THREE ESSAYS ON SHILLING, HERDING AND SCORING, IN POSTED-PRICE AND AUCTION MARKETS

Siqi Song

Doctor of Philosophy

University of York
Economics and Related Studies

March, 2018

Abstract

In this thesis, I focus on the purchasing environments of eBay's online platform and Christie's live auctions. In chapter 2, I investigate the presence of seller squeeze-shilling behaviour on eBay's smart-phone auctions. First, I present theoretical proofs to show that opportunistic sellers have the robust incentive to place shill bids in their own auctions, under both a private-valuation and an affiliated-valuation paradigm. Then I empirically prove the existence of squeeze shilling on the eBay platform. My findings show that the eBay proxy bidding system and the user agreement of retraction provides an excellent environment for opportunistic sellers to conduct squeeze shilling and extract extra profits. In chapter 3, I examine consumer herding behaviour in eBay posted-price listings using data from the iPhone screen protector market. I find that the cumulative number of total historical sales shown on the listings page is the essential trigger of herding behaviour. I find that, for listings with the advantage of the 'herding effect', the sellers of these listings have the incentive to manipulate the purchasing environment by increasing the posted price to extract extra profits. In chapter 4, I investigate the determinants of bidder's valuation in wine auctions at Christie's. I estimate the bidder's valuation distribution using an indirect inference approach. I find that the parameter, wine score, is the crucial structural parameter that characterizes the bidder's valuation distribution. By using the structural estimator, I estimate the optimal reserve price for each auction and simulate winning bids and find that the reserve price set by Christie's is not optimal.

Contents

List of figures

List of tables

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisors, Professor Zaifu Yang and Professor Francesco Bravo, for their continuous and overwhelming support of my Ph.D study and research, for their patience, motivation and immense knowledge. I would also like to thank my TAP member, Professor Peter Simmons, for his insightful comments.

I thank my fellow colleagues and friends at York, for the stimulating discussions, for the laughs and the cries, for the fun we have had in the last five years.

Last, but not least, I would like to thank my family for their love, warmth and spiritual support throughout writing this thesis and my life in general.

Declaration

I declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. All sources are acknowledged as references.

Chapter 1

INTRODUCTION

1.1 Shilling, herding and scoring

This thesis presents three empirical essays of rational consumer-seller behaviour in different trading platforms and environments.

1.1.1 Shilling

The recent era has seen the rise of e-commerce, which has undoubtedly transformed the way consumers and businesses trade. A combination of lower transaction costs and fewer space and time restrictions has given rise to an abundance of business opportunities on the web. Online auctions have emerged as one of the principle ways for people to trade on the web.

The eBay online platform was one of the first websites to introduce auctions for the sale of everyday products. It is now one of the leading online auction platforms, with operations in around 30 countries. When the site was founded in 1995, it initially sold mainly collectable items, which is a category that is typically sold by auction. The company has since expanded, and today, popular categories on the site include Fashion; Home and Garden; Electronics; Sports and Leisure; Collectibles; Health and Beauty; and Motors. The online platform uses a second-price ascending auction with a proxy bidding system.

There has been extensive research into the strategies that opportunistic sellers use to manipulate online auctions. As users are encouraged to submit their true valuation via the proxy bidding system, sellers are tempted to bid in their own auction under false identities in an attempt to increase their revenue; this phenomenon is known as shilling. The anonymity of users in the online environment has made manipulation of online auctions easier to conduct.

1.1.2 Herding

In addition to the original auction listings, eBay has evolved to include ‘Buy It Now’ listings, which are a type of posted-price listing. Inventory listings allow the seller to continuously list multiple identical items in one single listing. Online consumers can observe any previous sales that the listing has made, and thus, how many others have made the decision to purchase the exact same item. This information allows people to easily identify which items have been popular with other consumers. In some ways, ‘historical sales’ can be seen as an advertisement in itself.

Herding behaviour has been recognised in many social studies. There are several early theoretical papers that discuss rational herding behaviour in sequential settings, see for example, Banerjee (1992), Bikhchandani et al. (1998), and Bikhchandani and Sharma (2000). The American social writer and philosopher, Eric Hoffer (1995), noted, "when people are free to do as they please, they usually imitate each other." When there is an abundance of choice, it may be rational for the people to imitate each other. In the eBay online environment, people often prefer buying items that are considered ‘popular’, so we expect that herding behaviour would be present in inventory listings. However, herding behaviour poses a problem when the effects lead to far more adverse outcomes than when individuals make their decisions independently. Moreover, herding of listings can squeeze out good competition and higher quality products on the site. It may not be a coincidence that eBay has traditionally, and reputationally, had fewer branded products for sale on the site compared to its competitor, Amazon.

1.1.3 Scoring

Traditionally, auction houses facilitate the sale of precious or rare items to the affluent aristocrats of society. They feature categories that range from art to wine, to historical artefacts. Due to the nature of items, the value of items for auction can be widely influenced by the comments and views of experts. It has been found that wine drinkers are persuaded to buy wines from the retail market with good expert reviews and 'excellent' wine scores. Due to the complexity of the product, wine experts assign wine scores which can be a reliable estimator of a wine's quality.

Christie's, a British auction house, is one of the leading auction houses. It has a long history of auctions that dates back to their first auction in 1766. Their wine auctions offer some of the finest and rarest wines of the world. Whilst Christie's has sales that eclipse that of its competitor, Sotheby's, Christie's has traditionally had lower profits, which indicates large overhead costs, such as, advertising and specialist fees, which are both major influencers of public opinion. Specialist reviews are a vital tool which can be used to mould a person's opinion. So, perhaps it is not so surprising that Christie's offer free specialist consultations for all auction items.

1.1.4 Motivations

The emergence of e-commerce has generated an abundance of data that was not previously accessible, which has induced a surge of extensive economic research in the topics of e-commerce and online auctions. Shilling is a widely discussed problem of trading on online auction platforms, where opportunistic sellers, under false identities, bid in their own auctions for their economic self-interest. In online platforms, the anonymity of users can facilitate manipulation of the auction environment, which makes fraudulent behaviour difficult to identify. There are differing theoretical results for shilling behaviour which has been seen as analogous to a dynamic and adjustable reserve price. Several theoretical papers consider shilling in auctions and show that an adjustable reserve price can increase the seller's revenue, see for example, Izmalkov (2004), Myerson (1981), Milgrom and Weber

(1982), and Graham et al. (1990). However, other theoretical papers suggest that shilling may have negative consequences for the seller if shilling is detected, see Sinha and Greenleaf (2000) and Chakraborty and Kosmopoulou (2004).

Shilling can have a detrimental impact on the trustworthiness of auction platforms, however, due to the anonymity of users on the site, the existence and impact of shilling are hard to capture. Therefore, my motivation for chapter 2 is to investigate the existence and impact of squeeze shilling in eBay auctions, which contributes to the literature of shilling in online auction platforms.

Another challenge on online platforms is that there exists the problem of asymmetric information in the quality of items between buyers and sellers. Consumers have to make a rational purchase decision, but they face an overwhelming amount of choice for products and services. Sellers have the incentive to upload detailed descriptions and photos of their items to signal the quality of their item. However, the abundance of choice on an online platform inherently leads to a problem of information overload. Understandably, buyers have limited amounts of time to browse through the available options, so, they find a predictable indicator of quality: the number of historical sales. As the eBay platform seems to provide a good environment for herding, my motivation for chapter 3 is to empirically test whether herding behaviour exists in eBay inventory listings and whether herding has a negative impact on the online market, which contributes to the literature of herding behaviour.

The wine market seems to be a widely researched topic of interest. Several papers examine the determinants of price in wines, using hedonic price regressions, which include Oczkowski (1994), Schamel et al. (2001), Jones and Storchmann (2001), San Martin et al. (2008), Bombrun and Sumner (2003), and Landon and Smith (1998). They find a positive relationship between wine score and wine price in the primary market of wine. It is interesting to note that there seem to be few papers that focus on the secondary market for wine. The wines sold at auction are usually the finest and most expensive wines in the world. Besides the non-sensory characteristics of a wine, which includes, chateau, vintage year and location among others, wine score and tasting notes from prestigious wine experts are often taken into consideration for the first-time buyer. Therefore, it would be interesting to investigate how wine score affects the bidder's valuation. My motivation for chapter 4 is to estimate the bidder's private valuation distribution and reveal the effect of wine scores on bidder's valuation, which contributes to the literature of wine auctions. I also assess whether Christie's reserve prices are set at the optimal level, which contributes to the literature of auction mechanism design.

1.1.5 Outline of the thesis and contributions

In chapter 2, I identify and prove the existence of seller's squeeze shilling behaviour in eBay auctions with a proxy bidding system. The notion of squeeze-shilling originates from the idea that sellers try to extract or 'squeeze' profit from the consumer. I present some theoretical results and some empirical evidence about squeeze-shilling behaviour on eBay. First, I show that under both the private and affiliated valuation paradigm, sellers have the incentive to inflate the ending price of auctions through squeeze shilling in eBay auctions. I also show that opportunistic sellers always have the incentive to conduct squeeze shilling to achieve higher revenue when the detection rate of shilling is low, and that, under the interdependent or common valuation paradigm setting, early shilling is a good strategy for opportunist sellers to induce other bidders to submit their bids. Then, I test for the existence of squeeze shilling behaviour and characterise the common features of squeeze shilling behaviour. My findings show that the eBay proxy bidding system and the user agreement of retraction provides an excellent environment for opportunistic sellers to place shill bids and extract extra profits. I find that retractions are a good indicator to detect squeeze shilling behaviour on eBay.

In chapter 3, I empirically prove the existence of herding behaviour on eBay's inventory listings. My findings indicate that the figure of 'total historical sales' on each listing page is the crucial element that instigates herding behaviour on eBay's screen protector market. From both static and dynamic aspects, I show that the number of 'total historical sales' have a positive effect on future buyers' purchasing decision. My empirical work sheds a light on the price strategy for eBay sellers on this market. Upon entering the market, novice sellers should set a very competitive price to accumulate a sufficiently high figure in 'total historical sales' to trigger a herding effect in future sales. My empirical results indicate that, once the herding effect is formed, sellers can increase the listing price up to a certain price level and total sales continue to increase. Therefore, experienced sellers have the incentive to increase their listing prices to extract higher profits.

In chapter 4, I empirically estimate the bidder's valuation distribution using an indirect inference approach and analyse the impact of wine scores on bidders' private valuation in Christie's Fine Wine Auctions. I find that the parameter 'wine scores' is the crucial structural element to characterise the bidder's private valuation distribution. I then discuss the auction optimal design problem and find that the seller's total revenue is not maximized. The simulation outcomes indicate that a higher reserve price may lead to higher total revenue for sellers.

Chapter 2

EVIDENCE OF SHILLING BEHAVIOUR ON EBAY

2.1 Introduction

In the past two decades, the rise of E-commerce has transformed how consumers and businesses interact. The growth of online market platforms has broken down the barriers of time and geography, where online sellers can trade globally with millions of potential customers, transaction costs are lower, and business hours are more flexible than ever before. As a result, the eBay online auction platform has become one of the largest successes of the dot-com boom.

The auction format that eBay employs is a second-price ascending auction with a proxy bidding system and a hard closing time. Bidders are encouraged to submit their maximum bids and let the proxy bidding system bid on behalf of them. The bidder who submits the highest bid wins the auction and pays a price equal to an increment above the second highest bid. In the event of a tie, the bidder who submits their bid first wins the auction.

One of the downfalls of trading on an online platform is the inherent asymmetric information problem between buyers and sellers, where consumers often fall victim to manipulation from opportunistic sellers who are motivated by economic self-interest. In eBay auctions, shilling is a type of fraudulent behaviour that is considered to be prevalent but difficult to identify due to the anonymity of users on the site. On the eBay forum, many users express their surprise and puzzlement at suspicious bidding patterns. They suspect that the bidding patterns are a form of seller shilling behaviour - they call it "squeezing". The name originates from the concept that the seller extracts profits from bidders. However, bidders often find that it is difficult to provide sufficient evidence to support their conjecture. According to the descriptions on eBay forum, suspicious squeeze shilling is conducted in the following way: a seller uses a false bidding account to outbid the current high proxy bid, which reveals the current high bid. Then, the shill bidder retracts his/her bid and employs a different bidding identity to increase the auction standing price to the proxy bidder's maximum bid. This allows sellers to use shill bids to maximise their revenue. Many bidders share on the forum that, in the event of being outbid in an auction, they choose to participate in other auctions that sell the same item, not realising that they are still liable for their bids if the highest bid is retracted in the first auction. As a result, bidders are sometimes left with the dilemma of winning two auctions.

There are several reasons why it is difficult to identify seller shilling behaviour on an online platform, they are concluded as follows. First, there are various strategic and non-strategic reasons behind different bidding patterns, which arise from the differences in bidder's reactions and decisions. The variety of bidding patterns shields shilling behaviour, which makes it all the more difficult to identify the real honest bidders from the false shill bidders. For example, bidders who submit multiple bids below a current proxy high bid could be shill bidders with the intention of inflating the auction ending prices. However, these bids may also be due to inexperience or irrationality. Also, bidders who outbid a proxy bidder, but follow with an immediate retraction after becoming the highest bidder, could be flagged as shill bidders, who intend to reveal the proxy bidder's highest willingness to pay. However, these retractions could also be due to: typographic errors, item marking, winning in an adjacent auction or even getting cold feet. Suspicious behaviour is met with widespread speculation from sharp-eyed buyers, but even the most watchful of users find it difficult to track evidence of shilling behaviour - especially without inside information, such as the relationship between a bidder and a seller, or the physical identities of each bidder and seller. A shill bidder could be the sellers themselves, or they could be confederates of the seller, for example, friends or relatives who use genuine accounts to submit shill bids. Also, the cost of creating a new bidding account is low, which allows sellers to create multiple bidding accounts. In online bidding environments like eBay, the anonymity of online users have made it even easier for sellers to shill. Bidding accounts can be easily registered with false information to hide sellers' true identities, which makes it harder for honest bidders to identify shill bidders.

The aim of this chapter is to identify and prove the existence of sellers' squeeze shilling behaviour in eBay auctions. Theoretically, I show that, under both the private and affiliated valuation paradigm, eBay sellers have the robust incentive to inflate their auction ending prices by conducting squeeze shilling in eBay auctions. Furthermore, I show that, under the private valuation[1] paradigm, when a bidder submits their highest bid using the eBay proxy bidding system, the seller's strategy to not shill is dominated by the strategy of conducting squeeze shilling at the last minute. On the other hand, under the interdependent or common valuation[2] paradigm setting, I show that early shilling is a good strategy for opportunistic sellers, as shill bids send a signal of an item's high quality to uninformed bidders, who are uncertain about an item's value. Shill bids can induce uninformed bidders to submit higher bids and hence, increase the seller's expected revenue.

To empirically test the existence of squeeze shilling behaviour, I employ auction data sample from eBay's smart-phone category, specifically, iPhone. I characterise

[1] Private valuation paradigm: each of N potential bidders get an independent and identically distributed draw from a specific distribution.

[2] Common valuation paradigm: Bidding environments in which the value of the object is the same to all potential N bidders, for example oil mining.

the common features of squeeze shilling behaviour and use these as a criteria to distinguish and test the existence of squeeze shilling behaviour on eBay platform. The findings of my empirical analysis confirms the prediction of my theoretical results which strongly supports the existence of squeeze shilling behaviour on the eBay platform. I find that the eBay proxy bidding system and the user agreement of retraction provides an excellent environment for opportunistic sellers to place shill bids and extract extra profits, and that retractions can be regarded as a good indicator to detect and trace squeeze shilling behaviour on eBay.

The rest of this chapter is organized as follows. In section 2.2, I discuss the related theoretical and empirical literature of shilling in auctions. In section 2.3, I describe the shill-bidding model in eBay's bidding environment. In section 2.3.1 and 2.3.2 I also provide some theorems and proofs about shilling behaviour. In section 2.4, I explain the empirical methodology and discuss the empirical outcomes, with a description of data and summary statistics in section 2.4.1. Finally, in section 2.5, I conclude the chapter and give some general comments about shilling and eBay's auction environment.

2.2 Literature review

There are a number of theoretical papers that consider the effect of a shilling in English ascending-price auctions. Izmalkov (2004) shows that there exists an equilibrium with shilling that is equivalent to the optimal mechanism shown in Myerson (1981), who showed that a higher public reserve price leads to higher auction revenue under the independent-private value paradigm. Milgrom and Weber (1982) derive a similar result under the common-value setting. Therefore, it may also be interesting to consider whether eBay and sellers have a common interest, as sellers who extract more profits from higher auction prices contribute to higher commissions collected by the online platform. Hence, eBay might be reluctant to investigate and penalise shill bids.

There are some theoretical papers that consider online shilling to be analogous to a dynamic and adjustable, public reserve price, i.e. the reserve price can be adjusted throughout the auction after bidding has started. Graham et al. (1990) consider the situation where the reserve price is dynamic. They show that the longer the seller can observe bidder, the more knowledgeable s/he becomes about the bidder's valuations. Therefore, a dynamic reserve price can lead to an increase in the seller's revenue. In the absence of penalties for shilling, they suggest that the seller should weakly prefer an adjustable reserve price, or a shill bid, to an ex-ante fixed reserve price.

There are crucial informational differences between setting an ex-ante fixed reserve price and a dynamically adjustable reserve price or shill bid. Reserve prices are public information which are static and set at the start of the auction; however, shill bids are concealed from bidders and can be placed throughout the auction. It is also important to note that shilling in many countries is considered to be illegal and can have serious consequences, which are often not considered in the theoretical literature. Nevertheless, due to the reasons outlined above, the online environment has made shilling substantially easier to conduct.

There are also theoretical papers that show more scepticism about shilling and show that shilling is not always advantageous for the seller. Sinha and Greenleaf (2000) employ a discrete-bidding model, in which they consider the number of active bidders, their 'aggressiveness' and the sequence of bidder's and shill bidder's bids. Their paper demonstrates that there is a payoff between shilling to extract revenue and bidders shading their bids when they expect shilling to occur. Bid shading describes the practice that bidders intentionally submit their bids which are lower than their valuation.

Chakraborty and Kosmopoulou (2004) derive a similar result in their theoretical model, under a common-valuation setting. They show the seller's dilemma: when bidders do not have the knowledge about the presence of shilling, sellers have the incentive to shill to increase their final revenue; however, once bidders have the knowledge of shilling behaviour, bid shading will increase and revenues will decrease. Furthermore, shilling comes with the additional risk that the seller wins his/her own auction and must pay an additional expense to the auctioneer. Taking into account the two negative consequences of shilling, they predict that the seller should prefer not to shill bid to decrease the possibility of bid shading.

There are a number of empirical papers that document the existence of seller shilling behaviour. Kosmopoulou and De Silva (2007) run a laboratory experiment to verify the theoretical prediction of Chakraborty and Kosmopoulou (2004). In the first part of the experiment, participants are asked to bid in a number of auctions where sellers are not allowed to participate and shilling is not permitted. In the second part of the experiment, sellers are permitted to join the auctions and hence, shilling is possible. They find that when sellers are permitted to submit bids, the level of bidder's bids decreases and the average profit falls from 97.5% to 88.9% of the item value. The seller's dilemma is that if bidders believe there is no possibility of shilling, sellers make more profit but still have an incentive to shill to increase their revenue. However, once bidders realise that sellers have the ability to shill, they increase their bid shading and bid levels drop, resulting in a lower overall profit. This experiment emphasises an important motive for eBay to deter shilling - to encourage higher bid levels in their auctions and yield higher auction revenues. The eBay site has a reporting system in place, which encourages users to report any suspicious listings that seem to violate their policies. However, the penalties are not usually harsh and may include: warning, removal of listing, and/or limiting the user privileges. Moreover, seller accounts are not usually suspended unless there have been multiple reports with 'sufficient' evidence of the seller violating eBay policies.

Hoppe and Sadrieh (2007) examine shilling behaviour through a series of field experiments where they run auctions for three different scenarios, each with different treatments for the sale of identical products, consisting of a) blockbuster-movie DVDs - a 'thick-market' item[3] and b) collectable-coin bundles - a 'thin-market' item[4]. All auctions in the first treatment have a reserve price that is set at the lowest pos-

[3]A thick market is a market with a large number of buyers and sellers.

[4]A thin market is a market with a small number of buyers and sellers.

sible value. All auctions in the second treatment have a reserve price set at the lowest possible value, but with a secret reserve price at around 60% of the book value. All auctions in the third treatment have a reserve price set to the lowest possible value, but the seller later submits a shill bid at approximately 60% of the item's book value. They find that the revenues - the ending prices - are indistinguishable between the three treatments, regardless of whether the item had a thick or thin market. However, profits for sellers are considerably lower for the second treatment, where sellers have the additional cost associated with setting a higher reserve price.

Kauffman and Wood (2003) study empirically the behaviour of reserve price shilling in eBay's rare coin market. They distinguish two types of shilling: reserve price shilling, where they find that shilling can be used to avoid paying auction insertion fees; and competitive shilling, where shilling can be used to induce bidders to pay more for an item. In particular, they investigate reserve price shilling, where they use 'questionable bids' as the crucial instrument to determine shill bids. Assuming bidders are rational, they define a questionable bid as a bid that is submitted in auction B, even though an equal/lower bid would have been the highest standing bid in auction A, which sells the same item but ends at an earlier time. Their criteria for shilling consists of: 1) shill bids are 'questionable bids' that are submitted early in the auction and increases the standing bid by an above-average amount; 2) shill bidders only bid on a few select sellers; and 3) shill bids are less likely to win the item at the end of the auction. They examine over 10,000 rare-coin auctions and find that 622 auctions meet their criteria for shilling. They find that shilling is more likely to be observed in auctions that feature low reserve prices, high-value auction items, and long auction durations. Also, when shill bidders are identified in one seller's listing, other listings from the same seller are more likely to be found to have shill-bidding behaviour.

Engelberg and Williams (2005) empirically study a specific type of seller shilling behaviour on eBay's event-ticket market. The shill bidder uses a shilling strategy to inflate the auction standing price incrementally to the highest bidder's maximum proxy bid - they call this strategy, 'discover-and-stop'. The discover-and-stop strategy allows the shill bidder to 'squeeze' the surplus from the highest bidder. They have two criteria to identify the discover-and-stop bidder: 1) the bidder submits at least two incremental bids within 10 minutes, and 2) the bidder ceases bidding as the second highest bidder, when they find that the automatic proxy bid of the highest bidder is less than one increment higher than their penultimate bid, as the proxy bidding system never exceeds the maximum proxy bid. Engelberg and Williams employ a dataset of 40,000 event - ticket auctions on eBay in September 2004. They find that 3% of all bids in the dataset meet the criteria of the discover-and-stop strategy, and they estimate that half of these bids (i.e. 1.5%) are intentional shill bids. They conclude that shill bidders believe that eBay bidders tend to submit bids that end in round numbers, therefore, discover-and-stop bidders tend to submit bids that end in odd numbers to increase their chances of discovering the maximum bids of the proxy bidders. Also, bidders who submit bids in multiple auctions that are run by the same seller, have a higher likelihood of being discover-and-stop shill bidders. A disadvantage of discover-and-stop shill bids is that there is a risk of the seller winning the item if the bidders do not submit bids that end in round numbers.

2.3 A strategic shill-bidding model in eBay's bidding environment

I introduce a strategic model of standard eBay auction; it employs a second-price ascending auction with time priority, where there is one seller who wants to allocate one indivisible item to N risk-neutral bidders. Ex-ante, the seller sets an auction reserve price, S, and auction duration in days: one, three, five or seven days. When the auction commences, bidders submit open bids or proxy bids, which are 'maximum secret reservation prices', where the proxy bidding system automatically bids on behalf of the bidders. Bidders subsequently bid against each other and each bid that is newly submitted must be higher than the current standing bid by at least one increment. The minimum bid increment, M, is set by the eBay system.

At the end of the auction, the bidder who submits the highest bid wins the auction and pays the second highest bid plus an additional minimum bid increment. The amount of the highest bid remains secret. If only one single bidder participates in the auction and places a bid that is higher than or equal to the reserve price, then the standing bid is the reserve price. Automatic proxy bids cannot exceed the bidder's maximum proxy bid (i.e. the bidder's secret reservation price). It is important to note that if the difference between the two highest secret reservation prices is smaller than the minimum increment, the standing bid is equal to the highest secret reservation price. Bidders can submit bids at any time interval, $t \in [0, 1]$. Bidders have the knowledge of each other's bidding throughout the auction, where $t < 1$. At the last minute, $t = 1$, bidders only have the time to submit one additional bid - they can observe the bidding history up to $t < 1$, but cannot observe or react to other bidders' bids at $t = 1$. Only bids that are placed prior to $t = 1$ are transmitted with certainty. Due to the complexities of internet traffic and varying connection times, last minute bids are not always transmitted with success on eBay. So I assume that a bid that is placed at the last minute, $t = 1$, is transmitted successfully with a probability equal to $p < 1$.

Potential bidders can gather a plethora of information from looking at the auction bidding history, which includes: the previous bids on the item up to the current standing bid, complete with bidder identities and feedback scores; the auction reserve price; the auction duration; and the auction end time. Automatic proxy bids in the bidding history are highlighted in grey, as opposed to black for open bids - excluding the standing bid, which is black - which makes them easily identifiable. Any previous bid retractions and cancellations are included in the bidding history, which comprises of the bid amounts, bid submission times and retraction or cancellation times. In this chapter, retraction is the action of a bidder retracting their own bid, single or multiple bids. Cancellation refers to a seller cancelling a single bid or multiple bids in their own auction.

The eBay regulations state that bidders should only retract their bids in the following three circumstances: the bidder made a typographical error, in which they should immediately retract and resubmit their bid, the item description has changed or the seller's identity cannot be authenticated. However, it should be noted that bid retractions can be easily made, even for reasons other than the ones stated in the regulations, such as, if the bidder suspects the presence of fraudulent shilling behaviour. The probability that a bidder detects the shill bid and retracts their bid is equal to $q < 1$. For retractions made prior to the final 12-hours of the auction, all the preceding bids submitted by the bidder are also retracted from the auction. For retractions within the last 12-hours of the auction, the retraction must be made within an hour of the bid being submitted and only the final bid is retracted. There is speculation that eBay's retraction rule can be manipulated by sellers. If sellers can reveal the highest bidder's maximum proxy bid without risk or penalty, they have the incentive to conduct squeeze shilling in order to achieve higher expected revenue. Moreover, shill bids can be used to initiate bidding wars and attract more bidders to participate in the auction, which results in an increase in the auction ending price.

The following section discusses some theoretical observations that show that in the eBay bidding environment, sellers have the incentive to conduct shilling, under both a private and an affiliated valuation paradigm.

2.3.1 A private valuation of equilibrium model with squeeze shilling on eBay

I model the eBay auction under a private valuation paradigm, where each bidder j knows one's private valuation, v_j, at the beginning of each auction, $t = 0$, and each bidder's valuation is subject to a certain distribution with a probability density function, $f(v_j)$, and a cumulative distribution function, $F(v_j)$. A bidder wins the auction at the price of L, s/he obtains a payoff equal to $v_j - L$.

Theorem 1: [Under a private valuation paradigm, when the detection rate of shilling is low, shill bidders always have the incentive to conduct squeeze shilling at $t < 1$]

Proof. For simplicity, I consider the case where one seller wants to sell an indivisible item in an auction with two bidders. The first bidder is proxy bidder j, whose strategy is to submit his/her true valuation, $v_j > M + S$, using the eBay proxy bidding system at the beginning of the auction. The second bidder is shill bidder i, the seller who employs a false identity to participate the auction. When the auction commences, shill bidder i observes that proxy bidder j submits his/her bid through a proxy bidding system. As bidder j is the only bidder to submit a bid, the standing price remains at the reserve price, S. In order to conduct squeeze shilling, the shill bidder i reveals bidder j's maximum bid by submitting a bid, B_i, which is larger than v_j + M. Bidder i becomes the highest bidder and the standing bid rises from S to v_j + M, which is the bidder j's maximum proxy bid plus one minimum increment. After successfully revealing v_j, bidder i retracts B_i, and the standing price falls back to the reserve price, S. Subsequently, bidder i submits another shill bid at v_j, prior to $t = 1$, and the bid is transmitted with certainty, $p = 1$. However, the shill bidder runs the risk of bidder j detecting the shilling behaviour with probability, $q < 1$. If shilling is detected: bidder j retracts his/her proxy bid and bidder i pays the reserve price, S, with probability q. If shilling is not detected: the shill bidder yields revenue equal to v_j, with probability $1 - q$. Hence, the shill bidder's expected revenue is equal to $v_j(1 - q)$ + $(-S)q$. For comparison, if the bidder i - the seller - does not conduct shilling, bidder j submits his/her bid at v_j by proxy bidding system, and wins the auction and pays the reserve price, S.

The following derivation shows the condition that a seller can yield a higher expected revenue through squeeze shilling, at any time before $t = 1$, where bidders transmit their bids with certainty.

For sellers to have the incentive to shill: the expected revenue with shilling must be higher than the expected revenue without shilling, which is expressed as:

$$\begin{aligned}
&v_j(1-q) - Sq > S \\
&v_j(1-q) - Sq - S > 0 \\
&v_j(1-q) - S(1+q) > 0
\end{aligned} \tag{2.1}$$

which derives the following relation:

$$\frac{v_j}{S} > \frac{1+q}{1-q} \geq 1 \tag{2.2}$$

In any specific auction, the valuation of bidder j, v_j is constant, therefore, $\frac{v_j}{S}$ is also constant. The probability of detection, $0 < q < 1$, is constant, and the fraction, $\frac{1+q}{1-q}$, is a monotonically increasing function of q. The problem the shill bidder faces is to determine whether there is an incentive to shill. As the reserve price increases, the incentive to shill gets smaller. The closer q is to 1, then the higher the probability of detection and the riskier it is to shill. As v_j is fixed, the reserve price, S, must be sufficiently low to make shilling more profitable than not shilling for the opportunistic seller. For example, let's say the probability of detection is 50% in equation 2.2, then $\frac{1+q}{1-q} = 3$. The reserve price must be below a third of the highest bidder's valuation for the seller to have an incentive to shill. If the proxy bidder's highest valuation is revealed as 600 GBP, the seller only has an incentive to shill up to the highest bidder's valuation if the current standing price is less than 200 GBP to guarantee that squeeze shilling has a higher expected revenue than not squeeze shilling. It follows that when the shilling detection rate is low, shill bidders always have the incentive to conduct squeeze shilling at $t < 1$, under a private valuation paradigm.

It is worth noting that the theorem can be applied not only to the reserve price, but also to the standing bid. For example, if the standing bid is currently at 600 GBP but the highest bidder, say bidder k, submits a proxy bid of 800 GBP, and if the probability of detection $q = 0.5$, then, $\frac{1+q}{1-q} = 3$. Since $\frac{800}{600} < 3$, the seller would not have an incentive to shill any further as the expected revenue would be lower to shill than to not shill.

Theorem 2: [Under the private valuation paradigm, in the case where a proxy bidder submits a high bid before $t = 1$, the seller's strategy of not shilling is strictly dominated by the strategy of last minute squeeze shilling]

Proof. Assume the same auction setting as in Theorem 1. Consider that after revealing proxy bidder j's highest bid, shill bidder i waits until the last minute and submits a bid equal to bidder j's valuation, v_j, at $t = 1$, instead of $t < 1$, where the last minute bid is transmitted successfully with a probability of $p < 1$. If the shill bid is transmitted successfully, proxy bidder j has no time to react, which allows the shill bidder to achieve revenue equal to v_j. If the shill bid is not transmitted successfully, the shill bidder achieves revenue equal to the reserve price, S. Therefore, the expected revenue for a last-minute shill bid is equal to:

$$v_j p + S(1-p) \tag{2.3}$$

The following derivation shows that a seller yields a higher expected revenue through last-minute squeeze shilling, again, comparing to a non-shill revenue of S.

$$v_j(p) + S(1-p) - S = v_j(p) - S(p) = (v_j - S)(p) \tag{2.4}$$

Since, $1 \geq p > 0$ and $v_j - S > 0$, then the following must be true,

$$v_j(p) + (S)(1-p) > S \tag{2.5}$$

Therefore, when an opportunistic eBay seller observes a bid from a proxy bidder j, whose strategy is to submit his/her true valuation v_j at $t < 1$, the strategy of not shilling is strictly dominated by the strategy of squeeze shilling at $t = 1$.

2.3.2 An affiliated valuation of equilibrium model with squeeze shilling on eBay

In this section, I present an example that squeeze shilling can also take place in eBay auctions with an affiliated or common valuation setting. I introduce a 'dealer/expert' model similar to that of Roth and Ockenfels (2002) but introduce a seller who uses a false identity to submit shill bids. Bidders are symmetric in valuation apart from their private signals about the common value. The problem of asymmetric information arises from the online bidding environment, where some bidders are better informed than other bidders, especially in auctions of arts and antiques. From their own private information, bidders who have perfect knowledge about the value of the auctioned item are called informed bidders. In contrast, uninformed bidders are unsure about the value of the auctioned item but may try to gather information about the value of the auctioned item through observing early bids submitted by other informed bidders. Therefore, in order to prevent giving extra information about the value of the auctioned item and avoid a bidding war, many eBay bidders choose to submit their bids at the very end of the auction, which is widely known as ‘sniping’. I present that eBay sellers have the incentive to submit early shilling to give a signal to the uninformed bidders that the item is genuine and induce the uninformed bidders to submit their bid. This leads to an equilibrium with a higher expected auction ending revenue where the seller prefers to shill bid than not shill bid. The worst situation of seller shilling may induce a bidder with imperfect information to purchase a fake item.

Example: [Example of an equilibrium where shill bidders submit an early shill bid in eBay auctions with an affiliated valuation setting.]

Example: Consider the eBay bidding environment with an affiliated valuation component. There is one seller who wants to allocate one indivisible item. Ex-ante, the seller sets a reserve price, S. There are three bidders: an informed bidder k, an uninformed bidder j, and a shill bidder i.

- The first bidder is the informed bidder k, who can identify whether an item is fake or genuine. Bidder k has the valuation, $v_k(F) = 0$ if the item is fake, and $v_k(G) > M + S$ if the item is genuine. If the item is fake, bidder k will not place a bid on an item with no value, as it will yield a negative payoff, thus, the strategy of bidding any positive amount is dominated by no bid. If the item is genuine, bidder k submits a bid of $v_k(G)$ at the last minute, $t = 1$, to prevent giving extra information to other bidders.

- The second bidder is the uninformed bidder j, who is a proxy bidder with an uncertain valuation towards the auctioned item and cannot distinguish whether the item is fake or genuine. Bidder j has the valuation, $v_j(F) = 0$ if the item is fake, and $v_j(G) > v_k(G) + M > M + S$ if the item is genuine. Bidder j knows that: a) informed bidders will not bid on an item that is fake, and b) other uninformed bidders will not bid if there are no other bids. Therefore, if and only if, the uninformed bidder j observes another bidder who submits a positive bid, bidder j will conclude that the auctioned item is genuine and submit a bid of v_j through the proxy bidding system.

First, consider an equilibrium without shilling. If the item is genuine, the informed bidder k chooses to bid his/her valuation, v_k, at the last minute, $t = 1$, to protect his/her private information and avoid a bidding war. The probability of a last minute bid being transmitted successfully is equal to p. The uninformed bidder j does not bid unless s/he observes another bid. Since bidder k submits a bid at $t = 1$, bidder j does not observe the bid submitted by bidder k and does not have time to react to it. Then, if bidder k's last minute bid is transmitted successfully, s/he wins the auction with a single last-minute bid and pays the reserve price S. If the bid is not transmitted successfully, the item is not sold. Therefore, when the item is genuine the seller's expected revenue is equal to pS.

If the item is fake, the equilibrium without shilling is where neither the informed bidder nor the uninformed bidders place bids, so the item is not sold and the expected revenue for the seller is equal to 0.

Now I introduce a shill bidder i.

- The third bidder is the shill bidder i, who is the seller under a false identity. Initially the shill bidder places a shill bid at the reserve price, S, at $t = 0$, and conducts squeeze shilling if they observe any proxy bid submitted by a proxy bidder at $t < 1$.

First, consider the case with shilling where the item is genuine. After observing the initial shill bid S, the uninformed bidder j is led to believe the item is genuine and places a bid of v_j using the eBay proxy bidding system. The current standing bid rises to $M + S$. The shill bidder i observes the bid from proxy bidder j and conducts last-minute squeeze shilling at $t = 1$. Since the item is genuine, the informed bidder k places a last-minute bid of $v_k(G) < v_i(G) < v_j(G)$, at $t = 1$. If the shill bidder i's last-minute bid is accepted, the expected ending price is not affected by bidder k's last-minute bid. Therefore, the auction ending price is v_j with probability p. If the shill bidder i's last-minute bid is not accepted, but the informed bidder k's last-minute bid is accepted, then the auction ending price equals $v_k + M$, with probability $(1-p)p$. If both last minute bids from bidder i and bidder k are lost, the auction ending price equals $S + M$, with probability $(1-p)^2$.

Therefore, the expected revenue for sellers who submit early shills and conduct last-minute squeeze shilling is equal to

$$pv_j + (1-p)p(v_k + M) + (1-p)^2(S + M) \tag{2.6}$$

Since each of the three terms are positive and $v_j > S$, it follows that $pv_j > pS$ and the expected revenue with shilling is greater than the seller's expected revenue pS without shilling.

$$pv_j + (1-p)p(v_k + M) + (1-p)^2(S + M) > pS \tag{2.7}$$

When the item is genuine, the expected revenue with shilling generates a higher expected revenue than without shilling.

Now, consider the case with shilling where the item is fake. As above, the shill bidder i submits an initial shill bid at the reserve price, S, at $t = 0$. The informed bidder k does not submit a bid since the item is not genuine. The uninformed proxy bidder j observes the early bid of S and submits a bid of $v_j(G)$ using the proxy bidding system. Then, the shill bidder then conducts squeeze shilling at $t = 1$. If the last minute bid is successfully transmitted, the seller yields a revenue of pv_j. If the last minute bid is not successfully transmitted, the seller yields a revenue of $(1-p)(S + M)$. It follows that the seller's expected revenue is equal to $pv_j + (1-p)(S + M)$. It can be compared to the expected revenue without shilling, equal to 0.

$$pv_j + (1-p)(S + M) > 0 \tag{2.8}$$

When item is fake, the expected revenue with shilling generates a higher expected revenue without shilling.

Equations 2.7 and 2.8 show that eBay sellers that sell both genuine and fake items have the incentive to conduct early shilling. Even without any prior bids, shilling induces the uninformed bidder to join the auction and can provoke a bidding war, hence yielding a higher profit. Early bidding induces the uninformed bidder to submit their high bid and gives the opportunistic seller the incentive to conduct squeeze shilling. The worst situation that arises from shilling is the possibility of inducing an uninformed bidder to win a fake item, which yields a negative payoff for the bidder. Previous explanations of observed early bidding on eBay were item book-marking and irrational bidding behaviour, but my example gives an alternative explanation to early bidding behaviour in eBay auctions - early bids could be shill bids submitted by shill bidders in order to provoke a bidding war or to induce other bidders to join the auctions.

In general, the proofs above show that eBay opportunistic sellers have the incentive to submit shill bids to inflate the auction ending price in eBay auctions under both a private and an affiliated valuation paradigm. It also sheds a light on how squeeze shilling affects the auction bidding history from the aspects of auction ending price and number of bidders. I predict that auctions that have shilling should have higher auction ending prices and higher numbers of bidders, due to the fact that opportunistic sellers use multiple false bidding accounts to submit shill bids in order to decrease the detection rate of shilling, which contributes to a higher number of participants in auctions with squeeze shilling. Due to the nature of shill bidders, who do not have the intention of purchasing the item that they bid on, the shill bidders' accounts should have very low feedback. I employ these predictions to empirically test the existence of squeeze shilling on eBay platform.

2.4 An empirical test of squeeze shilling on eBay

The theoretical analysis in section 2.3 suggests the possible effects of squeeze shilling on auction bidding history and characterises the typical features of shill bidders, which are summarised as follows. First, as the aim of shilling is to increase the auction ending price, auctions with squeeze shilling should have higher auction

Table 2.1: Predictions based on the results of the theorems and example

Non-strategic reasons for retraction:	**Standing bid**	**Number of bidders**
Bidder changes mind	Decreases	Decreases
The item description changes	Decreases	Decreases
Naive bidding behavior	Decreases	Decreases
The bidder is unable to contact seller	Decreases	Decreases
Strategic reasons for retraction:	**Standing bid**	**Number of bidders**
Seller squeeze shilling	Increases	Increases

ending prices than those listings without squeeze shilling behaviour. Second, to decrease the detection rate of squeeze shilling, shill bidders tend to employ multiple false identities to make the shill bids look legitimate, as one false user identity is used to reveal the proxy bidder's maximum bid and other false user identities are used to inflate the auction standing price up to the proxy bidder's high bid. Therefore, auctions with squeeze shilling should have a higher number of bidders than the listings without squeeze shilling. Third, shill bidders frequently bid on items from the same seller but have a low probability of winning the auction, the user identities of shill bidders' accounts should have lower total feedback. These predictions are used to test for the existence of squeeze shilling on the eBay platform.

The process of successful squeeze shilling is described as follows. A shill bidder becomes the highest bidder in an auction, which inherently reveals the maximum bid of the second highest proxy bidder. Then, the shill bidder retracts the bid before the end of the auction, and shills up the standing price to the second highest bidder's valuation. I define the behaviour of bidder's bid and retraction as two types, the first type is called a strategic bid-retraction, which is conducted by a shill bidder, who specifically bids and retracts a high bid in order to reveal the proxy bidder's maximum bid. The second type is a non-strategic bid-retraction, made by honest bidders due to other reasons described in Table 2.1. The predictions in the table are based on the results of the theorems and the example. The difference is that the probability that other bidders observe a revealed high bid is very low, and hence it is unlikely that a shill bidder observes a non-strategic retraction.

In contrast to strategic bid-retractions, non-strategic bid-retractions are unlikely to reveal any useful information to facilitate the seller to conduct squeeze shilling behaviour. Therefore, strategic and non-strategic bid and retractions have different effects on auction bidding patterns and ending prices, see again, Table 2.1. I predict that strategic retractions are a good indicator of squeeze shilling behaviour.

From the auction bidding history observations, I mark the auctions with retractions and separate my dataset into two groups. The first group contains listings that are suspected to have squeeze shilling with strategic-retractions, and all listings belonging to the first group must have at least one retraction which reveals a proxy bidder's bid. The second group contains all the other auctions in the sample that are not suspected to have squeeze shilling.

2.4.1 Data and summary statistics

To empirically test the existence of squeeze shilling behaviour on the eBay platform, I collect data from the completed eBay auction listings of new condition iPhones, under the category of 'smart-phones'. The definition of new condition items given by eBay is, 'brand-new, unused, unopened and undamaged items in original retail packaging'. The data is collected between the 26th September to 24th November 2014; the short 60-day time interval for data collection reduces the effect of discounting for electronic items. Observations that are excluded from my dataset are: auctions with missing data or non-standard descriptions; auctions in which the seller is not registered with eBay; the time that an auction ends is not the same as stated at the time the auction commences[5]; auctions that include the 'buy it now' option; auctions where the secret reserve price is not met. The resulting dataset consists of 1075 observations of two models of iPhone (iPhone 6 and iPhone 6 plus), in total, 897 of these auctions are sold and the rest are unsold. It is interesting to note that a moderately high proportion of auctions (7.91%) in the sample have at least one retraction. If a high proportion of these listings with retractions prove to

[5]Sellers can pay an extra fee to eBay to specify or change the time that the auction ends.

have evidence of seller shilling, then squeeze shilling behaviour takes place more frequently on eBay than other known seller shilling behaviours. In comparison, in citeengelberg2005license, only 1.5% of auctions had the discover-and-stop seller shilling behaviour.

I collect a number of control variables which consist of product characteristics, seller characteristics, price information and auction information. Table 2.2 describes the variables in my dataset and Table 2.3 provides the summary statistics.

- Product characteristics include: phone generation, phone storage and phone condition.

- Seller characteristics include: the user identities, the lengths of time as an eBay member and total feedback scores since user registration.

- Price information includes: the reserve price, the ending price, postage, and the book value, which I obtain for each item specification from Apple's official store.

- Auction information includes: the duration of the auction, the number of bidders, the auction ending day of the week, the number of days in return policy, the number of retractions and the number of cancellations.

To protect the privacy of bidders, eBay does not disclose the full user identities, which can be used to identify and trace individual bidders. However, different bidders in the auction can be distinguished from a combination of user reputation scores and the times of user registration.

Figure 2.1 illustrates all the observed bid retractions in my dataset subject to a standardised auction timescale. The auction durations in my dataset vary between 1, 3, 5, and 7 days. For the purpose of illustration and comparison, the timescale of each auction is standardised to 1, where 1 indicates the end of the auction. For the purposes of clearer illustration, figure 2.2 excludes the two observations with

Table 2.2: Variable description

Variable	Description
Reserve price	The pre-determined price at which an auction commences.
Ending price	The price at which an auction ends.
Duration	The duration of an auction: 1 day, 3 days, 5 days or 7 days.
Duration dummy	Dummy variables for each auction duration, for example D1 takes the value 1 for a 1-day duration auction and 0 otherwise.
Retraction	A dummy variable equal to 1 if a listing has at least one retraction and zero otherwise.
Cancellation	A dummy variable equal to 1 if a listing has at least one cancellation and zero otherwise.
Storage	Three dummy variables for the phone storage, 16GB, 64GB or 128GB, for example S16 is equal to 1 for a phone storage of 16GB and 0 otherwise.
Book value	The retail price of each different mode of iPhone device, collected from the Apple's official store.
Bidders	The number of actual bidders who participate a specific auction
Total feedback	The seller's overall feedback rating since seller registration.
iPhone Plus	A dummy variable equal to 1 if an auctioned phone with a mode 'iphone 6 plus', and zero otherwise.
Day of week	Dummy variables for the last day of the auction, Monday to Sunday, for example, Monday, equal to 1 for an auction ending on Monday, and 0 otherwise.
Return	A dummy variable equal to 1 if a seller accepts returns and zero otherwise.
Postage	A dummy variable equal to 1 if the seller provides free delivery and zero otherwise.

Figure 2.1: Retractions on a standardised auction timescale, full dataset

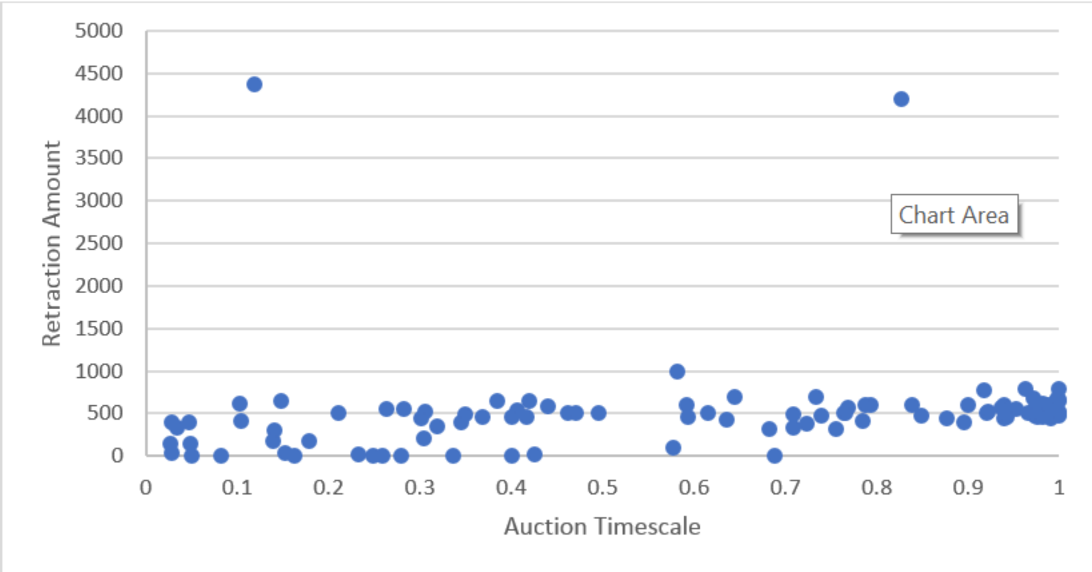


Figure 2.2: Retractions on a standardised auction timescale, excluding two data points

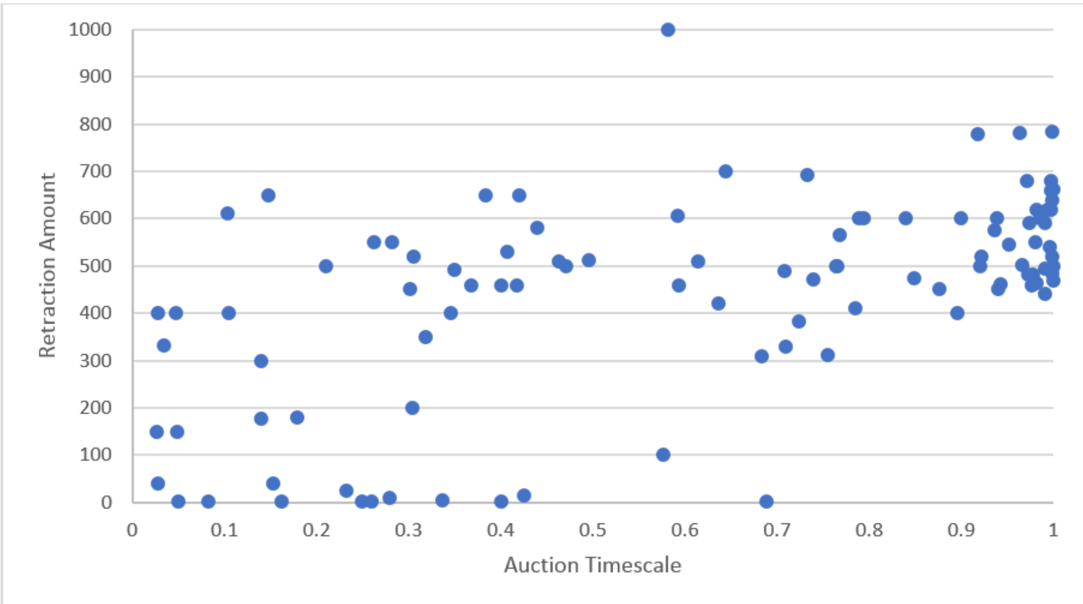


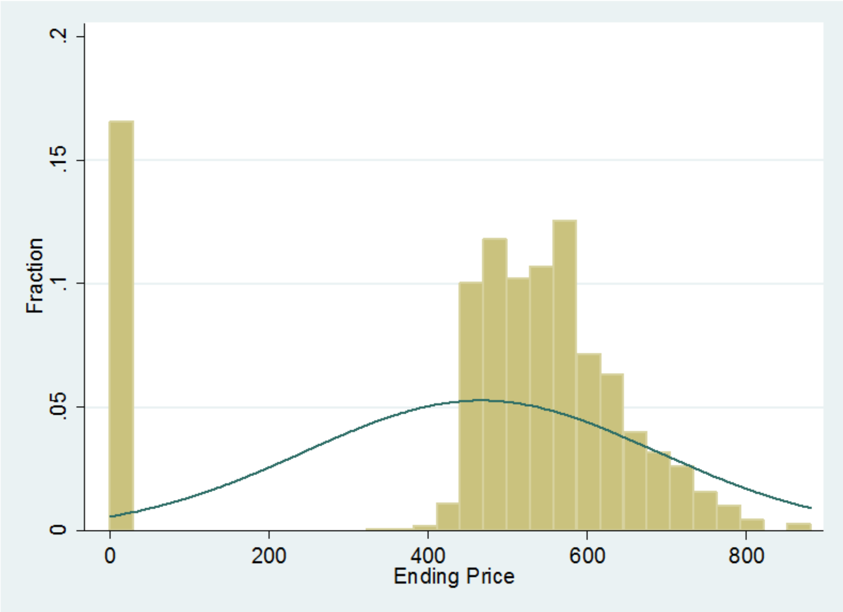
Table 2.3: Summary statistics

	1075 Observations				897 Observations			
	Include unsold				Exclude unsold			
Variable	**Mean**	**Std.Dev.**	**Min**	**Max**	**Mean**	**Std.Dev.**	**Min**	**Max**
Ending price	466.12	221.96	0	880	558.61	85.59	346	880
Reserve price	215.44	237.02	0.01	1000	178.74	219.91	0.1	800
Bidders	7.27	5.26	0	26	8.66	4.58	0	26
Bids	15.87	14.21	0	85	18.90	13.55	0	85
Duration	2.83	1.97	0	7	2.92	2.02	1	7
Retraction	0.09	0.37	0	5	0.11	0.38	0	5
Cancellation	0.14	0.76	0	11	0.08	0.47	0	7
Storage	44.90	39.00	16	128	46.40	39.43	16	128
Postage	5.87	3.22	0	8.75	5.94	3.18	0	8.75
Book value	609.60	74.00	539	789	612.88	74.06	539	789
Total feedback	689.18	4803.64	-2	89441	757.07	5237.68	0	89441

retraction amounts of 4365 GBP and 4206 GBP, which seem to be irrational bids. It is observed from the bidding histories that these two bidders resubmitted their bids and these retractions are typographic errors. Figure 2.2 shows that a disproportionately large amount of auctions congregate around the end of the auction. As the majority of bidders do not resubmit new bids after retracting, it seems that these retractions are not a consequence of typographic errors.

In my dataset, a total of 87 listings receive in total 104 retractions, in which: 77 listings have one retraction, 6 listings have two retractions, 2 listings have three retractions, 1 listing has four retractions and 1 listing has five retractions. By analysing the bidding histories, I deduce that 39 listings have strategic retractions and the rest have non-strategic retractions. It is also interesting to note that the listing with five retractions had one bid retraction that revealed the highest proxy bidder's bid. The bidder who made a retraction resubmitted his/her bid up to the amount of the highest bidder's maximum proxy bid. However, this suspicious bidding pattern led to all the bidders who submitted their bids in the last hour prior to the auction ending, successively retracting their bids. As a result, the seller cancelled the auction

Figure 2.3: Histogram of auction ending prices

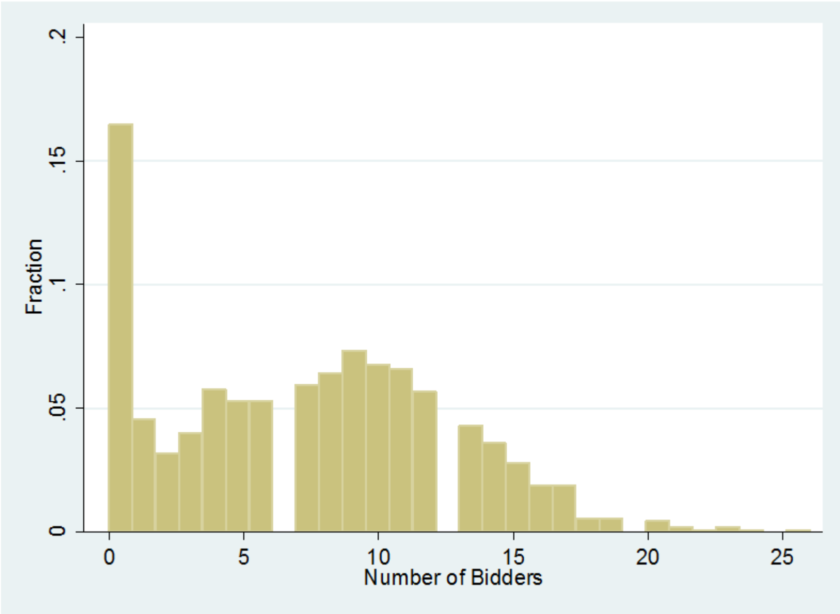


and the item was unsold.

Figure 2.3 illustrates the histogram of auction ending prices, which consists of 1075 auctions: 897 of them receive at least one bid (auction ending price > 0) and 178 auctions do not receive any bids (auction ending price = 0). One interpretation is that zero is a censored observation. Suppose bidder i has a latent valuation, denoted by v_i, s/he only places a bid if his/her valuation is higher than the reserve price S. Then, the winning bidder's bid can only be observed, when it is larger than the reserve price S. Therefore, auctions that have an ending price of zero can be interpreted as a left-censored variable, that equals zero when bidders' valuations are smaller than S. Figure 2.3 shows that the auction ending prices are heavily skewed with a clear non-normal kurtosis $(skewness = -1.28, kurtosis = 3.50)$.

Figure 2.4 illustrates the histogram of the number of bidders since the number of bidders can only take a non-negative integer value, it is a typical count variable. There is a wide variation in the number of unique bidders in each auction in my data sample. 28.93% of observations exceed 10 bidders, and the highest number of bidders in the sample is 26 bidders. Over 97.7% of the values are under 17. In the sample, 16.47% of auctions have no bidder participation.

Figure 2.4: Number of bidders



2.4.2 The impact of squeeze shilling on auction ending prices

In this section, my interest lies primarily on whether listings with strategic bid-retractions can generate higher auction ending prices compared to listings without strategic bid-retractions. The dependent variable, auction ending price, has continuous positive values for a wide range of values from 346 GBP to 880 GBP, but a nontrivial fraction of the dependent variable takes on a value of zero with positive probability. A linear regression model is not suitable for my data, as it is possible to obtain negative fitted values that would lead to negative predictions of the dependent variable, which is inconsistent with the fact that all auction ending prices must be non-negative. Additionally, linear regression models have a constant partial effect on the conditional expectation of the dependent variable which does not hold in my sample. Figure 2.3 shows that auction ending prices can be seen as a censored random variable with 0 being the censoring indicator. This suggests that a Tobit model with a left censoring point at zero is suitable for these purposes. However, the Tobit model relies on the normality of error terms in order to get consistent estimators, which might not be appropriate given the skewness and kurtosis of the dependent variable. I solve this problem by taking the natural logarithm of the dependent variable, auction ending prices, which reduces the skewness and non-normal kurtosis of the dependent variable (skewness = -0.38 and kurtosis = 2.53).

The Tobit model presents the observed response, y_i in terms of an underlying latent variable:

$$y^* = x'\beta + \varepsilon, \quad \varepsilon|x \sim \mathcal{N}(0, \sigma^2) \tag{2.9}$$

$$y = \max(0, y^*). \tag{2.10}$$

The latent variable, y^*, satisfies the classical linear model assumption; in particular, it has a normal, homoskedastic distribution. x' is the vector of control variables, which includes: $ln(reserve\ price)$, $ln(book\ value)$, $ln(total\ feedback)$, $number\ of\ bidders$, $duration$ and a number of dummies for: $day\ of\ the\ week$, $retraction$, $cancellation$, $storage$, $plus$, $returns$ and $postage$. The key independent variable for my Tobit regressions is the dummy for strategic retractions and I control for different covariates in four different Tobit specifications.

Table 2.4 reports the results for the Tobit regressions for four different specifications. Tobit (1) presents the Tobit regression estimates controlling only for ln(reserve price), retraction, cancellation, ln(total feedback), number of bidders and auction duration. Tobit (2) presents the estimates with added control for ln(book value). Tobit (3) presents the estimates with added dummies for the day of the week and for duration. Tobit (4) presents the estimates with added dummies for phone characteristics, returns and postage. Robust standard errors are reported in parentheses. Across all regressions, the estimates for the dummy of strategic retraction has a significant positive effect ($p - value < 0.01$) on auction ending prices, which is consistent with my prediction that listings with squeeze shilling behaviour tend to generate higher auction ending prices than listings without squeeze shilling behaviour. An increase in the reserve price results in an increase in auction ending price. These results are consistent with auction theory, which shows that under both a private and common-valuation paradigm, an increase in reserve price increases the auction ending price, as in Myerson (1981) and Milgrom and

Table 2.4: Tobit and CLAD regression: the impact of squeeze shilling on auction ending price

	Dependent variable ln(auction ending price)				
Independent	**Tobit(1)**	**Tobit(2)**	**Tobit(3)**	**Tobit(4)**	**CLAD**
ln(reserve price)	.0107101	.0120593	.0124564	.0114851	0.00999
	$(.0026959)^{***}$	$(.002142)^{***}$	$(.0021515)^{***}$	$(.002092)^{***}$	$(-.0001)^{***}$
Retraction	.1888335	.0680169	.0661065	.0595965	.0448933
	$(.0229884)^{***}$	$(.0224718)^{***}$	$(.0222546)^{***}$	$(.0226471)^{***}$	$(.0174611)^{**}$
Cancellation	-.1423231	-.1142384	-.1146215	-.1042091	-.085769
	$(.0285306)^{***}$	$(.0234926)^{***}$	$(.0235641)^{***}$	$(.0236122)^{***}$	$(.0286508)^{***}$
ln(book value)		1.005168	.9976186	-2.626379	1.082188
		$(.0543344)^{***}$	$(.0540814)^{***}$	$(1.371376)^{*}$	$(.0357588)^{***}$
ln(total feedback)	.0200954	.0093315	.0103389	.0112182	.0073983
	$(.0042637)^{**}$	$(.003826)^{**}$	$(.0038025)^{***}$	$(.004795)^{**}$	$.0028072^{***}$
Bidders	.0287668	.0281913	.0282304	.0279492	.0158843
	$(.0019758)^{***}$	$(.0017446)^{***}$	$(.0017268)^{***}$	$(.0017141)^{***}$	$(.0019537)^{***}$
Duration	.0127779	.0084009			.0066961
	$(.0031799)^{***}$	$(.0026192)^{***}$			$(.0020932)^{***}$
D3			-.0091056	-.0082247	
			$(.0119684)$	$(.0117416)$	
D5			.0341096	.0312218	
			$(.0158118)^{**}$	$(.01599)^{*}$	
D7			.0538603	.0540493	
			$(.0181301)^{***}$	$(.0176466)^{***}$	
Fri			.043049	.0447948	
			$(.0157826)^{***}$	$(.0155128)^{***}$	
Sat			-.0018176	.0024883	
			$(.0192216)$	$(.0190317)$	
Mon			.028892	.0293954	
			$(.0179857)$	$(.0176779)^{*}$	
Tue			-.0008639	-.0004157	
			$(.0207253)$	$(.0205338)$	
Wed			.0147807	.0148416	
			$(.0194022)$	$(.0191554)$	
Thu			.0116789	.013382	
			$(.0181295)$	$(.0177604)$	
S64				.5278515	
				$(.1840561)^{***}$	
S128				.9053034	
				$(.3495027)^{**}$	
Plus				.4661141	
				$(.1809749)^{**}$	
Return				.0212718	
				$(.0224309)$	
Postage				.0023701	
				$(.0180837)$	
cons	5.841108	-.5246957	-.4799905	22.29107	-.8627024
	$(.0328815)$	$(.3385496)$	$(.3373076)$	(8.628218)	$.2265677$
$Pseudo R^2$	0.8737	1.5869	1.6161	1.6644	0.3961

Notes. *p < 0.10, **p < 0.05, ***p < 0.01. Robust standard errors are given in parentheses

Weber (1982). However, it is also shown that a higher reserve price can reduce the number of bidders that participate in the auction if the reserve price exceeds the bidders' valuations. The regression results also indicates that listings with a longer auction duration, a higher number of bidders and higher total feedback scores are positively correlated with action ending prices. Auctions that end on Friday tend to generate higher ending prices than listings that end on other days of the week. The dummy for cancellation has a significant negative effect on auction ending prices. All these coefficients are significant at conventional levels. Free postage and a 14-day/60-day return policy have no effect on auction ending price.

Table 2.5 reports the marginal effect estimated at sample mean for all variables after the Tobit regression. The result in Tobit (1), which does not control for ln(book value), indicates that the auction ending prices are around 17.27% higher for listings with strategic bid-retractions than listings without strategic bid-retractions. After I control ln(book value) for Tobit (2), (3) and (4), the marginal effect of strategic retraction drops from 17.27% in Tobit (1) to 6.46%, 6.28% and 5.68% respectively. However, this effect is still substantial. In other words, a seller who conducts squeeze shilling can yield around 6% higher revenue than a seller who does not conduct squeeze shilling. These results are consistent with my predictions based on my theorems, shown in Table 2.1. The results also shows that for opportunistic sellers have the robust incentive to conduct squeeze shilling, given that the penalties of reported shilling on eBay are not harsh. This is the first piece of empirical evidence to support the existence of squeeze shilling behaviour in eBay auctions. Table 2.5 also indicates that listings that have bid cancellations tend to have auction ending prices that are around 10% lower, compared with listings that do not have bid cancellations. It is observed that many of the bids that were cancelled by sellers were made by bidders with very few or no feedback scores, which could be due to the sellers' concerns about incompletion in final transactions. It is also noted that listings that end on Friday tend to generate higher total ending prices of around 4% compared with listings that end on other days of the week. Table 2.5 indicates that an increase of one bidder from the sample mean of bidders, increases

Table 2.5: Tobit regression: the impact of squeeze shilling on auction ending price (marginal effects at mean)

	Marginal effect at mean			
Independent	**Tobit(1)**	**Tobit(2)**	**Tobit(3)**	**Tobit(4)**
ln(reserve price)	.0090261	.0111874	.011589	.0107396
	(.002268)***	(.0019869)***	(.0019994)***	(.0019538)***
Retraction	.1726629	.0645643	.0628631	.0568062
	(.0220956)***	(.0216222)***	(.0214495)***	(.021847)***
Cancellation	-.1090881	-.0999157	-.1005956	-.0926596
	(.0196072)***	(.0189505)***	(.0190692)***	(.0195973)***
ln(book value)		.9324986	.9281531	-2.455914
		(.0573834)***	(.056738)***	(1.283418)*
ln(total feedback)	.0169358	.0086569	.009619	.0104901
	(.0036123)***	(.00352)**	(.0035178)***	(.0044776)**
Bidders	.0242437	.0261532	.0262647	.0261351
	(.0015615)***	(.0014578)***	(.0014517)***	(.0014458)***
Duration	.0107688	.0077936		
	(.002687)***	(.0024237)***		
D3			-.008462	-.0076833
			(.0885272)	(.0756789)
D5			.0320666	.0294647
			(38.7937)	(34.82817)
D7			.0509015	.051305
			(6566.596)	(6137.172)
Fri			.0404914	.0423425
			(.0149307)***	(.0147397)***
Sat			(-.00169)	.0023287
			(.0178616)	(.0178253)
Mon			.0271075	.0277133
			(.0169581)	(.0167423)*
Tue			-.0008035	-.0003886
			(.0192716)	(.0191962)
Wed			.0138147	.0139396
			(.0182076)	(.0180605)
Thu			.0109043	.0125621
			(.0169821)	(.0167308)
S64				.5005562
				(.1704685)***
S128				.8849758
				(.3396856)***
Plus				.4399281
				(.1666408)***
Return				.019731
				(.0206089)
Postage				.0022149
				(.016887)

Notes. *p < 0.10, **p < 0.05, ***p < 0.01. Delta-method standard errors are given in parentheses

Table 2.6: Prediction from Tobit (1)

Prediction from Tobit (1)					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
$EndingPrice$	1,075	467.8707	220.1328	0	880
$Endin\hat{g}Price$	1,075	495.3747	108.1264	126.3493	990.5732

Table 2.7: Prediction from Tobit (2)

Prediction from Tobit (2)					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
$EndingPrice$	1,075	467.8707	220.1328	0	880
$Endin\hat{g}Price$	1,075	498.9988	129.123	59.62115	1123.093

Table 2.8: Prediction from Tobit (3)

Prediction from Tobit (3)					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
$EndingPrice$	1,075	467.8707	220.1328	0	880
$Endin\hat{g}Price$	1,075	499.108	129.6006	53.58391	1109.767

Table 2.9: Prediction from Tobit (4)

Prediction from Tobit (4)					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
$EndingPrice$	1,075	467.8707	220.1328	0	880
$Endin\hat{g}Price$	1,075	499.2611	130.0753	63.41147	1091.156

Table 2.10: Test of normality of Tobit specifications

Test of normality for Tobit		
	NR^2	**p-value**
Tobit (1)	1.8337021	.39977594
Tobit (2)	118.91111	1.509e-26
Tobit (3)	112.86363	3.104e-25
Tobit (4)	128.83002	1.059e-28

Table 2.11: Test of homoskedasticity for Tobit

Test of homoskedasticity for Tobit		
	NR^2	**p-value**
Tobit (1)	188.95479	9.311e-42
Tobit (2)	339.90367	1.552e-74
Tobit (3)	338.88117	2.588e-74
Tobit (4)	349.94298	1.025e-76

the auction ending price by around 2.6%, across all the Tobit specifications.

The prediction outcomes in Tables, 2.6 2.7, 2.8, 2.9 indicate that the Tobit model fits my data well, for example, the prediction from Tobit (2) shows that the sample mean of ending price is equal to 467.87 GBP, which is very close to the value of the predicted mean of ending price which is equal to 490.00 GBP. Also, the minimum of the sample is 0 GBP which is considerably close to the minimum of the predicted ending price of 59.62 GBP.

I conduct tests of normality and homoskedasticity by following the Stata user written coding given in chapter 16 of Cameron and Trivedi (2010). To implement the test, I compute and store various components of the test statistic, including the inverse of the Mill's ratio, the generalised residuals and the likelihood scores. Table 2.10 presents the results of the normality tests for the Tobit regressions in Table 2.4. I test the null hypothesis, H_0: the disturbances in the Tobit model have a normal distribution, against the alternative hypothesis, H_1, the disturbances in the Tobit model is not a normal distribution. The test results of Tobit(1) leads to an accept of the null hypothesis of the disturbance in Tobit model have a normal distribution ($p - value > 0.1$). The test outcomes of Tobit(2), Tobit(3) and Tobit(4) reject the normality hypothesis ($p - value < 0.01$), even though the dependent variable is transformed to logarithms.

Table 2.11 presents the results for the test of homoskedasticity for Tobit regressions in Table 2.4. I test the null hypothesis, H_0 : the variance is homoskedastic, against the alternative hypothesis, H_1 the variance has heteroskedasticity of unknown form. The test results of Tobit (1), Tobit (2), Tobit (3) and Tobit (4) reject the null hypothesis that the variance is homoskedastic ($p - value < 0.01$). Due to concern of heteroskedasticity, I report robust standard errors in my Tobit regressions. However, if the errors are heteroskedastic, the estimator might be biased and/or inconsistent. To solve the problems that arise from heteroskedasticity and non-normal error distributions, I employ a semi-parametric estimation to calculate the censored least absolute deviations estimator (CLAD), and bootstrap estimates of its sampling variance. The CLAD estimator is robust to heteroskedasticity and is consistent and asymptotically normal for a wide class of error distributions. I report the result for the CLAD regression in Table 2.4 and bootstrap standard errors are given in parentheses.

The model of CLAD (censored least absolute deviations estimator) is:

$$y = \max(0, y^*) + \epsilon. \tag{2.11}$$

$$Med(y^*|x) = x'\beta \tag{2.12}$$

The dependent variable in my CLAD model is $ln(auction\ ending\ price)$ and the independent variables include $ln(reserve\ price)$, $ln(book\ value)$, $ln(total\ feedback)$, $number\ of\ bidders$, $duration$. The estimate for the dummy of strategic retraction has a significant positive effect ($p - value < 0.05$) on auction ending prices and the marginal effect falls from 6.5% to 4.5%, in comparison to the marginal effect at mean for Tobit(2), however the effect is still substantial and it is consistent with my prediction that sellers who conduct squeeze shilling can generate higher revenue than those sellers who do not conduct squeeze shilling.

2.4.3 The impact of squeeze shilling on bidder entry decision

In this section, I identify the determinants of the bidder entry decision to investigate whether the auctions that have strategic retractions have a higher number of bidders compared to listings without strategic bid-retractions. The dependent variable is the number of bidders in each auction, which is a count variable, which only takes on a non-negative integer value. In my sample, 16.47% of auctions have no bidder participation. For the same reasons discussed for Tobit responses, an OLS model is not suitable for my data structure due to the possibility of obtaining negative fitted values and negative predictions for the dependent variable, which conflicts the fact that the number of bidders visiting an auction are non-negative. Therefore, I employ a Poisson regression model to investigate the determinants of bidder entry.

I assume that the dependent variable, the number of bidders, y_i, given a vector of covariates, x_i, has a Poisson distribution. However, Poisson maximum-likelihood estimation might not be appropriate in my case, since in my application, the sample mean of the bidders for each auction is 7.266047, which is much smaller than the sample variance, 27.7149 (5.2645^2), which seems to contradict the fact that a Poisson random variable satisfies $E(y) = v(y)$. Therefore, I use a quasi-maximum-likelihood approach. The vector of covariates, x_i, includes: $ln(reserve\ price)$, $ln(total\ feedback)$, $ln(bookvalue)$, $duration$ and dummies for $duration$, $retraction$, $cancellation$ and $day\ of\ the\ week$. The key independent variable for my Poisson regressions is the dummy for the strategic $retractions$ and I control for different covariates in three different Poisson specifications.

Table 2.12 reports the results for the Poisson regression model for three different specifications. Poisson (1) presents the Poisson regression estimates controlling only for $ln(reserve\,price)$, $duration$, $ln(total\ feedback)$, $ln(book\ value)$ and dummies for $retraction$ and $cancellation$. Poisson (2) presents the estimates with added dummies for $duration$. Poisson (3) presents the estimates with added dummies for $day\ of\ the\ week$. Robust standard errors are given in the parentheses. Across all regressions, the estimates for the dummy of strategic retraction has a significant positive effect ($p - value < 0.05$) on bidder's entry. These results are consistent with my theoretical predictions where opportunistic sellers use multiple false identities to submit fraudulent shill bids in their own auctions to inflate the auction ending prices. Shill bidders use multiple identities to prevent genuine bidders from detecting a shill bid account. Frequently used shill bid accounts are highly suspicious to other bidders, as bidders can observe from the user profile of a shill bidder that they have submitted multiple bids on items from the same seller. It is also possible to see the number of retractions, so a high number of retractions would also cause suspicions. Therefore, multiple bidding accounts can reduce suspicions. Cancellations have a significant negative effect ($p - value < 0.01$) on bidder's entry. One possible reason is that cancellations are usually targeted at bidders with low feedback scores, so newly-registered bidders choose not to participate in the auctions where they observe that the seller cancels bid from another new user. However, it is possible that the newly registered bidders have a high valuation and they are genuinely interested in purchasing the item. As a result, auctions with cancellations have a lower ending price. As expected, auctions with longer durations tend to have a higher number of bidders. It is interesting to note that the seller's total feedback has no effect on bidder's entry decision. A lower reserve price induces more bidders to participate in an auction. In general, the day of the week that the auction ends does not seem to effect bidder's entry, apart from Wednesday which has fewer bidders in an auction than the other days in the week ($p - value < 0.1$).

Table 2.13 presents the marginal effect at the mean value of covariates. The table shows that the dummy variable for strategic retractions has a positive effect on bidder entry decision. Listings with strategic bid-retractions have 1.2 bidders

Table 2.12: Poisson and NB2 regressions assessing the impact of squeeze shilling on bidder entry decision

Dependent variable: number of bidders				
Independent	**NB2**	**Poisson(1)**	**Poisson(2)**	**Poisson(3)**
ln(reserve price)	-.1708235	-.1537368	-.1539674	-.1542728
	$(.0075103)^{***}$	$(.00664272)^{***}$	$(.0064506)^{***}$	$(.0064456)^{***}$
ln(total feedback)	-.0013233	.0010486	.0009678	.001543
	$(.0146196)$	$(.0121436)$	$(.0121484)$	$(.0122438)$
Retraction	.252401	.1811935	.1789075	.1752677
	$(.1146641)^{**}$	$(.0760327)^{**}$	$(.0757636)^{**}$	$(.0747101)^{**}$
Cancellation	-.3073345	-.3188986	-.3169084	-.3148474
	$(.0926306)^{***}$	$(.1101521)^{***}$	$(.1101309)^{***}$	$(.1092399)^{***}$
ln(book value)	-.2048086	-.2289599	-.2220415	-0.2132123
	$(.1970827)$	$(.1578818)$	$(.1586299)$	$(.1575879)$
Duration	.0371982	.0374546		
	$(.0110129)^{***}$	$(.0081392)^{***}$		
D3			.0917779	.0924773
			$(.0416053)^{**}$	$(.041419)^{**}$
D5			.1531999	.1501714
			$(.051031)^{***}$	$(.0511948)^{***}$
D7			.2200171	.2284871
			$(.0533721)^{***}$	$(.0544063)^{***}$
Mon				-.0966479
				$(.0692495)$
Tue				-.0936071
				$(.0803296)$
Wed				-.1373889
				$(.0738855)^{*}$
Thu				-.0003827
				$(.0668328)$
Fri				-.0396018
				$(.0614133)$
Sun				-.0799332
				$(.0617351)$
constant	3.599236	3.705891	3.694135	3.697195
	(1.253151)	(1.004227)	(1.011277)	(1.009739)
Pseudo R2	0.0642	0.2080	0.2079	0.2098

Notes. *p < 0.10, **p < 0.05, ***p < 0.01. Robust standard errors are given in parentheses.

Table 2.13: Poisson and NB2 regressions assessing the impact of squeeze shilling on bidder entry decision (marginal effects at mean)

	Marginal effect at mean			
Independent	**NB2**	**Poisson(1)**	**Poisson(2)**	**Poisson(3)**
ln(reserve price)	-1.090371	-.9946682	-.996205	-.9970415
	(.0370139)***	(.0357181)***	(.0358189)***	(.0355748)***
ln(total feedback)	-.0084467	.0067845	.0062618	.0099722
	(.0607782)	(.0785606)	(.0785957)	(.0791177)
Retraction	1.815939	1.276815	1.259383	1.230209
	(.6166278)***	(.5814116)**	(.5783324)**	(.567565)**
Cancellation	-1.721089	- 1.801553	-1.791858	-1.779684
	(.5756381)***	(.5332239)***	(.534185)***	(.5306286)***
ln(book value)	-1.307298	-1.481358	-1.43666	-1.377958
	(1.121139)	(1.022942)	(1.027913)	(1.020188)
Duration	.2374367	.242329		
	(.0607782)***	(.05231)***		
D3			.6018176	.605775
			(25.0293)	(25.0811)
D5			1.052735	1.029501
			(5372.199)	(5270.5)
D7			1.553999	1.61756
			(829402.1)	(880055)
Mon				-.6033066
				(.4173465)
Tue				-.5834307
				(.4820345)
Wed				-.8442818
				(.4307124)*
Thu				-.0024728
				(.4318123)
Fri				-.2526547
				(.3871049)
Sun				-.5043773
				(.3805476)

Notes. *p < 0.10, **p < 0.05, ***p < 0.01. Delta-method standard errors are given in parentheses.

Table 2.14: Tests of overdispersion

Tests of overdispersion					
	y*	**Coef.**	**Std.Err**	**t**	**P>|t|**
Poisson (1)	$\hat{\mu}$	.1504687	.013324	11.293	0.000
Poisson (2)	$\hat{\mu}$	.1503828	.0133216	11.289	0.000
Poisson (3)	$\hat{\mu}$	.1475191	.0134527	10.966	0.000

more than listings without strategic bid-retractions. The results are consistent with the prediction that opportunistic sellers bid in their own auctions using false user accounts, in order to conduct squeeze shilling, which results in a higher number of bidders. Hence, this is the second piece of evidence to support the existence of shilling. The regression results also indicate that when the reserve price increases by 1%, the number of bidders decreases by around 0.01% in the estimation results of Poisson(1), Poison(2) and Poisson(3). Poisson(1) also indicates that when duration increases 1 day, the number of bidders increase 0.242. Listings with at least one cancellation tend to have around 1.8 number of bidders less than listings without cancellations, as shown in the estimation results of Poisson(1), Poisson(2) and Poisson(3).

Table 2.14 present the results of the overdispersion tests for Poisson (1), (2) and (3). To test for overdispersion in count data, I test the null hypothesis of equidispersion, $Var(y|x)$ = $E(y|x)$, against the alternative of overdispersion, $Var(y|x)$ = $E(y|x)$ + $\alpha^2 E(y|x)$. I test the null hypothesis, H_0: $\alpha = 0$, against the alternative hypothesis, H_1: $\alpha > 0$. An overdispersion test statistic can be computed by estimating the Poisson model, constructing fitted values $\hat{\mu}$ = $exp(x'\hat{\beta})$, and running an auxiliary OLS regression of the generated dependent variable, $y^* = \frac{(y-\hat{\mu})^2-y}{\hat{\mu}}$ on $\hat{\mu}$, without an intercept term and performing a t-test of whether the coefficient of $\hat{\mu}$ is zero, see Cameron and Trivedi (2005). The test results indicate there is significant overdispersion in my Poisson specifications, since all the p-values are smaller than 0.01, so I reject the null hypothesis and accept the alternative. This problem can be resolved by relaxing the equivalence assumption by obtaining a robust estimate of the variance-covariance matrix of the estimator (VCE), which can be implemented by the stata command 'vce(robust)'. I report robust standard errors in parentheses in Table 2.12. Another approach to model count data in my case is to use the NB model (negative binomial). The NB2 model with a quadratic variance function is a suitable functional form for count data with overdispersion as NB2 model is consistent under overdispersion that is generated by a Poisson-gamma mixture. I report the results of NB2 regressions in Table 2.12. In Comparison to the Poisson specification (1), the parameter estimates and standard error in NB2 model are similar. The key estimate for the dummy of strategic retraction has a significant positive impact ($p - value < 0.05$) on bidder's entry and those listing with strategic bid-retractions have 1.8 bidders more than listings without strategic bid-retractions, which is consistent with my prediction. The Pseudo R2 is equal to 0.0642, which is much smaller than that in Poisson(1), which equals 0.2080. However this difference does not indicate a worse fit for the Poisson model as the Pseudo R2 is not suitable to compare different types of models.

2.4.4 The shill bidder's profile and bidding history

Finally, the third piece of evidence of strategic squeeze shilling can be found from a shill bidder's bidding profile, which can be found by clicking on the user identity on an item's bidding history page. The bidding profile provides the number of retractions, the bid activity with a certain seller, as well as a 30-day bid history. An example of a bidding profile of a shill bidder is given as a screen capture in figure 2.5. Evidently, shill bidder, d***v, has zero feedback but has 6 bid retractions within the last 30 days and 14 bid retractions in the last six months. In total, s/he bid on 11 items in the last 30 days, but the zero feedback score indicates that s/he did not buy any items or complete any of the transactions. I observe that in the past 60-day history, this eBay member does not win any auctions under the category of 'mobile and smart-phones'. In the 30 day bid history, 80% of the total 36 bids were placed in listings from a particular seller, seller 1. From the 30-day bid history in figure 2.6, it can be observed that this particular shill bidder bids only once in any particular auction, and generally on items from the same seller (seller 1). The last

Figure 2.5: An example of shill bidder’s bidding profile

Bidding Details

Bidder Information		30-Day Summary	
Bidder:	d***v (0)	Total bids:	36
Feedback:	0% Positive	Items bid on:	11
Item description:	iphone 6 plus 64gb space grey	Bid activity (%) with this seller:	80%
Bids on this item:	1	Bid retractions:	6
		Bid retractions (6 months):	14

Figure 2.6: An example of a shill bidder’s bidding history

30-Day Bid History

Category	No. of Bids	Seller	Last Bid
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Garden & Patio > Garden Sheds	1	Seller 2	4d 23h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	2h
Mobile Phones & Communication > Mobile & Smart Phones	1	Seller 1	<1h

column "Last Bid" indicates the amount of time between the bidder’s last bid on the item and the end of the listing. The bid history shows that for bids related to seller 1’s listings, the last bid is usually placed less than an hour before the auction ends, however the bid related to seller 2’s listing is placed nearly 5 days before the end of the auction. It seems as though the bid in seller 2’s listing was possibly made to reduce the statistic of ‘bid activity’ with seller 1. The first column shows that the bids are mainly from the same category of ‘mobile and smart-phones’, where there exists a vast array of choices for the standardised products listed, at any given time. This behaviour would seem to be irrational for genuine buyers, hence, it can be deduced that this is the account of a shill bidder.

The common characteristics amongst shill bidders with strategic bid retractions are summarised as follows: a) few or no feedback in their personal profile, as purchases are rarely made on the account b) high activity levels with one particular seller, which suggests that there exists an association between the two accounts and c) relatively high numbers of retractions, since shill bidders have no motive to win in their own auction.

2.5 Conclusion

There are different interpretations of shilling in the literature. Whilst some argue that shilling is analogous to a dynamic reserve price, others find that it can cause bidders to shade their bids when they expect it to happen. In this chapter, I empirically prove the existence of squeeze shilling behaviour on eBay platform and show that retractions are a good indicator of squeeze shilling in eBay auctions. Firstly I theoretically show that the opportunistic seller has an incentive to conduct squeeze shilling using a bid-retraction mechanism under both private and affiliated valuation paradigms. I use the theoretical predications as a basis to empirically test the existence of squeeze shilling in eBay's auctions. The empirical results show that auctions with squeeze shilling tend to have higher end prices and a higher number of bidders than those listings without shilling behaviour. I also find that eBay bidders are relatively unaware of shilling behaviour, which allows opportunistic sellers to discretely shill up their auction prices. This chapter highlights a potential conflict of interest for eBay. On one hand, eBay has the incentive to allow shilling to happen, as shilling is related to higher auction ending prices, which is evident through their minimal penalties given to shill bidders. On the other hand, they should actively keep shilling under control to prevent bidders from shading their bids.

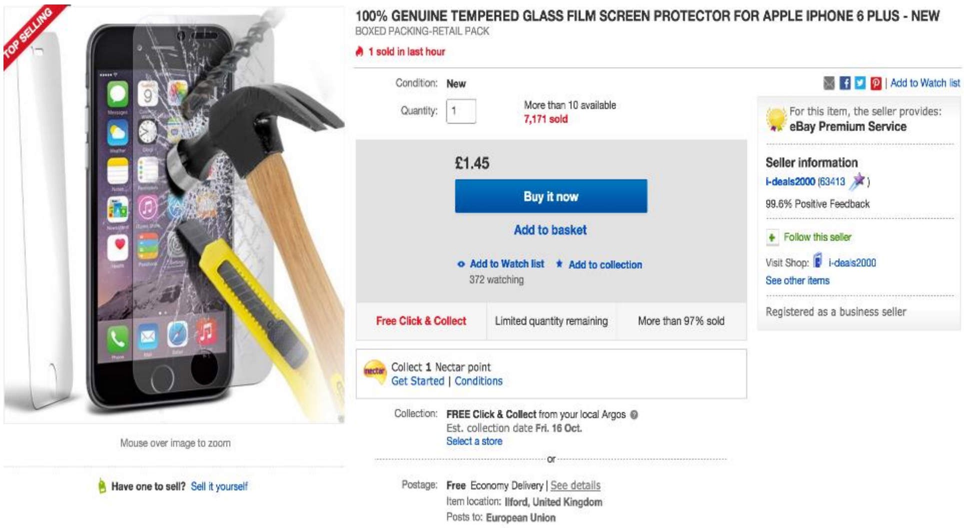
Chapter 3

EVIDENCE OF HERDING BEHAVIOUR ON EBAY

3.1 Introduction

The eBay platform is known to be one of the major online marketplaces in the world, where the items listed on the website spans almost every possible category. The average browsing time for an individual visitor is strikingly higher than any other major competitor. On the eBay platform, rational buyers have the problem of making an optimal purchase decision amongst hundreds of competing items. There exists an inherent asymmetric information problem between buyers and sellers in the online purchase environment. Without ex-ante information prior to a purchase, buyers have incomplete knowledge about the real quality of the items; moreover, online shoppers often have limited time to compare the long item descriptions and images provided by sellers. Under asymmetric information and the problem of information overload, rational buyers are induced to follow the choices of their predecessors, which can prove to be the most efficient and rational approach to decision-making. As a result, online shopping platforms provide a good environment that promotes the existence of herding behaviour.

Figure 3.1: eBay inventory listing



Whilst eBay was initially an auction platform, the site has evolved to include posted-price listings. A distinct type of posted-price listing is the 'inventory listing', where sellers can sell multiple identical products from one listing. In this chapter, I investigate inventory listings and empirically test whether herding behaviour exists in the online market platform.

Inventory listing is a type of selling method, which allows eBay sellers to continuously list single or multiple identical items in a single listing with no-time limit [1]. In order to help online buyers minimize the effect caused by the asymmetric problem of an online environment, eBay inventory listings offer detailed descriptions of item characteristics and seller features. Inventory listings include the cumulative total of historical sales, which is a crucial piece of information that is unique to inventory listings. This cumulative total updates automatically after each successful transaction is made, which provides precise information about the choices of preceding buyers. Figure 3.1 displays a screen-capture of an inventory listing and on the right side of the quantity indicator is the volume of total historical sales.

In real-world settings, such as, the eBay platform, people make choices sequen-

[1] This is a 30-day listing which can be automatically renewed every 30 days until sellers choose to end it. Each 30-day cycle incurs a new listing charge.

tially so that those who make decisions later can observe the choices of their predecessors. In sequential choice settings, people obtain signals in two main ways prior to making a purchase decision. On one hand, people have their own private knowledge or experience about a product or service. On the other hand, they can observe the actions of their predecessors. Herding behaviour occurs when the influence of the actions of predecessors become more dominant in the decision-making process than the signals from their own private information. As a result, people tend to follow the actions of their predecessors, which is known as the herding effect. A common example of herding is that of a tourist who has to make a decision about the choice of restaurant for dinner, on his/her first visit to the town. S/he compares the number of existing customers in a number of restaurants on a busy street, which forms part of his/her private information. If the tourist happens to be knowledgeable about the quality of food and services provided by those restaurants, the influence placed by the decisions of other customers is negligible. However, herding arises when the tourist makes a decision purely on the number of customers in the restaurant. Finally, he decides to enter the busiest-looking restaurant and disregards any private knowledge. This chapter empirically tests whether herding behaviour exists on the eBay platform and whether the addition of the element of total historical sales is the major instigator of herding behaviour on eBay inventory listings.

In order to accurately capture the characteristics of herding behaviour, I should choose an item that satisfies the following four criteria. First, the item must have a relatively low dispersion in price, which eliminates the effect of the price being a signal of quality. Second, the item should be homogeneous or standardised to minimize the effect of unobserved heterogeneity across different items and to eliminate the possibility of any popular trends that could cause a surge in sales. Third, the period of market entry should be relatively short, so that listings with higher historical sales should not be attributed to early entry into the market. Fourth, sellers should not have distinctive differences in their total feedback scores and quality of services, for example, delivery times or complementary gifts should be similar or controlled. These seller characteristics can be a contributor to high sales, which could form a false perception of herding behaviour.

In this chapter, I focus on the iPhone 6 plus screen protector. I collect data from eBay inventory listings under the category of 'mobile phone' and 'PDA Accessories'. I observe 1005 inventory listings under a single search result [2] in the period, 28th February to 12th March 2016.

The iPhone 6 plus screen protector meets all the criteria described above. The key criterion is that the item should be homogeneous or standardised. As the screen protector is a relatively standardised item, I also consider the effect of branding. It is noted that all the observations include the keywords in the listing title and have similar styles of images. Among the 1005 observations, 99.7% of listings are non-branded. In order to minimise the effect of unobserved heterogeneity and the effect of branded items, I exclude the branded listings that comprise 0.03% of my dataset [3]. Therefore, without ex-ante information, it is reasonable to assume that eBay buyers would regard all listed items as homogeneous items.

Another characteristic of my item of research is that the majority of sellers in the eBay screen protector market provide high-quality services and competitively low prices. So the likelihood that the herding effect is attributed to faster delivery or low listing prices is small. Total feedback scores are the main indicator of consumer satisfaction towards seller services. In my sample, the average total feedback score of sellers is 50887 and 98% of sellers offer high-quality services, that include, free delivery and a 14-day return policy. The price dispersion amongst screen protectors is low. A high proportion of sellers, 83% to be exact, list their items at a competitively low price that lies within the price interval, 0.99 to 3.99 GBP, where the average listing price stands at 2.88 GBP. I exclude two listings with inexplicably high prices of 299 and 399 GBP.

[2] I observe listings under the following search: '100% genuine, tempered glass screen protector, Apple iPhone 6 plus'

[3] I identify brands from the item description.

Another feature of the iPhone 6 Plus screen protector is that it allows for the control for the time of seller entry into the market. As the screen protector is an accessory of the smart-phone, iPhone 6 Plus, the earliest time of market entry is the release date of iPhone 6 Plus [4], which is only 18 months prior to the time of my data collection period. I traced the listing with the highest historical sales in my sample, with 54098 sales, and found that the listing had only been on the market for 13 months. I also observe that this listing is not one of the earliest listings on the market.

The chapter is aims to empirically prove the existence of herding behaviour on eBay inventory listings and analyse the effect of herding on the market structure. Herding is formed when successors have incomplete information and are uncertain about their choices. In a sequential choice setting, successors can observe the actions of their predecessors, but successors also have incomplete knowledge about how predecessors made their choices. Non-rational herding occurs when successors imitate the actions of their predecessors even when the choices of predecessors are biased. The aggregate outcome of successors' choices becomes far less diverse and deviates from the situation where buyers made their choices independently.

The results suggest that the eBay online platform offers an ideal environment to form herding, where the element of total historical sales is the main determinant that triggers herding behaviour. This suggests that the abundance of available information does not benefit online customers. The introduction of the total historical sales element in inventory listings encourages non-rational herding behaviour.

This chapter empirically shows that experienced eBay sellers manipulate the early purchase environment, by lowering prices, to form a biased herding effect. Successors in this herding environment inevitably pay more than their predecessors. Therefore, non-rational herding can damage the interest of buyers by increasing

[4]The iPhone 6 Plus was released in UK market on 19th of September 2014.

the market price. It also influences the market structure by crowding out competitors and deterring entry of new sellers.

The remainder of this chapter is organised as follows. In section 3.2, I discuss some related literature for online herding behaviour. In section 3.3, I describe the data and my research hypotheses. In section 3.4, I empirically investigate the existence of herding behaviour on eBay, from both static and dynamic aspects. I then discuss the methods that sellers use to gain an early advantage and influence herding behaviour to achieve higher overall profits. Finally, in section 3.5, I conclude my work with some general remarks.

3.2 Literature review

Herding behaviour relates to when people ignore their own private information and do what others do, as suggested by Banerjee (1992). There are a few early theoretical papers that discuss rational herding behaviour (or informational cascades) in sequential settings. The basic models of herding are proposed by Banerjee (1992), Bikhchandani et al. (1998), and Bikhchandani and Sharma (2000), who assume a perfectly elastic price in which all consumers or investors have the same available investment opportunity at the same price.

There are numerous empirical papers that give valuable insight into herding behaviour in financial markets. However, there are very few empirical papers that are found to study herding on online platforms. Simonsohn and Ariely (2008) examine non-rational herding behaviour on eBay's auction market, employing data from DVD auctions for their empirical analysis. They find that online bidders prefer to herd on auctions with more existing bids, despite it being an unreliable indicator of higher quality. Herding on auctions with higher bid counts leads to higher end prices and a lower probability of winning the auction. However, they also find that auction reserve prices only amount to trivial differences in expected revenue. Duan et al. (2009) empirically study information cascades on an online software-download platform. After controlling for other effects, they found that the download ranking of software products has a significant impact on online user choices in software products. As the sales or download rankings on online platforms are easily available to consumers, they can easily observe which items are more popular. They suggest that herding may be especially prominent in online markets as online users often have a problem with information overload, and so imitating the decisions of predecessors may be the most efficient and rational method of decision-making. However, they suggest that the effect of 'informational cascades' can lead to the adoption of inferior products.

3.3 Data and summary statistics

I collect data using an e-commerce market analysis spider program, Terapeak, which is the authorized analytics provider of eBay market data. In order to investigate whether historical sales is the major instigator of herding behaviour, I observe the daily differences in total historical sales. I collect a number of control variables, which consists of seller information and listing information. Table 3.1 summarizes the variables in my dataset. The seller characteristics include: user identity; total feedback scores since user registration; and a dummy for premium service sellers [5]. The listing information includes: listing price; postage price; the number of images in the description; the number of delivery days; dummy for a 'best offer' option; a dummy for returns; the number of days for return; and a dummy for a 'click and collect' service. Table 3.2 provides a table of summary statistics.

Figure 3.2 illustrates the distribution of total historical sales in listings of iPhone 6 Plus screen protectors in my data sample. The graph shows that total historical sales have a large dispersion; the distribution skews to the left and has a long right tail. The average number seller has 690 historical sale, where the highest seller has a total of 54098 sale, which is 78.40 times the average value. The majority of inventory listings have total historical sales that are lower than 1000, 89.2% to be exact. However, only 0.41% listings have total historical sales of more than 20,000.

[5] eBay Premium Service helps buyers identify the sellers that offer the best services, such as, minimum 14-day returns policy, free delivery, items sent within one working day and a seller with that is rated as having excellent service by buyers

Table 3.1: Variable description

Variable	**Description**
Total historical sales	The cumulative volume of total historical sales for each listing.
Image	The number of images posted on the listing
Click & collect	A dummy variable, equal to 1 if the item is available to collect from a local Argos store, and zero otherwise.
Returns	A dummy variable, equal to 1 if a seller accepts returns, and zero otherwise.
Premium service	A dummy variable, equal to 1 if a seller is a recognised premium seller, and zero otherwise.
Best offer	A dummy variable, equal to 1 if a seller provides a best offer option, and zero otherwise.
Listing price	The listing price set by a eBay seller
Total feedback	The seller's overall feedback rating, the sum of all positive negative and neutral feedback scores since seller registration.
Special offer	A dummy variable equal to 1 if a seller provides a special offer, for example three for two, and zero otherwise.
Postage	The price of postage added to the ending price
Delivery days	Number of delivery days

Table 3.2: Summary statistics

Variable	**Obs**	**Mean**	**Std.Dev**	**Min**	**Max**
Total historical sales	13,789	690.076	2,858.928	0	54,098
Images	13,775	3.392	2.830	0	16
Click & collect	13,775	0.388	0.493	0	1
Returns	13,775	18.233	9.671	0	60
Premium service	13,789	0.319	0.466	0	1
Best offer	13,779	0.084	0.278	0	1
Listing price	13,789	2.883	6.309	0.99	399.99
Total feedback	13,789	50,877.4	105,896.5	0	1,299,410
Special offer	13,775	0.2741	0.393	0	1
Postage	13,789	0.039	0.530	0	11.88
Delivery days	13,775	6.97	11.400	11	44

Figure 3.2: Distribution of total historical sales

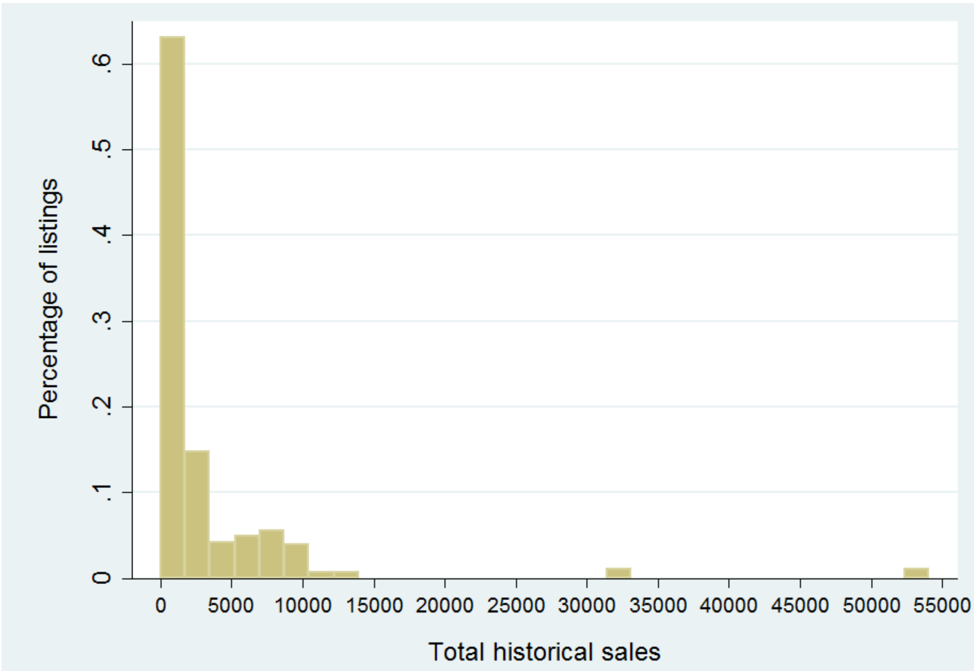
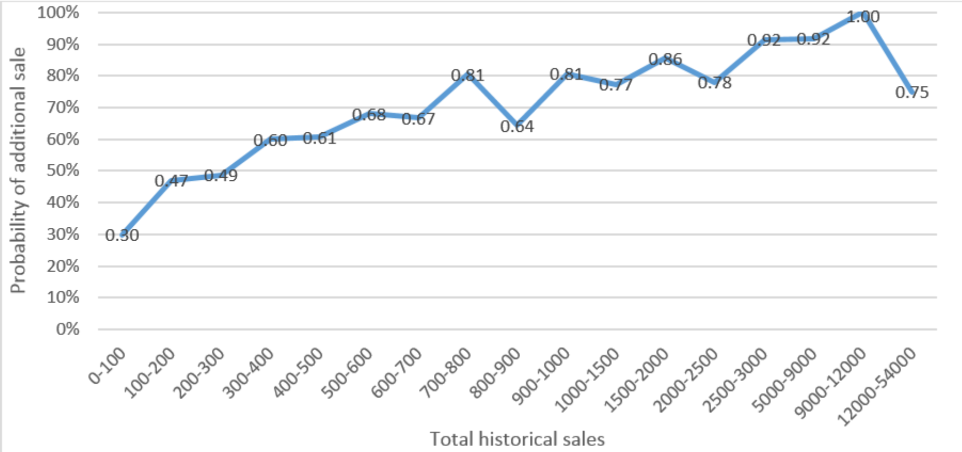


Figure 3.3: How do historical sales affect the probability of future sale?



It can be inferred that buyers on eBay prefer to buy items from a small group of listings, as evident in the high total historical sales. Assuming that all other items and seller characteristics are identical, such as item and service quality and the time of seller market entry, I take this as my first piece of evidence to support my conjecture of existing herding behaviour on eBay platform. The question to ask is whether high historical sales lead to even higher sales in the future, which forms a herding effect.

In my data sample, 97.13% of listings have made at least one successful transaction on the eBay inventory listing, which is shown through the number of historical sales. Figure 3.3 plots the probability of a sale at different levels of total historical sales intervals. The probability of future sale is the probability that I observe at least one additional successful transaction for a listing after 24 hours. The graph shows that the probability of future sale has a positive correlation with total historical sales. It illustrates that the listings with total historical sales that lie within the interval, 9,000 to 12,000, have a 100% probability of future sale, in contrast to listings with fewer than 100 historical sales, which only have a 30% probability of future sale.

Figure 3.4 illustrates the distribution of posted-prices for listings in my sample. The distribution skews to the left, which shows that sellers prefer to set their prices

Figure 3.4: Distribution of posted prices

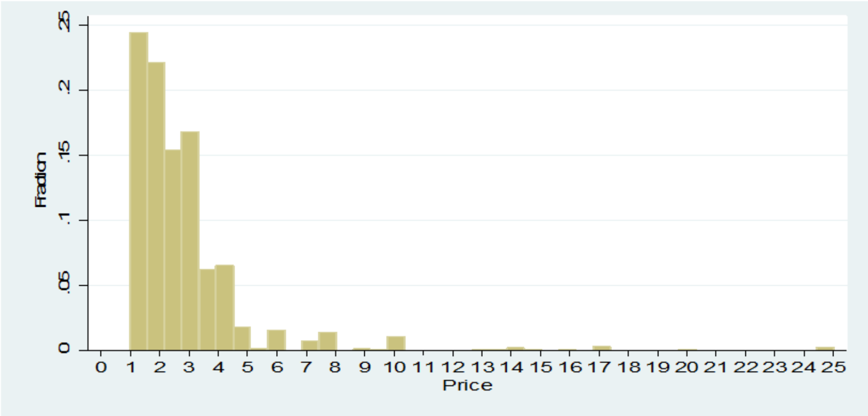
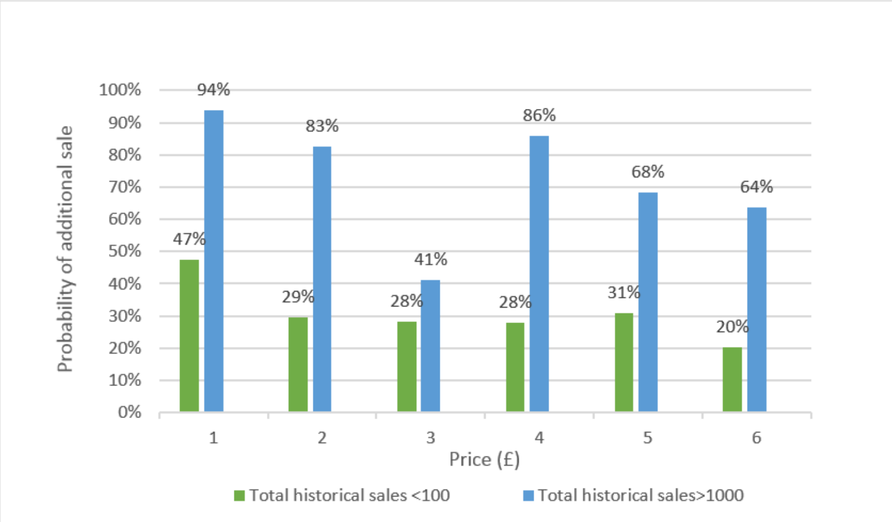


Figure 3.5: The probability of future sales for different price intervals

competitively. Statistically, 95.12% of sellers set their prices below 6 GBP, 93.24% below 5 GBP, and 77.79% below 3 GBP. The average price is set at 2.88 GBP. The lowest price is 0.99 GBP and 13.17% of sellers set their prices at this level.

Figure 3.5 plots the probability of future sale at different levels of listing prices. It shows that the probability of future sale is negatively correlated with listing price. Listings with a price of 0.99 GBP have the highest probability of future sales, at around 60%, whereas listings with a price between 15 and 15.99 GBP have a zero probability of future sales. This result highlights that buyers prefer to buy the competitively priced items in the market. Therefore, in order to capture convincing evidence of herding behaviour, the effect of competitive prices should be controlled. As I mentioned earlier, if higher sales are caused by buyer preferences in purchasing competitively priced items, it cannot be regarded as herding behaviour.

Figure 3.6: Comparing the probability of a future sale for low and high total historical sales listings



Conditional on current listing prices, I predict that herding behaviour exists on the eBay market, where listings with high historical sales are more likely to achieve more future sales.

In order to test my prediction, I control the effect of different price levels on the probability of future sale. I divide my data sample into two groups: I define the listings with total historical sale that are less than 100 as the low historical sale group, and the listings with total historical sales over 1000 as the high historical sale group. Then, I employ a method of pairwise comparison between the two groups, and calculate the probability of achieving at least one additional sale at different price levels of 1, 2, 3, 4, 5, and 6 [6] in GBP.

At a given price level, figure 3.6 shows the pairwise comparisons for high and low historical sales, and plots the probability that an inventory listing generates at least one additional transaction in 24 hours. It is evident that the group with high historical sales are more likely to complete at least one additional sale in the next 24 hours than the group with low historical sales. Take as an example, the

[6] Few sellers with high historical sales set their listing price above 6 GBP, so, these listings are excluded from my comparison.

price level at 2 GBP, where 83% of listings in the high historical sale group achieve at least one additional sale in 24 hours, in contrast to 29% of listings in the low historical sales group. It follows that buyers on eBay prefer to purchase items with high total historical sales.

3.4 Empirical analysis

In this section, I investigate whether there is evidence of herding behaviour in eBay inventory listings, by employing a cross-sectional Probit model and a dynamic panel data model. First, I prove that when controlling for all other observable heterogeneity, listings with a higher number of total historical sales tend to have a higher probability of future sales.

3.4.1 Static analysis

Statically, my interest lies primarily in the response probability of whether the number of total historical sales has a positive impact on the probability of future sale. I observe each listing at 12:00, on the 28th and 29th February 2016.

I employ a Probit model, where the dependent variable equals one if a listing has at least one additional sale on the 29th February 2016 and zero otherwise. In my sample, I have 1048 observations, where 51.1% listings have no successful transactions on the second day, 29th February 2016.

The regression model is as follows:

$$Pr(y = 1|x) = \Phi(x'\beta) \tag{3.1}$$

where,

$$y = \begin{cases} 1, & \textit{if at least one more item sold} \\ 0, & \textit{if no items sold} \end{cases}$$

The key variable in the Probit model is the number of total historical sales. Other observable heterogeneities are controlled, which include: ln(price), ln(total feedback), number of images, number of delivery days, postage, and dummies for click and collect, best offer, premium service, return policy and special offer.

Table 3.3: Probit estimates for the effect of total historical sales on future sales

Explanatory	Probit (1)	Probit (2)	Probit (3)	Probit (4)
Total historical sales	.000473	.0004462		.0004341
	(.0000736)***	(.0000751)***		(.0000752)***
ln(total feedback)	.0299228	.0389242	.0740297	.0392026
	(.0170412)*	(.0192871)**	(.0185025)***	(.0195566)**
ln(price)	-.2706176	-.3719266	-.4605992	-.3904474
	(.072778)***	(.0790679)***	(.0770421)***	(.0798852)***
Images		.0276704	0.0298594	0.0296359
		(.0161074)*	(.0157264)*	(.0161494)*
Click & collect		-.0771809	-0.0239835	-0.0490421
		(.0949955)	(.0918672)	(.0970663)
Delivery days		-.0174134	-0.0234673	-0.0176155
		(.0045611)***	(.0044591)***	(.0046227)***
Returns		-.0016148	.0018585	-.0014729
		(.0047769)	(.0045771)	(.0048496)
Best offer option		-.5011419	-.4225475	-.500588
		(.1528683)***	(.1501508)***	(.1533318)***
Postage		.067615	.080959	.0642692
		(.0928199)	(.0908505)	(.0932334)
Premium service				-.0645943
				(.094556)
Special offer				0.1791782
				(.1054776)*
Constant	-.2327302	-.0930374	-.2312916	-.1099622
	(.1640476)	(.1869393)	(.1837261)	(.1875542)
Pseudo R^2	0.0817	0.1018	0.0056	0.1040

Notes. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 3.3 reports the results for four Maximum Likelihood Probit coefficient estimates, where standard errors are reported in parentheses. Probit (1) presents the estimates for the baseline specification, controlling for total historical sales, ln(total feedback) and ln(price). Probit (2) presents the estimates with added controls for the number of images, the number of delivery days, and dummies for click and collect, returns, best offer, and postage. Probit (3) presents the estimates with the same controls as Probit (2) except it does not include the total historical sales estimate. Probit (4) has the same controls as Probit (2) with added controls for premium service and special offer dummies.

For Probit (1) (2) and (4), the results confirm that there is a significantly positive impact of total historical sales on the probability of future sale ($p - value < 0.01$), when I control all the other observable heterogeneity. These results are consistent with my prediction that listings with high historical sales attract more future sales, due to the effect of consumer herding behaviour. The results reveal that the coefficient of price is highly significant ($p - value < 0.01$) where higher listing prices correspond to lower rates of future sale, which is consistent with my findings in the last section. The regressions show that consumers are more likely to purchase items from sellers with high total feedback scores, which is also consistent with my expectations. It is also noted that listings with a higher number of delivery days and listings with a best offer option have lower probabilities of future sale ($p - value < 0.01$). Consumers purchase decisions are not affected by return policies, postage or a click and collect option, as the estimates are not significant in my result, which is intuitive as screen protectors are an inexpensive item.

Table 3.4 compares fitted and actual values for the Probit model. The table shows that 692 of the 1048 observations are correctly classified, and the remaining 356 observations are misclassified. There are 130 observations that are misclassified as 1 when the correct classification is 0, and there are 226 observations that are misclassified as 0 when the correct value is 1. The sensitivity measure ratio, which shows the rate of observations with $y = 1$ that are correctly specified, is equal to 55.8%. The ratio of specificity, which indicates the fraction of observations with

Table 3.4: Fitted and actual values

Probit model for sold dummy			
	True		
Classified	D	$\sim D$	Total
+	286	130	416
-	226	406	632
Total	512	536	1048
Classified + if predicated $Pr(D) \geq .5$			
True D defined as sold dummy $\neq 0$			
Sensitivity		$Pr(+\|D)$	55.86%
Specificity		$Pr(-\| \sim D)$	75.75%
Positive predictive value		$Pr(D\|+)$	68.75%
Negative predictive value		$Pr(\sim D\|-)$	64.24%
False + rate for true $\sim D$		$Pr(+\|D)$	24.25%
False - rate for true D		$Pr(-\| \sim D)$	44.14%
False + rate for classified +		$Pr(D\|+)$	31.25%
False - rate for classified -		$Pr(\sim D\|-)$	35.76%
Correctly classified			66.03%

$y = 0$ that are correctly specified, is equal to 75.7%. These two ratios indicate that the dataset fits my model well, as both of these ratios are higher than 50%.

Table 3.5 shows the marginal effects at the mean value for all the variables and standard errors are reported in parentheses. It indicates that the probability of future sales is very sensitive to the change of total historical sales. In my dataset, the average of total historical sales is 690 and marginal effect at mean in Probit specification 2 is equal to 0.0001772, so as the number of total historical sales increases by 1000, the probability of future sales increases by 17.72%. The number of total historical sales in my sample range from 0 to 54098, with a standard deviation of 2859. The regressions analysis supports the impression given by figure 3.5.

It is informative to show in a graph how different levels of total historical sale affect the probability of future sales. Figure 3.7 plots the fitted values from the Probit regression against total historical sales. It is evident that the probability of future sale increases when historical total sale increases. It also depicts that the marginal effect of total historical sales on future sale initially increases dramatically and then gradually decreases when total historical sale reaches a certain point. Listings with over 1600 total historical sales have a 100% probability of future sales. However, for listings with less than 100 total historical sales, the probability of future sales drop to 40%. This is consistent with my conjecture of how herding is formed on eBay. A large number of sellers enter the market with a minimal number of historical sales. In order to gain an advantage of the herding effect to maximise their profit, experienced sellers employ various strategies to boost their sales, such as better services, or low prices. Some sellers may even consider a strategy where they initially buy a number of their own items, in order to gain both positive feedback and historical sales, which is at the cost of listing and final value fees.

The increasing marginal effects of total historical sales on the probability of future sale can be explained by herding behaviour. After herding is formed, an increasingly large number of buyers follow the purchase decisions of other buyers and herd on a few select listings. Without ex-ante information or adequate knowledge,

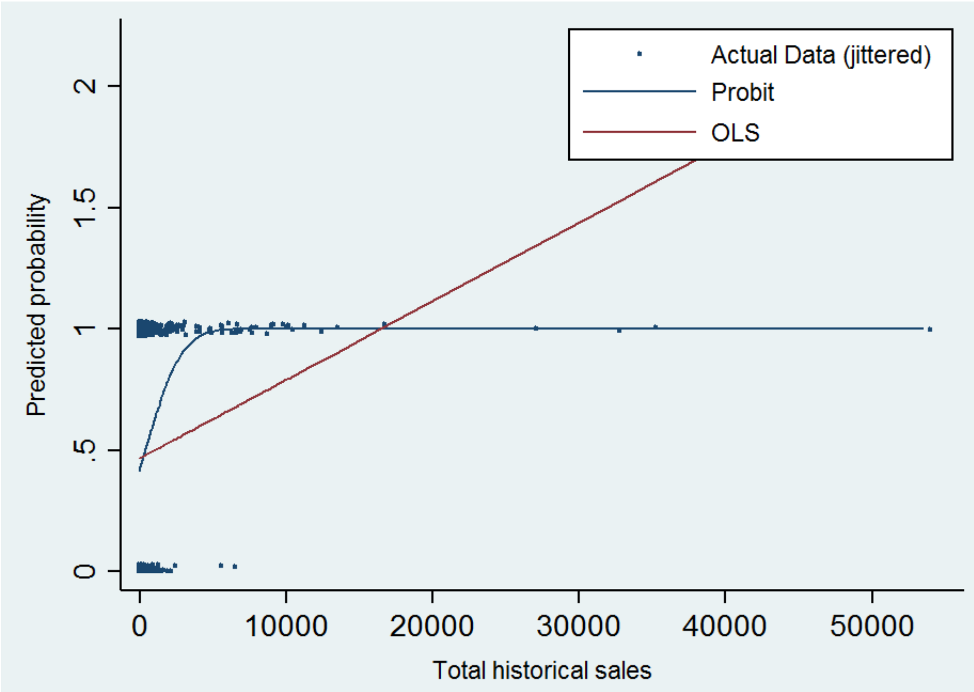
Table 3.5: Probit estimates for the effect of total historical sales on future sales (marginal effect at mean)

Explanatory	**Probit (1)**	**Probit (2)**	**Probit (3)**	**Probit (4)**
Total historical sales	.0001877	.0001772		.0001724
	$(.0000287)^{***}$	$(.0000293)^{***}$		$(.0000294)^{***}$
ln(total feedback)	.0118721	.0154558	.029518	.0155722
	$(.0067675)^{*}$	$(.0076689)^{**}$	$(.0073834)^{***}$	$(.0077768)^{**}$
ln(price)	-.10737	-.147682	-.1836553	-.1550947
	$(.028898)^{***}$	$(.0315037)^{***}$	$(.0307946)^{***}$	$(.0318412)^{***}$
Images		0.0109872	0.0119059	.011772
		$(.0063969)^{*}$	$(.0062713)^{*}$	$(.0064164)^{*}$
Click & collect		-0.0306464	-0.009563	-.0194806
		$(.0949955)$	$(.0918672)$	$(.039106)$
Delivery days		-0.0069144	-0.0093572	-0.0069973
		$(.0018154)^{***}$	$(.0017801)^{***}$	$(.0018398)^{***}$
Returns		-0.0006412	0.000741	-0.0005851
		$(.0018957)$	$(.0018244)$	$(.0019251)$
Best offer		-0.19899	-0.1684829	-0.1988451
		$(.0606868)^{***}$	$(.0598769)^{***}$	$(.0609005)^{***}$
Postage		0.0268481	0.0322809	0.0255292
		$(.0368449)$	$(.0362199)$	$(.0370201)$
Premium service				-.0256584
				$(.0376567)$
Special offer				.0711737
				$(.041146)^{*}$

Notes. *p < 0.10, **p < 0.05, ***p < 0.01.

Delta-method standard errors are given in parentheses.

Figure 3.7: Predicted probability of future sales against total historical sales



rational buyers can efficiently choose a high-quality item by quickly skimming listing results and taking note of the total historical sales element. For the time-restricted buyer, historical sales can be a good indicator of item quality. As the sales gap between competing sellers becomes more exaggerated, the herding effect becomes more prominent. As a result, seller behaviour becomes analogous to an equilibrium realized in oligopoly markets, where sellers can take advantage of their market share by increasing prices. Moreover, the sales gap is becoming increasingly large, not only amongst the two listing groups of high and low total historical sales but also amongst the top sales listings, as illustrated in graph 3.7. It can be observed that to the right that there are four listings with particularly high total historical sales of over 20,000 historical sales. The difference between the two listings with the highest sales is a substantial 18673 sales, despite that both listings are available on the market for the same length of time.

3.4.2 Dynamic analysis

In this section, I present empirical evidence that shows that the number of total historical sales on an eBay inventory listing has a significant positive impact on

Table 3.6: Summary statistics of panel dataset

Variable	**Obs**	**Mean**	**Std.Dev**	**Min**	**Max**
Total historical sales	1312	3075.66	6774.59	1	54098
Images	1312	3.067	2.60	1	12
Click and collect	1312	0.39	0.489	0	1
Returns	1312	20.09	12.02	14	60
Premium seller	1312	0.46	0.50	0	1
Best offer	1312	0.07	0.26	0	1
Price	1312	2.50	1.66	0.99	14.2
Total feedback	1312	68954.15	166109	31	1299410

future sales, from a dynamic aspect. Accordingly, I extend my investigation from a cross-sectional context to dynamic panel data. In order to examine the dynamic impact of historical sales on future sales, total historical sales should depend on past realizations. I observe 98 Inventory listings on eBay for 14 days, between 28th February 2016 and 12th March 2016. The following regression analysis shows how herding is formed on eBay, in a dynamic way.

Table 3.7 shows that I have an unbalanced short 14-day panel dataset with 98 observations, where 83 of listings have exactly 14 days of data, and 15 listings have less than 14 days of data. The 'pattern' describes the structure of the dataset, where the '1' represents that the data is available and the '.' represents missing data, for example, the first pattern represents an item sample where all 14 days are collected and the second pattern represents an item sample where one or more variables in the sixth day is missing. There are three main reasons that explain the unbalanced data structure throughout the 14 day data sample: a) some listings reached their 30-day listing end date and listing owners chose not to renew their listings, noted as 'listing ended', b) some listings were out of stock during my data collecting period, noted as 'out of stock', c) sellers were away and unable to complete transactions for personal reasons, such as, illness or holidays, noted as 'seller away'. In my case, I concentrate on approaches suited to a short panel, as

Table 3.7: Panel-data description

Seller:	1, 2, ..., 100						n = 98
Date:	28-Feb-2016, 29-Feb-2016, ..., 12-Mar-2016						T = 14
	Delta(date) = 1 day						
	Span(date) = 14 periods						
Distribution of T_i	Min	5%	25%	50%	75%	95%	max
	2	10	14	14	14	14	14

Frequency	**Percentage**	**Cumulative**	**Pattern**
83	84.69	84.69	11111111111111
2	2.04	86.73	11111.11111111
2	2.04	88.73	1111111111111.
1	1.02	89.8	11.11111111111
1	1.02	90.82	111...........
1	1.02	91.84	1111..........
1	1.02	92.86	1111.111111111
1	1.02	93.88	11111....11111
5	5.1	94.9	(other patterns)
98	100.00	100.00	xxxxxxxxxxxxxx

my dataset has several individual observations and short time periods, i.e. a small T and a large N.

Table 3.8 provides information of within variation and between variation of the regressors. The time-invariant variables include: premium service seller dummy, returns dummy, click and collect dummy, number of images and seller identity. These variables have zero within variation and are eliminated after the first differencing process.

I commence my dynamic analysis with an autoregressive model:

$$THS_{it} = \gamma_1 THS_{i,t-1} + \cdots + \gamma_p THS_{i,t-p} + x'_{it}\beta + \alpha_i + \epsilon_{it}, \quad t = p+1, \ldots, T \quad (3.2)$$

The dependent variable, THS_{it}, on the left of equation 3.2, is the total number of historical sales, at time t, for seller i. On the right side of the equation, α_i is the fixed effect and I include the first lag of the dependent variable $THS_{i,t-1}$ as one of the regressors, so p is equal to 1 in equation 3.2 and other independent variables, x_{it}, which include: $premiumservice$, $returns$, $clickandcollect$, $images$, $ln(totalfeedback)$, $price$, and $price^2$. In this case, I have a short panel data sample to fit into a fixed-effects model using the lags of the dependent variable as regressors. First, I need to remove the fixed effect, α_i, since it is associated with the problem of omitted variable bias. There are two methods to eliminate the fixed effect α_i. The first method is mean-differencing estimation and the second is first-differencing estimation. However, both of these two methods lead to inconsistent outcomes in this case. In mean-differencing estimation, the reason for inconsistency is due to the within model having the first independent variable $THS_{i,t-1} - \overline{THS_i}$, once the lagged independent variables are introduced it would be correlated with the error term $\varepsilon_{it} - \overline{\varepsilon_i}$, as $THS_{i,t-1}$ is correlated with $\varepsilon_{i,t-1}$, hence, it is also correlated with $\overline{\varepsilon_i}$, which generates an inconsistent outcome. On the other hand, the approach of Instrumental Variables (IV) estimation with lags is not feasible in this case, due to any lags, THS_{is}, being correlated with $\overline{\varepsilon_i}$, and therefore also correlated with the term $\overline{\varepsilon_{it}} - \overline{\varepsilon_i}$.

Table 3.8: Within variation and between variation of variables

Variable		Mean	Std.Dev	Min	Max	Observation
Total historical	overall	3075.66	6774.59	1	54098	N = 1312
sales	between		6679.77	1.36	53768.5	n = 98
	within		127.70	1938.02	3844.73	T-bar = 13.39
Postage	overall	0	0	0	0	N = 1312
	between		0	0	0	n = 98
	within		0	0	0	T-bar = 13.39
Seller	overall	49.32	28.92	1	100	N = 1312
	between		29.15	1	100	n = 98
	within		0	49.32	49.32	T-bar = 13.39
Image	overall	3.07	2.60	1	12	N = 1312
	between		2.71	1	12	n = 98
	within		0	3.07	3.07	T-bar = 13.39
Click and	overall	.39	.49	0	1	N = 1312
collect	between		.49	0	1	n = 98
	within		0	.39	.39	T-bar = 13.39
Returns	overall	20.09	12.02	14	60	N = 1312
	between		12.12	14	60	n = 98
	within		0	20.09	20.09	T-bar = 13.39
Premium	overall	.46	.50	0	1	N = 1312
	between		.50	0	1	n = 98
	within		0	.46	.46	T-bar = 13.39
Price	overall	2.50	1.66	.99	14.2	N = 1312
	between		1.66	.99	14.2	n = 98
	within		.21	1.50	3.96	T-bar = 13.39
Total feedback	overall	68954.15	166109	31	1299410	N = 1312
	between		163639.6	105.43	1297689	n = 98
	within		8467.29	37510.51	362111.50	T-bar = 13.39

The second method of first-differencing removes the fixed effect, α_i, but produces inconsistent estimations due to the independent variable, $\Delta THS_{i,t-1}$, in the first-difference equation 3.3, which is correlated to $\Delta\varepsilon_{it}$, even if I assume that error terms are serially uncorrelated. However, it also indicates that the error $\Delta\varepsilon_{it}$ in the first difference model is uncorrelated with $\Delta THS_{i,t-k}$ for $K \geq 2$, which allows for the use of lagged variables as instruments.

$$\Delta THS_{it} = \gamma_1 \Delta THS_{i,t-1} + \cdots + \gamma_p \Delta THS_{i,t-p} + \Delta x'_{it}\beta + \Delta\varepsilon_{it}, \quad t = p+1, \ldots, T \tag{3.3}$$

Even though the first difference estimation leads to an inconsistent outcome, I can still achieve consistent estimators if I use IV estimators in the first-difference model with appropriate lags of regressors as the instruments. Anderson and Hsiao (1981) suggest that it is feasible to conduct an IV estimation by employing an additional lag of the dependent variable, $y_{i,t-2}$, as an instrument for $\Delta y_{i,t-1}$, since, $y_{i,t-2}$ is uncorrelated with the first differencing error term, $\Delta\varepsilon_{it}$. The other lagged dependent variables can be instruments for themselves. Arellano and Bond (1991) implement the estimation and propose tests of the crucial assumption that ε_{it} are serially uncorrelated. The Arellano-Bond estimator uses an IV estimator based on the assumption that $\mathbb{E}(y_{is}, \Delta\varepsilon_{it}) = 0$, for $s \leq t-2$, so that the lags $y_{i,t-2}, y_{i,t-3}, \ldots$, can be used as instruments in the first-difference equation 3.3. Several papers suggest that it is feasible to employ additional moment conditions to obtain an estimator with improved precision and better finite-sample properties. In particular, Arellano and Bover (1995) and Blundell and Bond (1998) consider using the additional condition $\mathbb{E}(\Delta y_{i,t-1}, \varepsilon_{it}) = 0$ so that it also incorporates equation 3.2 and use as an instrument $\Delta y_{i,t-1}$. This method is known as the Arellano-Bover/Blundell-Bond system regression.

I implement an Arellano-Bover/Blundell-Bond model to fit my short panel dataset and Table 3.9 reports the regression estimates with robust standard errors. The time-invariant regressors, premium service, returns, click and collect, and images are eliminated after the first-differencing procedure. In order to prevent a poor performance of asymptotic result from generating too many instruments in the

Table 3.9: Arellano-Bover/Blundell-Bond for a dynamic panel model

Dynamic panel-data estimation		Number of observations: 1205		
Group variable: seller		Number of groups = 98		
Time variable: date		Observations per group:		
		Minimum = 1		
		Average = 12.29592		
		Maximum = 13		
Number of instruments = 39		Chi2 (4) = 38507.49		
		Prob > Chi2 = 0.0000		
Two-step results				
Total sales	**Coefficient**	**Std.Err.**	**z**	**P > |z|**
		Prob > F = 0.095		
Total historical sales L1	0.998	0.0066	151.73	0.000
	0.985	1.010		
ln(total feedback)	0.691	3.576	0.19	0.847
	-6.320	7.000		
Price	118.210	45.438	2.60	0.009
	29.153	207.268		
Price^2	-22.507	9.074	-2.48	0.013
	-40.292	-4.721		
constant	-81.210	62.315	-1.30	0.193
Instruments for differenced equation				
GMM type: L(2/3).totalsales				
Standard: D.ln(total feedback) D.price D.price^2				
Instruments for level equation				
GMM type: LD.totalsales				
Standard: _cons				

Arellano-Bond method, I employ the Stata command, maxldep(2), to restrict the number of instruments being generated. The Stata command, maxldep(2), indicates that I choose a maximum of two lagged variables to be used as instruments in my model and in this case, there are 39 instruments in total. Table 3.9 also illustrates which specific lags are employed as instrument variables in the regression model, for instance, L(2/3). Total sales specifies that two lagged variables $totalsale_{i,t-2}$ and $totalsale_{i,t-3}$ are employed as instruments conditional on that they are available and I have four standard instruments, In(total feedback), price, price^2 and constant.

The regression result in Table 3.9 reveals a dynamic relationship between the total historical sales at time t and $t-1$, The coefficient of total sales at time $t-1$ has a positive significant effect on the total sales at time t $(p-value < 0.01)$, which indicates that the current buyers' purchase decisions are significantly influenced by previous buyers' choices. However, the coefficient of total sales at time t-1 is equal to 0.998, which is close to 1, Which suggests that the process THS_{it} may have a unit root. Given this, I repeat the regression 3.2 with ΔTHS_{it} as a dependent variable and report the regression estimates with robust standard errors in Table 3.10 .

The result in Table 3.10 indicates that when the total historical sales at time $t-1$ increases by 1000, the sales at time t would increase by 18. Thus a higher number of total historical sales have a positive effect on buyers' purchasing decisions. Total feedback scores have a negative effect on buyers' purchasing decision, however it is only statistically significant at the 10% significance level. My model does not capture the dynamic effect of price on buyer's purchasing decision, it also supports the existence of herding behaviour as buyers ignore the price effect and herds on items with high total historical sales. Combined with the findings in the Probit regressions, this result also suggests the price strategy that eBay sellers use in this market. When sellers first enter this market, they should set a very competitive price, for example, 0.99 GBP, to accumulate a sufficiently high amount of total historical sale to form a herding effect to boost their future sales. After the herding

Table 3.10: Arellano-Bover/Blundell-Bond for a dynamic panel model

Dynamic panel-data estimation		Number of observations: 1205		
Group variable: seller		Number of groups = 98		
Time variable: date		Observations per group:		
		Minimum = 1		
		Average = 12.2959		
		Maximum = 13		
Number of instruments = 42		F(3,97) = 2.18		
		Prob > F = 0.095		
Two-step results				
FD total sales	**Coefficient**	**Std.Err.**	**z**	**P > $\lvert z \rvert$**
Total historical sales L1	.0181	.0074	2.44	.016
ln(total feedback)	-13.6253	8.0090	-1.70	.092
Price	31.2261	30.13027	1.04	.303
constant	35.0508	87.45177	.40	.689
Instruments for differenced equation				
GMM type: L.totalsales ln(feedback) price				
Instruments for level equation				
Standard: _cons				

Table 3.11: Arellano-Bond test

Log (Arellano-Bond Test)		H_0: no autocorrelation
order	**z**	**Prob>z**
1	-2.41	0.016
2	-1.02	0.308
3	-.24	0.813

Table 3.12: Sargan test of overidentifying restrictions

Sargan test	H_0: overidentifying restrictions are valid
Chi2(38)	52.25
Prob>chi2	0.062

effect is formed, experienced sellers can increase their listing prices up to a certain amount to maximise their future profit.

In order to get consistent estimations, Arellano-Bover/Blundell-Bond estimators require that the error term, ε_{it}, is serially uncorrelated. Table 3.11 illustrates the result of the Arellano-Bond Test to check if there exists a problem of serial correlation. The null hypothesis of the Arellano-Bond Test is $Cov(\Delta\varepsilon_{it}, \Delta\varepsilon_{i,t-p}) = 0$ for $p = 1, 2, 3$. If the error terms ε_{it} in my model are serially uncorrelated, then I would expect that order 1 $Cov(\Delta\varepsilon_{it}, \Delta\varepsilon_{i,t-1})$ is not equal to zero, since $Cov(\Delta\varepsilon_{it},\ \Delta\varepsilon_{i,t-1}) = Cov(\varepsilon_{it} - \varepsilon_{i,t-1}, \varepsilon_{i,t-1} - \varepsilon_{i,t-2}) = -Cov(\varepsilon_{i,t-1}, \varepsilon_{i,t-1}) \neq 0$ but I would expect the second order to be, $Cov(\Delta\varepsilon_{it}, \Delta\varepsilon_{i,t-2}) = Cov(\varepsilon_{it} - \varepsilon_{i,t-1}, \varepsilon_{i,t-2} - \varepsilon_{i,t-3}) = 0$ and the third order to be, $Cov(\Delta\varepsilon_{it}, \Delta\varepsilon_{i,t-3}) = Cov(\varepsilon_{it} - \varepsilon_{i,t-1}, \varepsilon_{i,t-3} - \varepsilon_{i,t-4})\ = 0$.

From the result of Table 3.11, I reject the null hypothesis at order 1, since the p-value is equal to 0.016, which is smaller than 0.05. However, at order 2 and order 3 I am unable to reject the null hypothesis since both p-values are larger than 0.05. So, I conclude that the error terms ε_{it} are serially uncorrelated in the model.

The next step requires a test of over-identifying restrictions. I run a Sargan test of over-identifying restrictions, with the null hypothesis, over-identifying restrictions are valid. The results are shown in Table 3.12. In my estimation, I employ 42 instruments to estimate 4 parameters, so I have 38 over-identifying restrictions and I am unable to reject the null hypothesis since $p = 0.062 > 0.05$. Therefore, I conclude that in this model the over-identifying restrictions are valid.

The result of the Arellano-Bover/Blundell-Bond system regression provides convincing evidence to support the existence of buyer herding behaviour on eBay inventory listings with high total historical sales. I also find that the number of total historical sales has a significantly positive impact on the future sales. My estimation results provide empirical evidence to support the conjecture that eBay sellers are motivated to manipulate information in the early stages to influence the purchasing decisions of buyers and form non-rational herding, as increasing the price within an appropriate interval after a formed herding effect will increase total revenue.

On the other hand, the existence of herding behaviour will not only harm the interest of buyers but also change the whole market structure in the long run. First, a top seller who already has a high market share will take advantage of the herding effect on their listing. New sellers are deterred to enter the market as they must set competitive prices to compete with top sellers. In order to encourage buyers to herd on their listing and accumulate a high volume of total historical sales, the common strategies for new sellers are to lower their costs of production and create advertisements to attract more attention to their listing. However, if profits cannot cover the operating costs, it would further increase the rate that new sellers leave the market. Additionally, in order to minimise the cost, some sellers choose to sacrifice the item quality to stay in the market, therefore the herding effect may also squeeze out branded items with higher prices and better quality, which leads to the market equilibrium price and the quality of products converging to a lower level. It also explains why the brand composition in the eBay screen protector market is quite different from other online shopping platforms, such as, Amazon, which do not include information about historical sales on their listing pages. In the eBay screen protector market, there is approximately 3% items are branded, and 97% of items are non-branded, in comparison, on Amazon, around 50% of items are branded in this market.

3.5 Concluding remarks

This study provides the first empirical evidence for the existence of herding behaviour on eBay inventory listing. My findings indicate that the element of total historical sales on the eBay inventory listing page plays an important role to instigate herding behaviour in this market, as I find that it has a significantly positive effect on the purchase decisions of future buyers, from both static and dynamic aspects. I also find that experienced sellers encourage the herding effect by initially offering low prices to increase the cumulative number of total historical sales. After herding behaviour is formed, they raise their prices to extract more surplus, since the dynamic analysis shows that the probability of future sales increases when the price increases within a certain price interval. This chapter also suggests that non-rational herding behaviour will not only damage the interest of a buyer but also change the whole market structure in the long run.

Chapter 4

The hidden importance of wine scores in fine wine auctions

4.1 Introduction

It is a simple and effective method for wine critics to capture their opinions about the real but unknown sensory quality of a wine. Wine scores are widely used as a benchmark for wines, by first-time wine buyers, who face a purchase decision with hundreds of options. Several empirical studies, such as San Martin et al. (2008), Bombrun and Sumner (2003), Landon and Smith (1998), Oczkowski (1994), Schamel et al. (2001) and Jones and Storchmann (2001), document that wine scores, as a quality measurement of the sensory characteristics of wine, have a positive impact on the consumers' willingness to pay in wine retail markets. However there is little research that analyses the effect of wine scores on bidders' valuations in the secondary market for wine, such as, in the wine auctions at Christie's, where the most expensive wine transactions take place.

Christie's is one of the world's leading auction houses for the fine wines and spirits category and has a long history of achieving exceptional auction prices for fine wine. Christie's employ a traditional English auction format, an auctioneer sells an indivisible item to N potential bidders, commencing with a suggested reserve price.

There are many differences between the primary and secondary market for wines, not only from the aspects of purchasing environment and selling format but also from the disparity in the consumer's purchase purpose, the contrast in the values of wine and the differing approaches for consumers to gather information. In the secondary market for wine, most wines are produced and sold directly from prestigious chateaus, such as, Romanee-Conti and Cheval Blanc of France, but there are also a few wines that have exquisite historical references, such as, wines that were found in the cabin of a sunken shipwreck. Due to the scarcity and uncertain value of wine, sellers prefer to use the auction format, as opposed to posted price, with the expectation of achieving higher profits. In Christie's wine auctions, bidders are either wine collectors or experienced wine investors who may have the intention of reselling the wine in the future. Contrastingly, in the retail market, the purchase purpose for a consumer is generally for private usage. Both collectors and investors seek out age-worthy wines, wines that improve with age and almost certainly increase in value, however, it is interesting to note that most wines (95% in fact) are not destined to age. It seems that the scarcity of age-worthy wines make wine scores and tasting notes all the more prominent in the fine wine market. Buyers of fine wine rely on reviews and scores from prestigious critics to form their opinion on wine. Therefore, it is interesting to investigate how wine scores affect a bidder's valuation in an auction, and whether bidders regard wine scores as an important quality indicator to reveal the unknown sensory characteristics of wines.

There are two major challenges for this chapter. Firstly, to analyse the effect of wine scores and other variables on bidders' valuations, the bidder's private valuation distribution is required. However, the bidder's valuation is a latent variable, which is not revealed during the course of an auction. Therefore, it is difficult to get access to a detailed dataset for live auctions, especially from leading auction houses, like Christie's and Sotheby's. Auction houses will normally publish a simple report on their web-page which summarises the auction outcomes the day after the sales are completed. It only documents the sales title, date, the location of salesroom and the amount of winning bids. Therefore, numerous unobserved variables also cause problems in estimating the bidder's valuation distribution.

In this chapter, I employ auction data that I collect from completed Christie's wine auctions. I adopt an indirect inference approach to estimate the bidder's valuation distribution and analyse the effect of wine scores on bidders' valuations. There are three major findings according to the results of my empirical analysis. First, wine score is an essential element that characterises the underlying private valuation distribution of bidders. I find that wine score has a predominant effect on bidders' valuations. The results from my regression indicate that as wine score increases by one point, the expected private valuation increases by around 8.02%. Second, the regression outcomes contradict other studies of retail wine market that use the hedonic price equation, as certain non-sensory characteristics vineyard location and vintage year of wine do not have an effect on bidders' valuations. It suggests that bidders in Christie's wine auctions have no distinct differences in valuation for wines from the most popular vineyard locations in France, namely, Bordeaux, Burgundy, and Rhone. Third, by successfully estimating the parameter that characterises the bidder's valuation distribution, it allows me to calculate the optimal reserve price following the framework proposed by Riley and Samuelson (1981). I find that the reserve price set by Christie's is not optimal; a higher reserve price can yield higher revenue.

This chapter is organised as follows. Section 4.2 introduces some background about wine scores. Section 4.3 discusses the literature surrounding the relationship between wine scores and consumer's willingness to pay. Section 4.4 provides the set-up of the model for Christie's wine auction under private valuation paradigm. Section 4.5 presents the structural model for the indirect inference approach. Section 4.6 presents the dataset and summary statistics. Section 4.7 provides a detailed description of the two-stage indirect inference method. Section 4.8 reports the results from my indirect inference analysis and tests whether the reserve prices are set at the optimal level. Finally, section 4.9 concludes the chapter with some general remarks.

4.2 The importance of wine scores

The value of a bottle of wine depends on both sensory characteristics and non-sensory characteristics. The sensory characteristics yield utility directly to the wine consumer, in the form of appearance, smell and taste. On the other hand, non-sensory characteristics do not yield utility directly to wine consumers and are attributed to: grape variety, chateau, location, vintage (viticulture), wine-making technique (vinification) and storage – these specifications can normally be found on the label of a bottle of wine. A bottle of wine can be regarded as a complex good, which varies in quality as it matures, as the sensory of characteristics of wine will vary with time elapses and storage conditions. Therefore, first-time wine buyers, who do not have the opportunity to taste-test the wines, all encounter the same dilemma in the market – how do they make a rational purchase decision when faced with hundreds of different wines, which vary in taste and are produced in different locations, chateaus, and vintage years.

In order to make a rational purchase decision, wine buyers must collect a plethora of information to maximise the probability that they buy a wine that is well matched to their desired sensory characteristics. Buyers in a wine future market also need a reliable predictor of the quality of a wine, as they need to minimise the risk caused by uncertainty about the quality of sensory characteristics of the wine. Due to imperfect information of sensory characteristics and excess wine options in the market, wine shoppers find they have insufficient time and knowledge to make complex comparisons amongst different wines and make a rational purchase decision. Wine critics, such as Jancis Robinson and Robert Parker, assign scores and drinking notes for wines. Wine scores help first-time buyers form expectations about the unknown quality of the sensory characteristics of a wine. A higher wine score indicates a higher quality of the wine sensory characteristics, which leads to a higher market price. Therefore, an unbiased wine score, which accurately reflects the wine buyers' taste preference, is an essential indicator to measure the wine shoppers' willingness to pay for one additional unit of sensory characteristics. However, it is important to consider that wine scores are only useful to consumers if the scores (and drinking notes) precisely capture the average preferences of wine consumers in the market. Wine scores give consumers accurate information about the wine so they can act and bid as if they have already tasted the wine and have clear knowledge about the wine's sensory characteristics. Therefore, wine scores given by the wine experts are a useful tool to guide first-time wine buyers, who are uncertain about the sensory characteristics of a wine, when they are making a purchase decision.

4.3 Literature review

Several empirical studies use wine scores as a measure of the potential quality of sensory characteristics of wines but there are very few that analyse the effect of wine scores on bidder's valuation in the secondary wine market. To analyse the effect of wine score on consumers' willingness to pay in the retail market, these studies use hedonic price regressions, which control for a number of observable heterogeneities, such as, vintage year and regions. The main finding of these studies is that there exists a positive relationship between wine score and wine price in the wine retail market.

Oczkowski (1994) find that the price of Australian table wine increases with the score given by a popular Australian wine guide. Schamel et al. (2001) use a time-series model to study the effects of the variables, wine score, regional vineyard location and grape variety, on the price of Australian and New Zealand wine. They find that the species of grape and regional vineyard locations from the late nineties have a significant effect on price. Jones and Storchmann (2001) investigates twenty-one popular Bordeaux wines and regresses the price of wine on vintage, grape composition and wine score, measured by Parker-points. They find that wine score has a larger effect on the price of wine than the wine's dominant grape variety. Also, the vintage year has a larger positive effect on the price of wine for Merlot than Cabernet Sauvignon.

There are also three studies that analyse the impact of wine sensory quality on wine prices in the American wine retail market, which all employ the 100-point wine score system given by Wine Spectator, a popular lifestyle magazine and major wine-buyers guide. Two studies proposed by San Martin et al. (2008), who studied Argentinian wine, and Bombrun and Sumner (2003), who focused on Californian wines, report similar results. They find that an increase of one point in wine score increases wine price by 4%. Landon and Smith (1998) study red wine from the Bordeaux region, and find that the purchase decision of a typical consumer is heavily affected by Chateau reputation as opposed to wine score. The impact of wine scores assigned by the wine guide, Wine Spectator, is much smaller; a one-point increase in score only leads to an increase in wine price of less than 1%.

4.4 Bidding at Christie's wine auction with private valuation paradigm

The wine auctions in Christie's employ a traditional English ascending price auction, which I model as a non-cooperative game where there is one auctioneer who wants to allocate one indivisible item to N risk neutral bidders under a private valuation paradigm. I assume that the bidders' valuation distribution is exponential, with a probability density function, $f_v(v)$, and a cumulative distribution function, $F_v(v)$. The reason is that the pdf of an exponential distribution is monotonically decreasing, which is consistent with the reality that the probability of bidders who have a high valuation is decreasing, which facilitates my later computation when I estimate bidder's optimal bid. I assume each bidder knows their own valuation but they do not know their opponents' valuations.

The auctioneer commences the auction at a predetermined reserve price and asks the bidders to raise their bids. In Christie's, the reserve price for each lot is not published in the auction catalogue and it only becomes known to the bidders as the auction takes place in the salesroom. Only potential bidders who have valuations above the reserve price submit their bids. If no bidder submits a bid, the item goes unsold. If only one bidder submits a bid, the sole bidder wins the auction at the reserve price, conditional on his/her bid being higher than the secret reservation price of that auction. The secret reservation price is the minimum price a seller is willing to sell his/her item, which is equivalent to the seller's valuation. If there are two or more bidders who submit bids, they submit their bids sequentially and the bidder with the highest bid wins the auction at the valuation of the second highest bidder. Therefore, the distribution of the winning bid is equal to zero when the lot is unsold, the reserve price when there is only one bidder and the second highest valuation when there is more than one bidder.

Potential bidders can bid in different ways: they can submit a written absentee bid before an auction commences, which represents the bidder's highest willingness to pay or they can bid during an auction, either in person or via telephone. It is important to note that all potential bidders must complete a registration form before the auction commences. Therefore, although Christie's allow bidders to submit their bids online, only bidders that are already pre-registered can submit their bids online, so the bidding environment is closed with a fixed number of potential bidders after auctions commence. This bidding environment is different from other online auction bidding platforms, such as, eBay, where the number of potential bidders is unknown and new bidders can enter the auction at any time during the auction.

Potential bidders can gather information about the wine in upcoming auctions in three ways: using the information provided in a published catalogue, participating in the pre-auction viewing section, (which may include a consultation with a wine specialist), and having past experience or their own private valuation. A wine auction catalogue provides detailed information about each wine lot, such as lot size, bottle size, name of chateau, wine vintage year, wine condition and price estimates for each lot. It also contains conditions of sale and a section on how to place a bid in an auction. However, given the information provided by the wine catalogue, potential bidders are unable to make inferences about the uncertain sensory characteristics of a wine. Bidders who make bids must accept the condition and description of wines, and accordingly, Christie's allow bidders to personally inspect each wine lot before the auction and make a consultation with a wine specialist from Christie's.

4.5 Structural model for indirect inference approach

To analyse the effect of wine scores on bidders' valuation, the first step is to characterise the distribution of the latent variable, bidder's private valuation[1]. I need to specify a theoretical model that links the observed variable, winning bids, with the latent variable, bidders' valuations, and propose an empirical method to estimate the unknown parameters.

The model proposed by Milgrom and Weber (1982), widely known as the clock model, is prevalently used in modelling English auctions. In the clock model, the auctioneer sets the clock at a pre-determined reserve price. In this model, each bidder knows their valuation and holds a button to signal that they are still in the auction. As the price rises continuously and exogenously, bidders drop out of the auction as the price reaches their valuation, until only one bidder remains in the auction. The valuation of the second highest bidder is revealed and the last remaining bidder wins the auction with a bid equal to the second highest bidder's valuation, the second-order statistic $v(2 : N)$. Hence, the Milgrom-Weber clock model is often referred to as a form of second-price auction. The dominant strategy of each non-winning bidder, B_i, is to stay in the auction until the price reaches his private valuation; as a result, the non-winners' equilibrium bidding strategy is $b_i = b(v_i) = v_i$. Therefore in principle, it is possible to employ the probability density function of the second-highest order statistic to construct the likelihood function and estimate the underlying bidders' valuation distribution.

However, this method has some potential problems. First, the observed winning bids do not uncover the winning bidders' true valuation. Second, due to the ex-

[1] Several studies estimate the bidder's valuation distribution in English auctions. See Paarsch (1997), Haile and Tamer (2003) and Brendstrup and Paarsch (2006). I use a different way to estimate the bidder's valuation distribution.

istence of a reserve price in Christie's auctions, only bidders who have private valuations above the reserve price submit bids, so the empirical distribution of observed winning bids is truncated with different truncation points for each observation. Third, in many auctions the items for sale are not identical: covariate heterogeneity is essential. To overcome the problems above, I employ the framework proposed by Riley and Samuelson (1981) as a structural model combined with an indirect inference method to estimate the probability law of valuations.

The auction theoretical works proposed by Riley and Samuelson (1981) and Myerson (1981) prove that all auctions with certain properties generate the same expected revenue for the seller, which is known as the revenue equivalence proposition. The model assumes risk-neutral bidders with an independent private paradigm, bidders' bidding functions that are increasing functions of their private valuations and the bidder with the highest valuation wins the auction. These assumptions are satisfied by the Christie's wine auction. Thus I am going to use the Riley and Samuelson model to link the observed variable, winning bids, with the latent variable, bidders' valuations, and propose an empirical method to estimate the bidder's valuation distribution.

Following the framework in Riley and Samuelson (1981), the equilibrium winning bid for each auction in Christie's is derived as follows.

As in Riley and Samuelson (1981), I derive the optimal wining bid for a specific bidder, winning bidder 1, in auction k, with a total of N potential bidders. Bidder 1 has a private valuation, v_1, and chooses to report their valuation as w. I assume the bidder's bidding strategy is solely determined by their valuation, hence, the auctioneer is able to calculate their bid according to their reported valuation. Due to the assumption of risk-neutrality, the bidders' utility can be presented as a linear relationship in their monetary payoffs, which makes it possible to separate the probability of winning from the term of expected payment. The expected payoff for bidder 1 is,

$$\Pi(w, v_1) = v_1 \times Pr(winning) - Expected\, Payment \tag{4.1}$$

As bidding behaviour is non-cooperative, an equilibrium bidding strategy for bidder i is, $b_i = b(v_i)$. I assume that all bidders will place a bid, $b_i = b(v_i)$, except bidder 1, who decides to report w. Therefore, bidder 1 will submit a bid equal to $b(w)$, so the payment function of bidder 1 depends not only on w, but also on $(v_2, \ldots, v_N)$.

$$Payment\,[b(w), b(v_2), \ldots, b(v_N)] \tag{4.2}$$

Since the valuation of all the other bidders are unknown to bidder 1, his/her expected payment, $P(w)$, is,

$$P(w) = \mathbb{E}\{Payment\,[b(w), b(v_2), \ldots, b(v_N)]\} \tag{4.3}$$

If bidder 1 is the winning bidder in an auction with his reported value w and tenders $b(w)$, his reported value must be higher than the rest of $N-1$ potential bidders, so the probability of winning bid is,

$$Pr(v_j < w, j \neq 1) = F_v(w)^{N-1} \tag{4.4}$$

So bidder 1's expected profit is,

$$\Pi(w, v_1) = v_1 \times F_v(w)^{N-1} - P(w) \tag{4.5}$$

Under truth-telling, the first-order condition for bidder 1's expected profit maximization problem is,

$$\frac{\partial \Pi(w^*, v_1)}{\partial w} = v_1(N-1)F_v(w^*)^{N-2} f_v(w^*) - P'(w^*) = 0 \tag{4.6}$$

Only when w^* is equal to bidder 1's true valuation, the $bid(v_1)$ is optimal in equilibrium,

$$P'(v_1) = v_1(N-1)F_v(v_1)^{N-2} f_v(v_1) \tag{4.7}$$

The expected payment for a bidder who has a valuation equal to the reserve price, r, is,

$$P(r) = rF_v(r)^{N-1} \tag{4.8}$$

and the expected payment from bidder 1's perspective becomes,

$$\begin{aligned} P(v_1) &= rF_v(r)^{N-1} + \int_r^{v_1} P'(u)du \\ &= rF_v(r)^{N-1} + \int_r^{v_1} udF_v(u)^{N-1} \\ &= rF_v(r)^{N-1} + v_1F_v(v_1)^{N-1} - rF_v(r)^{N-1} - \int_r^{v_1} F_v(u)^{N-1}du \\ &= v_1F_v(v_1)^{N-1} - \int_r^{v_1} F_v(u)^{N-1}du \end{aligned} \tag{4.9}$$

Bidder 1 wins the auction with his report value, v_1, if and only if, his valuation v_1 is higher than the rest $N-1$ potential bidders, therefore the probability of winning bid is equal to,

$$Pr(v_j < v_1, j \neq 1) = F_v(v_1)^{N-1} \tag{4.10}$$

Therefore the equilibrium bidding strategy for the winning bidder is,

$$b_1 = v_1 - \frac{1}{F_v(v_1)^{N-1}} \int_r^{v_1} F_v(u)^{N-1}du \tag{4.11}$$

Equation (4.11) links the latent variable, the bidder's valuation and observed variable, the winning bidder's bid and indicates that the bidding strategy for a winning bidder in a wine auction is determined by the bidder's private valuation, the bidder's valuation distribution and the number of potential bidders, an exogenous variable.

Another important implication of the Riley and Samuelson model is that it allows sellers to estimate the optimal reserve price for their auctions, assuming that the bidder's valuation distribution is common knowledge. As the bidder's valuation distribution is unknown, I use the indirect inference approach to estimate the parameter that characterises the bidder's valuation distribution, which I employ to estimate the optimal reserve price and empirically test whether the reserve price set by Christie's is optimal.

The following derivations show how to construct the optimal reserve price for auctions. If the item is sold, sellers have utility that is equal to the sum of the expected revenue; if the item is not sold, sellers have the expected utility of retaining the item.

The item is not sold if the winning bid is lower than the seller's valuation, v_0, which is equal to the secret reservation price. The sum of the seller's utility is shown in equation (4.12).

$$v_0 F_v(r)^N + N \int_r^{\overline{v}} \left[u f_v(u) + F_v(u) - 1 \right] F_v(u)^{N-1} du \tag{4.12}$$

In order to maximise the seller's expected gain, differentiate equation (4.12) with respect to the reserve price, r, to give the following first-order condition that holds when r equals the optimal reserve price ρ^*:

$$N v_0 F_v(r)^{N-1} f_v(r) - N \left[r f_v(r) + F_v(r) - 1 \right] F_v(r)^{N-1} = 0 \tag{4.13}$$

By taking out the common factors, N and $F_v(r)^{N-1}$, the first order condition becomes,

$$v_0 f_v(r) - r f_v(r) - F_v(r) + 1 = 0 \tag{4.14}$$

So, if buyers are risk neutral and the assumption of a symmetric independent private value paradigm (IPVP) holds, the seller's expected gain is maximized when the optimally-chosen reserve price, ρ^*, solves equation (4.15):

$$\rho^* = v_0 + \frac{[1 - F_v(\rho^*)]}{f_v(\rho^*)} \tag{4.15}$$

Equation (4.15) shows that to estimate the optimal reserve price ρ^* requires information about the bidder's valuation distribution $F_v(v)$ and seller's valuation for the item at auction. Equation 4.15 also indicate that the optimal reserve price for a specific auction is independent of the number of bidders.

4.6 Data and summary statistics

I collect data from Christies' online auction report, the catalogue book, and video of the auction salesroom provided by Christie's live wine auctions. The event was categorised under the title, 'Fine and Rare Wine', and the auction was held on the 16th March 2017 at London King Street, with sale number 14365. The total revenue from this sale is equal to 918,426 GBP (excluding the buyer's premium of 17.5%). The buyer's premium varies according to the location of salesroom across different countries. For consistency, all the auction realized prices, reserve prices, and presale estimation prices in my dataset exclude the buyer's premium and applicable taxes.

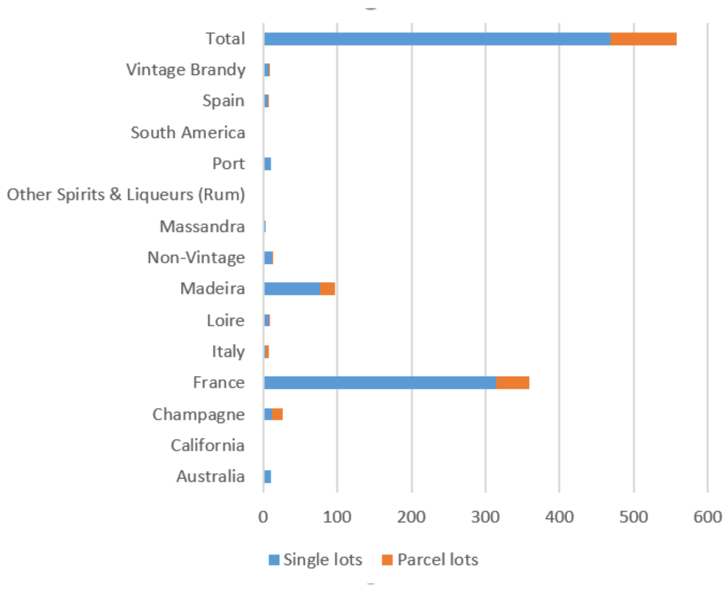
The dataset includes auction ending prices, reserve prices, the number of bids, and presale estimation prices with an upper and lower bound level. Presale estimation price of each wine lot is based on the condition, rarity, expected quality, provenance and recent auction realised prices of similar items. Lot information collected includes the number of wine bottles, carried in each lot, bottle sizes, drinking notes and wine scores.

The dataset consists of 558 wine lots that are categorised under 18 headings according to the type of wine, spirits or liqueurs, and provenance. The headings are namely: Australia, California, Champagne, Claret, Italy, Loire, Madeira, Non-Vintage, Massandra, Other Spirits & Liqueurs (Rum), Port, Red Burgundy, Rhone, South America, Spain, Vintage Brandy, White Bordeaux, and White Burgundy. Figure 4.1 shows the distribution of wines under different headings. It also indicates that wines from France account for a disproportionately high amount of lots, 359 out of 558 lots, which is around 68% of total lots. I focus only on the wine lots from the most popular wine category, the chateau location of France. This eliminates the complication arising from both the wine master's and bidders' location preferences.

Parcel lots are prevalent in Christie's wine auction, as shown in my dataset. Figure 4.1 illustrates that 90 out of the total 558 lots are parcel lots, which account for around 16%. A parcel lot is a sequence of several lots which are sold in order, all wine lots in the same parcel are identical and contain wines of the same quantity, condition and bottle size with an identical estimation price for each lot. However, the wines carry in the same lot do not have to be identical and a lot may be composed of several different wines.

When the parcel is auctioned, bidding starts with the first lot in the parcel. At the discretion of the auctioneer, the winner of the first lot can take any or all further

Figure 4.1: Distribution of lots in Christie's wine auction categories



lots in the parcel at the same price. Any remaining lots will continue to be sold by auction, starting at the previous second highest bidder's bid and the next winning bidder will also be able to exercise the option to take any or all the remaining lots in the parcel for the same price. Bidding will continue in this manner until all lots are declared sold or unsold. Christie's recommend bidders to bid on the first lot of the parcel. Absentee bids that are superseded in any lot in a parcel will be submitted in the next lots in the sequence until the absentee bid is successful or until the end of the parcel.

It is important to note that a number of empirical studies document the existence of a price declining anomaly for parcel lots in various auction settings, such as, Buccola (1982) for livestock auctions, Burns (1985) for wool auctions, and Ashenfelter (1989), McAfee and Vincent (1993) and Di Vittorio and Ginsburgh (1994) for wine auctions. In order to avoid the distortion caused by a price declining anomaly, I only include the auction information from the first sold lot in the 90 observations of parcel lots in my dataset.

It is also important to note that there are several lots that carry a mixture of wines with different characteristics, such as location, chateau, vintage year, and bottle

Figure 4.2: Example of mixed French wines lot

MIXED FRENCH WINES

Lying in Corsham, Wiltshire (Octavian)

Château Lynch-Bages 1964
Slightly bin-soiled, marked labels. Levels: top-shoulder (1)
Château Léoville-Barton 1990 (2)
Marquis d'Angerville, Volnay Champans 1998 (2)
Alex Gambal, Chambolle-Musigny Les Amoureuses 2006 (2)
Château de Beaucastel, Châteauneuf-du-Pape,
1981 (1)
1983 (1)
1985 (1)
The Beaucastel with corroded capsules. Slightly damaged labels. Levels: 1cm below base of corks

319 10 bottles *per lot* £500-600
€590-700

sizes. These mixed wine lots are not ideal for my empirical analysis. To illustrate this point, consider for example a mixed wine lot which contains two bottles of wine that have different vintage years; it would be illogical and misleading to take an average value for the covariate 'vintage year' for this lot, and similarly for other covariates such as price per litre and score. Since it is difficult to find suitable values for these observed heterogeneous covariates, I exclude the observations of mixed wine lots and only focus on the lots with either a single bottle or multiple identical bottles. Figure 4.2 captures lot No. 319, an example of a mixed wine lot which contains 10 regular bottles of French wine that are from different Chateaus in France and with different vintage years.

Table 4.1 reports the descriptive statistics of my sample, which includes 153 cross-sectional observations that are either single bottle lots or multiple identical-bottle lots. Observations also include the first lot of each parcel lot. The number of bids in each lot differs. Lots that only have one bid indicate that the lot was sold at the reserve price. In the case where only a single bidder participates in the auction, a higher reserve price that is closer to the bidder's valuation can lead to a higher ending price, which highlights the importance of setting an optimal reserve price.

Table 4.1 also shows that the average auction realised price per litre is 43.33% higher than the average auction low estimation price per litre, that the average reserve price per litre is around 22% higher than the average low estimation price per

Table 4.1: Summary statistics

Variable	**Obs**	**Mean**	**Std.Dev.**	**Min**	**Max**
Ending price (with premium)	153	2914.67	12149.90	106	146875
Ending price per litre (with premium)	153	768.01	2533.36	20.89	18016.67
Ending price (without premium)	153	2480.57	10340.34	90.21	125000
Ending price per litre (without premium)	153	653.62	2156.05	17.78	15333.33
Reserve price	153	2127.71	8359.30	50	100000
Reserve price per litre	153	651.00	2169.09	16.67	13333.33
Bids	153	3.12	2.64	1	14
Vintage year	153	1994.95	15.37	1947	2012
Wine score	153	17.54	1.03	13	20
Size (in litres)	153	6.70	5.72	0.75	45
Low estimate	153	2031.24	9163.01	60	11000
Low estimate per litre	153	534.91	1984.17	18.52	13333.33
High estimate	153	2532.09	10849.42	70	130000
High estimate per litre	153	686.50	2531.74	25.93	20000

litre, and that 69% of auctions have a reserve price higher than the low estimation price. This suggests that the low estimation prices for many lots are set deceivingly low. It can be speculated that auction houses purposefully set at an attractively lower estimation price to attract more potential bidders to register their interest, but set a higher reserve price in order to achieve higher revenues. In Christie's auctions only registered bidders are allowed to submit bids, so the number of potential bidders in an auction is fixed from the start of the auction. Therefore, it is important to attract more potential bidders to join the salesroom before the auction commences to increase competition.

Figure 4.3: Distribution of lot size

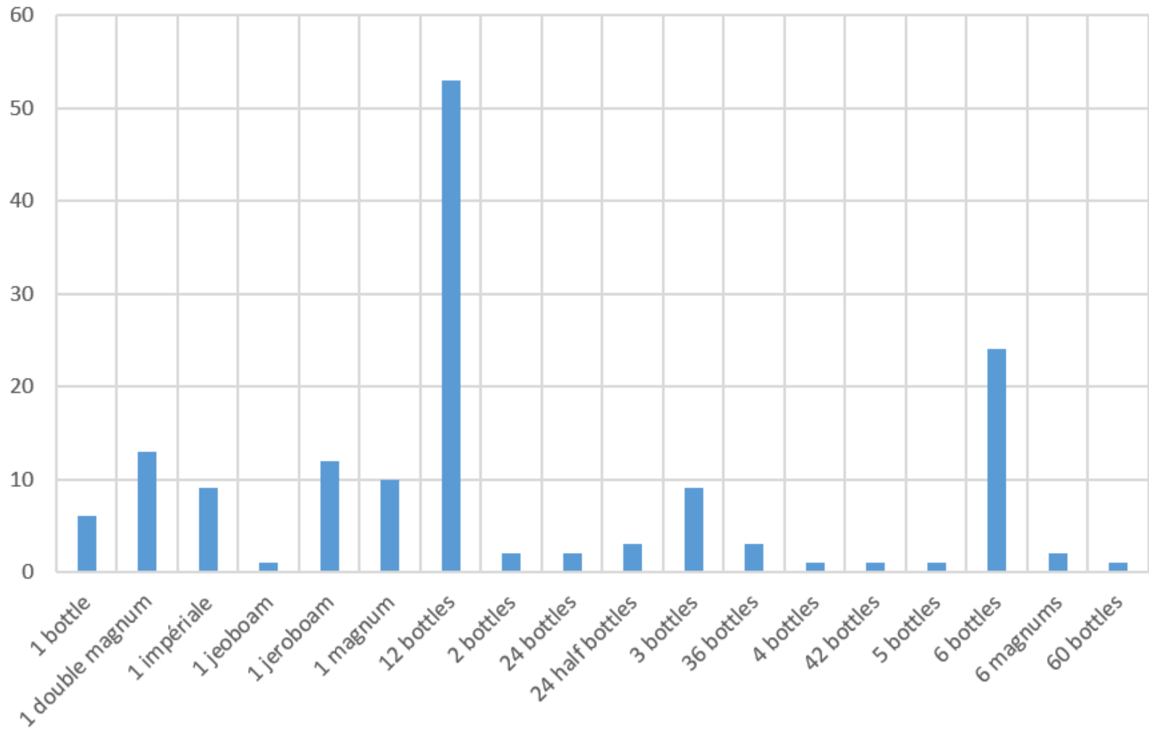
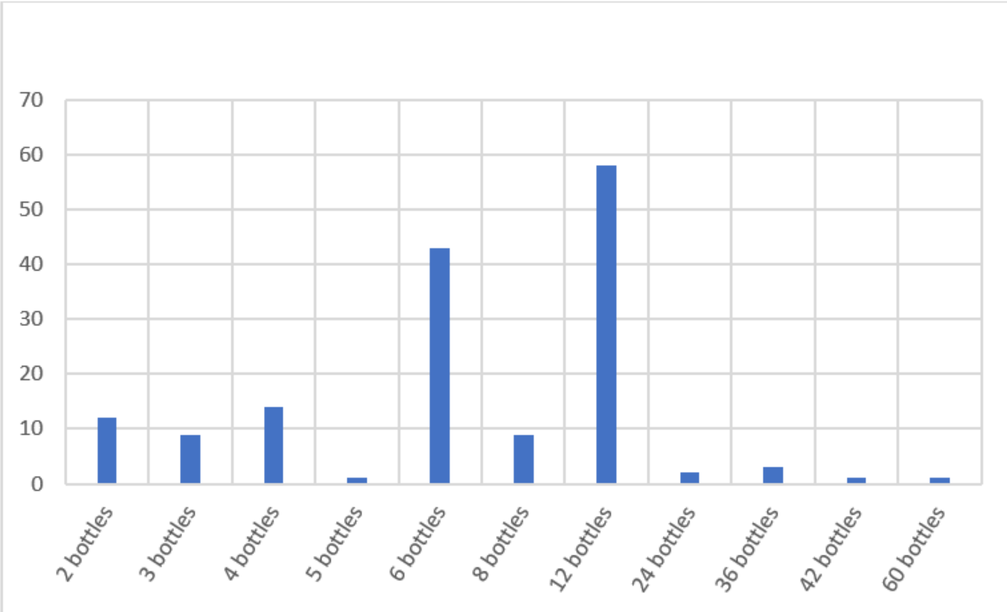


Figure 4.3 illustrates the frequency distribution of lots according to different lot sizes. To get the better idea of distribution of wine lot size, I convert the unit of each wine lot into bottles and re-ordered the horizontal axis accord to the size of wine lot, see Figure 4.4. One wine lot may contain several bottles of wine, thus the lot size varies among different auctions, which makes it difficult to conduct a comparison across lots. In order to facilitate the comparison, I standardize the ending prices, reserve prices and lot estimations to GBP per litre.

Drinking notes and scores for each wine are given by Jancis Robinson, a renowned wine master from the UK, which are taken from her website JancisRobinson.com. Her scores use a 20-point scale which allows for half-scores. Points are given based on the specific sensory characteristics of wine, such as colour, aroma and flavour, as well as more technical qualities including the balance of sugars, acids, tannins and volatile acidity. According to Jancis Robinson's historical scores and rating records, her wine scores usually fall within the range of 9.5 and 20, with an average score value that is equal to 16.4 points, where around half of the wine

Figure 4.4: Distribution of lot size (bottle)



tastings have scores between 16 and 17 points.

Table 4.2 gives an insight into what the numerical scores mean in Jancis Robinson's 20-point Scale, as described on her website. It is important to note that wines scores are given in the context of the particular wine in question, as it is difficult to compare two completely different wines on a linear scale. On her website, she gives the example that a red Burgundy simply cannot be scored on the same scale as a New World Pinot Noir.

Figure 4.5 illustrates the distribution of wine scores in my sample. The minimum value of scores in the sample is equal to 13, the mean value of scores is equal to 17.5 with a standard deviation 1.03. Both the average and minimum wine scores in the sample is higher than the average and minimum scores listed on the Jancis Robinson website, which suggests that the quality of wine auctioned in Christie's auction is above average.

There are several reasons why Jancis Robinson's wine scores are distinguishable from others. Jancis Robinson is one of the most respected and renowned wine critics in the UK, with wine buyers worldwide who consider her scores when making

Table 4.2: Jancis Robinson's 20-point scale

Score	**Explanation**
20	Truly exceptional
19	A humdinger
18	A cut above superior
17	Superior
16	Distinguished
15	Average
14	Deadly dull
13	Borderline faulty or unbalanced
12	Faulty or unbalanced

Figure 4.5: Distribution of wine score

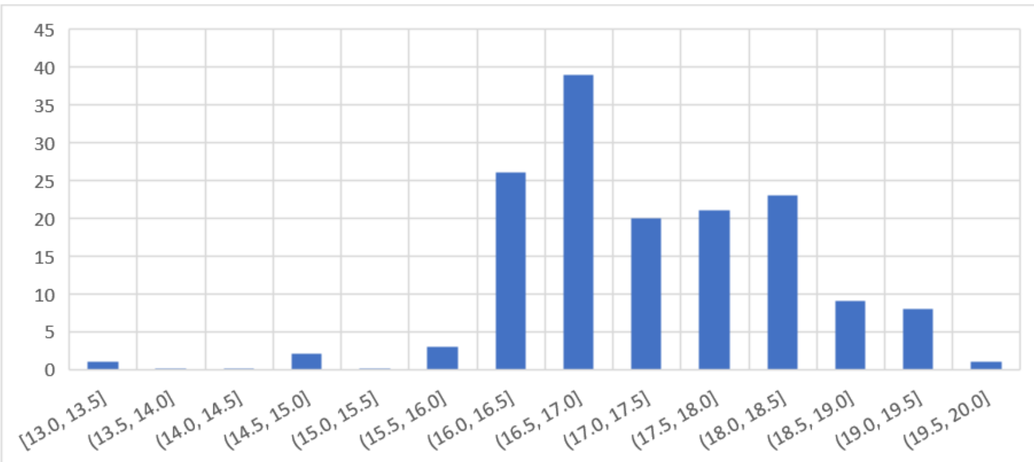
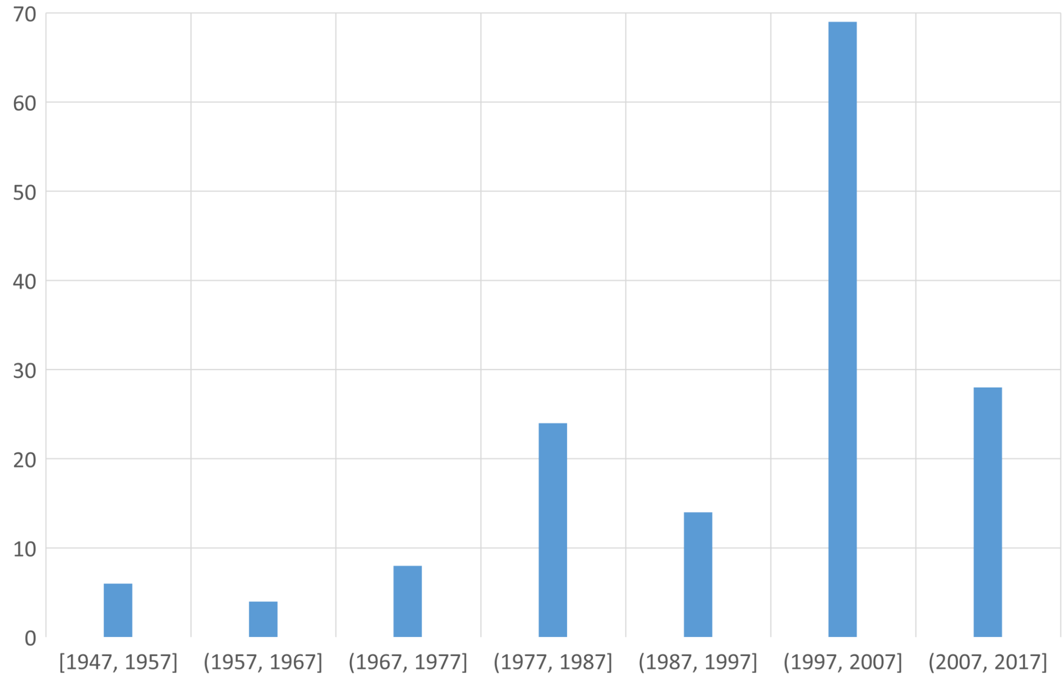


Figure 4.6: Distribution of wine vintage year



a primary purchasing decision. Wine score records on her website are updated frequently, and the majority of wines in my dataset are tasted and scored within the short time interval of 2 years. Due to the varying quality of wine characteristics as time elapses, it is essential that the tasting date is close to the time of wine score data collection to better capture the real quality of wines. Also, the quality score given by Jancis Robinson is based on the sensory characteristics of wine, such as taste, aroma and colour, and if the wine has not yet reached its peak maturity, it is also combined with perceived potential. Therefore, a wine that is on the way up to its peak of maturity, such as, a new Burgundy, would likely be assigned a high score due to its high potential performance in the future; however, a wine that is on its way down from its peak will be given a score that denotes the sensory characteristics on the day of the taste testing. Finally, all the wine scores are exclusively given by Jancis Robinson which avoids the problem of measurement error in the empirical analysis. These four characteristics of Jancis Robinson's wine score system enable me to make reliable and consistent inferences on the impact of wine score on bidders' private valuations.

Figure 4.6 illustrates the frequency distribution for the vintage year of wines in my dataset. The wine lots in this dataset have vintage years spanning seven decades, with a standard deviation of 15.37. By far, the most common interval for wine vintage year is 1997-2007 with 69 lots, accounting for 45.10% of the total lots in my dataset; contrastingly, wine lots from the first five vintage year intervals, 1947 to 1997, only account for 56 lots. It can be noted that the distribution of wine vintage year is skewed heavily to the right with a long left tail, and the reason behind this is that most wines sold at auction have a peak maturity of around 10 to 20 years. The age of peak maturity differs accordingly with wine types: longer for sweet wines as the sugar acts as a preservative, but shorter for white wines as the breakdown of some components can make the wine bitter.

4.7 Empirical methodology: indirect inference

I estimate the impact of wine scores on bidders' private valuations using the structural model of the winning bidder's equilibrium bidding strategy, derived in section 4.5. I adopt the indirect inference approach proposed by Li (2010), to estimate the parameters of the structural model and characterise the underlying bidder's private valuations.

The indirect inference approach is a simulation-based method, which is used to estimate the parameters of a model that has latent variables, incomplete data or an analytically intractable likelihood function. Early theoretical work related to indirect inference was proposed by Smith (1993), who used an indirect inference approach in a time series model. Gourieroux et al. (1993) developed the approach further in full generality by using parameter calibration. Gallant and Tauchen (1996) suggested a similar approach for the moment conditions models, known as the efficient method of moments (EMM).

The approach that I adopt has two stages. The first stage uses ordinary least squares to estimate the parameters of an auxiliary model. In the second stage, the estimated parameters from the first stage are used to simulate the latent variable in the structural model. The parameters of the model are estimated by minimising the distance between the simulated variable from the structural model and the same variable in the auxiliary model. Computations are carried out using Mathematica.

4.7.1 Indirect inference: stage one

In the first stage, I run an ordinary least squares estimation to get the estimations of $\hat{\beta}_n$. I consider the following linear regression specification,

$$Log(y_i) = x_i'\beta_n + \varepsilon_i \quad where\ \varepsilon \sim \mathcal{N}(0, \sigma^2) \tag{4.16}$$

where y_i is the auction ending price of wine per litre in lot i and x_i is a vector of observed covariates, that includes vintage year, wine score, and location dummies related to the provenance of France (Burgundy, Rhone, and Bordeaux). I use Bordeaux as a baseline in the regression.

4.7.2 Indirect inference: stage two

In the second stage, the estimated parameters of wine characteristics from the first stage, $\hat{\beta}_n$, are used to simulate the latent private valuation of winning bidders for each auction, v_i^s. Then, I simulate the winning bidder's equilibrium bid, $\tilde{b}_i^s$, according to the structural model derived in section 4.5. Then the parameters of the structural model, θ, are estimated by minimising the distance between the simulated variables from the structural model and the same variables from the auxiliary model.

I consider 100 [2] simulation paths of the winning bidder's valuation, v_i^s, for 153 auctions: $[v_i^s(x_i, \theta), i = 1, ..., 153 | s = 1, ..., 100]$, where the winning bidders' valuations, v_i^s, are independently drawn from the bidder's valuation density function, equation 4.17, by substituting the first-stage estimates of the coefficients, $\hat{\beta}_n$, and the observed heterogeneity covariates, x_i.

$$f_v(v_i|x_i) = \gamma \exp(-\gamma v_i), \quad \gamma = \frac{1}{\exp(\hat{\beta}_0 + \hat{\beta} x_i)}, \tag{4.17}$$

In the presence of a reserve price, the distribution of observed bids in my sample is truncated as only bidders who have valuations higher than the reserve price submit

[2] I tried higher level of simulation path for example 150, 200, the estimates do not vary much.

bids. Therefore, each simulated winning bidder's valuation must be higher than the reserve price.

Then, I simulate the winning bidder's equilibrium bid, $\tilde{b}_i^s$, by substituting the observed covariates, x_i, number of potential bidders, N, and the corresponding simulations of the winning bidders' valuations, v_i^s, into the structural model derived in Section 4.5 and keep the structural parameters θ_0 and θ unknown. The structural model of the equilibrium bidding strategy for winning bidder i is,

$$
\begin{aligned}
& b_i = v_i - \frac{1}{F_v(v_i|x_i)^{N-1}} \int_r^{v_i} F_v(u|x_i)^{N-1} du \\
& \textit{where} \quad F_v(v_i|x_i) = 1 - \exp(-\gamma v_i), \quad \gamma = \frac{1}{\exp(\theta_0 + \theta x_i)}
\end{aligned}
\tag{4.18}
$$

In auction theory, the number of potential bidders N is assumed to be known. However, in the context of this chapter, it is a latent variable. It would be misleading to assume that the latent variable, the number of potential bidders, is equal to the observed variables, the number of bids or the number of actual bidders, as the relationships between the number of potential bidders and each of the latter two observed variables are ambiguous. Levin and Smith (1994) find that the number of observed bidders in an auction is endogenously determined, given a set reserve price. Only potential bidders who have private valuations above the announced reserve price will place bids. Also, in English auctions, bidders are allowed to submit more than one bid so the number of observed bids tends to be higher than the number of potential bidders. Bajari and Hortacsu (2003) find that in eBay online auctions, the number of potential bidders is large, however, only a comparatively small amount of bids are observed, which implies that only a low proportion of potential bidders submit a bid in an auction. To overcome this problem, I employ the results from Guerre et al. (2000) who show that the maximum number of actual bidders is a good proxy for the number of potential bidders, with the assumption that every auction has the same number of potential bidders. In my model, the maximum number of registered bidders in a single auction is equal to 6, therefore the number of potential bidders N is assumed to be 6 in the structural model. I generate 100 winning bidders' equilibrium bids, $[\tilde{b}_i^s(x_i, \theta), i = 1, ..., 153|, s = 1, ..., 100]$, for each of the 153 lots using equation 4.18

Then, I regress the simulated winning bids, $\tilde{b}_i^s$, on the observed heterogeneity covariates, x_i, for each of the 100 simulations, $[s = 1, \ldots, 100]$, and get the estimators, $\hat{\beta}_n^s$. Note that $\hat{\beta}_n^s$ is a function of the structural parameters, θ.

The indirect inference estimator $\hat{\theta}$ for the structural parameter θ is now defined as the solution to the following minimum distance problem,

$$\min_{\theta \in \Theta} [\hat{\beta}_n - \frac{1}{s} \sum_{s=1}^{100} \hat{\beta}_n^s(\theta)] A [\hat{\beta}_n - \frac{1}{s} \sum_{s=1}^{100} \hat{\beta}_n^s(\theta)] \tag{4.19}$$

where the positive definite matrix A serves as a weighted matrix.

The problem of finding the indirect inference estimator, $\hat{\theta}$, is simplified when the number of the auxiliary parameters is the same as number of the structural parameters. So, the reduced equation of equation 4.19 is,

$$\hat{\beta}_n - \frac{1}{s} \sum_{s=1}^{100} \hat{\beta}_n^s(\theta) = 0 \tag{4.20}$$

By solving the reduced equation (4.20), I get the indirect inference estimators, $\hat{\theta}$, that determine the bidder's valuation distribution.

4.8 Empirical results and estimates of the optimal reserve price

4.8.1 First stage: OLS estimation of auxiliary model

Table 4.3 illustrates the first stage auxiliary model regression results, and all the coefficients are statistically significant at the 5% level. The regression results suggest that wine scores serve as a particularly good quality indicator for wine sensory characteristics; consequently, wine scores have a significant impact on auction ending prices. From my OLS regression, an increase of one point in wine score results in an increase in auction end prices by a substantial 16.19%.

Table 4.3: First stage auxiliary model estimation for Christie's wine auctions

Linear regression						
			Number of observations: 153			
			$F(4, 148) = 47.20$			
			$Probability > F = 0.0000$			
			$R - squared = 0.4245$			
			$RootMSE = 0.3757$			
Log (price per litre)	**Coef.**	**Robust Std.Err.**	**t**	**p>|t|**	**[95% Conf. Interval]**	
Burgundy	.2172	.0543	4.00	.0000	.1010	.3245
Rhone	.3479	.1717	2.03	.0444	.0087	.6872
Scores	.1619	.0347	4.66	.0000	.0932	.2305
Log(year)	-69.5953	8.1806	8.51	.0000	-85.7611	-53.4296
Constant	231.0542	27.0582	8.54	.0000	177.5839	284.5246

The regression result suggests that there are distinct preferences for the vineyard location of wine. The outcomes indicate Rhone is the most preferred location in my sample, as the auction ending prices of wine from vineyards in Rhone are around 35% higher than Bordeaux, while the wines from Burgundy have prices that are around 22% higher than Bordeaux.

The vintage year of wine is also a crucial determinant of auction ending prices; an increase in the vintage year by one percentage point will decrease auction ending prices by 69.60%. The negative sign is consistent with my expectations; amongst the top quality age-worthy wines in the fine wine market, newer wines are relatively more common and generally of lesser value, in comparison to the older collectable wines.

4.8.2 Second stage: structural model analysis

Table 4.4 presents the estimates obtained from the one hundred simulations in stage two of my indirect inference analysis. The indicator for sensory characteristics, wine scores, is significant at the 5% level with a positive effect on bidders'

Table 4.4: Second stage: structural model results adopting OLS as an auxiliary model

Log (price per litre)	**Coef.**	**Bootstrap Std. Err.**
Burgundy	.0117	$.2131$
Rhone	-.0760	$.1222$
Scores	.0802	$.0343^{**}$
Log(year)	-16.3907	13.1712
Constant	53.1253	41.0726
Notes. *p < 0.10, **p < 0.05, ***p < 0.01.		

valuations. However, the magnitude decreases from 16.19% to 8.02%, so an increase in wine score by one point will increase the expected private valuation of a typical bidder in Christie's by around 8.02%.

The structural model estimates in Table 4.4 suggest that when bidders are uncertain about the quality of sensory characteristics of age-worthy wines, they may heavily rely on the wine scores. In particular, the wine scores given by Jancis Robinson well capture the average preferences of wine collector and investor in the fine wine market. Wine scores seem to be a good guide for the first time buyers, who are uncertain about the quality of sensory characteristics of a bottle of wine.

Table 4.4 also shows that the estimates for the non-sensory characteristics, chateau, location and vintage are not significant at the 5% significance level, in the second stage of the indirect inference approach. These results contradict those obtained by the auxiliary model and also to a number of empirical studies of the wine market. One possible explanation for this result is that all the wines in my dataset are from most famous regions and prestigious vineyards of France and all vineyards in the dataset are located in a relatively small area. Therefore, the differences in the non-sensory characteristics are less important to potential bidders. Another explanation is that wine scores provide a strong enough indicator of quality to affect and shape the consumer's preferences and persuade consumers to acquire a taste for certain sensory characteristics.

Due to the calculation of the asymptotic variance-covariance matrix being computationally complicated in this case, I employ the bootstrap method to get an estimate for the asymptotic variance and covariance matrix. I compute the standard errors in Table 4.4 using the following bootstrap procedure.

First, I fit the model by substituting the structural estimators $\hat{\theta}$ and observed heterogeneous covariates, x, into the auxiliary model in stage 1. I retain the fitted values, $\hat{y}_i$ = $\hat{\theta}_0$ + $\hat{\theta}_1 x_{i,1}$+$\hat{\theta}_2 x_{i,2}$ + $\hat{\theta}_3 x_{i,3}$ + $\hat{\theta}_4 x_{i,4}$, and the corresponding residuals for each of the 153 auctions, $\hat{\varepsilon}_i$ = y_i - $\hat{y}_i$, for i= 1,. . . ,153.

Second, I resample from the centered residuals ($\hat{\varepsilon}_j$ - $\mathbb{E}(\hat{\varepsilon}_j|x_i)$), $[j = 1, \ldots, 153]$ to get bootstrap response variables y^* = $\hat{y}_i$ + ($\hat{\varepsilon}_j$ - $\mathbb{E}(\hat{\varepsilon}_j|x_i)$).

Third, I regress y^* on the vector of covariates, x to obtain $\hat{\beta}^*$, in the view of my previous discussions, those OLS estimates can be treated as the estimates of the first stage auxiliary model of the indirect inference approach. I repeat the indirect inference estimation approach proposed in section 4.7 to obtain the bootstrap structural parameters $\hat{\theta}^*$.

Fourth, I repeat the second and third steps 800 times and calculate the standard errors for each estimate.

The above indirect inference estimation allows me to recover the underlying bidder's private valuation distribution, which, in turn, can be used for estimation of the optimal reserve price for each auction at a given level of seller's valuation, v_0. Equation 4.15 defines the optimal reserve price and is derived in section 4.5. For reference, it is given again as follows:

$$\rho^* = v_0 + \frac{[1 - F_v(\rho^*)]}{f_v(\rho^*)} \tag{4.21}$$

assuming the bidder's valuation distribution is exponential, and the probability density function and the cumulative distribution function is respectively,

$$\begin{aligned} f_v(v_i|x_i) &= \gamma \exp(-\gamma v_i) \\ F_v(v_i|x_i) &= 1 - \exp(-\gamma v_i) \quad \textit{where} \quad \gamma = \frac{1}{\exp(\theta_0 + \theta x_i)} \end{aligned} \tag{4.22}$$

By substituting for $f_v(\rho^*)$ and $F_v(\rho^*)$ into equation 4.21, it follows that,

$$\begin{aligned} \rho^* &= v_0 + \frac{[1 - (1 - \exp(-\gamma\rho^*)]}{\gamma \exp(-\gamma\rho^*)} \\ \rho^* &= v_0 + \frac{\exp(-\gamma\rho^*)}{\gamma \exp(-\gamma\rho^*)} \\ \rho^* &= v_0 + \frac{1}{\gamma} \quad \textit{where} \quad \gamma = \frac{1}{\exp(\theta_0 + \theta x_i)} \end{aligned} \tag{4.23}$$

where p^* is the optimal reserve price for each wine lot, v_0 is the seller's valuation, θ is a vector of structural parameters, and x_i is a vector of observed heterogeneous covariates for each wine lot. Since my structural analysis yields estimates for the structural parameters in equation 4.23, it is possible for me to estimate optimal reserve price for each wine lots for a given seller's valuation.

However, in my case, the seller's valuation, v_0, is a latent variable, as it is not revealed during the course of the auction. In many empirical studies of sealed-bid and English auctions, such as, timber auctions, auctioneers set the reserve price equal to the seller's valuation, which is the lowest price a seller is willing to sell their item, usually at a fair market price. However, I find that this is not the case in Christie's wine auctions. Christie's catalogue book provides some information about the secret reservation price, it states that, 'unless otherwise indicated, all lots are subject to a secret reservation price which cannot be more than the lot's low estimate. Here, the secret reservation price is the lowest price the seller is willing to sell the item, so if the highest bid in a auction does not exceed the secret reservation price, the lot is unsold. This statement allows me to make inferences about the possible range of seller's valuations. According to Christie's catalogue book each wine lot has a secret reservation price which cannot be higher than the lot's low estimate. Therefore, if the reserve price of a wine lot is higher than its low estimate, it must also be higher than both the secret reservation price and the seller's valuation.

I compare the reserve price for each lot with the corresponding low estimates and find that 69% of the wine lots in my sample have reserve prices that are above the lot's low estimate, and only the remaining 31% of lots have reserve prices that are below or equal to the low estimate. As a result, only the remaining 31% of the lots can possibly have a functioning secret reservation price, where the reserve price is below the upper bound of the secret reservation price. This contradicts the statement in Christie's catalogue book, which suggests that all auctions commence at a price below the low estimate.

The prevalent existence of absentee bids is a possible explanation behind the reserve prices that are above the low estimate, observed in a large proportion of lots in my dataset. There are several papers that recognise the existence of absentee bids in live auctions, for example, Ginsburgh (1998), report that a large number of bidders in Christie's and Sotheby's auctions submit absentee bids. Some bidders choose to submit absentee bids before the start of the auction and the auctioneer bids on behalf of them. The reasons why bidders choose to submit absentee bids differ, but it is usually a combination of the value of the lot being relatively inexpensive and the high added costs of the bidder participating in the auction in person, such as travel and time costs.

It would be interesting to explain how absentee bids affect the reserve price. Christie's allow bidders to submit absentee bids up to 24 hours before the auction commences, therefore, the reserve price can only be confirmed the day before the sale takes place. There are three ways in which absentee bids can affect the reserve price of the auction. I assume that the auctioneer originally sets the reserve price at the seller's valuation. First, if the auctioneer only receives one absentee bid which is lower than the original reserve price, the reserve price does not change. Second, if the auctioneer receives an absentee bid that is higher than the original reserve price, the new reserve price is equal to the original reserve price plus a minimum increment. Finally, if the auctioneer receives more than one registered absentee bid that is higher than the original reserve price, the auction will normally commence at the price equal to the second highest bidder's absentee bid plus a minimum increment. Therefore, the existence of absentee bids explains why there are a number of observed reserve prices that are considerably higher than the seller's valuation. If I take all the observed reserve prices as a proxy for the seller's valuation, the results would be biased. However, according to the analysis above, it is reasonable to believe that the reserve prices of the remaining 31% of lots (46 lots) are closer to the seller's valuation, v_0, so it is interesting to test whether the reserve prices in these lots are high enough.

I estimate the optimal reserve prices for the 46 wine lots using the structural estimates from Table 4.4 for given levels of seller's valuation. I use the estimated optimal reserve prices and estimated structural parameters to simulate the winning bidders' bids. In order for the simulation outcomes at different reserve levels to be comparable, all the winning bidders' simulated valuations are drawn from the same distribution that is truncated at the observed original reserve price, for each wine lot. The simulated bidders' valuations are identical at given levels of optimal reserve price.

Initially, I assume the original reserve price announced by Christie's is equal to the seller's valuation and estimate the optimal reserve price for each wine lot. However, I find that most of the estimated optimal reserve prices are higher than the simulated winning bidders' valuations, so most of the wine lots are unsold. Therefore, I decrease the level of the seller's valuation and estimate the corresponding optimal reserve prices. The optimal reserve prices are used to simulate the average revenue and sell through rates for the 46 wine lots. Table 4.5 reports the average simulated revenue and average sell-through rate at given levels of optimal reserve prices, for 800 simulations. The table reports the estimated optimal reserve prices at given levels of sellers' valuations at 60% and 55% of the reserve prices. The revenue and sell through rate at the observed reserve price are given

Table 4.5: Simulations of auctions with optional reserve price at different levels of seller's valuations

V_0 as a percentage of reserve price	Average increase in optimal reserve price, given as a percentage of original reserve price	Average total revenue, in GBP	Average sell through rate
100% (original)	-	418,455	100%
60%	41.6%	570,241	65.60%
55%	36.6%	474,165	70%

for comparison.

I find that an increase in the reserve price tends to increase the frequency of some lots being unsold, since the higher estimated optimal reserve prices tend to exceed the simulated winning bidders' valuations. However, I find that the average total revenue increases as the reserve price increases. As an example, when I assume the sellers' valuations are at 55% of the original reserve price, the estimated optimal reserve price level is 36.6% higher than the original reserve price level. At the optimal reserve price, I calculate the average of the 800 simulated total revenues of 46 lots to be 474,165 GBP, however, the sell-through rate decreases to 70%. The average total revenue of the 46 lots increases by 55,710 GBP, which is an increase of around 13%, in comparison to the average total revenue of 418,455 GBP yielded from auctions with the original reserve price level. Evidently, this suggests that the reserve price for Christie's auctions is set below the optimal level, since, for each auction, an increase in the reserve price level increases the auction ending price.

4.9 Summary and conclusion

In this chapter, I empirically analysed the effect of wine scores on the bidder's private valuation in Christies' wine auctions. Firstly, due to the bidder's private valuation being a latent variable, I used an indirect inference approach to estimate the bidder's private valuation distribution. I found that wine score is the essential structural element that determines the underlying private valuation distribution of bidders in Christie's wine auction. The expected private valuation of bidders increases by 8.02% for every one-point increase in wine score, based on the 20-point scale by Jancis Robinson. This result indicates that, when bidders face a purchase decision, they heavily rely on wine scores when they are uncertain about the quality of fine wines in the auction market. It also indicates that wine scores capture the average preferences of buyers in the fine wine market well.

I also found that the coefficients of two of the most popular vineyard locations, Bordeaux and Burgundy, are not significant, which conflicts with other empirical findings in wine retail market. After revealing the bidder's private valuation distribution, I estimated the optimal reserve price for each auction and found that the original reserve price set by Christie's is not optimal; an increase in reserve price may lead to a higher total revenue.

Bibliography

Anderson, T. W. and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association*, 76(375):598–606.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Ashenfelter, O. (1989). How auctions work for wine and art. *Journal of Economic Perspectives*, 3(3):23–36.

Bajari, P. and Hortacsu, A. (2003). The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics*, pages 329–355.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, 12(3):151–170.

Bikhchandani, S. and Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff papers*, 47(3):279–310.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Bombrun, H. and Sumner, D. A. (2003). What determines the price of wine. *AIC Issues Brief*, 18:1–6.

Brendstrup, B. and Paarsch, H. J. (2006). Identification and estimation in sequential, asymmetric, english auctions. *Journal of Econometrics*, 134(1):69–94.

Buccola, S. T. (1982). Price trends at livestock auctions. *American Journal of Agricultural Economics*, 64(1):63–69.

Burns, P. (1985). Experience and decision making: A comparison of students and businessmen in a simulated progressive auction. Technical report, The Field Experiments Website.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Cameron, A. C. and Trivedi, P. K. (2010). *Microeconometrics using stata*, volume 2. Stata press College Station, TX.

Chakraborty, I. and Kosmopoulou, G. (2004). Auctions with shill bidding. *Economic Theory*, 24(2):271–287.

Di Vittorio, A. and Ginsburgh, V. (1994). Red wines of medoc vintages from 1949 to 1989 at christie's auctions. *Unpublished paper*.

Duan, W., Gu, B., and Whinston, A. B. (2009). Informational cascades and software adoption on the internet: an empirical investigation. *MIS quarterly*, pages 23–48.

Engelberg, J. and Williams, J. (2005). A license to shill: Incremental bidding and ebay's proxy system. *Kellogg School of Management working paper*.

Gallant, A. R. and Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12(4):657–681.

Ginsburgh, V. (1998). Absentee bidders and the declining price anomaly in wine auctions. *Journal of Political Economy*, 106(6):1302–1319.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, 8(S1).

Graham, D. A., Marshall, R. C., and Richard, J.-F. (1990). Phantom bidding against heterogeneous bidders. *Economics Letters*, 32(1):13–17.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574.

Haile, P. A. and Tamer, E. (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51.

Hoppe, T. and Sadrieh, A. (2007). An experimental assessment of confederate reserve price bids in online auction. Technical report, Otto-von-Guericke University Magdeburg, Faculty of Economics and Management.

Izmalkov, S. (2004). Shill bidding and optimal auctions. In *International Conference on Game Theory*, pages 12–16.

Jones, G. V. and Storchmann, K.-H. (2001). Wine market prices and investment under uncertainty: an econometric model for bordeaux crus classés. *Agricultural Economics*, 26(2):115–133.

Kauffman, R. J. and Wood, C. A. (2003). Running up the bid: detecting, predicting, and preventing reserve price shilling in online auctions. In *Proceedings of the 5th international conference on Electronic commerce*, pages 259–265. ACM.

Kosmopoulou, G. and De Silva, D. G. (2007). The effect of shill bidding upon prices: Experimental evidence. *International Journal of Industrial Organization*, 25(2):291–313.

Landon, S. and Smith, C. E. (1998). Quality expectations, reputation, and price. *Southern Economic Journal*, pages 628–647.

Levin, D. and Smith, J. L. (1994). Equilibrium in auctions with entry. *The American Economic Review*, pages 585–599.

Li, T. (2010). Indirect inference in structural econometric models. *Journal of Econometrics*, 157(1):120–128.

McAfee, R. P. and Vincent, D. (1993). The declining price anomaly. *Journal of Economic Theory*, 60(1):191–212.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

Oczkowski, E. (1994). A hedonic price function for australian premium table wine. *Australian Journal of Agricultural and Resource Economics*, 38(1):93–110.

Paarsch, H. J. (1997). Deriving an estimate of the optimal reserve price: an application to british columbian timber sales. *Journal of Econometrics*, 78(1):333–357.

Riley, J. G. and Samuelson, W. F. (1981). Optimal auctions. *The American Economic Review*, 71(3):381–392.

Roth, A. E. and Ockenfels, A. (2002). Last-minute bidding and the rules for ending second-price auctions: Evidence from ebay and amazon auctions on the internet. *American Economic Review*, 92(4):1093–1103.

San Martin, G. J., Troncoso, J. L., and Brümmer, B. (2008). Determinants of argentinean wine prices in the us. *Journal of Wine Economics*, 3(1):72–84.

Schamel, G., Anderson, K., et al. (2001). *Wine Quality and Regional Reputation: Hedonic Prices for Australia and New Zealand*. Centre for International Economic Studies.

Simonsohn, U. and Ariely, D. (2008). When rational sellers face nonrational buyers: evidence from herding on ebay. *Management science*, 54(9):1624–1637.

Sinha, A. R. and Greenleaf, E. A. (2000). The impact of discrete bidding and bidder aggressiveness on sellers' strategies in open english auctions: Reserves and covert shilling. *Marketing Science*, 19(3):244–265.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S1).