing a threshold that is highly likely (according to the known probability distribution of T), and that also maximises the information gain.

The algorithm can also be extended to incorporate the volume of EBIs on offer at a unit bid price. Be default, the algorithm ignores the volume information in order to avoid biasing the reserve price towards large-volume bids. Let $w_i \in (0, 1)$ denote the proportion (of the cumulative total EBIs) that are on offer at x_i per unit, such that $\sum_{i=1}^{n} w_i = 1$. The estimations in Lines 12 to 14 in Algorithm 1 could then use the weighted unit bids (e.g. $\mu = \sum_{i=1}^{n} w_i x_i$).

It is also possible to change the assumptions about the model class of the distributions of X, X_{low} and X_{high}. In this paper, we assumed that they are normally distributed because that is the simplest assumption without prior information. Given information about the likely shape of the distributions of these random variables, other models can be used in the algorithm instead.

7 Empirical results from EcoTender in Victoria

The bid threshold algorithm has been trialed in a number of tenders throughout Victoria. These tenders include three EcoTenders and multiple river, wetland, grassland and woodland tenders across different catchments in Victoria.

The first EcoTender run in the Corangamite Catchment (CC) was under-subscribed (low competition). The budget could not be expended if all bids were accepted. Funding all bids in this circumstance would have resulted in the purchase of some relatively high cost sites (when compared to other bids in the tender), foregoing the benefit of separating high and low cost suppliers of environmental benefit. The funding of all bids also signal to the community that government is willing to accept all (including very costly bids) without discrimination. This has the potential to encourage rent seeking behaviour in future tenders. Information on prior grants schemes was not useful in setting a reserve price in this instance as there was no information on the expected outcomes obtained under these schemes.

The second EcoTender in the Port Phillip and Western Port (PPWP) was also under-subscribed. There was also an absence of data from alternative grant schemes to inform the setting of a reserve price. In contrast, the third EcoTender in the West Gippsland (WG) had very high participation. The EcoTender budget could have funded less than half of the bids received.

The table in Figure 7 shows information on the costs, budgets and environmental benefits obtained using the bid threshold algorithm in the three EcoTenders [4]

[4] The method used to estimate the environmental benefits differs across each of the tenders. Therefore the data cannot be compared to one another meaningfully and has been omitted from the table.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
EcoTender	**Budget Allocation**	**Algorithm cost**	**Total cost all bids**	**Percentage of total cost**	**Percentage of budget**	**Marginal cost (Budget)**	**Marginal cost (Algorithm threshold)**	**Marginal cost of the last bid**	**Percent of total available EB (Budget)**	**Percent of total available EB (Algorithm)**	**Percentage change in unit cost (Algorithm over Budget)**	**Percentage change in unit cost (Total cost over Algorithm)**
CC	$1.5M	$1,068,680	$1,402373	76%	71%	N/A	$242.92	$2025.18	100%	98%	27%*	27%
PPWP	$1.5M	$874,275	$1,30,2031	67%	58%	N/A	$6572	$57851	100%	94%	40%*	40%
WG	$2.5M	$2,701,037	$5,195,234	52%	108%	$0.55	$0.78	$46	97%	98%	-7%	88%

1. The a priori allocated budget
2. The total dollar allocation to bids if the algorithm threshold is used
3. Total cost of all bids
4. The algorithm budget (2) as a percentage of the total cost (3).
5. The algorithm budget (2) as a percentage of the allocated budget (1).
6. The marginal cost of the last bid if the budget constraint were applied. This figure cannot be calculated for CC and PPWP because the tender was undersubscribed and the budget could not be spent.
7. The marginal cost of the last bid if the algorithm threshold is applied
8. The marginal cost of the last bid
9. Percentage of total EB if the budget were used to select bids
10. Percentage of total EB if the algorithm threshold were used to select bids
11. Percentage change (increase) in average unit cost using the budget (1) rather than the algorithm threshold to select bids. *) The values for CC and PPWP are the same as (12) because the budget could not be spent.
12. Percentage change in average unit cost using the total cost (3) to select bids rather than the algorithm threshold

From the table it can be seen that the bid threshold algorithm cuts off at a different percentage of the total cost in each EcoTender demonstration (see Column 4). This is consistent with the theory behind the algorithm, as the cut-off is determined by the point that results in a best fit of two normal distributions to the unit bid data there is no requirement that this point be similar for different unit bid data.

The tender budget is not relevant to the determination of the bid threshold, although in the event that the cut-off is higher than the budget, the budget may be used as the cut-off. In the West Gippsland EcoTender, the bid-threshold determined by the algorithm was indeed higher than the budget. In this case additional funds were secured enabling the acceptance of more bids.

In the Corangamite and Port Phillip and Westernport EcoTender rounds, the cost effectiveness gains from using the threshold determined by the algorithm over the traditional budget method are significant. The unit cost was 27% higher in Corangamite and 40% higher in Port Phillip and westerport using the budget in comparison to the algorithm (see Column 11). In the West Gippsland EcoTender demonstration using the algorithm results in a 7% increase in the average unit cost for environmental benefit units. This is to be expected as increasing the tender budget results in purchasing additional sites, moving to the right along the supply curve. There are reasons why this might be cost-effective and these are discussed in the next section.

8 Location of the bid threshold

The algorithm places few constraints on the location of the bid threshold. There must be at least one bid on each side of the threshold. Therefore the threshold must be strictly greater than the smallest unit bid and strictly less than the largest unit bid. As the set of unit bids often displays some normal characteristics, the bid threshold is more likely to be centrally located.

The location of the threshold in comparison to the budget is likely to change depending on participation. If a tender is undersubscribed (as in Corangamite and Port Phillip and Westernport), the cut-off is necessarily to the left of the budget-determined threshold. Similarly, when there is little competition and high cost sites are funded using the budget determined cut-off, the algorithm threshold is likely to be to the left of the budget threshold. When participation is high relative to the budget (for example funding all bids in the West Gippsland EcoTender would cost in excess of twice the budget), the bid threshold is more likely to be to the right of the budget-determined threshold.

When a tender is undersubscribed, or there is little competition, there are clear advantages to applying the bid threshold algorithm. In these cases there is an increase in cost effectiveness obtained from shifting the funding cut-off to the left on the supply curve (see the results for Corangamitte and Port Phillip and Westernport). There is also a price signal argument that favours the use of the bid algorithm over the budget when faced with inadequate competition. When an agency accepts and rejects bids in a tender a price signal is sent back to landholders regarding the agencies willingness to pay. If an agency purchases all bids, the agency may signal to the community that they are willing to pay a very high price.

This price signal is weaker in a discriminatory price auction than a uniform price auction because each landholder is only aware of the status of their bid. However, if landholders communicate (as may occur in smaller communities) they may gain more information on the agencies willingness to pay. In a uniform price auction the price signal is stronger as all successful landholders are paid, and therefore aware of, the marginal bid price. If insufficient competition results in the funding of high cost bids, this may lead to higher bids in the subsequent tenders.

When a tender has higher participation relative to the budget it is less clear whether the budget or algorithm threshold is more beneficial to the agency. In particular, when the algorithm determined threshold lies to the right of the budget determined threshold, the average unit cost will be higher using the algorithm than the budget method. The increase in cost per unit may be perceived as a fall in cost effectiveness of the tender as whole. However, this decrease in cost effectiveness relates only to on-site costs. The fixed costs of running a tender are not factored into this cost-effectiveness estimation.

Given a pool of funds to be allocated across several tender rounds, funding more bids per round reduces the total number of tenders run leading to decreased transaction costs. In other words, there is a trade-off between keeping on-site costs low and keeping transaction costs low. Funding additional sites in a tender where there is a large amount of competition may be more cost effective once these transaction costs are taken into account. Moreover when participation is high, the investment has already been made in assessing the quality of each site. Using funds to purchase additional bids in the current tender has the benefit of making use of the sunk costs of site assessments in the current round rather than paying to undertake additional site assessments in a future tender.

Tender budgets are often announced before the tender is implemented. As levels of participation are unknown *ex ante*, it is difficult to set the budget optimally – ensuring enough competition to keep the price low without pushing transaction costs too high. Using an algorithm that is executed *ex post* instead of the budget method has the advantage of implicitly responding to participation levels. For example, consider two tenders with the same budget where participants are drawn from the same independent and identically distributed normal distribution. In tender one, one hundred bidders are drawn and in tender two, two hundred bidders are drawn. The expected bid threshold in each case is the mean. Consequently the expected marginal cost is equated between the two tenders. Using the budget threshold, the marginal cost in tender one is expected to be higher than tender two as the greater participation in tender two should result in more low cost options.

9 Conclusion, limitations, and further work

This paper has demonstrated that the bid threshold algorithm can be used to determine an *ex post* reserve price in conservation tenders. The algorithm is designed for use in situations where information needed to apply standard economic theory to developing a reserve price is absent. The algorithm is useful when there are no alternative programs running concurrently and there is no data available on opportunity costs. The algorithm is particularly useful when a tender is under-subscribed.

The bid threshold algorithm presented in this paper is one method for determining an *ex post* reserve price in conservation tenders. The algorithm is designed for use in situations where information needed to apply standard economic theory to developing a reserve price is absent. In cases where reliable information on opportunity costs exists, these opportunity costs are likely to provide a reserve price that results in a more efficient allocation of resources. As discussed previously, situations where there is insufficient information to form a reasonable estimation of agency opportunity costs may occur when:

- there is no information on prior tenders or alternative schemes;
- there is some information from different areas, however there is an expectation that opportunity costs are highly heterogeneous across regions; and,
- the scoring system for a tender has changed significantly and there is no method for comparing new and old units of environmental benefit.

Further to these situations, depending on the timing between tender rounds, a tender implemented twice in one area may incur a higher average unit cost to the agency. Here good value for money landholders may have been contracted in the first tender, leaving higher cost landholders to bid in the second round. The bid threshold algorithm could provide an alternative cut-off method in this case.

The algorithm can be useful in tenders with both low and high levels of competition. Use of the algorithm in a tender with low levels of competition results in two potential benefits. Firstly, using the algorithm determined threshold results in a lower average unit cost for the procuring agency. Secondly, using the algorithm to determine a cut-off will prevent the community receiving signals that the agency is likely to accept high bids in future tenders. Using the algorithm in high competition tenders identifies bids that exceed the budget, but represent good value for money relative to other bids in the tender. It may be beneficial to the agency to spend additional funding on these bids where transaction costs are already sunk, than to use this funding to procure environmental benefit via an alternative means.

There are limitations associated with the use of the bid threshold algorithm. The algorithm selects the bid threshold based solely on information contained in the unit bid data with no regard to transaction costs. If a fixed pool of funds can be allocated between multiple tenders, if less bids are funded per tender, more tenders will be run increasing the total transaction costs. Consequently, the algorithm may determine a threshold that results in the acceptance of too few bids to be cost effective when considering allocating a pool of funds across an undetermined quantity of multiple tenders. This problem is not unique to situations involving the bid threshold algorithm. The trade-off between keeping on-site and transaction costs low exists when the threshold is determined by budget also. Work is needed to investigate the relationship between transaction costs and on-site costs when allocating funds over an undetermined quantity of tenders. This information could be used to investigate modifications to the algorithm that take transaction costs into account when determining the threshold.

In circumstances where some information is available from a prior tender, but the reliability of this information is unknown, it may be desirable to use a combination of both prior information and information contained in the unit bid data set to determine a reserve price. Further empirical work could be undertaken to investigate extensions to the algorithm presented in this paper that take prior information into account.

References

Bishop, C. M. (2007). *Pattern recognition and machine learning.* New York: Springer.

Christofides, A., Tanyi, B., Christofides, S., Whobrey, D., & Christofides, N. (1999, October). The optimal discretization of probability density functions. *Computational Statistics & Data Analysis*, *31*(4), 475-486.

Connor, J., Ward, R., & Bryan, B. (2008). Exploring the cost-effectiveness of land conservation auctions and payment policies. *Australian Journal of Agricultural and Resource Economics*, *51*, 303-319.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). Wiley-Interscience.

Eigenraam, M., Strappazzon, L., Lansdell, N., Beverly, C., & Stoneham, G. (2007). Designing frameworks to deliver unknown information to support market-based instruments. *Agricultural Economics*, *37*(S1), 261-269.

Hajkowicz, S. (2007). Allocating scarce financial resources across regions in queensland, australia. *Ecological Economics*, *61*, 208-216.

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

Latacz-Lohmann, U., & Hodge, I. (2003). European agri-environmental policy for the 21st centuary. *Australian Journal of Agricultural and Resource Economics*, *47*, 123-139.

Nemes, V., LaNauze, A., Walsh, C., Fletcher, T., Bos, D., RossRakesh, S., et al. (2010). Saving a creek one bid at a time: A uniform price auction for stormwater retention in urban catchments.

Shannon, C. E. (2001). A mathematical theory of communication (reprint). *SIGMOBILE Mob. Comput. Commun. Rev. (Special issue dedicated to Claude E. Shannon)*, *5*(1), 3–55.

Stoneham, G., Chaudhri, V., Ha, A., & Strappazzon, L. (2003). Auctions for conservation contracts: an emperical examinaitoin of victoria's bushtender trial. *Australian Journal of Agricultural and Resource Economics*, *47*(4), 477-500.

Vukina, T., Zheng, X., Marra, M., & Levy, A. (2008). Do farmers value the environment? evidence from a conservation reserve program auction. *International Journal of Industrial Organization*, *26*, 1323-1332.

Wallace, C. (2005). *Statistical and inductive inference by minimum message length.* New York: Springer.

Windle, J., & Rolfe, J. (2007). Exploring the efficiencies of using competitive tenders over fixed price grants to protect biodiversity in australian rangelands. *Land Use Policy*, *25*, 388-398.