
Publisher's PDF, also known as Version of record

Link to published version (if available):
10.5176/2010-2283_1.2.50

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html
Estimating Demand for Dynamic Pricing in Electronic Markets

John Cartlidge*
Department of Computer Science
University of Bristol
Bristol, UK, BS8 1UB
john.cartlidge@bristol.ac.uk

Steve Phelps
Centre for Computational Finance
& Economic Agents, University of Essex
Colchester, UK, CO4 3SQ
sphelps@essex.ac.uk

ABSTRACT
Economic theory suggests sellers can increase revenue through dynamic pricing; selling identical goods or services at different prices. However, such discrimination requires knowledge of the maximum price that each consumer is willing to pay; information that is often unavailable. Fortunately, electronic markets offer a solution; generating vast quantities of transaction data that, if used intelligently, enable consumer behaviour to be modelled and predicted.

Using eBay as an exemplar market, we introduce a model for dynamic pricing that uses a statistical method for deriving the structure of demand from temporal bidding data. This work is a tentative first step of a wider research program to discover a practical methodology for automatically generating dynamic pricing models for the provision of cloud computing services; a pertinent problem with widespread commercial and theoretical interest.

Keywords
Demand modelling, dynamic pricing, eBay, cloud computing

1. INTRODUCTION
An understanding of supply and demand (S&D) is fundamental to microeconomics, finance and marketing. Yet, the theoretical and statistical tools necessary for a detailed empirical analysis of S&D in real markets has historically remained elusive [6, 8]. More recently, however, techniques have been developed to estimate S&D in financial markets [7]; and in electronic auction markets such as eBay [9, 10]. These models are able to recover S&D curves by analysing high-frequency trading data, thus allowing an analysis of the marketplace in sufficient detail to be of use to traders.

Previous studies by other authors have outlined the principles by which S&D could be analysed in a retail electronic auction marketplace [12]. In this paper we apply these principles to eBay, demonstrating how S&D can be estimated from actual empirical trading data; and exemplifying how structural S&D estimates can be used to determine optimal sales strategy: how many additional units to sell; and at what price.

Although this paper concentrates on eBay, the statistical estimation model we employ is general enough to be used in any online auction venue. In future, the authors aim to extend and apply the model to the dynamic pricing of cloud computing—the online utility provision of hardware and software services. Cloud computing is a burgeoning market that is set to flourish in the near future; presenting a paradigm shift in information technology that may have a profound positive disruption on society [5]. Yet, there is currently little theory on advanced pricing models for cloud computing. The authors aim to address this deficiency.

The outline of this paper is as follows. Section 2 describes online auctions and the demand estimation problem for sellers. Section 3 introduces the demand estimation model; and is followed by an exemplar application on real empirical eBay data in Section 4. Section 5 discusses future work extensions; in particular dynamic pricing for cloud computing services. Finally, Section 6 provides a brief summary and conclusion.

2. ELECTRONIC AUCTION VENUES
Over the past 15 years there has been a phenomenal rise in the volume of trade executed in online auctions. Founded in 1995, eBay alone now has a global presence in 37 markets and a customer base of 233 million.2 Online auctions offer rich pickings for individuals and corporations wanting to exploit the potential liquidity of the global marketplace.

2.1 Power Selling in Online Auctions
While customer to customer (C-to-C) trade still accounts for a large proportion of online auction volume, increasingly there has been a rise in the number of businesses selling to individual customers (business to customer, or B-to-C, trade). Corporations with business models incorporating the sale of large quantities of stock in online auctions are known as power sellers. Rather than offload individual items or one-off shipments, power sellers regularly ship large volumes in online auctions as part of their core sales strategy.

*Contact author.
1That is, data sampled at a frequency higher than one day.
Sellers have a strong incentive to understand the behaviour of their potential customers. Estimating the parameters—eBay market, volume, time, listing format, price—that will maximise revenue requires knowledge of the dynamics of consumer demand. For example, a seller that anticipates a large surge in demand for a particular marketplace will have a better understanding of how and when to increase supply in that market and the price they should expect to achieve. Although all sellers benefit from a better understanding of demand, it is power sellers—those that supply the greatest quantity per unit time—that have the most to gain (or lose).

### 2.2 The Demand Estimation Problem

In online auction venues, power sellers not only have access to more traditional methods of estimating demand—personal experience, market research, trial and error—they also have access to the bid history of each auction; the time-stamped succession of highest bids. This valuable resource enables observation of how often and at what price bids are posted during the auction period. By observing the highest bid registered by each user one can begin to estimate the maximum (limit) price of each individual. Once the limit price of each bidder is known, it is an easy step to calculate the demand function—the volume demanded at each price.

Unfortunately, a problem exists: to estimate demand it is necessary to know the limit price of all potential buyers. However, since an auction’s bid history records the succession of highest bids, it is not a full record of all potential bidders—individuals that arrive at an auction once the auction price has already exceeded their own limit price are not recorded. Such potential bidders are forced to leave an auction without registering a bid and thus do not appear in the bid history of an auction—their bids are censored. This leads to an under-estimation of demand (see Fig. 1).

The problem is: how to recover censored bids from observed bid history for more accurate estimation of market demand? To tackle this problem, the following section introduces a model to recover censored bids.

### 3. MODELLING DEMAND

We introduce a statistical model for recovering censored demand in electronic auctions: that is, demand that may not be directly observed from the set of auction transaction data. Using [12] as a foundation, the model relies on time-stamped auction data. We use the arrival times of market participants to estimate arrival rates—demand—across different price levels. While the model is general enough to be applied to any electronic auction, we use eBay—a prominent online market with high liquidity—as our exemplar.

#### 3.1 Model Assumptions

To make it easier to work with observed bid history data, we first segment price into discrete intervals. Then, across all auctions we measure the time until first bid arrival in each price segment. We should expect that occasionally there will be long time-intervals before a first bid is registered, but that more often these intervals will be shorter. If we suppose that bids arrive independently and that no two bids arrive simultaneously, then we may assume that the time until first arrival of a bid in each price segment follows an exponential distribution.

The value we wish to estimate is the most likely expected number of bids, $\lambda_i$, that will be posted in each price segment, $i$, during a given time interval; thus allowing us to evaluate the relative proportion of bidders, or demand, in each price segment.

We model $\lambda_i$ using a Poisson distribution by assuming the following:

- **A** Bids arrive at random in continuous time.
- **B** Bids arrive singly: the probability of two bids arriving simultaneously is zero.
- **C** Bids arrive uniformly: the expected number of bids in a given interval is proportional to the size of the interval; and arrival rates do not vary over time.
- **D** Bids arrive independently: the probability of a bid arriving with price $i$ in any small interval is independent of the probability of a bid arriving with price $i$ in any other small interval.

We make some further assumptions about the strategic behaviour of bidders:

- **E** Bidders bid at exactly their limit price.
- **F** Bidders attempt to post a bid upon arrival: they do not strategically wait.

Finally, we assume the following is true of the auction mechanism:

- **G** Posted bids must be greater than current auction price.

---

\[3\] In reality, arrival rates rapidly increase towards the end of an auction as bidders attempt to *snipe*. Preliminary analysis has shown that as many as 25% of all bidders arrive during the final hour. Hence, the model is likely to underestimate arrival rates.
3.2 Estimating Arrival Rates
Following [12], let us segment prices into \( K \) equally sized bins and let \( X_i \) denote the time of arrival of the first bid in price segment \( i \), where \( i = 1, 2, \ldots, K \). Then, \( X_i \) is an exponentially distributed continuous random variable with probability distribution function:

\[
f(x) = \begin{cases} 
\lambda_i e^{-\lambda_i x} & : x \geq 0 \\
0 & : x < 0
\end{cases}
\] (1)

Hence, \( X_i \) has mean \( \frac{1}{\lambda_i} \) and expected arrival rate \( \lambda_i \). The probability of \( X_i \) occurring near time \( t = \tau \) is:

\[
P(t < X \leq \delta t) = f(t) \cdot \delta t
\] (2)

The probability of \( X \) occurring after time \( T \) is:

\[
P[X > \tau] = 1 - P(X \leq \tau) = 1 - \int_0^\tau \lambda_i e^{-\lambda_i t} dt = 1 - \left[-e^{\lambda_i t}\right]_0^\tau = 1 + e^{-\lambda_i \tau} - e^0 = e^{-\lambda_i \tau}
\] (3)

We demonstrate how to estimate arrival rates \( \lambda_1, \lambda_2, \ldots, \lambda_K \) through a worked example with \( n = 2 \) bidders; however the result can be generalised to \( n \) bidders.

Assume there are two potential bidders with limit prices \( i \) and \( j \) \((i < j)\) and corresponding arrival times \( X_i \) and \( X_j \). Then, when an auction is complete, it is possible that the bid history may contain: (a) no bidders; (b) one bidder of type \( i \); (c) one bidder of type \( j \); or (d) two bidders. Let \( \langle x_1, x_2, \ldots, x_n : a_i \rangle \) denote the recorded bid history of auction \( a \) end time \( \tau \), then for an auction \( A_T \), we can calculate the following likelihoods:

(a) Probability no bidders appear in bid history:

\[
P(\emptyset : A_T) = P[X_i > \tau] \cap (X_j > \tau) = e^{-\lambda_i \tau} \cdot e^{-\lambda_j \tau}
\]

(b) Probability only bidder \( i \) appears in bid history:

\[
P([x_i : A_T]) = P[(x_i > \tau) \cap (X_j > \tau)] = \lambda_i e^{-\lambda_i x_i} \cdot e^{-\lambda_j \tau}
\]

(c) Probability only bidder \( j \) appears in bid history:

\[
P([x_j : A_T]) = P[(X_i > \tau) \cap (x_j > \tau) \cap (X_j > \tau)] = e^{-\lambda_i \tau} \cdot \lambda_j e^{-\lambda_j x_j} \cdot e^{-\lambda_j \tau}
\]

(d) Probability both \( i \) and \( j \) appear in bid history:

\[
P([x_i, x_j : A_T]) = P[(x_i < X_i \leq x_i + \delta x_i) \cap (x_j < X_j \leq x_j + \delta x_j)] = \lambda_i e^{-\lambda_i x_i} \cdot \lambda_j e^{-\lambda_j x_j} \cdot e^{-\lambda_j \tau}
\]

Assume that we have observed three auctions with bid histories as follows: two bids \( \langle x_i, x_j : A_T \rangle \); no bids \( (\emptyset : A_T) \); one bid \( \langle x_j : A_T \rangle \). Then the likelihood function is:

\[
L(\lambda_i, \lambda_j) = \lambda_i e^{-\lambda_i x_i} \cdot \lambda_j e^{-\lambda_j x_j} \cdot e^{-\lambda_j \tau} \cdot e^{-\lambda_i \tau} \cdot \lambda_i e^{-\lambda_i x_i} \cdot \lambda_j e^{-\lambda_j x_j}
\]

Taking natural logarithm gives the log-likelihood function:

\[
\ln(L(\lambda_i, \lambda_j)) = \ln(\lambda_i + 2 \ln \lambda_i + \ln \delta x_i + 2 \ln \lambda_j) - \lambda_i x_i + T + x_j - \lambda_j x_j + T + x_j
\]

Then, maximum likelihood values of arrival rates are:

\[
\frac{\partial l}{\partial \lambda_i} = \frac{1}{\lambda_i} - (x_i + T + x_j) = 0 \Rightarrow \hat{\lambda}_i = \frac{1}{x_i + T + x_j}
\]

\[
\frac{\partial l}{\partial \lambda_j} = \frac{2}{\lambda_j} - (x_j + T + x_j) = 0 \Rightarrow \hat{\lambda}_j = \frac{2}{x_j + T + x_j}
\]

Thus, we see that the maximum likelihood arrival rate for type \( i \) is the number of auctions in which we observe a bid of type \( i \) (in example, above) divided by the sum of the arrival times in each auction of either: the first bid to arrive of type \( i \); or in auctions where no bids of type \( i \) appear, the first bid to arrive of the next highest type; or if no higher bids arrive, the auction close time, \( T \).

Similarly, the arrival rate for type \( j \) is the number of auctions in which we observe a bid of type \( j \) (in above example) divided by the sum of the arrival times in each auction of either: the first bid to arrive of type \( j \); or in auctions where no bids of type \( j \) appear, the first bid to arrive of the next highest type; or if no higher bids arrive, the auction close time \( T \).

For brevity, let \( x_{i+}^n \) be the arrival time of bidder \( i \) in the \( n^{th} \) auction if bidder \( i \) is recorded in the \( n^{th} \) auction, or the arrival time of the next highest bidder, or the auction duration if no higher bidder arrives. Then, we get the general result:

\[
\hat{\lambda}_i = \frac{\text{# auctions in which type } i \text{ bid appears}}{x_{i+}^1 + x_{i+}^2 + \ldots + x_{i+}^n}
\]

Let us call the divisor in the above equation the effective opening time for bidders of type \( i \); i.e., the total time type \( i \) bidders have available to place a bid across all auctions. Once an auction price has surpassed a bidder’s limit price, the auction is effectively closed to that bidder. If limit price is never surpassed, then the effective auction close equals actual auction close. Using this terminology, the above equation can be rewritten in words:

\[
\hat{\lambda}_i = \frac{\text{total number of bids from type } i \text{ bidders}}{\text{total effective opening time}}
\]

This is an intuitive result: the average arrival rate of bidders equals the number of bidders observed over the total time available for bids to be placed. Finally, to effectively consider parallel auctions, we must measure bid arrivals using absolute time rather than auction time. Then, the arrival rate of bidders of each bin is calculated as:

\[
\hat{\lambda}_i = \frac{\text{total number of bids from type } i \text{ bidders}}{\text{total time that at least one auction is effectively open}}
\]
3.3 Confidence Interval Estimation
The model developed in the previous section reduces to the standard formula for calculating survival rates with Type I censored data (see, for example, [1]). Under Type I censoring, the maximum likelihood for survival rate, $\lambda$, is:

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^{n} x_i}$$  \hspace{1cm} (4)

where $x_i$ is the $i^{th}$ data point (may be arrival or censoring point), $n$ is total number of data points, both censored and uncensored, and $r$ is number of failures. Using this, we can estimate the $100(1 - \alpha)$% confidence interval for $\lambda$ as:

$$\frac{2n}{\hat{\lambda} \cdot \chi^2_{(2n, \frac{1}{2})}} < 1 - \alpha < \frac{2n}{\hat{\lambda} \cdot \chi^2_{(2n, 1 - \frac{1}{2})}}$$  \hspace{1cm} (5)

Where $\hat{\lambda}$ is maximum likelihood estimation, $\lambda_i$ is true value, and $\chi^2_{(v, \alpha)}$ is the value of chi squared distribution with $k$ degrees of freedom that gives $x$ cumulative probability.

3.4 Multiple Bids from Individual Bidders
Once a bidder has placed a bid in an auction, there is nothing to stop them from bidding in other auctions. Indeed, it is likely that once a auction has effectively closed an unsuccessful bidder will move to an alternative auction venue. However, unless a bidder wins an auction and then subsequently bids in another auction, we assume that each bidder demands only one item. Thus, directly counting proxy bids will lead to an overestimation of demand.

The model estimates demand by resolving multiple bids across auctions as follows:

A Simulate each auction using proxy bid data to calculate the effective opening times of each price bin.

B Order all proxy bids across all auctions by time. Begin with the earliest bid and move down the list in order of time. For each bid, if bidder ID is new—ID does not appear in Demand list—add bid to Demand; otherwise, edit bid in Demand such that:

(a) Bidder ID: identity of bidder.
(b) Bid time: time of earliest bid.
(c) Bid price: price of highest bid.

C For each price bin, calculate the mean arrival rate of bidders by dividing the number of bids in Demand by the total effective opening time.

3.5 Model Validation
To validate the statistical model, we employ a simple simulation of eBay containing 4 parameter distributions: bid price, $B_p$; time to next bid arrival, $B_t$; time to next auction opening, $A_t$; and auction length, $A_l$. Let $X \sim N(\mu, \sigma)$ denote random variable $X$ is normally distributed with mean $= \mu$ and standard deviation $= \sigma$; and $X \sim U(\{i, j, \ldots, n\})$ denote $X$ is random variable uniformly distributed across the set of values $\{i, j, \ldots, n\}$. Then, set: $B_p \sim N(50, 15)$

3 Bidders have been observed strategically posting stub quotes in multiple simultaneous auctions; submitting very low bids in the hope that no other bidders join the auction.

Figure 2: Simulation of eBay. Demand estimation (solid line) is a good fit of true demand (dot).
will attain for increasing sales volume by a given quantity.

4.2 Revenue and Costs
The application has two revenue models. These are:

**Differential pricing:** assume each item is sold at the highest price it can attain. Prices will vary between sales.

**Fixed pricing:** assume all units are sold at the same fixed price. This is a more conservative estimator of revenue.

Fig. 4 shows anticipated revenue and costs for a seller, $X_s$, using a fixed price revenue model. These costs include eBay and shipping costs associated with each sale and rise with volume sold. Other costs specific to $X_s$, such as labour costs or OEM commission, are entered into the model as customized variables. These costs are not graphed, but are factored into profit calculations, below.

The revenue curve is calculated using competitive demand. In Fig. 4, we see anticipated revenue has a maximum around quantity 700, but then steeply falls. This is because we are using a fixed revenue model—all units sold at the same price. Thus, as volume increases and marginal price falls, the price of each unit also falls.

4.3 Profit
Profit equals total revenue minus total cost. Fig. 5 shows anticipated profit as a function of increased sales quantity. We see that a maximum profit of approximately $10,000 is achieved with approximately 700 sales. Once sales exceed 1000, profits become negative.

Assuming costs rise as volume is increased, there will always be a turning point in the profit curve; even under differential pricing. This turning point is the maximum possible anticipated profit available to a seller for a given product. Maximum profit and volume can be used to produce summary estimates of current production efficiency.

Note that Fig. 4 does not include costs that are specific to a particular seller, $X_s$. Hence, profit cannot be directly calculated from the revenue and costs curves displayed.

4.4 Extensions
We have described a model for estimating demand in electronic auction venues and applied it, using real eBay data, to infer optimal sales strategy and dynamic pricing. This is a promising first step. However, the model makes some unrealistic assumptions such as fixed arrival rates, which are likely to reduce estimation accuracy. In future, we would like to address this by modifying the model to incorporate more realistic behaviours, such as sniping; and use heuristic methods to calibrate the model against empirical data, e.g., [11].

5. PRICING CLOUDS
Over the last 50 years, roughly once a decade, the commercial provision of IT has undergone a step change—from mainframes to the Internet via mini-computers, single-user PCs and the client-server LAN model [4]. Following this pattern, we appear to be entering a new paradigm: the online utility provision of hardware and software services, or “cloud computing”. Offering on-demand availability to effectively infinite computing resources at prices kept low though scale economies of central supply, cloud computing has the potential to positively impact the whole of society, in a manner analogous to the impact of the utility provision of electricity during the 20th Century [5].

Computing behemoths such as Amazon, Google and Microsoft are investing heavily in the provision of cloud com-
computing infrastructure through large warehouse-scale computing facilities [2]. However, current understanding of advanced pricing mechanisms for cloud services is inadequate. Traditionally, hosted services operate on a simple amortized cost-plus charging model, with little price differentiation.

Research into the decentralised control and provision of utility computing through the use of auction markets has generated a large literature [3]. Such utility computing markets present a direct opportunity to apply techniques—structural estimation of S&D for dynamic pricing—developed in this paper. If future, we aim to extend our work toward dynamic pricing of cloud computing services—a pertinent research goal with highly lucrative commercial potential.

6. SUMMARY AND CONCLUSIONS

We have described a statistical model for estimating product demand by utilising the bid information available from electronic auction markets such as eBay. Further, we have applied this model to estimate supply and demand for a real marketplace using actual empirical data. To forecast future revenue accurately, sellers must have an accurate understanding of consumer demand. The more accurately a seller understands demand, the better they are able to maximise profit. Whilst the experience of sellers, trial and error, and market research reports each lend some insight into the behaviour of customers, each is a very poor alternative to the quantitative estimates produced by the model presented here. Producing a full demand curve, the model allows sellers to maximise profit by predicting sales volume and average price ahead of time. This accurate forecasting ability allows sellers to optimise their strategy ahead of time, rendering costly (and risky) trial and error strategies obsolete and presenting an opportunity for dynamic pricing.

The model is entirely general. As long as there is an ascending auction format, the model is able to build a representation of demand, whatever the product or where ever it is sold. By understanding demand, sellers are able to better optimise their sales strategy and reduce risk. As such, the model is of value to any company or individual that wants to sell, offering significant positive impact on the revenue-generating potential of all auction sellers.

7. ACKNOWLEDGEMENT

This work was supported by EPSRC grant: EP/H042644/1

8. REFERENCES


Author Profiles

Dr. John Cartlidge has a First Class Honours degree in Artificial Intelligence & Mathematics (2000) and a Ph.D. in Computer Science (2004) from the University of Leeds. He spent four years in industrial research working for Hewlett-Packard Labs and other smaller companies before returning to academia in 2008. He is currently a Research Associate in Cloud Computing at the Department of Computer Science, University of Bristol, UK.

Dr. Steve Phelps received his Ph.D. degree while he was with the Agent Applications, Research and Technology (Agent ART) Group at the University of Liverpool. He is currently a Lecturer with the Centre for Computational Finance and Economic Agents (CCFEA) at the University of Essex, UK. His research centres on agent-based modelling in finance and economics.