An Empirical Analysis into the Determinants of eBay Auction Market Closing Prices

Conall Halton

Department for Economics

The University of Essex

EC831: Undergraduate Final Project

27/04/2018

Abstract

Using a unique dataset of auctions for the iPhone X we conduct an empirical analysis to understand the main determinants of prices within eBay auctions. We show that reputation and the number of bidders have an influential role in price determination, while auction lengths, minimum bids, and bid timing have no impact in determining prices. We expand traditional analysis by introducing additional unique variables for study. This indicates that description length, using stock photos and delivery estimates also play an important role in determining prices, something previously overlooked. We also demonstrate a method to control for external price determination in our sample due to time, by calculating a days from release variable.

Table of Contents

1 Introduction 2

2 Literature Review 3

3 Theoretical Models 5

3.1 Shepario's Price-Quality Schedule 5

3.2 Akerlof's Market For Lemons 11

4 Data 13

4.1 Data Description 13

4.2 Variable Description 16

4.3 Descriptive Statistics 18

5 Empirical Specification 20

5.1 Model 20

5.2 Adequacy Tests 21

6 Empirical Findings 25

6.1 Results 25

6.2 Comments 27

7 Conclusion 33

8 References 34

9 Appendix 36

1 Introduction

Electronic commerce (e-commerce) has seen significant expansion over the past two decades, with retail e-commerce alone contributing 36bn to the UK economy ONS (2016). A key player in UK online e-commerce is eBay, which ranked as the second largest e-commerce website in the UK by number of visitors Statista (2016). The continued success of eBay has spurred a growing academic interest in the empirical analysis of data which the site produces, posing the possibility to collect detailed microeconomic data across a more diverse array of auction markets (compared to traditional government auctions) and empirically test economic auction theories.

There is a developed range of literature interested in determining the factors which influence prices within eBay auctions, with a range of goods chosen for analysis. Findings are unanimous in agreeing that a positive reputation correlates with higher prices and that a negative reputation results in lower prices, the extent to which however varies widely between papers. Additional variables which impact price have also been studied, although to a lesser extent, the inclusion of item images and increased unique bidders have demonstrated a price premium, whereas auction lengths, alongside relatively underrepresented auction aesthetics, are less conclusive.

Using data describing 432 auctions for the iPhone X we contribute to the literature by examining the impact of these variables on price within our dataset, while also making unique contributions to the literature by proposing additional variables not analysed previously. In doing so, we bring the analysis up to date to represent the aesthetics of the current eBay website, present the first study conducted on the UK version of the website, and are enabled to make inference about the wider determinants of prices on eBay by referencing our results to the previous literature.

Our paper begins by presenting a review of the current literature. We then introduce two economic models which are relevant to our analysis, providing a theoretical framework on the effects of reputation and information asymmetry. An overview of our dataset and data collection method is then provided, with an econometric model then specified and estimated based on this dataset. Findings are then assessed with reference to the current literature and theoretical models outlined.

2 Literature Review

Existing research focused on eBay online auction markets follow a similar empirical method; auction data is extracted from the website for a specific product, with regression analysis then used to determine the impact of variables on the closing price.

A common theme throughout papers is investigating the impact of eBay's buyer-seller reputation system on closing price. The main divergence between papers then occurs in the additional variables which are studied and the product selected for analysis, a relatively homogeneous (new item) or heterogeneous (used item) provides a general classification for products analysed.

Lucking-Reiley et al. (2007) base their study on 461 auctions for pennies sold on eBay US between July and August 1999, extracting data using a "spider-program". They show that the feedback ratings of sellers have a measurable effect on final auction prices. In particular, they discover that negative feedback ratings relate to a statistically significant decrease in price, whereas positive feedback ratings increase the final auction price, the later finding with low statistical significance. Furthermore, they include variables which show that higher minimum bids have a positive impact on the final auction price and also show that longer auctions lead to higher prices on average.

Eaton (2005) also analyses the impact of negative seller feedback on the final sales price, examining four models of used Paul Reed Smith guitars across 361 individual eBay auctions. The relative heterogeneity of the good analysed implies greater information asymmetry in the auctions as compared to Lucking-Reiley et al. (2007), especially as coins come with independent certificates of grade. Eaton (2005) proposes there is therefore a greater need to signal "the quality of the instrument and also the quality of the seller" within these listings.

Concurring with Lucking-Reiley et al. (2007), Eaton (2005) finds negative causation between the existence of negative feedback and the auctions closing price. As an additional finding, Eaton (2005) finds the interaction between negative feedback with whether an auction included pictures to have a positive coefficient across all guitar models. This suggests that aspects of the auction that reduce information asymmetry between buyers and seller, such as the inclusion of images, results in an increase in closing prices.

Bajari and Hortaçsu (2003) focus their study on US coin auctions using a dataset of 407 individual auctions. They capture an "overall reputation" variable which is positive minus negative feedback, this variable exhibits a positive impact on price, supporting the findings of Lucking-Reiley et al. (2007) and Eaton (2005) regarding reputation. Within their analysis they also investigate the number of unique bidders who participate in the auction; they find a positive, statistically significant effect on price through this variable. As with Lucking-Reiley et al. (2007), minimum bids are also investigated, findings concur, with a higher minimum bid showing a positive impact on closing price. To extend this analysis, Bajari and Hortaçsu (2003) also demonstrate that higher minimum bids discourage bidder entry, which if too high could result in a non-sale.

Houser and Wooders (2006) dedicate their paper examining 94 auctions of Intel Pentium III 500 Mhz processors to the effects of reputation. Their analysis affirms the idea that positive reputation leads to an increase in sales price, while negative reputation leads to a decrease.

More recent analysis has been conducted by Depken and Gregorious (2010), they select a brand new, network locked, iPhone 4, 8GB model for analysis, with data based on 192 auctions. Again, they find that positive reputation correlates positively with price, although they find a significantly lower price premium than other papers. Depken and Gregorious (2010) fail to investigate the impact of a seller exhibiting negative feedback, instead looking at the impact of a seller exhibiting no feedback at all; this unique contribution finds that no feedback will decrease the auction price on average by 25%, a similar impact to the existence of negative reputation.

In a similar fashion to Lucking-Reiley et al. (2007), Depken and Gregorious (2010) investigate the impact of differing auction lengths on price. In contrast to Lucking-Reiley et al. (2007), who found an increase in price as the auction length increases, Depken and Gregorious (2010) find no similar result, with longer listings reducing prices. Potential explanations for this anomaly between papers may be due to the limited sample size that Depken and Gregorious (2010) analyse, with only 192 auctions. The literature is less consistent on this topic leaving scope for further investigation.

The most unique contribution from Depken and Gregorious (2010) surrounds greater analysis into auction aesthetics which are notably vacant from preceding literature (excluding the brief analysis from Eaton (2005) on the use of item images). Depken and Gregorious (2010) investigate a number of variables, such as bold text in the title and the use of eBay's premium gallery display image option, finding no statistical significance from these variables. On the other hand, the variables for the inclusion of a premium border and listing title length do show a statistically significant impact on closing prices; increasing and decreasing closing prices respectively. The findings of Depken and Gregorious (2010) surrounding reputation seem robust and correlate with other literature on the topic. Their results on auction characteristics seem less robust however, especially as there is limited literature available for comparison on this topic.

3 Theoretical Models

Two theoretical models are derived which are relevant to our empirical analysis. The first model outlined by Shepario (1983), focuses on the reputation of the seller. The second model, an adaptation of "Market for Lemons" by Akerlof (1970), examines how prices are determined in a market with the presence of incomplete information. The predictions of these economic models will be assessed via our empirical analysis.

3.1 Shepario's Price-Quality Schedule

Our theoretical interest in reputation relates to the price implications of reputation. Shepario (1983) develops a model which explores the price implications of firm-specific reputations in a perfectly competitive environment with imperfect information. Although sellers on eBay may not be viewed in the traditional sense of a firm, their role in selling products and maintaining a reputation endow them with many similar characteristics, making the modelling undertaken by Shepario (1983) relevant and applicable. The model derived by Shepario (1983) is outlined below, amended to account for our non-traditional take on the firm as a seller on eBay.

The price a seller can charge is determined by his reputation, R:

$$R = p(R)$$

The cost of production, C, is dependent on quality, q:

$$C = c(q)$$

It is assumed that $C' > 0 \ and \ C'' > 0$. Sellers choose quality over time to maximize the present value of their profits, we view quality as a seller providing a quality service within their auction, such as selling a genuine item, shipping the item on time and not behaving in a dishonest manner, rather than just the physical qualities of the good sold.

The formation of reputation, R, is an equation based on quality from the previous period:

$$R = q_{t-1} \tag{1}$$

This reflects the fact that quality cannot be observed prior to purchase by buyers, and hence sellers can, at least for one period, cheat on their customers by reducing quality.

It is assumed that a minimum quality, q_0, is imposed. This may be given several interpretations, but the simplest is that it is prohibited to sell items below quality q_0. In our setting we can view this minimum quality as regulation imposed by eBay to protect buyers from dishonest sellers. A firm has a good reputation if consumers believe their products to be of high-quality $q > q_0$.

Finally, entry is permitted, but new sellers must prove themselves in order to build up a reputation. Initially, they must therefore sell their product at price $p(q_0)$.

Equilibrium in this model is a price-quality (or, equivalently, a price-reputation) schedule $p(q)$ such that:

A. Each consumer, knowing $p(q)$, chooses their most preferred product on the schedule to consume
B. Markets clear at every quality level (this determines the number of active sellers in equilibrium)

C. A seller with reputation R finds it optimal to produce quality $q = R$ rather than to deviate (that is, consumers' expectations regarding quality are fulfilled)

D. No new entry is attractive

Derivation of the equilibrium schedule begins with condition C; this explains how a seller with reputation q will not want to cheat its reputation. One way in which a seller in the market could cheat on its reputation is to cut quality to the minimum, take short-run gains, and exit the market. This would yield cheating profits π_c of:

$$\pi_c = p(q) - c(q_0)$$

The alternative strategy is to behave and maintain quality q forever, yielding present value behaving profits π_b of:

$$\pi_b = \frac{1+r}{r}\big(p(q) - c(q)\big)$$

In order for cheating to be unattractive we must have that:

$$\pi_b \geq \pi_c$$

$$\Rightarrow \quad \tfrac{1+r}{r}(p(q) - c(q)) \geq p(q) - c(q_0)$$

$$\Rightarrow \quad p(q) \geq c(q) + r\big(c(q) - c(q_0)\big) \tag{2}$$

Price must exceed cost to prevent the existing seller from cheating and degrading quality for short-term gains. This condition puts a lower bound on the price of items of a certain quality.

Moving to condition D, which explains that in equilibrium new entry must be unattractive (non-positive profit). The profit of a new entrant who maintains quality q forever is:

$$\pi_n = p(q_0) - c(q_0) + \tfrac{1}{r}(p(q) - c(q))$$

Condition D therefore states that:

$$p(q_0) - c(q_0) + \tfrac{1}{r}\big(p(q) - c(q)\big) \leq 0$$

$$\Rightarrow \quad p(q) \leq c(q) + r\big(c(q) - p(q_0)\big) \tag{3}$$

Finally, it must be that:

$$p(q_0) = c(q_0) \quad (4)$$

As:

- If $p(q_0) < c(q_0)$, no seller would supply quality q_0
- If $p(q_0) > c(q_0)$ any new entrant could profitably undercut sellers of quality q_0 by simply offering a product/service of quality q_0 at a price between $p(q_0)$ and $c(q_0)$, since consumers of quality q_0 know they will not face lower quality than q_0, they will be happy to buy from new entrants at the lower price

Equation (4) is equivalent to stating that new sellers begin with a reputation $R = q_0$ before they undertake establishing a reputation, only being able to command prices of items at minimum quality.

By substituting (4) into (3) we obtain:

$$p(q) \leq c(q) + r\big(c(q) - c(q_0)\big)$$

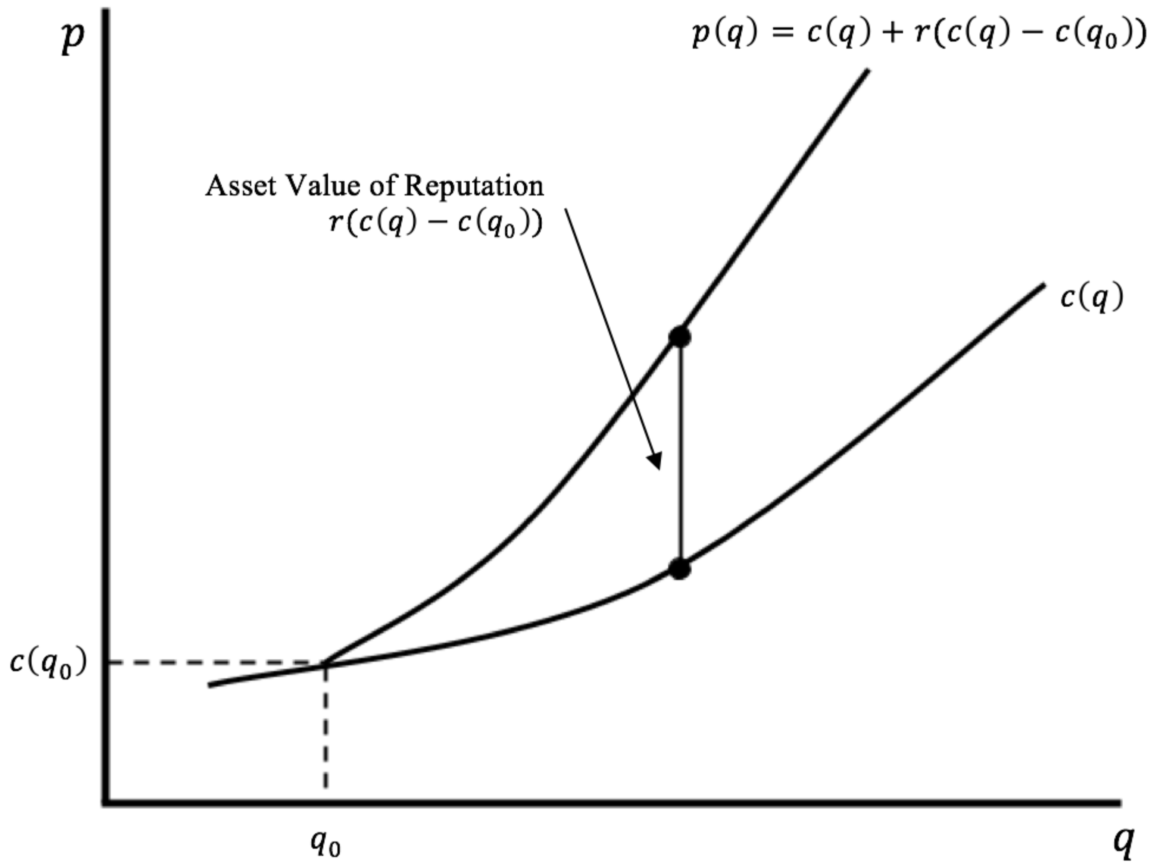
We see that this equation is the reverse inequality of (2). These two equations combined fully determine the price-quality schedule:

$$p(q) = c(q) + r(c(q) - c(q_0)) \quad (5)$$

The price-quality schedule demonstrates that the cost of providing items of quality q is the per-unit production cost $c(q)$, plus a onetime information cost $c(q) - c(q_0)$. Which can be viewed as the cost of establishing a reputation for quality q.

Graph 1 displays the price-quality schedule $p(q)$ alongside the schedule that would prevail under perfect information $c(q)$:

GRAPH 1 – PRICE-QUANTITY SCHEDULE – $p(\mathbf{q})$



The graph demonstrates equation (4), where new sellers beginning with a reputation $R = q_0$ can command only prices of items at the minimum quality; shown by the point $(c(q_0), q_0)$. Divergence then occurs as sellers with $R = q > q_0$ have the ability to command higher prices.

Notice that these higher prices, which we will define as the asset value of reputation q, $v(R_q)$, can be defined as:

$$v(R_q) = r(c(q) - c(q_0)) \tag{6}$$

The value attributed to reputation is the difference in cost of a seller maintaining a high-quality reputation, q, as compared to the cost of a reputation/quality standard q_0, which earns a rate of return, r, in competitive equilibrium.

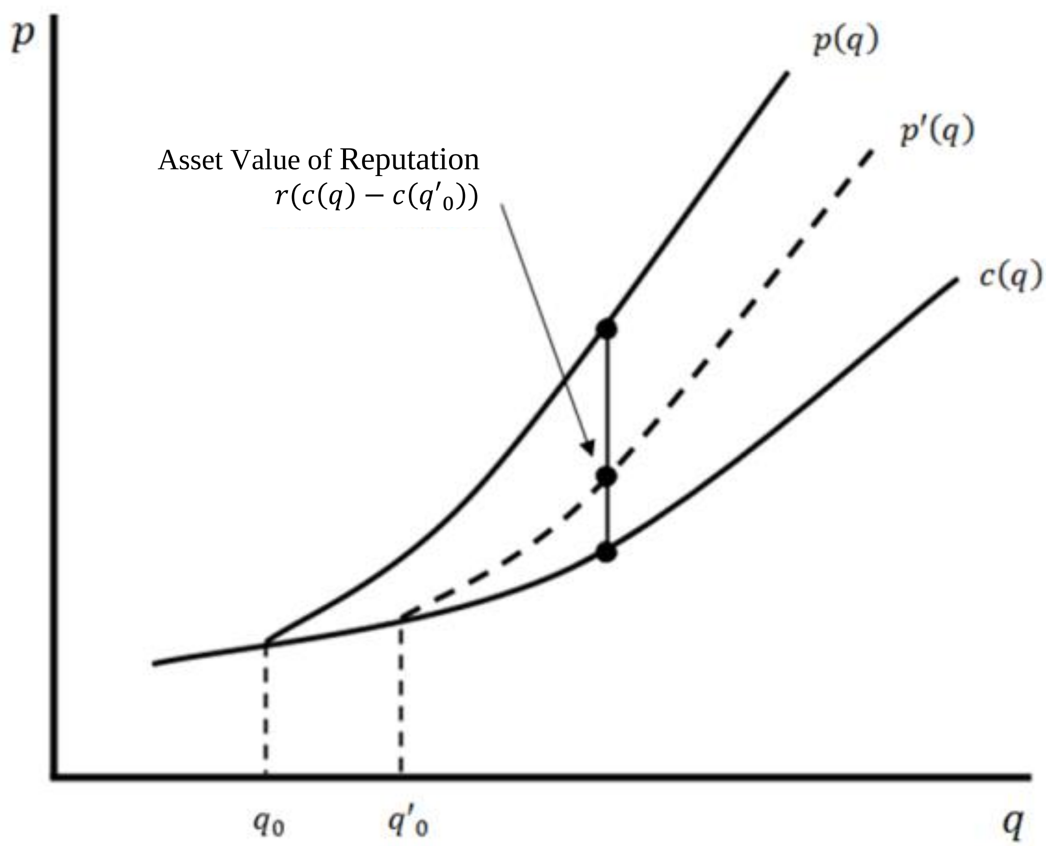
This indicates that sellers who establish a positive reputation, q, will see a price premium over less established sellers; shown on the graph as the vertical difference between $c(q)$ and $p(q)$. These premiums represent only a fair rate of return on the investment the seller has made in their reputation i.e. not a supernormal profit.

With reference to equation (5), we can see that not only does the price-quality schedule depend on q, but also on q_0, the minimum quality standard set. In assessing an increase in the minimum quality standard to $q'_0 > q$ we can define a new price-quality schedule $p'(q)$:

$$p'(q) = c(q) + r(c(q) - c(q'_0)) \tag{7}$$

Graph 2 displays the $p'(q)$ schedule, alongside the original schedule $p(q)$ and the $c(q)$ schedule:

GRAPH 2 – PRICE-QUANTITY SCHEDULE – $p'(q)$



An increase in q_0 shifts the $p(q)$ schedule to the right, this increase raises the minimum price at which new entrants to the market can charge. In doing so the investment required to establish a good reputation is reduced, eroding away some of the premiums attributed to the asset value of reputation. This decrease is demonstrated by the reduction in vertical distance between the $p(q)$ and (q) schedule.

The equilibrium price-quality model proposed by Shepario (1983) has important predictions for our empirical analysis. We should expect that sellers who have invested in establishing a positive reputation (offering a higher quality than q_0), should be rewarded a premium through higher prices, representative of the sellers asset value of reputation $(v(R_q))$.

The size of this premium will be dependent on the minimum quality standard (q_0) which is imposed by eBay, with higher quality standards reducing the asset value of reputation $(v(R_q))$. In an extreme case, with strict quality standards imposed such that $p(q) = c(q_0)$, we expect to see no premium for the existence of a quality, q, reputation.

3.2 Akerlof's Market For Lemons

The presence of information asymmetry and its effects have been well documented in economic literature. One such paper, Akerlof (1970), the "Market for Lemons", demonstrates how incomplete information can cause low prices to crowd out quality goods within a market.

The auctions we analyse are beset by information asymmetry; buyers are unable to gather complete information about the products being offered, with sellers holding significantly more information on the quality of the product offered. Using an adapted version of the "Market for Lemons", we develop a theoretical understanding of the impact that information asymmetry can have on prices within eBay auctions. The model is an adaptation of Shy (1995) in his derivation of the "Market for Lemons".

A consumer has two choices, purchasing a new iPhone or a used iPhone from the market. Each purchase comes with a degree of uncertainty surrounding the quality of the phone:

N^G = Value of new, good quality iPhone, occurring with probability α

N^B = Value of new, bad quality iPhone, occurring with probability $1 - \alpha$

U^G = Value of used, good quality iPhone, occurring with probability β

U^B = Value of used, bad quality iPhone, occurring with probability $1 - \beta$

One may expect that purchasing a new phone would come with no uncertainty about quality, implying $\alpha = 1$. However, due to the nature of eBay, there may still be uncertainty around quality, such as if the new iPhone sold is authentic (which would imply bad quality).

We can expect with certainty that $\alpha > \beta$, as common quality faults in iPhones such as a depleted battery are much more likely in used iPhones which develop over time with use.

The expected utilities relating to the purchase of a new, $EU(N)$, and used, $EU(U)$, iPhone are defined as:

$$EU(N) = \alpha N^G + (1 - \alpha)N^B$$

$$EU(U) = \beta U^G + (1 - \beta)U^B$$

Assume that $N^B = U^B = 0$, and let p^N and p^Urepresent the prices of new and used iPhones, respectively.

A buyer will therefore be indifferent between purchasing a new and a used iPhone given that:

$$EU(N) - p^N = EU(U) - p^U$$

$$\Rightarrow \qquad p^N = EU(N) - (EU(U) - p^U)$$

$$\Rightarrow \qquad p^N = \alpha N^G - (\beta U^G - p^U) \qquad (8)$$

Equation (8) demonstrates that the price of a new iPhone will increase as α, the probability of a new, good quality iPhone, increases. It is apparent then that the accuracy of the buyers' ability to determine the value of α will play a central role in the final pricing of the iPhone being sold.

In order to raise the price of the iPhone for sale, sellers are expected to undertake in activities which better inform consumers about α. In doing so sellers erode away information asymmetry, allowing the buyer to make an informed assessment about the quality of the iPhone. An example of which is including images of the item.

This theoretical model predicts that any activity undertaken within an eBay auction which reduces information asymmetry and increases the buyers' belief that the iPhone is of good quality, α, will result in higher prices.

The extent of this effect, which will be investigated via our empirical analysis, will depend on pre-existing beliefs that buyers have regarding the quality of new iPhones sold on eBay. Under the intense assumption of $\alpha = 1$, where buyers are always certain of quality, any measures by the seller to reduce information asymmetry would be redundant in increasing prices.

4 Data

4.1 Data Description

To conduct our analysis, observations were collected from the online auction website eBay UK during the 15-week period between 20th November 2017 and 2nd March 2018. The item of focus is a brand new, network unlocked, space grey, iPhone X, 256GB model, which was released on 3rd November 2017. In total 432 observations are included in the cross-sectional dataset, with each observation detailing a unique auction for this item which finished during the specified time period (a detailed description of the individual variables which make up an observation are covered in section 4.2).

A manual extraction method was used to remove data, relying on the "completed listings" section of the website; displaying the three month sales history of the queried item from the current search date. When accessed, each completed listing displays all relevant information required for the analysis, including structural auction characteristics, factual information regarding the bidders and seller involved (including a link the their "feedback profile"), and information on their behaviour during the auction.

The item we select for analysis follows closely to Depken and Gregorious (2010). It seems like a natural progression of the literature to analyse an updated version of this item within our analysis, as not only does this item have useful characteristics which simplify the data collection and analysis, it will also allow for a closer comparison between our findings and the most recent paper in the literature. Unlike Depken and Gregorious (2010) the unlocked version of the iPhone has been chosen for this analysis, due to the limited observations available for a locked version (likely due to the close proximity to the release date of this phone).

The most useful characteristic of the item is its homogeneity; ensuring that every auction analysed is for an identical product. In doing so, item uniformity is ensured across all auctions, which means price differentials will not occur from the condition of the item. This removes the necessity to obtain and include the book value of each item sold, a process which Lucking-Reiley et al. (2007) undertook in their study. Our choice of a homogeneous good is in contrast to Eaton (2005) who purposely selected a heterogeneous used good, he stipulates that prices in auctions he analyses will be more responsive to parameters which convey information to the buyer, due to the existence of greater information asymmetry for the items he analyses.

Although time-consuming, and restrictive on sample size when contrasted to an automated exaction process such as the one used by Lucking-Reiley et al. (2007), a manual method of data extraction allowed for a more accurate sample to be collected. It was also better suited to the relatively small number of completed listings available for analysis, especially when compared to Lucking-Reiley et al. (2007) who initially needed to survey over 20,000 completed auctions for pennies. The improved accuracy of this process stems from the need for human interpretation throughout much of the extraction process; examples of which are covered in section 4.2. Without this human interrogation we could introduce inaccuracy and bias into our sample, at a trade-off to an increased sample size. Manual extraction has been employed by many others in the literature, including Eaton (2005) and Depken and Gregorious (2010), and is well suited for the product being analysed, creating a refined and accurate dataset.

A unique feature of our data which is unseen in the current literature is that our data comes from the UK version of the eBay website, while all previous analysis has been conducted on data from US version of the website. The dynamics and setup of the UK website are identical to that of the US, with the only differences being prices listed in pounds, and domestic sellers being located in the UK. By using data based on the UK website we will therefore be making a unique contribution to the literature, helping to ascertain if findings from the US are consistent with that of the UK when investigating the determinants of price.

A final descriptive comment on the dataset relates to its sample size. In total 432 observations were captured, this is a relative strength of our data when compared to other papers; of papers in the literature base which use a manual extraction method the average sample size is only 264 observations. Having a larger sample size allows us to make more powerful inference about the true population values and will strengthen the reliability of results within the paper. Having captured the largest sample of all similar papers, we can be confident that our sample size is statistically adequate, giving confidence when interpreting our outputs.

Continues on next page.

4.2 Variable Description

A description of each variable which contributes to an observation is displayed in table 1. Indicator variables are binary variables which take the value of 1 to indicate the presence of the effect and 0 to indicate its absence.

TABLE 1 – VARIABLE DESCRIPTION

VARIABLE	DESCRIPTION
	CONTINUOUS
WIN_BID	Winning bid of the auction i.e. price the item sold for (£)
WIN_BID_SEC	Seconds between the listing end time and placement of the winning bid
WIN_BID_FEED	Total feedback of the winning bidder
NUM_BIDR	Number of bidders participating in the auction
NUM_BIDS	Number of bids placed in the auction
DAYS_RELEASE	Number of days from the launch of the item for sale
MIN_BID	Minimum bid set by the seller (£)
TITLE_CHAR	Number of characters in the title of the listing
DESC_CHAR	Number of characters in the description of the listing
NUM_PHOTO	Number of photos included in the listing
POST_COST	Cost of postage for the item (£)
DELIV_EST	Estimated delivery time in days
POS_FEED	Positive seller feedback received in the past 12 months
NEG_FEED	Negative seller feedback received in the past 12 months
NUM_ON_DAY	Number of auctions for the same item which ended on the same day as the observation
	INDICATOR
RES_SALE	If the auction resulted in a sale
RELIST	If the item was relisted
DAYS1/3/5/7/10	If the auction duration was one, three, five, seven or ten days
SUB	If the listing included a subtitle
RET_ACCEPT	If the seller accepted returns
STOCK	If the listing uses stock photos
COLLECT	If the seller specifies collection only
GLOBAL	If the seller offers global shipping
BUSI_SELL	If the seller is registered as a business seller
DOW_WEEKDAY	If the auction ended on a weekday

We present an extensive list of variables which data is collected on; this covers variables which are covered in the literature base, alongside additional unique variables we introduce for analysis (*DESC_CHAR, SUB, RET_ACCEPT, STOCK, DELIV_EST, GLOBAL, BUSI_SELL, COLLECT, RES_SALE and NUM_ON_DAY*). Information on a majority of the variables is extracted directly from the "completed listing" page on the website without amendment, however, some variables require human interrogation as they are not accurately displayed within these pages.

Feedback of the seller, *POS_FEED* and *NEG_FEED*, constitute the feedback percentage of the seller, calculated using the following formula: $\frac{Positive\ Feedback}{Positive\ Feedback + Negative\ Feedback}$. This is the primary way buyers become informed about the "quality" of the seller. As an example, a seller with 10 positive ratings and 2 negative would receive a feedback percentage of 83%. Houser and Wooders (2006) highlight an issue with extracting this data directly, as "eBay updates feedback profiles in real time, so a user's profile at the time we collect this data will not be the same as his profile at the time the auction ended if, in the interim, he has received additional feedback". We use the same method to correct for this as Houser and Wooders (2006); each seller's feedback profile was used to manually trace back more recent negative and positive feedback acquired since the end of the auction, and used to work out the true feedback values which bidders would have seen when the auction closed. This gives a more accurate representation of the feedback values buyers would have used when making their bidding decision.

Further variables of comment are *RES_SALE* and *RELIST*, a listing can fail to result in a sale for two reasons, receiving no bids, or completing and the buyer fails to pay/seller refuses payment. Auctions can often initially look like a sale has taken place however the transaction between buyer and seller hasn't occurred, the website does not explicitly display information regarding this and a manual classification is therefore required. The primary indicator that a listing didn't result in a sale is the item being relisted; this is displayed on the "completed listing" item page. All sellers have the option to use this "relist" feature, investigations must then be completed to establish if the item was relisted because the buyer failed to pay (usually indicated by negative feedback for that transaction, or a description on the relisted item containing "relisted due to time wasters"), or if the seller has an additional item for sale. Once this has been established an accurate classification on *RES_SALE* and *RELIST* can be conducted.

These amendments provide support to the claim made in section 4.1 regarding a manual extraction method being better suited; as these amendments, which improve accuracy, would be difficult to classify without human interpretation. Our data collection process and variable choice therefore demonstrate a comprehensive look into all auctions for the item in question, giving us a rich dataset with the ability to conduct a detailed and unique analysis.

4.3 Descriptive Statistics

Table 2 presents summary statistics for the variables used in our analysis. Two observations were excluded as they contained variables with extreme outliers when compared to the mean.

TABLE 2 – DESCRIPTIVE STATISTICS

VARIABLE	N. Obs.	Mean	Std. Dev.	Min	Max
WIN_BID	424	996.70	89.80	825	1,319
WIN_BID_SEC	424	25,166	82,944	0	604,800
WIN_BID_FEED	403	291.10	783.55	0	10,820
NUM_BIDR	432	9.206	5.106	0	25
NUM_BIDS	432	18.70	13.90	0	70
DAYS_RELEASE	432	61.52	29.75	13	115
MIN_BID	432	470.1	353.9	0.01	1,300
TITLE_CHAR	432	60.36	10.72	25	91
DESC_CHAR	432	365.2	675.0	0	10,922
NUM_PHOTO	432	2.394	1.486	1	7
POST_COST	412	5.098	3.830	0	9.5
DELIV_EST	412	3.951	1.189	2	9
POS_FEED	432	45.09	115.8	0	1,356
NEG_FEED	432	0.231	0.706	0	8
NUM_ON_DAY	432	5.692	2.510	1	12
RES_SALE	432	0.861	0.346	0	1
RELIST	432	0.0995	0.300	0	1
DAYS1	432	0.359	0.480	0	1
DAYS3	432	0.326	0.469	0	1
DAYS5	432	0.113	0.317	0	1
DAYS7	432	0.169	0.375	0	1
DAYS10	432	0.0324	0.177	0	1
SUB	432	0.0301	0.171	0	1
RET_ACCEPT	432	0.197	0.398	0	1
STOCK	432	0.169	0.375	0	1
COLLECT	432	0.0463	0.210	0	1
GLOBAL	432	0.257	0.437	0	1
BUSI_SELL	432	0.0671	0.251	0	1
DOW_WEEKDAY	432	0.704	0.457	0	1

Of all 432 auctions 424 received at least one bid with an average bid of £996.70, not all auctions received a bid, which can be seen by referencing the number of bids variable which has a minimum value of 0. There is a large variation in the number of seconds between when the winning bid was placed and the end of the auction, the shortest being within the final second before the end of the auction, and longest a week before the end of the auction. The minimum bids set by sellers ranged from £0.01 up to £1300.

There is only small variation in the number of characters in the title, this is likely due to the limit eBay imposes on title length. On the other hand, there is a wide spread in the number of characters in the description, on which eBay proposes no limit. The number of photos a seller uploads ranges from 1-7, this is a relatively restricted range and can be attributed to the good we have chosen to analyse; with a homogeneous good requiring less descriptive pictures. Only 17% of listings used stock photos rather than images of the item taken in person, while only 3% of auctions utilised the subtitle feature which comes at an additional cost to the seller.

Sellers indicating that their listing was available only for collection made up 5% of observations. This has implications on variables which capture information regarding postage, as missing values are generated for these variables when only collection is offered; seen by only 412 entries captured for the delivery estimate variable. Interestingly 25% of auctions offered global shipping, this practice is actively encouraged by eBay and sellers are regularly informed of the benefits of offering global shipping through messages from eBay.

Positive feedback of sellers has a much higher variation than that of negative, this is likely due to the smaller scale in which negative feedback is represented on; with a minimum and maximum negative value of 0 and 12, contrasted to a minimum and maximum value of 0 and 1356 of positive. 6% of auctions in our sample were posted by business sellers, an interesting finding within the data is that characteristics which are mandatory to business sellers (such as offering returns) are not highly correlated with business sellers exclusively, indicating that individual sellers are offering similar services to their business counterparts.

The reduced number of observations for the winning bidder's feedback (403) is due to sellers selecting "private listing". For listings such as this we could not capture this value and a missing value is generated. This choice appears to be made at random by sellers in our sample.

Shorter auctions are most common in our sample, with 36% of the auctions being for one-day and 32% at three-day. On average 6 auctions close each day, however, on some days only 1 auction closed, while others had as many as 12 auctions closing for the item. 70% of sampled auctions ended on a weekday.

Of all auctions sampled, 86% resulted in a sale; where a transaction took place between buyer and seller. Auctions that were relisted after finishing sits at 10% of all auctions sampled. These results indicate that some sellers are deterred from selling their item via eBay when a sale falls through – approximately 4% of auctions which fail to result in a sale are not relisted. There is a possibility however that some auctions slipped through the manual classification process for the *RELIST* variable, as the seller may not have relisted the item directly from the same, non-sale auction.

5 Empirical Specification

5.1 Model

We outline and explain the central multiple linear regression (MLR) model which will be estimated using ordinary least squares (OLS) during our analysis. The model is an extension of that proposed by Depken and Gregorious (2010); their empirical specification relates the winning bid of an auction to a number of explanatory variables, representing buyer, seller and auction characteristics. Our model takes a similar format and is of the form:

$$\begin{aligned} WIN\_BID = \beta_0 &+ \beta_1 WIN\_BID\_SEC + \beta_2 WIN\_BID\_FEED + \beta_3 NUM\_BIDS \\ &+ \beta_4 \mathrm{DAYS\_RELEASE} + \beta_5 DAYS1 + \beta_6 DAYS3 + \beta_7 DAYS5 + \beta_8 DAYS7 \\ &+ \beta_9 DAYS10 + \beta_{10} DOW\_WEEKDAY + \beta_{11} MIN\_BID + \beta_{12} DESC\_CHAR \\ &+ \beta_{13} TITLE\_CHAR + \beta_{14} SUB + \beta_{15} RET\_ACCEPT + \beta_{16} STOCK \\ &+ \beta_{17} NUM\_PHOTO + \beta_{18} POST\_COST + \beta_{19} DELIV\_EST + \beta_{20} GLOBAL \\ &+ \beta_{21} POS\_FEED + \beta_{22} NEG\_FEED + \beta_{23} BUSI\_SELL \\ &+ \beta_{24} NUM\_ON\_DAY + \beta_{25} COLLECT + u \end{aligned}$$

Here *WIN_BID* is the dependent variable, which depends upon twenty-five independent explanatory variables and corresponding coefficient parameters β_{1-25}, a constant term β_0, and an unobserved error term u.

The specification of this model allows us to investigate the size, direction, and significance of variables in determining auction market closing prices within our sample. All variables which have reoccurred in the literature base are included, alongside the additional unique variables which we outlined in section 4.2; supporting both our unique contribution to the literature and bolstering the explanatory power of our model.

The inclusion of *DAYS_RELEASE* in our model differs in justification when compared to the additional unique variables we include. This variable allows us to control for external depreciation of the iPhone, unrelated to factors contained within eBay auctions. This depreciation occurs as close to the release of the phone, the resale price initially increases due to the limited supply, however falls quickly as more products enter circulation (shown in appendix A). The inclusion of this variable therefore controls for this external influence in our model, allowing us to estimate the internal influence of eBay auction market characteristics with more accuracy, by controlling for time depreciation. This variable would also be applicable to studies of extended length, as items bought for investment (such as coins) may appreciate externally over the period of study.

5.2 Adequacy Tests

Before estimating our model, we undertake the preliminary activity of ensuring it is properly specified. The choice to use OLS to estimate our model relies upon the fact that OLS is the best linear unbiased estimator (BLUE) among the class of all linear estimators, so long as the Guass-Markov assumptions are satisfied Wooldridge (2012). The Guass-Markov assumptions are: linear in parameters, random sampling, no perfect collinearity, zero conditional mean and homoscedasticity (MLR.1-5, respectively).

Conditional on these assumptions being met, using OLS to estimate our model will provide unbiased ($E(\hat{\beta}) = \beta$) and efficient (lowest variance) estimates of the parameters β_{0-25}. Once combined with the additional assumption of normality (MLR.6), we are enabled to construct accurate confidence intervals and conduct hypothesis tests on these estimated parameters. The combined assumptions (MLR.1-6) are known as the classical linear model (CLM) assumptions.

Failure of any of the CLM assumptions could introduce bias and inconsistency into our results, or cause OLS to no longer be the best estimator. It is therefore essential that we investigate our model against these assumptions and take any corrective action necessary to ensure they are satisfied pre-estimation.

The random sampling assumption MLR.2 is of important consideration. Our sample selection method represents a random sample of observations for auctions of the iPhone X, with a large sample size enabling us to be representative of the population of auctions for the iPhone X on eBay; allowing for our inference to be relevant and unbiased. However, our main interest relates to the wider population of all eBay auctions, and for this our sample may be considered non-random as it is a specific subset of the entire population, being oversampled. We must then be mindful about inference in our results section, as to distinguish between variables which impact price in auctions for the sampled item, and the more general impacts on prices on the wider eBay website. The latter may only be achieved by comparing results to the literature base and searching for consensus, as inference from the non-random sample would create bias and inconsistent results with respect to the larger population.

Collinearity of variables was detected in our model using a correlation matrix and variance inflation factors. This process highlighted the variables *NUM_BIDS* and *NUM_BIDRS* as being highly correlated (with a correlation value of 0.86 and VIF of 5.58 and 5.45 respectively), demonstrated in appendix B. To avoid introducing imperfectly collinear variables into our model only *NUM_BIDS* has been included in our regression. This process will not harm the explanatory power of the model, as these variables are capturing similar influence, but will allow us to avoid bias in our estimators.

Measures to avoid violating MLR.3 relate to dummy variables used in our model, specifically *DAYS1-10*, estimating the model in its current form would introduce perfect collinearity into our model via the dummy variable trap; where the set of dummy variables is so highly collinear with each other that OLS cannot identify the parameters of the model. When estimating the model, we will omit *DAYS1*, this will allow us to understand the impact of auction length on price with respect to a one-day listing, with this impact then being captured in the constant.

A final adjustment must be taken with regards to the *COLLECT* dummy variable, this variable has a perfectly collinear relationship between postage variables *POST_COST and DELIV_EST;* as listings which offer collection only all have missing values for these variables (not offering postage at all). To avoid violating MLR.3 we choose to omit the *COLLECT* variable in our central model. This has the consequence of removing collection only observations from this model, supporting OLS to be unbiased, however as a consequence only investigating auctions which post the item. To ensure we do not leave the effect of offering collection only uninvestigated we present a variant of the above model in section 6.

An additional serious violation of the CLM assumptions could arise in our model due to endogeneity (MLR.4, zero conditional mean), where an explanatory variable is correlated with the error term. There is little precedent set within the literature regarding the endogeneity of variables, but we explain possible ways it could arise in our estimation. The omission of a relevant variable in our estimation is the most likely cause of endogeneity, causing the error term (containing this omitted relevant variable) to correlate with the explanatory and dependent variables. By ensuring we have included all variables which have been studied in the literature base, alongside including our own additional variables, we aim to have been able to control for enough factors to assume that those that are left in the error are unrelated to the explanatory variables, as guided by the literature; resulting in MLR.4 being satisfied.

However, it is still possible that relevant variables could have been excluded in our model, indeed it is rare for strict exogeneity to hold in data from a non-ideal sample (not an experiment in which all factors can be controlled for). Possible remaining variables which could cause endogeneity in our model may relate to characteristics of auctions which are close complements to the iPhone X; such as the number of alternative phones for auction that day. We could hypothesise that these variables may negatively correlate with *NUM_BIDS* and *WIN_BID*, as those distracted by auctions for other phones may not bid within our observation, which may cause the error term to be correlated with *NUM_BIDS*. Unfortunately, capturing much broader data such as this is beyond the scope of our paper, but we can be confident that we have included all explanatory variables contained in the literature, as well as those directly related to the dependent variable in our model. We therefore must simply acknowledge that the presence of endogeneity is still a possibility, which would result in bias and inconsistent estimators.

The assumption of homoscedasticity (MLR.5) relates to the variance of the error term, u, in our model, specifically, that its variance conditional on the explanatory variables is constant. Homoscedasticity fails whenever the variance of the unobserved factors changes across different segments of the population, a failure of this assumption will not bias our estimates, but will result in invalid standard errors for our estimators, no longer being "best". To test for homoscedasticity in our model there are a number of quantitative and qualitative methods we can take. A preliminary estimation of our model and plot of the residuals against the fitted values is given by appendix C. This graph shows a classic cone-shaped pattern of heteroscedasticity, with variance increasing at higher values of the dependent variable.

To further investigate the presence of heteroscedasticity we use the Breusch-Pagan (BP) test to check for linear forms of heteroscedasticity, and the White's test for non-linear forms; appendix item D and E report the output of these tests respectively. Both tests present the null hypnotises of homoscedasticity against the alternative of heteroscedasticity. Using the critical value of 0.01, a 1% level of significance, we reject the null under the BP test and fail to reject the null under the White's test. This indicates that our model suffers from linear forms of heteroscedasticity. As a remedy to this, we will run our estimations using heteroscedasticity-robust standard errors.

Assumption MLR.6 of normality states that the error term, u, should be normally distributed. So long as this assumption holds we can be confident in constructing accurate confidence intervals and conducting hypothesis tests on the estimated parameters, as we can then obtain the exact sampling distribution of the t and F statistics. Appendix item F presents a plot of the studentized residuals from the preliminary estimation of our model against a normal distribution, this plot indicates that our residuals maintain a distribution which is relatively close to the normal distribution.

A final amendment for statistical adequacy relates to the functional form of our variables. We propose a number of logarithmic (log) transformations to be taken on independent variables before we estimate the model. Variables *DAYS_RELEASE, WIN_BID_FEED, POS_FEED* and *NEG_FEED* exhibit a nonlinear relationship with *WIN_BID*, this relationship will not be detected by OLS due to the strict linearity assumption. The log transformation of these variables however improves the linear relationship with *WIN_BID*, as logging converts multiplicative relationships into additive relationships. Including the log transformation of these variables will therefore support the linear characteristics of OLS and better fit the model. Log transformations on *NUM_BID, WIN_BID_SEC* and *DESC_CHAR* are also proposed, here, multiplicative changes in the explanatory variables interpreted as marginal changes of the dependent variable are simply more meaningful for our inference, with the transformation not significantly affecting the relationship these variables have with *WIN_BID*. These transformations require a value of one to be added to some variables before the log can be taken, as to avoid missing values arising from calculating the log of zero. In imposing these transformations, we make OLS a more relevant estimator by encouraging linear relationships and allow ourselves to investigate relationships in percentage terms.

The tests and measures proposed in this section allow us to be confident our model will closely adhere to the CLM assumptions, resulting in a properly specified model. We must note however that it would be unrealistic to expect our model to meet all assumptions perfectly, as our data was not generated by an ideal experiment. Rather, the linear regression model under full ideal CLM assumptions should be thought of as the benchmark case with which we compare our model to. Our model adheres closely to the CLM assumptions and maintains many similarities with the models proposed in the preceding literature, resulting in a statistically adequate model.

6 Empirical Findings

6.1 Results

Table 3 presents our estimation results, incorporating the statistical adequacy amendments outlined in section 5.2. Each model takes *WIN_BID* as the dependent variable and includes *DAYS_RELEASE* as to account for time. The first five models separately relate groups of related variables to the dependent variable, (6) is our comprehensive model, and (7), a variant of (6), which accounts for the *COLLECT* variable.

TABLE 3 – ESTIMATION RESULTS

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(DAYS_RELEASE+1)	-125.9***	-128.7***	-126.6***	-122.3***	-126.6***	-124.3***	-123.2***
	(4.938)	(5.156)	(4.894)	(4.699)	(4.935)	(5.135)	(5.095)
ln(WIN_BID_FEED+1)	-3.262**					-3.520**	-3.052**
	(1.321)					(1.421)	(1.378)
ln(POS_FEED+1)	4.530**					1.350	1.770
	(2.173)					(2.061)	(2.056)
ln(NEG_FEED+1)	-13.81*					-12.36*	-14.31**
	(7.207)					(7.223)	(7.267)
BUSI_SELL	7.906					10.16	8.135
	(8.606)					(10.48)	(9.656)
MIN_BID		0.00828				0.0106	0.00550
		(0.0139)				(0.0140)	(0.0139)
ln(NUM_BIDS)		12.03**				15.60***	12.07**
		(5.679)				(5.786)	(5.577)
ln(WIN_BID_SEC+1)		1.218*				0.771	0.655
		(0.674)				(0.701)	(0.688)
DAYS1			-			-	-
DAYS3			-7.743			-6.353	-5.187
			(6.382)			(6.070)	(6.079)
DAYS5			-2.456			-2.784	-4.896
			(7.905)			(7.806)	(7.825)
DAYS7			-6.095			-2.621	-3.674
			(6.312)			(6.277)	(6.145)
DAYS10			-24.06**			-2.546	-12.47
			(11.76)			(11.83)	(10.66)
DOW_WEEKDAY			0.376			1.095	1.051
			(5.082)			(5.015)	(4.916)
NUM_ON_DAY			-0.965			-1.344	-1.150
			(0.944)			(0.929)	(0.915)
ln(DESC_CHAR+1)				4.870***		4.599***	4.778***
				(1.569)		(1.604)	(1.566)
TITLE_CHAR				0.469**		0.120	0.298
				(0.206)		(0.232)	(0.228)
SUB				14.26		4.534	6.790
				(16.88)		(17.77)	(18.20)
STOCK				-29.77***		-30.72***	-33.98***
				(6.273)		(6.750)	(6.560)
NUM_PHOTO				-3.500*		-3.906*	-3.851*
				(1.923)		(2.126)	(2.087)
RET_ACCEPT					1.816	0.506	0.947
					(5.907)	(5.712)	(5.540)
POST_COST					0.00602	0.0450	-
					(0.665)	(0.668)	
DELIV_EST					-6.872***	-4.876*	-
					(2.355)	(2.579)	
GLOBAL					8.426	4.791	6.163
					(5.554)	(6.044)	(6.096)
COLLECT					-	-	-9.839
							(12.71)
Constant	1,498***	1,464***	1,508***	1,441***	1,523***	1,460***	1,431***
	(23.38)	(32.88)	(22.73)	(25.30)	(23.69)	(43.11)	(40.98)
Observations	403	424	424	424	406	385	403
R-squared	0.692	0.693	0.686	0.708	0.685	0.732	0.730
Adjusted R-squared	0.688	0.690	0.680	0.704	0.681	0.714	0.715

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.2 Comments

We comment on the findings of our estimation, all analysis is conducted under a ceteris paribus assumption. Statistical significance indicates that using a t-test, at a given p-value, there is sufficient evidence to reject the null of the coefficient being statistically indifferent from zero at the given level of significance. Differences in the number of observations between models are due to missing values of *COLLECT* and *WIN_BID_FEED*.

As expected the *DAYS_RELEASE* variable exhibits a negative relationship with price across all models, indicating that a 1% increase in the days from release implies a £1.25 decrease in price on average. This finding affirms our justification to include this variable to control for time depreciation of the iPhone, having a statistically significant effect on price at the 1% level. The effect has a greater negative impact on price at lower values of *DAYS_RELEASE*, demonstrating that the iPhone experiences faster depreciation at points in time closer to the release of the phone. This unique variable is underrepresented in the current literature; this is likely due to the items being studied not being subject to large external depreciation/appreciation over the time period they are studied for.

Model (1) includes variables relating to the feedback of the buyer and seller. Positive seller feedback correlates with higher prices within model (1) and (6), however the strength of which diminishes in size and statistical significance in the latter; indicating that other factors are better attributed to higher prices than that attributed to positive reputation in model (1). Model (6) indicates that a 1% increase in a seller's positive reputation is associated with a £0.01 increase in price, although this is not statistically significant. This finding concurs with the empirical literature, with positive reputation shown to have only a minor impact on prices. In contrast, negative seller feedback demonstrates a significant negative effect, consistent across model (1) and (6). Model (6) shows a 1% increase in negative feedback reducing the final auction price by £0.12; this effect matches the literature, although our finding is marginally smaller than other papers. We affirm Lucking-Reiley et al. (2007) in stating that "negative ratings matter considerably more than positive ones" when buyers are considering a sellers eBay reputation, with negative feedback having a significant impact on prices.

With reference to the theoretical proposition of Shepario (1983), our findings support the idea of a price premium for sellers who have established a positive reputation. Although the positive reputation variable showed no statistically significant effect, sellers who maintain a reputation that is non-negative experience a premium over sellers with a negative reputation, implying there is value in maintaining a positive reputation. It appears however that new entrants into the market can sell at a price slightly above that of the minimum quality standard, q_0, then see a positive return on the accumulation of a positive reputation.

The significantly smaller parameter estimates we found in comparison to other papers in the literature base imply that the asset value of reputation, $(v(R_q))$, is lower within this market. This indicates that buyers are inherently more confident in the quality of the items we study, which we associate with attempts by eBay to raise the minimum quality standard on the website over the past eight years (such as introducing the "eBay money back guarantee"). These attempts appear to have been effective; demonstrated by the reduction in the value attributed to a positive reputation in our data. To conclude if this effect has been website-wide would require additional analysis across different items. We can, however, be confident that our estimation demonstrates that reputation is important, but of less value, in determining prices for the item we study.

A statistically significant effect on price also relates to the number of feedback the winning bidder has. Higher levels of feedback for the winning bidder result in lower prices, with model (6) showing a 1% increase in the winning bidders feedback relating to a £0.03 decrease in price. This finding suggests that more experienced eBay bidders, who will be more familiar with the dynamics of the website, are likely to pay less for an iPhone. A potential explanation for this finding is that buyers with low feedback are becoming members of the website with the exclusive intention of purchasing the iPhone, with less knowledge about how auctions on the website work, and an expedited desire for the phone, they are willing to bid higher in comparison to more established buyers, who understand that waiting for longer can result in lower prices.

The dummy variable which represents if the seller is a business seller has a positive coefficient, but is not statistically significant from zero, indicating there is evidence that simply being a business seller will not result in higher prices. As stated in the data section, characteristics which are mandatory to business sellers such as offering returns do not correlate highly with business sellers exclusively, with individual sellers therefore mimicking much of the services offered by their business-based counterparts. This variable has not been tested extensively in the literature and may have different interpretations across different markets on eBay.

Model (2) focuses on variables which describe characteristics of bids within auctions. Higher minimum bids present no statistically significant effect across all models, which contrasts with the findings of Lucking-Reiley et al. (2007) and Bajari and Hortaçsu (2003). The number of bidders who participate in the auction plays a significant role in determining the price; model (6) presents that for each additional bidder who participates, the winning bid rises by £15.60, at the 1% level of significance. This finding concurs with that of Bajari and Hortaçsu (2003), and has routes in the dynamics of "English" auctions, where the price can only rise when bidders compete with each other; more bidders therefore raises the level of competition in the auction and extracts more from their maximum willingness to pay. Although statistical significance couldn't be attributed to the minimum bid, we must not downplay the effect a low minimum bid has on bidder entry, with lower minimum bids attributed to an increased amount of bidders, as demonstrated in appendix G.

The final variable in model (2) describes the seconds between the winning bid and the end of the auction, a unique variable we have included. Although showing a statistically significant positive effect in model (2), this diminishes in the comprehensive model. This finding indicates that the time that a bid is placed has little impact on the selling price. Our analysis therefore provides little support to the method of "sniping" (bidding at the last second) in order for buyers to pay lower prices, which would predict a positive, statistically significant parameter.

Model (3) contains variables related to the length of the auction. As *DAYS1* has been omitted from the model we interpret the coefficients with this variable as the base category (included in the constant), all auction lengths show lower prices as compared to a one-day auction, with a ten-day auction having the most significant effect on price. Once other factors are controlled for in model (6) all auction lengths still show lower prices, now with three-day listings leading to the largest decrease in price, with closing prices being £6.35 lower as compared to a one-day auction. These findings have low statistical significance, which suggests the length of the auction may actually have little impact on the overall closing price. These findings somewhat concur with that of Depken and Gregorious (2010), who found little impact on price regarding the length of the auction, only demonstrating a statistically significant negative effect in five-day auctions. This finding disagrees with Lucking-Reiley et al. (2007) however, who demonstrated price premiums for longer auctions.

From model (6), listings for items that complete on weekends tend to see a price premium of £1.10, although it is hard to conclude that this difference is statistically different from zero. This is a similar finding to Lucking-Reiley et al. (2007), who initially proposed the rationale for price premiums being higher on weekends as "participation rates are higher on weekends", but through similar analysis concluded that this effect is not statistically significant.

The unique variable which we include for the number of auctions which completed on the same day suffers a similar fate. The negative coefficient on this variable supports our original hypothesis that more listings on the same day induce greater competition between auctions and lower prices, with each additional auction showing prices to be £1.34 lower. However, due to the low statistical significance this variable presents it is difficult for us to distinguish this effect from zero.

Model (4) includes structural characteristics of the auctions; these variables are of particular interest to our theoretical understanding as they are the main ways in which information is conveyed between buyer and seller. These variables have been well documented in the literature, excluding the use of stock photos and the number of characters in the description, which is a unique contribution of our analysis. The number of characters included in the items description shows a positive effect on closing prices across all models; with a 1% increase implying a £0.05 increase in price on average, significant at the 1% level. This is a significant finding within our paper, as this variable has seen little representation in the current literature but clearly plays a significant role in determining prices.

The number of characters in the title of the listing has a positive relationship with price in model (4), however in model (6) this effect reduces in size and statistical significance. Our findings therefore indicate the length of the title has a marginally positive effect on price, with each character contributing an additional £0.12 to the final auction price, but we should be cautious in interpreting this effect as conclusively different from zero. This finding is in contrast to Depken and Gregorious (2010), who found title length to have a statistically significant negative relationship with price. The small sample size which Depken and Gregorious (2010) used may have been an aggravating factor in this result however.

Variables relating to the photos of the item *NUM_PHOTO* and *STOCK* present consistent and statistically significant effects across all models. The use of stock photos for an item has a negative effect on price; with the use of a stock photo indicating a £30.72 decrease in price in model (6), at the 1% level of significance. Although the use of stock photos has not featured in the literature before, Eaton (2005) did study the effect of not including pictures at all in an auction; demonstrating that not including an image of the item decreased the bid by $183 (an 11% decrease based on the mean winning bid in his data, compared to a 3% decrease based on the mean winning bid of our data). Not including a photo of an item is prohibited on eBay now, a change since Eaton (2005) conducted his analysis, the use of a stock photo can therefore be seen as a very close substitution to not including a photo at all, and updates the literature to the current dynamics of the website. The difference in the magnitude of this effect likely relates to the homogeneous nature of the item we have selected to study, contrasted with the more heterogeneous item Eaton (2005) analyses. As stated there is a significantly larger information asymmetry between buyer and seller in auctions for used items, making initiatives to inform the buyer more valuable.

The number of photos an auction contains provides a puzzling result, with model (6) reporting each additional photo leading to a £3.91 decrease in price. As the item is new and sealed within a box our theoretical understanding would indicate that additional photos should have an insignificant impact on price, with simply the use of one non-stock photo sufficing in reducing information asymmetry (as the goods are homogeneous). This finding could therefore speak to the concerns expressed in section 5.2 regarding endogeneity, where the *NUM_PHOTO* variable has attributed a negative impact to itself which actually relates to an omitted variable in our model. This is only raised as it is difficult to think of a logical explanation as to why an increase in the number of photos would lead to a decrease in the price, it may simply however be a feature of the data we have captured.

The variables in model (4) support the prediction of the adapted version of the "Market for Lemons" model we presented. Factors which better inform the buyer about the product induce a price premium, which is especially true for the length of the description and inclusion of a non-stock photo of the item in our sample. This demonstrated price premium is lower compared to that of Eaton (2005), which implies that, α (the belief that the item is of good quality), is already relatively high within the market, although not high enough to make measures to reduce information asymmetry redundant in increasing prices. We attribute this to the fact that the item studied is new and homogeneous, so establishing the specific quality of the item is less necessary. It also implies that buyers are inherently more reassured in purchasing from eBay as compared to thirteen years ago when Eaton (2005) conducted his study, attributed to the numerous buyer protection improvements the website has made.

The final segregated model, model (5), includes variables regarding the postage offered within the auction. The most significant variable in this specification relates to the delivery estimate offered, with each additional day taken for the item to be delivered reducing the price by £6.87. This effect is reduced when considered in the comprehensive model (6), but is still statistically significant, at a reduced size. The other variables in this model are statistically insignificant from zero, although have the signs we would expect. This includes the unique variable we included to indicate if the seller offers global shipping, with our results contradicting the messages of eBay that offering this service will result in higher prices, something our data could not confirm.

Model (7) was derived to ensure we do not leave the *COLLECT* variable uninvestigated, as up to this point our analysis has largely excluded the effect of this variable. By dropping the *POST_COST* and *DELIV_EST* variables we are enabled to then include the *COLLECT* variable and associated observations. We do see some minor changes in other variables coefficients in this model compared to model (6), however none change in direction or statistical significance.

Using model (7) we fail to detect that offering collection only results in lower prices. This finding is not conclusive however, due to the limited representation collection only listings have in our sample (at only 4.6%), warranting further investigation.

7 Conclusion

This paper provides additional analysis to help understand the determinants of prices within eBay auctions; current analysis is not up to date with the current website and has not been conducted for the UK before, which this paper addresses.

Reputation plays a significant role in determining prices within our dataset and the literature. Negative feedback is more important than positive in determining prices on eBay, but sellers still see a price premium by holding a positive reputation. Additionally, we find that the reputation of the buyer impacts price, with higher reputation correlating with lower prices in our dataset.

How many bidders participate in the auction has a significant impact on prices in our study; this is a common theme in the literature which we support, clearly playing a role in determining prices on eBay. Low minimum bids encourage bidder entry, however, have no direct link to influencing prices in our study.

Auction lengths present little to no effect in determining prices. This finding concurs with Depken and Gregorious (2010), however directly contradicts Lucking-Reiley et al. (2007), who found longer auctions to result in higher prices in comparison to a one-day auction. This may be a feature of the item we both chose to study, or may have wider implications for listings on eBay as a whole, demanding additional analysis to be conclusive.

The number of days which have passed since the release of the product has a very large impact on determining price in our data, with additional days decreasing the price significantly. This is a unique variable not seen in the literature before, which is essential due to the product we analyse. Further analysis is required to determine if this impact has a website-wide determinant of prices, but we believe for certain items and studies of an extended length this will be true; as prices are being determined by external factors which are not specifically related to auction characteristics.

Initiatives by sellers to better inform buyers about the product have positive impacts on price, especially the length of the description and the use of non-stock photos. Our use of the stock photos variable is unique, but can be related closely to Eaton (2005) who studied not providing a photo at all. We generally find a lower price determination effect from these initiatives, indicating buyers are now more confident in purchasing from the website.

Variables which relate to postage are underrepresented in the literature. We introduce a range of new variables relating to postage but only find the delivery estimate to have a significant effect in determining prices. We fail to detect a price differential in collection only listings; although they are underrepresented within our sample.

We have demonstrated a number of variables which we can be confident are instrumental in price determination on the eBay website, with consistent effects throughout the literature. Our analysis contradicts the consensus at points and would benefit from further investigation across different markets. We introduce a range of new variables into our study, some of which help to determine prices within the auctions we analyse. As these variables are underrepresented in the literature it is difficult to conclude if these effects are website-wide. This paper provides a comprehensive update to the literature and creates a benchmark from which further study can now take place regarding price determination within eBay auctions.

8 References

Akerlof, G. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics,84(3), pp. 488-500

Bajari, P. and Hortaçsu, A. (2003). The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions. The RAND Journal of Economics, 34(2): pp.329-355

Depken, C. and Gregorius, B. (2010). Auction characteristics, seller reputation, and closing prices: Evidence from eBay sales of the iPhone. International Journal of Electronic Business, 8: pp. 170-186

Eaton, D. (2005). Valuing Information: Evidence from Guitar Auctions on eBay. Journal of Applied Economics & Policy, 24: pp. 1-19

Houser, D. and Wooders, J. (2006). Reputation in Auctions: Theory, and Evidence from eBay. Journal of Economics & Management Strategy, 15: pp. 353-369

Klemperer, P. (1999). Auction Theory: A Guide to the Literature. Journal of Economic Surveys, 13: pp. 227-286

Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2007). Pennies from eBay: The Determinants of Price in Online Auctions. The Journal of Industrial Economics, 55: pp. 223-233

Office of National Statistics Statistical Bulletin: E-commerce and ICT activity, 2016

https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/ecommerce andictactivity/2016

Shepario, C. (1983). Premiums for High Quality Products as Returns to Reputations, The Quarterly Journal of Economics, 98(4), pp. 659-679

Shy, O. (1995). Industrial Organization: Theory and Applications. Cambridge: MIT Press.

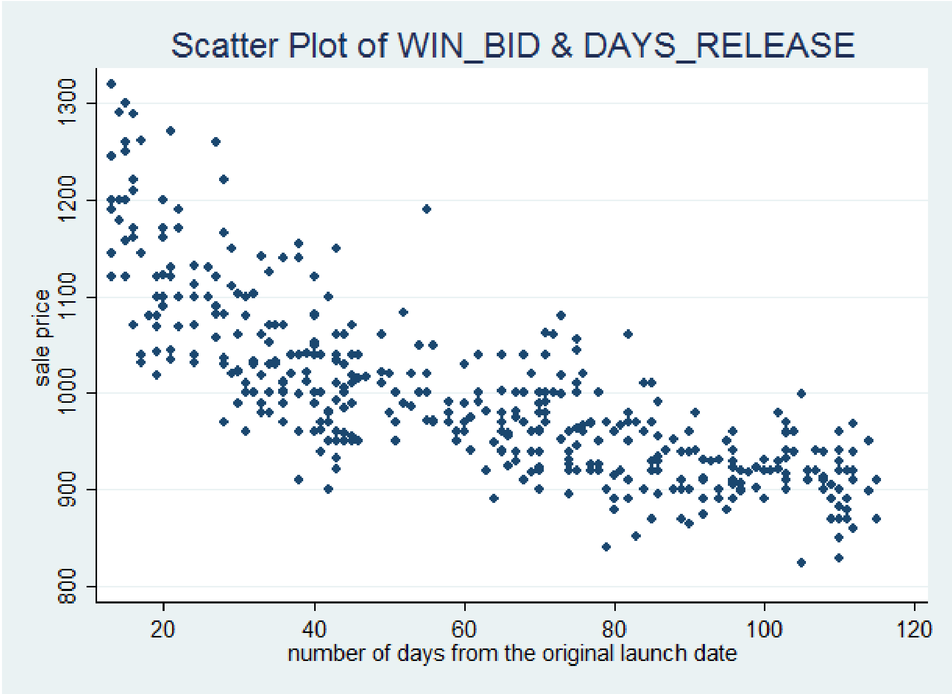
Statista: Commerce websites ranked by visitors in the United Kingdom (UK), March 2016

https://www.statista.com/statistics/286489/uk-visitors-to-mass-merchandiser-websites/

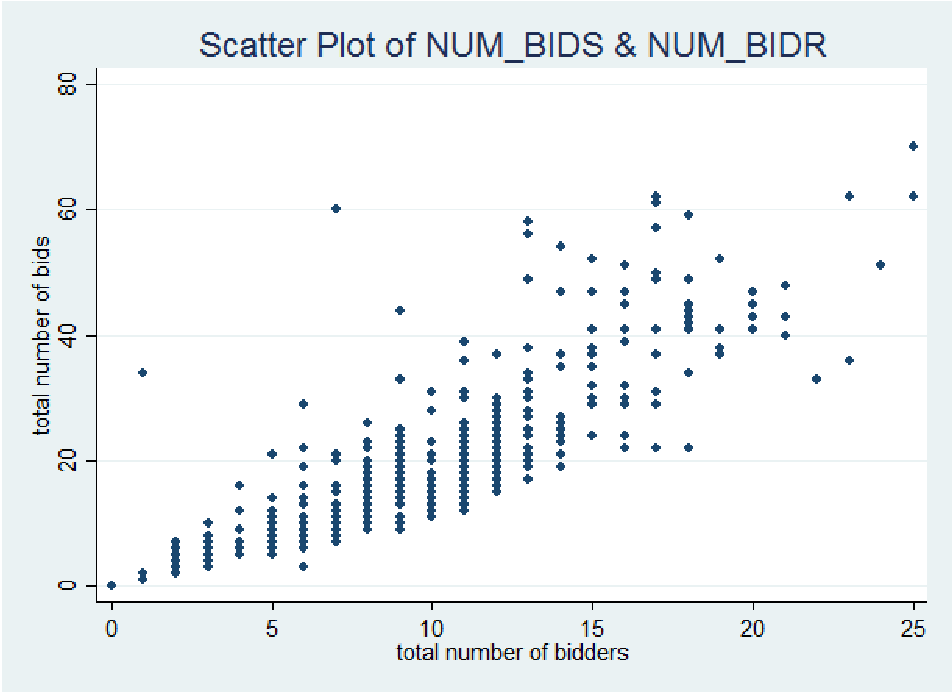
Wooldridge, J. (2012). Introductory Econometrics: A Modern Approach, Fifth Edition. Mason, Ohio: South-Western Cengage Learning,

9 Appendix

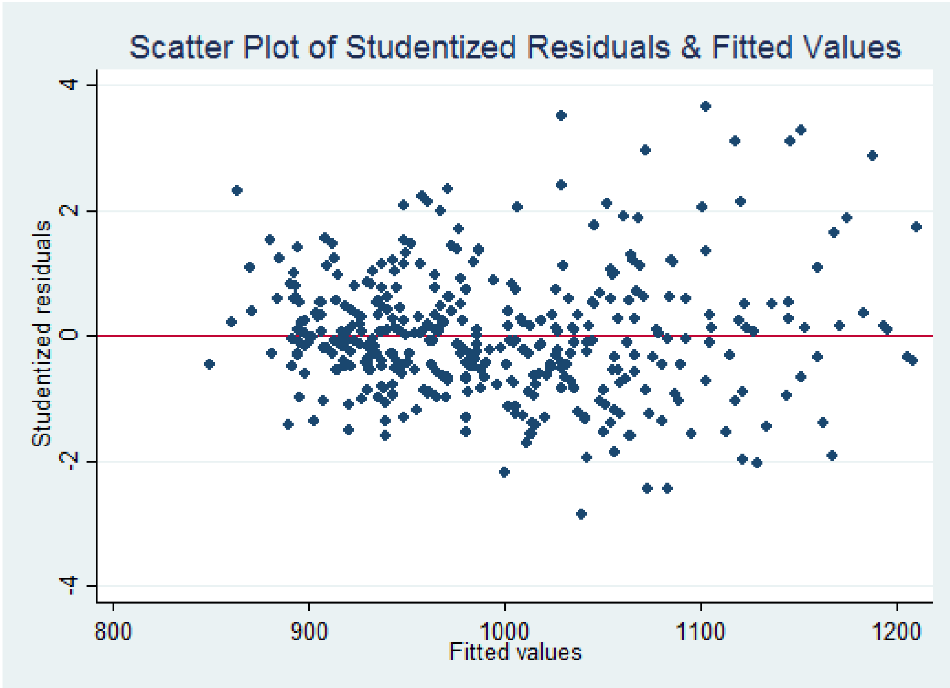
Appendix A



Appendix B



Appendix C



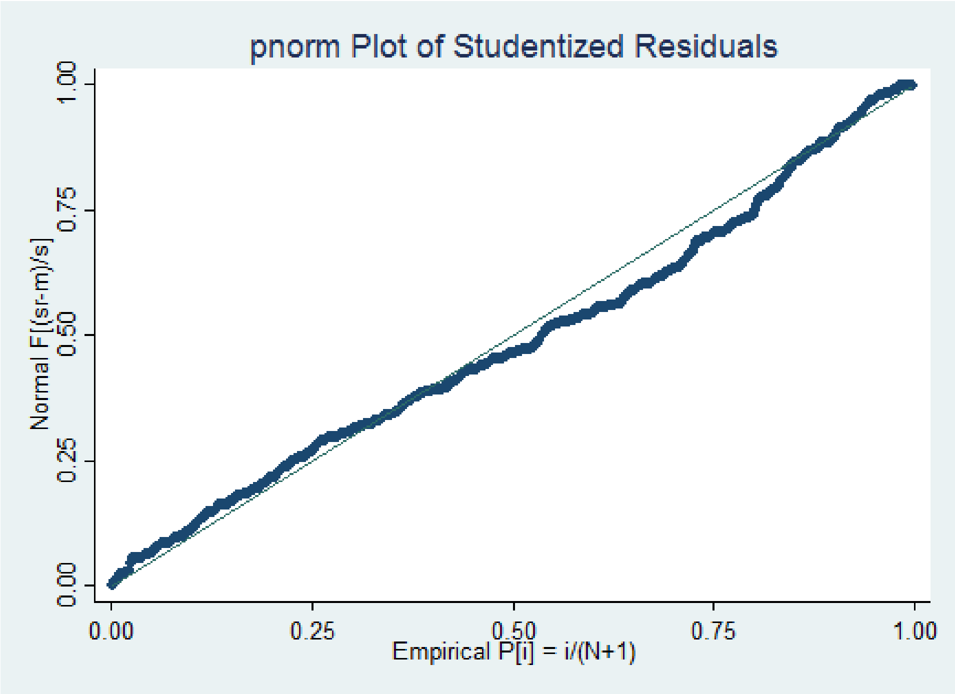
Appendix D

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity Ho: Constant variance Variables: fitted values of win_bid chi2(1) = 43.86 Prob > chi2 = 0.0000

Appendix E

White's test for Ho: homoskedasticity against Ha: unrestricted heteroskedasticity chi2(258) = 242.26 Prob > chi2 = 0.7512

Appendix F



Appendix G