



Cartlidge, J., & Phelps, S. (2010). Estimating consumer demand from high-frequency data. In K. Kumar (Ed.), *Annual International Academic Conference on Business Intelligence and Data Warehousing (BIDW 2010)*. (pp. 132-138). Singapore: Global Science and Technology Forum (GSTF).

[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>

### Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact [open-access@bristol.ac.uk](mailto:open-access@bristol.ac.uk) and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

# Estimating Consumer Demand From High-Frequency Data

John Carlidge<sup>\*</sup>  
University of Central Lancashire  
Preston, PR1-2HE  
United Kingdom  
jpcarlidge@uclan.ac.uk

Steve Phelps  
University of Essex  
Colchester, CO4 3SQ  
United Kingdom  
sphelps@essex.ac.uk

## ABSTRACT

Price discrimination offers sellers the possibility of increasing revenue by capturing a market's consumer surplus: arising from the low price elasticity segment of customers that would have been prepared to pay more than the current market price. First degree price discrimination requires a seller to know the maximum (reserve) price that each consumer is willing to pay. Discouragingly, this information is often unavailable; making the theoretical ideal a practical impossibility.

Electronic commerce offers a solution; with loyalty cards, transaction statements and online accounts all providing channels of customer monitoring. The vast amount of data generated—eBay alone produces terabytes daily—creates an invaluable repository of information that, if used intelligently, enables consumer behaviour to be modelled and predicted. Thus, once behavioural models are calibrated, price discrimination can be tailored to the level of the individual.

Here, we introduce a statistical method designed to model the behaviour of bidders on eBay to estimate demand functions for individual item classes. Using eBay's temporal bidding data—arrival times, price, user id—the model generates estimates of individual reserve prices for market participants; including hidden, or censored, demand not directly contained within the underlying data. Market demand is then estimated, enabling eBay power sellers—large professional bulk-sellers—to optimize sales and increase revenue. Proprietary software automates this process: analyzing data; modelling behaviour; estimating demand; and generating sales strategy.

This work is a tentative first step of a wider, ongoing, research program to discover a practical methodology for automatically calibrating models of consumers from large-scale high-frequency data. Multi-agent systems and artificial in-

telligence offer principled approaches to the modelling of complex interactions between multiple individuals. The goal is to dynamically model market interactions using realistic models of individual consumers. Such models offer greater flexibility and insight than static, historical, data analysis alone.

## Keywords

Consumer Modelling, High-Frequency Data, eBay

## 1. INTRODUCTION

An understanding of supply and demand is fundamental to microeconomics, finance and marketing. However, historically the theoretical and statistical tools necessary for a detailed *empirical* analysis of supply and demand in real-life markets remained elusive [2]. More recently, techniques have been developed to estimate the supply and demand curves in financial markets [3] and in electronic auction markets such as eBay [2, 4]. These models are able to recover supply and demand curves by analysing high-frequency trading data<sup>1</sup>, thus allowing an analysis of the marketplace in sufficient detail to be of use to not only to economists but also to traders.

Although such quantitative tools have recently been applied in financial markets, the availability of high-frequency data in markets such as eBay opens up the possibility for algorithmic trading in *retail* markets [6]. This paper outlines the first steps in building a high-frequency algo-trader in this domain. Previous studies by other authors have outlined the principles by which supply and demand could be analysed in a retail electronic auction marketplace [6]. In this paper we *apply* these principles and demonstrate that supply and demand can be estimated from actual empirical trading data. We also validate the estimation model by comparing its predictions against a Monte-Carlo simulation of the underlying model. This provides us with framework which can be extended allowing us to drop some of the more unrealistic assumptions of the original model.

The outline of this paper is as follows. In Sections 1.1 and 1.2 we give an overview of online auction marketplaces and the estimation problem. In Section 2 we describe our statistical model. In Section 3 we give the results from applying this model to real empirical data. In Section 4 we discuss how

<sup>1</sup>That is, data that are sampled at a frequency higher than one day. For example, high-frequency financial data is available at sub-second time scales.

<sup>\*</sup>Contact author.

this work can be taken forward and finally we conclude in Section 5.

### 1.1 Online Auctions and Power Sellers

Over the past decade 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 33 markets, a global customer base of 181 million registered users and worldwide trade of more than \$1,511 worth of goods every second. Online auction sites offer rich pickings for individuals and corporations that can exploit the potential of the global marketplace they encompass.

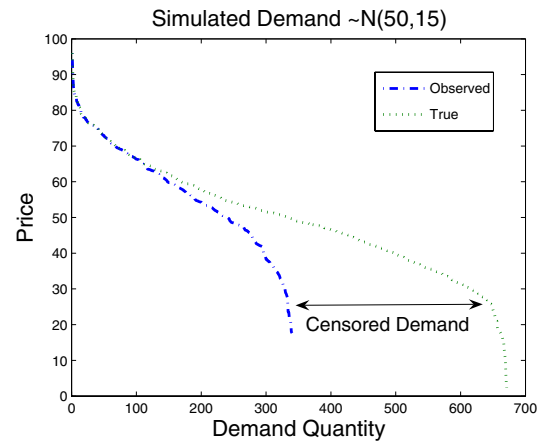
Although customer to customer (C-to-C) trade still accounts for a large proportion of online auction volume, increasingly there has been a rise in the number of businesses selling to individual customers (business to customer, or B-to-C, trade). Corporations whose business models incorporate the sale 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 supply large quantities of stock to online auctions as part of their ongoing sales' strategy.

When selling, it is essential to understand the behaviour of the consumers you are selling to. In order to estimate the parameters (which market to drop items into, what volume to supply, what time to list and what listing format, for example) that will maximise revenue, one needs working knowledge of the dynamics of consumer demand. For example, a seller that anticipates a large surge in demand in a particular marketplace will have a better understanding of how and when to increase supply in that market and at what price they should expect to achieve. By accurately determining a market's underlying consumer demand, sellers are able to significantly increase revenue and profit. Whilst every seller participating in an online auction stands to benefit from a better understanding of consumer demand, it is power sellers - those that supply the greatest quantity per time period - that have the most to gain (or lose).

### 1.2 Estimating Demand in the Marketplace

In order to optimise sales strategy it is important for sellers to be aware of the nature of demand. In online auction venues, power sellers not only have access to more traditional methods of estimating demand (personal experience, market research, trial and error, etc.), they may also make use of the bid history of each auction (the time-stamped record of each successive highest bid). This valuable resource enables sellers to observe how often and at what price bids are posted during the entire auction period. By observing the highest bid registered by each user one can begin to estimate the maximum or limit price of individuals in the market. Once the limit price of each potential buyer is known it is then an easy step to calculate the demand function - the volume demanded at any given price. The historical record of bids posted in online auctions offers an excellent method of estimating demand.

Unfortunately, however, a problem exists. In order to make an accurate estimation of demand, it is necessary to know the limit price of every individual. Since an auction's bid history only records successively higher bids, it does not



**Figure 1: Simulation of bidder demand using normal distribution of limit prices. Auction rules – new bids must be greater than the current highest – lead to censoring of bids. Thus, observed demand is much lower than the true demand.**

contain a full record of potential bidders. By observing bid history alone, the estimation of demand is likely to be too low.

Estimating market demand using bid history alone does not consider those individuals that arrive at an auction once the auction price has already exceeded their own limit price. Such potential bidders are forced to leave an auction without registering a bid and thus do not appear in an auction's bid history - their bids are censored. This leads to an underestimation of demand (see Fig 1). The problem is how to recover these censored bids in order to form a more accurate estimation of the underlying demand within a market? How to best estimate demand using observed bid history alone? To tackle this, the following section introduces a model to recover censored bids.

## 2. RECOVERING CENSORED BIDS

This section outlines a statistical model for recovering bids that are censored by the auction process - i.e., those bids that would have been submitted had the auction price not already exceeded the limit price of a newly arriving bidder. The model is based on the work of [6] and utilises the fact that an auction bid history displays the arrival time of each submitted bid. The model uses observed arrival times to formulate an estimation of the most likely arrival rate of bids across different bid price levels. These arrival rates are then mapped onto the observed bid history data to give a refined estimation of demand in the marketplace that takes into account not only observed bids, but also those bids that are censored.

### 2.1 Model assumptions

In order to make it easier to work with observed bid history data, let us first segment price into discrete intervals. Then, across all auctions for which we can observe bid histories, consider measuring the time until first arrival of a bid 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 are independent of each other and that all bids are greater than zero, 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 number of bids,  $\lambda_i$ , likely to be posted in each price segment,  $i$ , during a given time interval - this will allow us to evaluate the relative proportion of bidders in each segment and thus the relative proportion of limit prices.

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

- A Bids occur at random in continuous time.
- B Bids occur singly. The probability of two bids arriving simultaneously is zero.
- C Bids occur uniformly, i.e., the expected number of bids in a given interval is proportional to the size of the interval. Arrival rates do not vary over time.<sup>2</sup>
- D Bids occur independently, i.e., the probability of an arrival of a bid with price  $i$  in any small interval is independent of the probability of an arrival of a bid with price  $i$  occurring in any other small interval.

Let us make some further assumptions as to the strategic behaviour of bidders:

- E Bidders bid at exactly their limit price.
- F Bidders attempt to post a bid upon arrival. They will not strategically wait.

And finally, assume the following is true of the auction mechanism:

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

## 2.2 Estimating Bid Arrival Rates

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  $i = 1, 2, \dots, K$ . Then,  $X_i$  is an exponentially distributed continuous random variable with probability distribution function:

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

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

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

<sup>2</sup>In reality, arrival rates rapidly increase towards the end of an auction period as bidders attempt to *snipe*. Preliminary analysis has shown that as many as 25% of all bidders may appear in the final hour. Hence, the model is likely to underestimate arrival rates.

The probability of  $X$  occurring *after* time  $T$  is:

$$\begin{aligned} P[X > T] &= 1 - P(X \leq T) \\ &= 1 - \int_0^T \lambda_i e^{-\lambda_i t} \\ &= 1 - \left[ -e^{-\lambda_i t} \right]_0^T \\ &= 1 + e^{-\lambda_i T} - e^0 \\ &= e^{-\lambda_i T} \end{aligned} \quad (3)$$

We demonstrate how to estimate arrival rates  $\lambda_1, \lambda_2, \dots, \lambda_K$  by using an example that considers only  $n = 2$  bidders in an auction - the result may be generalised to  $n$  bidders.

Assume there are two potential bidders, each with limit prices  $i$  and  $j$  such that  $i < j$  with 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  $\langle x_i, x_j, \dots, x_n : a_t \rangle$  denote the recorded bid history of auction  $a$  with end time  $t$ , then for an auction  $A_T$ , we can calculate the following likelihoods:

- (a) Probability no bidders appear in bid history:

$$\begin{aligned} P[\langle - : A_T \rangle] &= P[(X_i > T) \cap (X_j > T)] \\ &= e^{-\lambda_i T} \cdot e^{-\lambda_j T} \end{aligned}$$

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

$$\begin{aligned} P[\langle x_i : A_T \rangle] &= P[(x_i < X_i \leq x_i + \delta x_i) \cap (X_j > T)] \\ &= \lambda_i e^{-\lambda_i x_i} \delta x_i \cdot e^{-\lambda_j T} \end{aligned}$$

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

$$\begin{aligned} P[\langle x_j : A_T \rangle] &= P[(X_i > X_j) \cap (x_j < X_j \leq x_j + \delta x_j)] \\ &= e^{-\lambda_i x_j} \cdot \lambda_j e^{-\lambda_j x_j} \delta x_j \end{aligned}$$

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

$$\begin{aligned} P[\langle x_i, x_j : A_T \rangle] &= P[(x_i < X_i \leq x_i + \delta x_i) \cap (x_j < X_j \leq x_j + \delta x_j)] \\ &= \lambda_i e^{-\lambda_i x_i} \delta x_i \cdot \lambda_j e^{-\lambda_j x_j} \delta x_j \end{aligned}$$

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

$$\begin{aligned} L(\lambda_i, \lambda_j) &= \lambda_i e^{-\lambda_i x_i} \delta x_i \cdot \lambda_j e^{-\lambda_j x_j} \delta x_j \cdot e^{-\lambda_i T} \\ &\quad \cdot e^{-\lambda_j T} \cdot e^{-\lambda_i x_j} \cdot \lambda_j e^{-\lambda_j x_j} \delta x_j \end{aligned}$$

Taking natural logarithm gives *log-likelihood* function:

$$\begin{aligned} l(\lambda_i, \lambda_j) &= \ln \lambda_i + 2 \ln \lambda_j + \ln \delta x_i + 2 \ln \delta x_j \\ &\quad - \lambda_i x_i + T + x_j - \lambda_j x_j + T + x_j \end{aligned}$$

Then, maximum likelihood values of arrival rates are:

$$\begin{aligned} \frac{\partial l}{\partial \lambda_i} &= \frac{1}{\lambda_i} - (x_i + T + x_j) = 0 \Rightarrow \hat{\lambda}_i = \frac{1}{x_i + T + x_j} \\ \frac{\partial l}{\partial \lambda_j} &= \frac{2}{\lambda_j} - (x_j + T + x_j) = 0 \Rightarrow \hat{\lambda}_j = \frac{2}{x_j + T + x_j} \end{aligned}$$

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  $i$  (1 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  $j$  (2 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:

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

Let us call the divisor in the above equation the effective opening time for bidders of type  $i$  - that is, the total time bidders of type  $i$  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:

$$\hat{\lambda}_i = \frac{\text{total number of bids from type } i \text{ bidders}}{\text{total time that at least one auction is effectively open}}$$

### 2.3 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:

$$\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{r} \quad (3)$$

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:

$$\frac{2n}{\hat{\lambda}_i \cdot \chi_{(2n; \frac{\alpha}{2})}^2} < \frac{1}{\hat{\lambda}_i} < \frac{2n}{\hat{\lambda}_i \cdot \chi_{(2n; 1 - \frac{\alpha}{2})}^2}$$

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

### 2.4 Multiple Bids from Individual Bidders

Once a bidder has placed a proxy bid in an auction, there is nothing to stop them from bidding in other auctions. Indeed, it is likely that once an auction has effectively closed the bidder will move to another auction and place a similar bid. Some bidders have been observed strategically bidding across multiple simultaneous auctions with very small bids, in the hope that no other bidders join the auction. In many cases, we see bidders bidding in multiple auctions. However, unless a bidder wins an auction and then subsequently bids in another auction, we can suppose that each bidder demands only one item. Thus, estimating demand by counting the number of proxy bids alone will likely lead to an overestimation. It is necessary to take account of bidders bidding across multiple auctions.

The model estimates demand by resolving multiple bids across auctions in the following way:

- 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 the bidder ID has yet to appear, i.e., if this is the bidder's first proxy bid, add the bid to the "demand" list. If the bidder ID has already placed a bid (if it appears on the demand list), then replace the bid on the demand list with the following pseudo-bid:
  - (a) Bid time: time of earliest bid (the time the bidder first entered eBay).
  - (b) Bid price: price of highest bid (the limit price of the bidder).
- C For each bin, calculate the mean arrival rate of bidders by dividing the number of pseudo-bids by the total effective opening time.

### 2.5 Model Simulation

Fig 2 shows the true and observed demand of bidders in a simulated eBay market. This data is the same as that shown in Fig 1. The solid red line displays the demand estimated by the model described above. There is clearly a good fit between the estimated and true demand, suggesting that the model works as anticipated; recovering the underlying demand that is censored through the eBay auction process.

## 3. SOFTWARE APPLICATION

In this section, we demonstrate how the model is used on real eBay data, to analyse supply and demand in the marketplace and inform sales strategy. All data is taken from the German eBay auction market of Lexmark X1155 printers, over a 6 month period between Nov 2005 and April 2006.

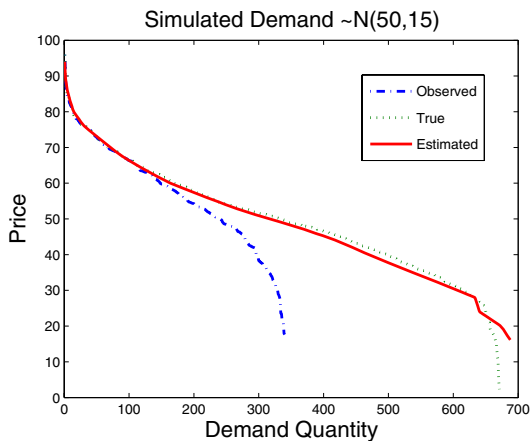


Figure 2: Simulation of bidder demand using normal distribution of limit prices. Observed demand (blue dash) is much lower than true demand (green dot) due to censoring of data. The model estimation of demand (red line) is a good fit of true market demand.



Figure 4: Application Screenshot. Demand (red) and competitive demand (blue dash) over same period as Fig. 4.1. Competitive demand is equal to the excess demand at a given price below the equilibrium price.

demand tends to the demand curve as price tends to zero. The competitive demand curve is used to estimate the price a seller will attain for increasing sales by a given volume.

### 3.2 Revenue and Costs

The application has two revenue models. These are:

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

**Fixed pricing** we assume that all units will be sold at the same fixed price.

The application also has two cost models. These are:

**No reserve pricing** the eBay costs associated with free auction with no reserve listings.

**Buy-it-now** the eBay costs associated with Buy-it-now listings.

Once a revenue and cost model is chosen, the application is then able to calculate the anticipated revenue and costs available to a seller in that market. Fig 5 shows the anticipated revenue and costs for X7170s using fixed price revenue model. These costs include all eBay costs, labour and shipping costs associated with each sale. Costs specific to a seller (such as labour costs) are entered into the model as customized variables for each individual seller. We see that costs (blue dash) rise as volume sold increases.

The revenue curve is calculated using the competitive demand curve (see Fig 4). Anticipated revenue is calculated as the total revenue expected from a given quantity of sales. In figure 4.3, we see anticipated revenue has a maximum at around quantity 600, but then steeply falls. This is because we are using a fixed revenue model. This implies that all units are sold at the same price. Thus, as volume increases and marginal price falls, the price of each unit also falls.

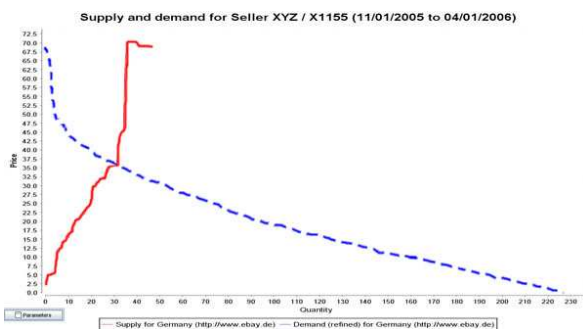


Figure 3: Application Screenshot: Supply (red) and Demand (blue dash) forecast for Lexmark X1155 over 6 month period in Germany.

### 3.1 Competitive Demand

In order to estimate the potential profits available to a seller in the marketplace. We first compute the competitive demand function. This is calculated as the difference between demand and supply (or excess demand) at all prices below the equilibrium price. The competitive demand curve shows the demand available to a seller wanting to push more volume into the market. There is a conservative assumption made here: that all the highest bidders will have already been taken by a competitor. This is the reason for using only the demand below the equilibrium. Having made this assumption, we see that the effect of increasing sales volume is to reduce price.

Fig 3 shows the demand and supply curves estimated by the application for Lexmark X1155 in Germany. Fig 4 shows the demand (red) and competitive demand (blue dash) curves. Here, competitive demand is directly calculable from the demand and supply curves of Fig 3. We see that competitive

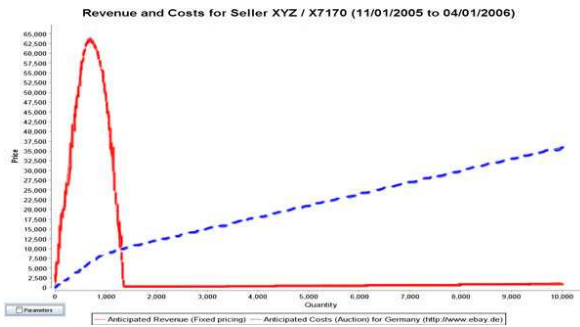


Figure 5: Application screenshot. Revenue and Costs of Lexmark X7170 in Germany with fixed pricing revenue model. Anticipated revenue (red line) quickly rises, then begins to fall around quantity 800. Costs (blue dash) steadily rise



Figure 6: Application screenshot. Revenue and Costs of Lexmark X7170 in Germany with differential pricing revenue model. Anticipated revenue (red line) quickly rises, then bottoms out once quantity reaches approximately 1200. Costs (blue dash) steadily rise.

Fig 6 plots the same data but uses a differential revenue model. We see here that revenue rises steeply until roughly 1200 units but then bottoms out. Since this is a differential revenue model, anticipated revenue never falls. That is because the marginal revenue of a new unit does not impact the price of other sales. This is the major difference between the fixed price and differential revenue models. As such, the fixed revenue model is a much more conservative estimator of revenue.

### 3.3 Profit

From our calculations of revenue and costs, we can plot estimated profits. Profits = revenue - costs. However, the cost curves shown in Fig 6 do not include all the costs that a seller may incur. Other costs may include OEM commission, for example, which are a percentage of sales. These costs are factored in on an individual seller basis. Using these additional costs and the revenue-cost curves shown in Fig 6, the anticipated profit curve can be calculated. Fig 7 shows anticipated profit as a function of quantity. This profit curve uses the same data as Fig 5 and includes a fixed revenue model. It can be seen that maximum profit is approximately 10,000

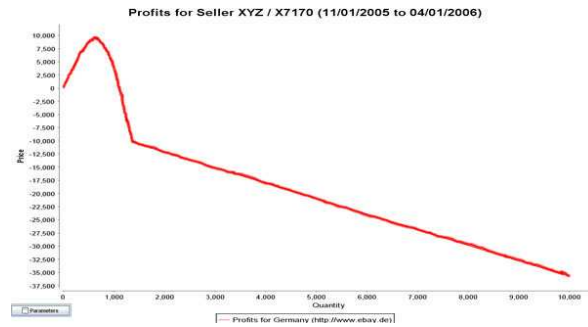


Figure 7: Application screenshot. Potential profits for Lexmark X7170 in Germany. The profits curve has a maximum at quantity 600 and profit 10,000. Profits fall as volume is increased beyond 600 and become a loss one more than 1000 items are sold.

and will be achieved with approximately 600 sales.

Since costs will always rise as volume increases, there will always be a turning point in the profit curve. This turning point is the maximum possible anticipated profit available to a seller for a given product. The maximum profit and volume can then be used to produce summary estimates of current production efficiency.

## 4. FUTURE WORK

We have described the first stages of a research program for automatically calibrating models of consumers from the increasing amounts of large-scale high-frequency data available on consumer transactions and preferences. In this paper we have focussed on calibrating consumers' demand functions for a single commodity. In future work we will apply similar estimation methods to calibrate a more general behavioural model of the consumer based on their transaction history.

A key part of this research will be the development of rigorous methods for estimating simulation models. One of the weaknesses of our existing model is that it makes several simplifying assumptions in order to obtain a closed-form likelihood function. In future work we will modify our simulation model so that incorporates more realistic behaviour, such as sniping, and use heuristic methods such as [5] to calibrate this model against empirical data.

## 5. SUMMARY AND CONCLUSIONS

We have described a statistical model which can estimate the demand for a given product by utilising the bid information that is available from online auction sites such as eBay, and 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 our model.

By producing a full demand curve, our 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.

The model is entirely general. As long as there is bid data available (which is true of all online auction venues, not just eBay), 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.

## 6. REFERENCES

- [1] L. J. Bain and M. Engelhardt. *Introduction to Probability and Mathematical Statistics*. Duxbury Press, 2nd edition, 1992.
- [2] J. Laffont. Game theory and empirical economics: The case for auction data. *European Economic Review*, 41:1–35, 1998.
- [3] W. Ng. Modelling dynamic demand and supply curves of electronic markets. Technical Report WP029-08, CCFEA, 2008.
- [4] K. Sailer. Searching the eBay Marketplace. Discussion Papers in Economics 1234, University of Munich, Department of Economics, 2006.
- [5] P. Winker. Applications of optimization heuristics to estimation and modelling problems. *Computational Statistics & Data Analysis*, 47(2):211–223, September 2004.
- [6] A. Zhang, D. Beyer, J. Ward, T. Liu, A. Karp, K. Guler, S. Jain, and H. Tang. Modeling the price-demand relationship using auction bid data. Technical report, Hewlett-Packard Labs, 2002.