UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

The Impact of Soft-Close Auctions on Bid Behavior:

A Randomised Controlled Experiment

Martin Kellgren

Supervisors: Johan Stennek and Juha Tolvanen

Master's thesis in Economics, 30 hec

Spring 2023

Graduate School, School of Business, Economics and Law, University of Gothenburg, Sweden

Abstract

This thesis studies the effect of longer extensions on bidding behaviour in online auctions. The study uses a randomised controlled experiment through *Pantbanken Sverige*, a Swedish pawnbroker. The results indicate positive effects on the seller's revenue due to longer extensions, although these effects were not statistically significant. The upper bounds of the confidence intervals do not rule out the possibility of price increases of ten percent when increasing the extensions from thirty to ninety seconds. Additionally, longer extensions significantly increase the probability of further bidding after triggering the first extension. This effect is stronger for gold items, where previous research has highlighted the strategic advantage of late bidding for common value goods. This study provides a first attempt at an empirical investigation of the dynamics of variations in the extensions in online auctions. The potential profitability increase in revenue emphasises the need for further research on the subject.

Keywords: Online Auctions, soft-close auctions, longer extension, randomised controlled experiment, private value, common value, seller's revenue.

Acknowledgments

I would like to extend my sincerest gratitude and appreciation to my supervisors, Johan Stennek and Juha Tolvanen, for their invaluable guidance, support, and encouragement throughout my research. Their expertise, patience, and insightful feedback have been instrumental in getting through the process of writing this master thesis.

I would also like to thank Peter Sundström at *Pantbanken Sverige* for generously allowing me to conduct my experiment on their auction platform. Without Peter Sundström and *Pantbanken Sverige*, this research would not be possible.

Martin Kellgren, August 2023

Contents

Chapter 1

Introduction

English auctions are commonly used online, and these online auctions often use different ending rules. Hard-close auctions have a strict deadline for submitting bids, where the highest bid wins the item. Another commonly used format is the soft-close auction, where a bid within a time limit of the preliminary deadline extends the auction beyond it. In theory, these auctions can go on forever.

Previous studies compare the differences between soft-close and hard-close auctions, e.g., Ariely et al. (2005) and Cao et al. (2019). These empirical and theoretical papers have mainly focused on differences in revenue between the two types of auctions. Some of them find that soft-close auctions bring higher revenue to sellers due to the strategic advantages of late bidding, giving less time to competitive bidders to respond (Glover & Raviv, 2012). However, none discuss the consequences of variations in the time added. With the most focus being on the differences between auctions using fixed or flexible deadlines, it shows a gap in the research on how economic performance is affected by different lengths of extensions in soft-close auctions. Additionally, there is a lack of formal economic theory on the differences between hard- and soft-close auctions, late bidding and the effects of different extensions. Therefore, this thesis is a first attempt to understand how variations in the extensions in online auctions affect auction outcomes. It aims to empirically investigate how changes to online auctions' extensions affect revenue, efficiency and bidders' behaviour through a randomised controlled experiment. More specifically, this thesis aims to answer the following research questions: How do extensions affect the economic performance in online auctions with a focus on the seller's revenue and efficiency? How do different extensions affect the bidding behaviour in extended auctions? Lastly, it examines whether there are variations in the bidding behaviour and economic performance for different types of goods, i.e., between common and private value items. The experiment has been done in collaboration with *Pantbanken Sverige*, a Swedish pawnbroker who sells their items on an online auction platform.

The thesis is organised in the following way; Chapter 2 introduces the related research on the subject, the theoretical framework and the proposed hypotheses. Then, chapter 3 describes *Pantbanken Sverige* in detail and explains the experimental design, data and empirical methods. Chapter 4 shows the experiment's findings and discusses its economic implications and validity. Lastly, chapter 5 concludes.

Chapter 2

Literature Review and Informal Theory

The differences in bidding behaviour, revenue and efficiency between hard and soft close auctions have been widely studied in the academic field. However, as far as I know, no formal economic theories exist on these topics or how extension length variations affect the auction outcomes. Therefore, the literature review focuses on auction theory, researchers' informal theories about the differences in behaviour between the two auction regimes and what the empirical data shows. This section reviews existing research and its implications for this paper.

2.1 Auctions

Auctions are used specifically when the auctioneer does not know the valuations of the individuals. Krishna (2009) describes the different forms of auctions, and the one format relevant to the scope of this thesis is the English auction, which is also the oldest form of auction. Historically, they are held with all participants in the same room, where the auctioneer starts at a low price, which increases as long as there are at least two interested bidders. By not following the increased price, the potential buyer signals that his valuation has been surpassed. It is a second-price auction when the winning bidder pays the price at which the second-last bidder dropped out.

There are different forms of values. There are private values for goods where their value comes from their consumption, and every user values them differently. In this case, each bidder's valuation of the good is independent of others' valuations. The other kind is interdependent values, when bidders do not precisely know the value of the item being auctioned. However, they can have some private signal of the value, e.g., the expertise of the item. This definition of an item's value is especially applicable when the item can be resold after the auction. In this case, bids signal to competitors the common value of the item, which makes competitors reevaluate as the bidding continues (Krishna, 2009). The extreme case of this definition is when an item is of common value when the exact value at the time of the auction is unknown, but the same for all bidders.

2.1.1 Revenue Equivalence Principle and the Efficiency of English auctions

There are two parts to focus on regarding auction outcomes: revenue and efficiency. The former is simply the auctioneer's revenue, and the latter is that the auctioned item is allocated to the bidder who values it the most (Krishna, 2009).

The Revenue equivalence principle was formed in the sixties and states that in a standard auction, i.e., the highest bidder wins the auction. When values are private, independently and identically distributed between risk-neutral bidders, then auctions have the same expected seller's revenue regardless of the format. In simpler terms, if the auction follows the standard rules, bidders follow a symmetric and increasing equilibrium strategy, i.e., within an auction format, bidders use the same strategy and higher valuations are associated with higher bids. Then, the expected revenue for the seller will be the same regardless of the specific auction format used (Krishna, 2009).

There are underlying assumptions for this principle to hold. The first one is the independence of the different bidders' values, i.e., one individual's value does not affect the value of others. The second regards risk neutrality; each bidder maximises their profit. The third one regards budget constraints; each bidder can pay their value. The fourth assumption is that the distribution of values follows the same distribution for all bidders and auctions, with the same number of bidders (Krishna, 2009).

In English auctions, bids are visible to all potential bidders. At every auction stage, each bidder only has to decide whether to drop out or continue bidding. Therefore, they should only stay in the auction as long as the current bid is below their estimated value under the given information; this leads to efficiency for similar reasons as the Revenue equivalence principle, that the one with the highest valuation should win the auction (Krishna, 2009). The implications of online auctions on revenue and efficiency will be discussed later in this section.

2.1.2 Informal theories of Hard versus Soft-Close Auctions

Roth & Ockenfels (2002) describe some strategic reasons why late bidding might occur in online auctions. One reason is that experienced bidders with better knowledge of an item's values might want to keep this information private from others and not reveal it immediately since a bid is information revealed on a person's valuation. The other explanation made by Roth & Ockenfels (2002) is that bidders want to avoid getting into a bidding war with either like-minded bidders or incremental bidders who raise the price as it goes without using a bidding agent. Late bidding is hypothesised by the authors to be more prominent in hard-close auctions than soft-close ones since a fixed deadline gives less time for competitive bidders to respond to a bid in the last seconds of an auction.

Dang et al. (2015) also discusses bidding behaviour. They argue that for private value items, it is the dominant strategy to bid according to their true valuation upon arrival to the auction, independent of the closing rule of the auction. However, it is argued that bidding late is a dominant strategy for common value items. The authors also argue that the reason behind this is that bidders want to keep the information private from their competitors and not reveal it immediately in the auctions.

Ely & Hossain (2009) argues that experienced bidders bid late in online auctions because the average bidder is naive. Instead of using a proxy bidder to place a bid equal to their valuation, they treat it as an English auction. They bid incrementally as the auction continues until they reach their dropout price. Therefore, it is better to counteract this by bidding late. Entering and bidding earlier in the auction would trigger a response and an escalating price, which the authors call an *escalating effect*.

Ariely et al. (2005) argues that short or no extensions in online auctions incentivise bidders to bid late. The authors also discuss the possibility, in these types of auctions, of using bidding bots, which are online services that bid at the last second or before the deadline for the buyer. The authors note that this service is not guaranteed to submit the bid due to network latency issues. However, it should increase the frequency of sniping in hard-close auctions, where late bidding plays a more critical role than in auctions with extensions.

Another thing to discuss regarding the role of extensions in auctions is the bidder participation. There are two ways to view the effect of an increased extension on bidder participation. The first one is that the process of new potential bidders being drawn to an auction is random, i.e., their awareness of the auction is random (Boatwright et al., 2010).

Therefore, different extensions affect bidder participation. An increase in the length of an auction in case of triggered extensions should make new bidders arrive and place their bids for items. The other view works in another direction in that the longer extension window increases the probability that a bidder places another bid. A longer extension makes it more likely for the bidder to attend the deadline. If bidders are drawn in randomly, auctions with longer extensions should have more unique bidders. In either of these cases, the longer extension should bring in more potential bidders or more bids, leading to higher revenues for the seller.

On the other hand, the increased extension leads to monitoring costs for potential bidders as more time is required in case of longer extensions. At the same time, with a longer extension, the monitoring cost of the auction goes down since the browser does not have to be refreshed as often to follow the auction.

In summary, late bidding in online auctions results from strategic decisions to protect valuation information, avoid bidding wars, and counteract naive strategies, with factors like item type and auction rules influencing bid timing. However, without a formal theory of the effects of extensions, it is unclear what direction the effect of longer extensions has on revenue and efficiency.

2.1.3 Previous Empirical Research

The focus now shifts from the informal theories in the literature to the empirical research on late bidding and different auction formats. In a paper by Roth & Ockenfels (2002), they look at auction data from eBay and Amazon, which employ two different auction regimes. The former uses a hard-close auction, meaning that it has a fixed deadline. While the latter utilises a soft-close auction regime, bids during the last minute of the auction extend the deadline by ten minutes. With these platforms, it is also possible to submit a proxy bid. A buyer can put in his reservation price, and the computer does the bidding for him. If no one places a higher bid, the bidder using the proxy will win the item at the second-highest bid price and a small bidding increment. *Pantbanken* uses its bidding agent similarly, as will be described later. The data they use are focused on computers, with widely available retail prices (private value goods) and antiques where values are more challenging to find (common value goods). In the case of the former category, an individual's willingness to pay is private information. While for the latter, expertise comes into play in the evaluation since retail prices are not widely available. Therefore others' bids convey information about the item's value. Their results point toward late bidding being more prominent in hard-close auctions but also occurring in soft-close ones. They complement the observational data with a survey of bidders from eBay, and the survey's answers point towards late bidding as a common strategy for bidders to avoid bidding wars in both types of auctions. The responses also indicate that knowledgeable bidders wait with their bids not to reveal information about the goods that aligns with Ely & Hossain (2009) and Dang et al. (2015). Another critical point is that experienced bidders might choose which site to go to participate in auctions. Those with more experience and knowledge prefer to bid in a hard-close auction where they can use the strategic advantage in late bidding.

In another paper by Ockenfels & Roth (2006), they find further evidence that inexperienced bidders tend to bid incrementally, causing so-called bidding wars. In addition to this, they find that relatively more experienced bidders tend to bid later in both eBay and Amazon auctions to avoid driving the price up with incremental bidders.

In a paper by Ariely et al. (2005), the authors use observational data and a laboratory experiment to analyse bidding behaviour. The observational data comes from eBay and Amazon online auctions with the same auction regimes as in the previous paper. The observational data confirms that the soft-close auctions have frequent early bids, and the eBay auctions have more frequent sniping behaviour, i.e., more frequent last-second bids are used to win auctions with less competitive bids.

To explain these different bidding behaviours in these types of auctions, Ariely et al. (2005) conducted a laboratory experiment. They held the auctions in different settings, but all used discrete periods such that they could identify late bids. Each bidder would randomly be assigned a private value, and the auction winner would receive the difference between the final price and their private valuation. All auctions used the second-price rule, where the final price was decided by the second-highest bid plus a 25-cent increment. The participants were randomised into different auction groups. One group used hard-close auctions where the probability of a successful late bid was a hundred percent sure to go through, and another one with some uncertainty where a late bid only had an eighty percent chance to be successfully submitted. The uncertainty was to simulate online bidding with the bidding bots and potential latency issues. The last auction regime was close to how Amazon auctions operate, where if a bid was submitted in the last period, it had an eighty percent chance of being submitted. If so, the auction was then extended for a new period. In all auctions, the requirement was to bid at least as high as the current bid, such that the current price increased in fixed increments. Therefore the low bidder was unaware of how high the other bid was, similar to how online auctions work if a proxy bid is used. The participants were randomly assigned another bidder within each group and auction, so the auctions were done in pairs. The auctions and randomisation were done twenty times. The results of the experiments showed that there was more late bidding in the hard-close auctions, especially since the auctions were done multiple times. Hence, the frequency of late bidding increased as the bidders gained experience in the format. They also found in the experimental environment that the revenue and efficiency were higher in the soft-close auctions. The reduced incentive to bid late in those auctions makes the buyers bid closer to their valuations. One crucial point the authors make is the risk of generalizability of the experimental results to real internet auctions, that they had a fixed number of bidders for each auction. While on the internet, the number of bidders is more random.

The bidding behaviour in the hard versus soft-close auctions is also examined by Cao et al. (2019). The authors argue that sniping has been confirmed as an equilibrium strategy in hard-close auctions but also examine the prevalence in soft-close auctions. They were looking at data from eBay, which uses a hard-close regime, and Overstock, which uses ten-minute extensions in their auctions. They used the data from a jewellery retailer who sold their goods on both websites.

They find that sniping exists in both auction formats, but the timing differs. In soft-close auctions, sniping takes place right before an extension is triggered. The authors argue that sniping is more profitable in soft-close auctions and offer two possible explanations. The first one is that sniping is a form of tacit collusion, in which bidders tacitly agree to submit bids lower than their actual valuations to increase their surplus in the last period of an auction. Where extensions that may follow are a method to retain the equilibrium, the other is that those with lower opportunity costs are more likely to snipe in soft-close auctions. Since it also includes a monitoring cost after the last-second bid, the bidder must stay in the auction and observe whether further bidding is done in case of extensions. This cost is almost non-existent in hard-close auctions since it can be done with bots and cannot be extended. They even argue that soft-close auctions do not counteract sniping. Contrary, it could damage the seller revenue more, as they argue it does for Overstock compared to eBay auctions.

Glover & Raviv (2012) evaluate differences in revenue in Yahoo! auctions for Ipods, where sellers themselves can choose whether to have a hard or soft-close auction for the items. They also argue that hard-close auctions incentivise late bidding as a form of tacit collusion among buyers and reduce expected revenue for sellers. Their estimations find a 13-20 percent increase in revenue for sellers with extensions in their auctions compared to those with a fixed deadline. Their findings go in line with the experimental results from Ariely et al. (2005) but contradict the findings of Cao et al. (2019).

One thing to remember regarding classifying items is whether they are private, common value, or a mix of both. In a paper by Boatwright et al. (2010), the problems of classifying items as private, common value or a combination of them are discussed. The authors argue that it is impossible to completely distinguish between private and common values in online auctions. They suggest that an item's valuation is made up of two components. Both a private value part and a common value, where the latter makes bidders partially adjust their valuation as competitors place their bids.

A gap in the literature that becomes apparent in these research articles is that they do not discuss the length of the extensions in soft-close auctions or if there is an optimal length. They mainly compare revenue, efficiency, and timing of bids between soft and hard-close auctions. This gap in the literature is what this thesis aims to fill by further examining the effects of variations in the extensions in soft-close auctions. One problem that Roth & Ockenfels (2002) point out is the fact that bidders, especially relatively more experienced ones, self-select into auctions where strategies like sniping might be more profitable for them. Therefore, looking at field data might make the comparison more difficult. While simultaneously, Ariely et al. (2005) complement this type of study with a laboratory experiment, where there are clear signs of sniping being mitigated by extensions, and that experience in different auction formats tends to shape bidder behaviour. Therefore, a randomised controlled trial is an opportunity to evaluate the role of extensions for different goods at a single auction house. That sniping is a more significant issue in hard-close auctions could be expected, while this type of bidding behaviour still seems to be a strategy that also exists for soft-close auctions. The motivation behind late bidding in both types of auctions is that more experienced bidders do not want to reveal their information about valuations for common value goods to end up in a bidding war with incremental bidders. Late bidding is relevant for the case of *Pantbanken*, where many items such as types of jewellery, watches, and art might not be pure private value items but common value items where expertise might be relevant to evaluate items. As mentioned before, Cao et al. (2019) argue that extensions lead to higher monitoring costs compared to sniping in hard-close auctions. The soft-close auctions that the authors bring up use extensions that are ten minutes long. An extension mitigates sniping whilst also putting off potential bidders with relatively too high opportunity costs if it is too long. Therefore, the dynamics of changed extensions in soft-close auctions are an exciting topic.

2.1.4 Implications for the Study

In the case of *Pantbanken*, the definition of the item values is ambiguous. Some items, such as art and accessories, have more properties of private value that potential bidders assign a value to according to the utility they will gain from owning the item. While with jewellery made out of precious metals, there is an interdependent component of the value of the metal. Since gold follows a market value and can be recast and made into new jewellery, it cannot only be consumed as a piece of jewellery. Therefore, for the scope of this thesis, gold items are considered to have more of a common value component, while the rest of the items sold are considered to have more of a private value component.

Regarding possible violations of the Revenue equivalence principle and efficiency, online auctions pose issues that could affect them. The first one is that in an online auction, not every potential bidder follows the auction constantly. As mentioned before, bidder awareness of online auctions is random. Therefore, there is potential for participants to miss that a new bid has been placed. Bidders in online auctions risk dropping out of the auction not because it has surpassed their valuation but simply because they cannot place a new bid, which would violate the increasing equilibrium. The second one is the potential for sniping that a bid is placed just before an extension is triggered, which limits the possibility for someone to react if they are not monitoring the auction, especially since no notification to the bidders happens at that stage. Late bidding that triggers the 30-second extension at *Pantbanken*'s auctions leaves little time for other bidders to react, as many of the papers discussed above look at soft-close auctions where extensions are ten minutes.

Another possible violation of the revenue equivalence principle is that the bidders' valuations are not independent, i.e., when the values are not private. Then, a strategic reason exists for bidders to bid late in the auction for common value items. As a way to hide their information or expertise on the item for sale, as argued in the paper by Ely & Hossain (2009) amongst others. Too short extensions could lead to a situation where those with a higher valuation miss out on the auction when they do not have time to respond to developments, leading to revenue and efficiency losses.

2.2 Hypotheses

As seen in the literature review, most empirical results suggest that extensions in auctions increase seller revenue compared to a hard-close auction format. However, what happens when the length of the extension is changed? A 30-second extension does not give bidders much more time compared to an auction with a fixed deadline, although by definition, it is a soft-close auction. In that case, an extension of 90 seconds should increase the selling price and, in that way, bring more revenue and allocate items more efficiently. Because it gives competing bidders more time to react to late changes in the auctions and mitigates the consequences of late bidding. Simultaneously, a longer extension increases the monitoring cost of the auction in case of extensions. Individuals with lower alternative costs will be more likely to monitor the auction, and if this does not correspond to those with the highest valuation, revenue and efficiency might go down instead. As previously mentioned, no formal theories exist on the role of late bidding and variations in the extension length. Therefore, the directional effect of the longer extension is ambiguous. Then, in this first attempt to empirically investigate the effect of longer extensions, this thesis uses the following two-sided hypotheses:

1. Longer extensions affect the selling prices in online auctions.
2. Longer extensions impact the likelihood of further bidding.
3. These effects are different for gold and non-gold items, i.e., common and private value goods.

Chapter 3

Experimental Design

3.1 Pantbanken Sverige

The field experiment was conducted in collaboration with *Pantbanken Sverige*. It is a Swedish pawnbroker with 22 offices around Sweden. They are physical offices where customers can go in person, but one conducts their business online and through phone calls. As a pawnbroker, they do the valuations of items over the counter for an appraisal of the loan value of the item. The pawnbroker then gives the customer a loan valuation of their item or items. If they accept the valuation, the money is transferred, and the pawnshop keeps the item as security. Many different things could be handed in, and the most common ones are gold jewellery and watches. However, technically, everything with a second-hand value can be used as security, like high-end fashion goods, art, porcelain items from famous producers and, to some extent, new electronics. The loan lasts for six months, during which the customer can come and collect their item by paying back the initial loan plus a monthly interest rate and administrative fees. Alternatively, if they come in and pay their outstanding interest rates and fees, the loan is extended for another six months starting from the payment date. The interest rate starts at a 3.75 percent monthly effective rate but is lower for bigger loans. If a loan defaults, the pawnbroker recovers the debt by selling the item on their auctions (Pantbanken Sverige, 2023b). Multiple items placed as a security for a single loan are separately sold unless they are more valuable together, e.g., a porcelain set.

3.1.1 Pantbanken's Auctions

The auctions are held online on their website, where buying customers can see and place bids on items online. They use a reserve or starting price, approximating the item's minimum market value. This price is based on their auctions of similar items and other selling prices at other Swedish auction houses. The starting price also comes to play in their initial loan valuations, such that there is a marginal between the initial loan value and the planned reservation price if the customer defaults on their loan. Another thing to note about *Pantbanken*'s auctions is that they have a buyer provision, meaning buyers must pay an extra fifteen percent on their winning bid. After a sale, *Pantbanken* can only keep the initial debt, the buyer provision, the accumulated interest and fees. If any surplus remains after that, it goes to the initial loan-taker, as Swedish law regulates.

Therefore, they are incentivised to make a valuation as close to the market value as possible since it generates more interest revenue and there is less risk of a surplus. On average, *Pantbanken* sells approximately 50 percent of the items they have up for auction, meaning that the rest they need to buy back themselves. The ownership is transferred to *Pantbanken* for unsold items, which they then sell in their stores and webshop. If they make a too low valuation, they earn less since the surplus goes to the initial loan taker. On the other hand, if they put a too high starting price, they have to buy it back themselves and probably sell it at a loss. Therefore, it is in *Pantbanken*'s interest to make their best approximation of the minimum market value of an item when it is going to be sold at an auction. One thing to note here is that the value of gold items fluctuates more following changes to gold prices, inducing more uncertainty about the market value when the initial loan is placed.

Each of the 22 offices has its auctions every fourteen days, except one office that has them weekly. The auctions are spread out over the weekdays, except Saturdays when no auctions are held. For the days when more than one office holds their auctions, their starting time is spread out. Therefore, auctions either start at 11, 14, 16 or 19. The variation in starting times is to always have items up on their auction page. The auctions close 30 seconds apart, and new auctions go up as long as there are items. The number of items in each auction depends on how many loans the pawnshop placed roughly six months earlier without being serviced.

The current auction rule states that if a bid is submitted within the last 30 seconds, another 30 seconds is added to the auction. Therefore a competing buyer always has 30 seconds to react to the last bid, i.e., if a new bid is submitted at 11:00:45 when the original end time is 11:01:00. Then the new end time will be 11:01:15. An auction could therefore go on forever as long as new bids are submitted. Bidding can be done in two ways either through a proxy bidder, where the buyer submits a maximum bid, and the computer does the bidding necessary for the buyer to win the auction. E.g. for an item with a minimum price of 500, the buyer is willing to pay 1000 for it. If there are no other bids, the buyer will win the auction at 500. If the competing bid is 800, the bidder using the bidding agent will win the auction at 850. If a buyer uses the bidding agent, their maximum bid is not visible to *Pantbanken* or other buyers. The only visible information is the current and previous bids in the auction. The other way of submitting a bid is through the live auction, where the following 16 items with the closest deadline appear.

Bids are made in fixed increments. For bids between zero and 500, bids are in steps of 25 SEK. After that, the increase is 50 SEK up to 1000. The subsequent increase is 100 SEK up to 3000, then 200 SEK until 10000. The last increases are 500 SEK up until 20000, and for items with prices above that, the increments are 1000 SEK. If someone is no longer the leading bidder, *Pantbanken* notifies the bidder by email at the earliest 30 minutes after the competing bid has been placed. In the live auction just before closing, no such emails are sent (Pantbanken Sverige, 2023b). When two bidders submit the same bid, the earliest bid leads the auction. If someone makes the same bid on the last day as someone who has used the bidding agent, then the first bidder at this level will win the auction if no more bids come in. Therefore *Pantbanken* utilises English second-price auctions since winners pay the value of the second highest bid plus the bidding increment.

3.2 Experimental design

The randomised controlled experiment was conducted as follows. Expired loans, i.e., the items handed in as security, will be sold on *Pantbanken*'s auctions. These are then prepared for auctions, where the pawnbroker writes an informative description for the potential buyers and does a new valuation to find an appropriate minimum price and photographs for the website. When this is done, the items get an auction number, so *Pantbanken* can match it with the buyer after the auction. When all items for one auction is prepared, they send this list with all information and photographs to their auction system, which an external firm handles. Items are published 14 days before the ending date, but this can vary slightly depending on the offices' workload. There they receive an auction id, which is separate from the auction number. This number is simply for the HTML code behind *Pantbanken*'s website. It was over this number that the randomisation was done, and this happens as the items are published to the auction system. Items randomised into the treatment group had a 90-second extension as opposed to the control, which had *Pantbanken*'s regular 30-second extension. Potential bidders did not know that there were two different extension rules [1]. Therefore, bidders could not know that the extension would differ from 30 seconds until an extension was triggered. When this happened, the auction's closing time for that item would change to the new deadline and only then could bidders know that something was different. The experiment ran for nine days when 2292 items finished their auction, from the 31st of January to the 8th of February 2023. Of all items, 1149 were randomised into the treatment and 1143 into the control group.

One thing to point out with the estimates of the experiment is the possibility of biases in the estimates. With the randomisation, there should be no differences on average between the control and treatment groups regarding the characteristics of the items up for auction, and potential differences were tested with t-tests and two-sample tests of proportions to ensure there were no issues with the randomisation. It should also be pointed out that if there is any bias in *Pantbanken*'s best approximation of the market value of items, this bias should not affect the treatment and control group diversely, this bias would be the same across both groups due to the randomisation. Another potential bias arises because only some items in the treatment group received the treatment, the longer extension. Items were unsold, or bidding stopped before an extension was triggered. If not every item in the treatment group receives treatment, it leads to attenuation bias driving the treatment effect estimates towards zero since it is a measurement error (Angrist & Pischke, 2009). However, in this case, bidders were unaware of the experiment a priori, meaning that the longer extensions could only affect their bidding behaviour

[1] This was a decision taken by Pantbanken and not a decided experimental design

after an extension was triggered. Therefore, looking at only extended items provides a reasonable estimate of the average treatment effect since it is unaffected by attenuation bias. I.e., it is a reasonable estimate of the effect of the longer extensions compared to the status quo in the absence of equilibrium effects.

3.2.1 Experimental Data

The data was collected from *Pantbanken*'s website using a Python web crawler because *Pantbanken* does not collect all the information needed for this study. Their data collection focuses more on the closing price, the starting price and the number of bids in total. Therefore, additional information had to be gathered. The variables collected were the auction ids, bidder identifications, size and timing of all bids. The data on when the auctions ended, the office, item descriptions and their categories were also collected. *Pantbanken* categorises items into electronics, photography equipment, glass/porcelain, watches, art/crafts, silver items, nine subcategories of jewellery and other items. The subcategories consist of rings, necklaces and amongst other types of jewellery.

Table 3.1: Summary Statistics of Variables

Variable	**Obs**	**Mean**	**SD**	**Min**	**Max**
Starting Price	2,292	2330	4330	1	79000
Sold	2,292	0.542	0.498	0	1
Closing Price	1,239	2110	4530	1	79000
Number of Bids	2,292	2.533	3.614	1	44
Unique Bidders	2,292	1.551	1.041	1	9
Gold	2,292	0.683	0.465	0	1

Some summary statistics of the sample can be found in table 3.1. The average starting price in the sample was approximately 2300 SEK. However, as can be seen in the range of the data, there is much variation in the prices. In the sample, 54.2 percent of the items were sold. The average closing price was 2100 SEK. One thing to note is that the minimum starting price of 1 SEK, some of these were dropped from the sample if their ratios between closing and starting price became outliers in the data set. Likely, these items that were sold came as a small part of items bunched together in a loan without contributing to the loan value since *Pantbanken* does not give out loans for these low amounts. Many of the items with low prices in the sample are silver items. Since silver is valued at 3.5 SEK/gram, several pieces of jewellery made out of silver are commonly placed as a single loan but separated and sold as a single item when they go to auction (Pantbanken Sverige, 2023a). That the average close price was lower than the average starting price is explained by a higher percentage of sold items of cheaper items than the most expensive ones in the sample. When looking at the sold items, the average number of bids for each item was approximately 2.5. However, this also shows a considerable variation ranging from only a single bid to 44. The unique number of bidders for each item in the auction was an average of 2 bidders, with a range from one to nine unique bidders. To get a variable for gold items, one thing that had to be done in handling the data was to text-mine the descriptions for the carat content of items. *Pantbanken* put the carat of items in the description, ranging from 14-24 carats, and not as a separate category. As shown in table 3.1, 68.3 percent of the sample was gold items—consequently, the main part of the sample categories comprised different kinds of jewellery.

3.3 Empirical Methods

3.3.1 Differences in Revenue

To get the point estimates of revenue for the treatment and control groups and to see if they differ in the ratio between closing and starting price, t-tests for two sample means were conducted on the data from the experiment. As described by Cortinhas & Black (2012), this test is used when the population variance is unknown, and hence the sample variance is used instead. The test is done assuming that the two samples come from populations where the variances are equal. The method is used to test the null hypothesis:

$$H_0 : \mu_0 - \mu_1 = 0 \tag{3.1}$$

against the alternative hypothesis that the difference in means is different from zero. It can also be used for one-sided tests, where one mean is larger or smaller than the other. The t-statistic used to test the hypothesis is computed as follows:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2(n_1-1)+s_2^2(n_2-1)}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{3.2}$$

Where $\bar{x}_k$ is the estimated sample means, μ_k the hypothesised population means, s_k the sample standard deviations and n_k the sample sizes. The t-statistic follows a probability distribution called the student's t-distribution when the underlying population parameters are normally distributed. It is obvious to see that the second term in the numerator of eq. 3.2 becomes zero if the hypothesised difference in population means is zero. The degrees of freedom $df = n_1 + n_2 - 2$, in combination with the before-decided significance level, defines the critical value for the test. To reject the null hypothesis, the observed t-statistic must be larger than the critical value in absolute terms (Cortinhas & Black, 2012).

A critical assumption for t-tests is that the variable studied is normally distributed. This assumption is usually challenging to prove, but the Central Limit Theorem often mitigates this. It states that a variable's mean in a large enough, identically and independently distributed sample is approximately normally distributed for the sampling distribution. If the population is normally distributed from the start, the sampling distribution of means will be normally distributed regardless of sample size. The Central Limit Theorem provides reasonable approximations if the sample consists of at least 40 observations, regardless of the shape of the distribution in the population and as the sample size increases towards infinity, the theorem becomes stronger (Welkowitz, 2012).

3.3.2 Comparing Proportions

A similar test to the t-test will be used when testing differences in proportions. Examples of these proportions tested are proportions of gold items in each group, and to test for differences in probabilities of further bidding after a first extension etc. The t-test does not apply to proportions since they are based on dummy variables in the sample. Therefore, another test is needed and the technical details are as follows. For proportions, the Central limit theorem states that for large samples where $n_1 \cdot \hat{p}_1, n_1 \cdot \hat{q}_1, n_2 \cdot \hat{p}_2$ and $n_2 \cdot \hat{q}_2 > 5$ with $\hat{q} = 1 - \hat{p}$, then differences in sample proportions are approximately normally distributed with a mean difference of $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2$ and the standard deviation of differences in sample proportions is $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}}$. Where $\hat{p}_k$ are estimated sample proportions and n_k are sample sizes. The test statistic used to test for significant differences is a z-statistic, and it is computed as:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{(\hat{p} \cdot \hat{q})(\frac{1}{n_1} + \frac{1}{n_2})}}. \tag{3.3}$$

Where $\hat{p} = \frac{x_1+x_2}{n_1+n_2} = \frac{n_1 \cdot \hat{p_1} + n_2 \cdot \hat{p_2}}{n_1+n_2}$ and $\hat{q} = 1 - \hat{p}$. x_k is the number of observed frequencies of the variable of interest. The denominator in eq. 3.3 uses a weighted average of the sample proportions, the estimated proportions, and sample sizes to compute the standard deviations in the difference in sample proportions. The z-statistic is used for the same type of hypothesis testing as the T-test (Cortinhas & Black, 2012).

The t-test and two-sample tests of proportions will be used to assess the validity of the randomisation. To ensure that factors that should not have been affected by the treatment do not differ between the control and treatment groups. It should also be noted that in the results, the sign of the estimated difference has been reversed to simplify interpretation.

3.3.3 Comparing Distributions of Unique Bidders

To examine whether the number of unique bidders for the extended items comes from the same distribution for the treatment and control group, an Epps-Singleton two-sample test was used, as a method to see if longer extensions brought in new bidders. The test uses the empirical characteristics functions $\phi_1(t)$ and $\phi_2(t)$. Usually, similar tests are done with the Kolmogorov-Smirnov test, but in this case, it is not applicable since the data it will be used with is discrete and not continuous. Similar to the Wilcoxon Rank-Sum test, the test's null hypothesis is that the two samples come from the same distribution. It uses the characteristics functions for each sample to test for differences, not their observed distribution functions, which is a transformation of the latter. It is a less intuitive method than a probability density function to derive the moments of the function. While it has the advantage that it can be used on discrete variables for which the empirical characteristics function is completely defined. In contrast, the distribution function is only defined in certain points. The Epps-Singleton has the condition that all observations are independent within and across samples when testing the null hypothesis that $\phi_1(t) \neq \phi_2(t)$. These empirical characteristic functions are defined in equation 3.4.

$$\phi_{n_k}(t) = n_k^{-1} \sum_{m=1}^{n_k} e^{itX_{km}} \tag{3.4}$$

Where the parameters $t_1, t_2, ...t_j$, which Epps and Singleton calibrated to be $t_1 = 0.4$ and $t_2 = 0.8(J)$. The other parts of the model are $i = \sqrt{-1}$, for a sample k with n_k observations X_{km} is the mth observation in the sample. The t_j is then standardized by the estimated scale $\hat{\sigma}$, which Epps and Singleton suggested being the interquartile range, i.e., the statistical dispersion around the median of the data. Therefore, the test is done with $\tilde{t}_j = \frac{t_j}{\hat{\sigma}}, j = 1{,}2$. The method then creates 4x1 vectors out of all observations in the following way:

$$g(X_{km}) = (cos\, t_1 X_{km}, sin\, t_1 X_{km}, cos\, t_2 X_{km}, sin\, t_2 X_{km})' \tag{3.5}$$

Then g_k is both the real and imaginary parts of the characteristics functions for both t_1 and t_2 summed up as:

$$g_k = n_k^{-1} \sum_{m=1}^{n_k} g(X_{km}) \tag{3.6}$$

The difference between these two characteristic functions is then the difference between the two vectors: $G_2 = g_1 - g_2$. If the null hypothesis is true, then $\sqrt{n_1 + n_2}\, G_2$ would be asymptotically distributed as multivariate $N(\overrightarrow{0}, \Omega)$. The estimator for this covariance matrix is: $\hat{\omega} = \frac{1}{\nu_1}\hat{S}_1 + \frac{1}{\nu_2}\hat{S}_2$ where $\nu_k = \frac{n_k}{n_1+n_2}$, i.e., it is the k share of the total sample. And $\hat{S}_k = \frac{n_k - 1}{n_k} Cov\, g(X_{km})$ is the covariance matrix of that sample. Then the generalised inverse of $\hat{\omega}$ is used to compute the test statistic: $W_2 = (n_1 + n_2)G_2' \cdot \hat{\Omega}^{+} \cdot G_2$.

The test statistic W_2 is distributed asymptotically as chi-squared with r ranks of freedom. The ranks of the generalised inverse of the covariance matrix give the ranks of freedom. W_2 is a measure that tells us how different the empirical characteristic functions of two samples are when we consider the variance and covariance between them (Goerg & Kaiser, 2009).

3.3.4 Comparing Distributions of Revenue

The Wilcoxon Rank-Sum Test was used to evaluate the effect of longer extensions on auction profitability, providing a non-parametric alternative to the t-tests. It was used to examine any differences in the distribution of ratios between closing and starting prices for items sold between the treatment and control groups, serving as a measure of revenue.

The method and its technicalities are described by Mann & Whitney (1947). It tests the variable x and y and their continuous cumulative distributions f and g to test whether f=g. The test statistic, rank-sum, which is based on the relative ranks of x and y, is used with ties in the data average ranks. The Wilcoxon Rank-Sum is computed first as the sum of all ranks in the first group, in this case for y:

$$T = \sum_{i=1}^{n} R_{1i}. \tag{3.7}$$

Under the null hypothesis that the two distributions of a sample size of n number from variable x and a sample size m from variable y are the same, it uses a recurrence relation to compute the probabilities of finding the observed U, given how many times n can be drawn from n+m observations. The authors also propose that the recurrence relation can be used to compute the mean, variance and fourth moment of U. The fourth moment has then been used to prove that as n and m goes towards infinity, the distribution of rank sums approaches a normal distribution. Therefore it is used to compute the expected rank-sum and variance with the following formulas:

$$E(T) = \frac{n((n+m)-1)}{2} \tag{3.8}$$

and

$$Var(T) = \frac{nms^2}{n+m}. \tag{3.9}$$

This is then used in the same way as would be done in a z-test, where the test statistic and p-values are given by:

$$z = \frac{T - E(T)}{Var(T)} \tag{3.10}$$

As briefly mentioned before it is also a non-parametric test meaning that assumptions of the shape of the underlying data are unnecessary due to the normality of the rank-sums. It can also be used as a one-sided test to test whether f > g and vice versa.

3.4 Limitations

The original experimental design that was proposed to *Pantbanken* had more than one treatment group, where the extension would differ between the treatment groups. The original proposal was to have the control group as it is and then have at least two treatment groups with three and ten-minute extensions. This was not doable with their current live auction website since it was deemed that even longer extensions would interfere with their business. Therefore, the 90-second extension was agreed upon as doable for both parties. One disadvantage of only adding an extra minute to the current extension could be that a longer extension is needed to affect the bidders' behaviour.

Another point to make is that with more than one treatment group, the optimal length of an extension to an online auction could be examined where this experiment can only compare the effect of the increased extension to the status quo. The number of observations also poses a limitation to this study. That roughly 50 percent of the items up for auction at *Pantbanken* are sold was to be expected, but not every sold item was extended. Since an extension only occurred if a bid was placed within 30 seconds of the deadline. The experiment only lasted nine days, meaning statistical power is lost when sub-sampling to only the extended items. Not having more observations limits the possibility of a more in-depth study of whether there were heterogeneous effects of a change in the extension to items with different starting prices and types, for example. The possibility of equilibrium effects could also be an interesting topic to examine if the experiment would be conducted over a much extended period. To see if the behaviour of bidders changes as they learn of the longer extensions. One could then examine differences in the first and second periods of the experiment. As mentioned in section 2.1, Ariely et al. (2005) noted that bidders changed their behaviour when auctions were repeated over time. In this experiment, no data on the bidders' experience was available due to their integrity, which means there was no information on whether they were first-time participants in *Pantbanken*'s auctions or experienced.

Chapter 4

Empirical Results

4.1 Quality of Randomisation

In the sample of 2292 items, 1239 of them were sold, approximately 54 percent. Of the sold items, 628 belonged to the treatment group and 614 to the control group. To ensure that the randomisation was done correctly, t-tests for differences in means in starting price and two-sample tests of proportions were done on twelve categorical variables. These tests were done on the complete sample, both sold and unsold items. The variables were the four most common product categories, as Pantbanken categorised, offices, starting price, whether it was sold or not, and the most relevant factor, whether the auction was extended or not. The results of the tests on the most important variables can be seen in table 4.1. The rest of the results can be found in appendix A.1 The first row is a t-test, and the rest are the tests of proportions. It should also be noted that standard deviations are reported for the t-test estimates, but it is the standard error for the proportions. As can be seen in the table in appendix A.1, approximately 68.3 percent of the items were made out of gold. Silver items made up six percent of the sample. Since most items in the sample consisted of precious metals, it was logical that the most common product categories comprised different kinds of jewellery, such as rings, pendants, necklaces, and bracelets. As expected for a correctly conducted randomly controlled experiment, the tests for the differences in means and proportions between the treatment and control groups could not find statistically significant differences for the thirteen variables. Showing that there were no differences between the two groups in the starting prices, which office the items belonged to, or what type of goods they were. No significant differences indicated that the randomisation worked without problems and that the experiment had balanced groups. That there were no differences in characteristics means that the only thing that should cause differences between the two groups should be the longer extensions in the treatment group for the items with bids placed in the last 30 seconds of auctions and the subsequent bidding.

Table 4.1: Balance table

	Control		Treatment			t-test	
Variable	Mean	SD/SE	Mean	SD/SE	Diff.	t-stat/z-stat	p-value
Starting price	2240	4170	2400	4480	160	(-0.92)	0.35
Sold	0.537	0.015	0.547	0.015	0.009	(-0.451)	0.652
Extended	0.041	0.006	0.043	0.006	0.002	(-0.284)	0.776
Observations	1143		1149		2292		

The variable extended is important to note in table 4.1. Only around four percent of the items up for auctions were extended beyond their initial deadline in the two groups. When the sample is restricted to the 1239 sold items, only 97 were extended. I.e., only 7.8 percent of the sold items were exposed to the actual treatment. Therefore, statistical power is lost since only a small part of the sample is actually treated.

Therefore, some analysis has been restricted to the sub-sample where actual treatment took place to examine the effect of the longer extensions. Differences over the whole sample of the sold goods will be evaluated to find the intention to treat.

4.2 Number of Unique Bidders

The Cumulative distributions function plotted for the treatment and control groups in the extended sample to examine if the longer extension brought more unique bidders to the auction. The plots can be seen in figure 4.1. It shows that there is a slightly higher concentration of items with only one bidder in the control group, i.e., that bidders placed a bid in the last 30 seconds of the auction but that there were no subsequent bids. Otherwise, the distribution is almost identical for both groups. Still, the treatment group has extended items with eight and nine unique bidders, which was not the case in the control groups.

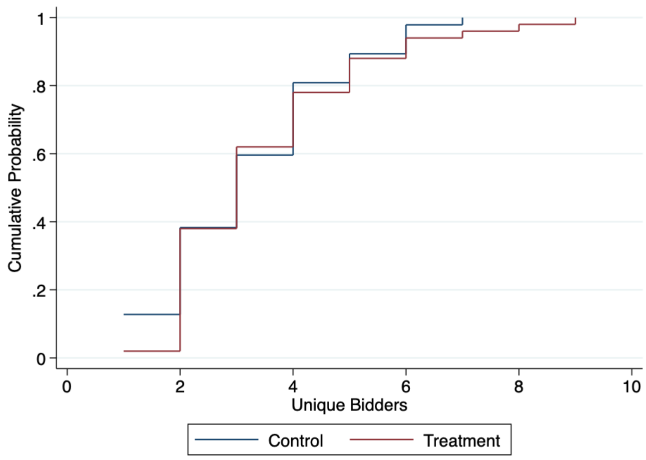

Figure 4.1: CDF of Unique Bidders by Treatment

The Epps-Singleton two-sample test using the empirical characteristic function was conducted to test whether the distribution of the unique number of bidders differed between the two groups. In table 4.2, the results are shown. The test statistic of $W_2 = 6.202$ was found, and with a p-value of 0.185, it was insignificant. Therefore, the test showed no significant proof of new bidders joining in the extended time.

Table 4.2: Epps-Singleton Two-Sample Empirical Characteristic Function test

Group	Observations	Test statistic	p-value
Control	47		
Treatment	50		
Total	97	6.202	0.185

4.3 Seller's Revenue

A ratio between the closing price and the starting price was used to see if there was an impact on the seller's revenue. Using the ratio normalised the prices and simplified the comparison. Two tests were used to test the effects of the longer extension on the ratio. Firstly t-tests to test the difference in means between the groups and the Wilcoxon Rank-Sum Test to see if there are differences in the distribution of the ratios. For both tests, the sample was split into different subgroups. First, to test the whole sample of sold items, both extended and the ones that were not, this was further split into gold and non-gold items. Since only 7.8 percent of the sold items were treated, the same tests and sub-sampling were done for the extended items. Figure 4.2 shows a histogram of the frequency distribution of ratios for the extended items in the treatment and control group. Visual inspection shows a higher frequency of observations around the lower bound of ratios just above one in the control group. Still, it also contains some of the higher values. For the treatment group, more frequent observations are found above the lower bound of the values. In the sample of extended items, one observation had a ratio above ten, which was deemed an outlier. Therefore, such observations have been excluded from the sample when doing the t-tests. Looking at the whole sample, extended and not, only two observations were above ten. A more in-depth description of this can be found in appendix A.2. These observations have not been excluded from the sample in the non-parametric tests. As single observations, they are not skewing the results as much when the size of the observation is not used when it comes to the distribution of the variable etc.

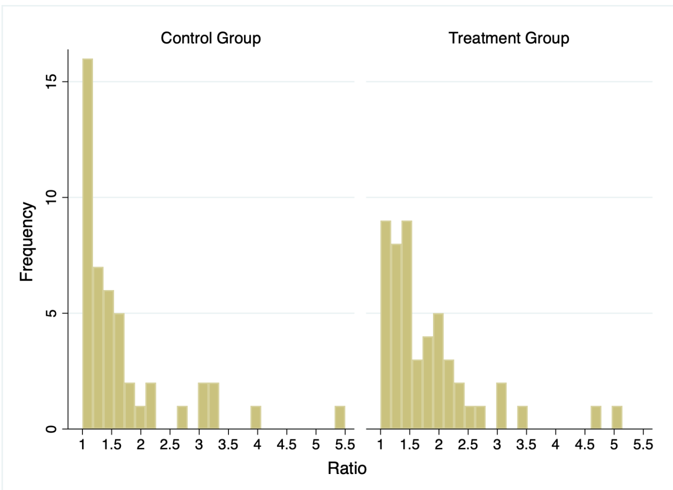

Figure 4.2: Frequency Distribution of Ratios

4.3.1 t-Tests of Ratios

The t-tests were done against the null hypothesis that there were no differences in the ratios between treatment and control. The results are in table 4.3. The first column shows the results over the whole sample of sold items. The items in the treatment group were, on average, sold for approximately 1.5 times their starting price. This mean was slightly higher than the control group, but it is only 1.3 percent higher, and the p-value of 0.764 shows that the null cannot be rejected. The second column shows the result of all gold items. On average, they were sold at lower ratios for both the control and treatment groups compared to the whole sample. These are just above 1.1 times the starting price. The lower ratio could be explained by the items being closer to common value goods, making it easier to value gold because it has a more defined market value than other items. The standard deviations were also lower, indicating less variation between starting and closing prices. Here, the mean for the control group was three percent higher than in the treatment. This result was still insignificant at a 10 percent significant level. However, when looking at the p-values for the one-sided test, i.e., half of the two-sided test, the p-value of 0.060 indicated that the control ratio was higher than for the treatment group. In the third column, a t-test has been carried out over all

Table 4.3: t-Tests Results: Testing the variable Ratio

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Mean (Treatment)	1.482	1.114	1.924	1.818	1.488	2.122
Std. Dev. (Treatment)	1.050	0.260	1.482	0.840	0.439	1.002
Obs (Treatment)	624	341	283	50	24	26
Mean (Control)	1.463	1.148	1.877	1.693	1.449	2.040
Std. Dev. (Control)	1.093	0.309	1.458	0.926	0.547	1.222
Obs (Control)	613	348	265	46	27	19
Estimated Diff. between Means	0.018	-0.034	0.047	0.125	0.039	0.083
Total Obs	1,237	689	548	96	51	45
t-Statistic	0.301	-1.555	0.374	0.694	0.280	0.250
p-Value	0.764	0.121	0.708	0.489	0.781	0.804
Sample	Whole	Gold	Non-gold	Whole	Gold	Non-gold
Only extended	No	No	No	Yes	Yes	Yes

Note: Tests 4-6 contain only extended observations

non-gold items. Here, the mean ratio of the treatment group was 1.92 times the starting price, which was higher than in the control group. The difference was only

2.5 percent, and neither this test could reject the null that the mean ratio differs between the groups with a p-value of 0.708.

In the last three columns, the t-tests were conducted on further restricted samples, i.e., on only those items where the treatment occurred. The sub-sampling otherwise followed the same pattern as columns one to three. The point estimates of the fourth test showed that the items were sold at slightly above 1.8 times the starting price. Again the treatment was sold at a slightly higher price than the control. The mean in the treatment group was 7.4 percent higher than in the control group. The p-value of the test was 0.489, meaning it could not reject the null. Looking at only the extended gold items the point estimate of the treatment group was slightly higher than the control. The mean of the treatment group was 2.7 percent higher than the control. Neither this test could reject the null with a p-value of 0.781. In the last test, when only extended non-gold items were included, the point estimate was four percent higher in the treatment group compared to the control. However, the null could not be rejected with a p-value of 0.804, i.e., that there were no statistically significant differences between the groups.

In both cases, when looking at the whole treatment and control groups and only the extended observations, they followed the same pattern. The samples restricted to only gold items were sold at a lower ratio than the whole sample. In comparison, non-gold items were sold at a higher ratio and had a larger deviation in the means of the ratios.

A sensitivity analysis of these tests was conducted when all ratios above ten were included, and the results can be found in appendix A.2. In short, the lack of significance of all six tests remained. Still, for test four and six the point estimates of the difference between the two groups indicated a negative effect on the revenue due to longer extensions. Therefore, it should be kept in mind that when including the outliers, it does not show the same positive effect of longer extensions for all tests as shown in table 4.3.

Table 4.4 shows the confidence intervals for the estimated differences between the two groups displayed in table 4.3. Looking at the upper bound of the confidence interval for test one, a hypothesised difference larger than 13.8 percent of the starting price would be needed to reject the null of no difference between the groups given the sample. Test four, which only examined the extended items, would need a difference larger than 48.3 percent of the starting price. The economic implications and further research suggestions will be discussed in section 4.5 .

Table 4.4: Confidence Intervals for t-Tests

Tests	**95% CI**
Test 1	0.138 -0.101
Test 2	0.009 -0.076
Test 3	0.294 -0.200
Test 4	0.483 -0.233
Test 5	0.321 0.242
Test 6	0.752 -0.586

4.3.2 Distribution of Ratios

The next thing that was tested was to check for differences in the distribution of ratios between the groups, and therefore Wilcoxon Rank-Sum Test was used. The sub-sampling and the results are displayed and were done in the same way as for the t-tests. This was done as another way of testing effects on revenue but with a test not requiring normality of the variables tested. The results are displayed in table 4.5. Of all the tests when testing the null hypothesis that the two groups

Table 4.5: Results: Wilcoxon Rank-Sum Test

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Treatment Rank Sum	378601	111596	78090	2667	686	644
Treatment Obs	625	341	284	50	24	26
Control Rank Sum	389579	126108	73434	2086	640	437
Control Obs	614	348	266	47	27	20
Combined Sum	768180	237705	151525	4753	1326	1081
Combined Obs	1239	689	550	97	51	46
z-Statistic	1.476	2.524	0.083	-1.567	-1.171	-0.732
p-Value	0.140	0.012	0.934	0.118	0.246	0.471
Sample	Whole	Gold	Non-gold	Whole	Gold	Non-gold
Only extended	No	No	No	Yes	Yes	Yes

Note: Tests 4-6 contain only extended observations. For these, exact p-values were computed due to n<200.

follow the same distribution, only test two showed a p-value of 0.012. Therefore, at a five percent significance level, the tests suggested that the distribution of ratios for gold items in the control was higher than in the treatment because the rank-sum was higher than in the treatment. For test one, the right-sided p-value was 0.070, half of the two-sided p-value displayed in the table. It indicates that the distribution of ratios in the control was larger than in the treatment when looking at the whole sample. Test four showed the opposite; a negative test statistic and a left-sided p-value of 0.059 suggests that the distribution of ratios was higher in the treatment, i.e., a suggestion that the treatment shifted the distribution of ratios to a higher one. For the rest of the tables, nothing could be statistically concluded about a difference in distributions between the two groups. Therefore, the data do not indicate any clear shifts in the distribution of the ratios due to the different extensions. The implications of tests two and four are further discussed in section 4.5

4.3.3 Heterogeneity

Wilcoxon Rank-Sum Tests were done again but for new sub-samples to test whether the treatment had a heterogeneous effect. Therefore, a new variable was created indicating whether an item belonged to one of the most frequently sold categories in the sample, to see if there was any differences in revenue for the items that were most frequently sold. During the data-generating process, the three most frequently sold item categories were rings, pendants and bracelets. These are

Table 4.6: Heterogeneity Results: Wilcoxon Ranked Sum Test

	Test 1	Test 2	Test 3	Test 4
Treatment Rank Sum	116236	75398	693	650
Treatment Obs	349	276	24	26
Control Rank Sum	116666	80005	582	477
Control Obs	333	281	26	21
Combined Sum	232903	155403	1275	1128
Combined Obs	682	557	50	47
Z-Statistic	1.180	0.905	-1.574	-0.568
p-Value	0.238	0.365	0.117	0.577
Common good	No	Yes	No	Yes
Only extended	No	No	Yes	Yes

Note: Common good denotes the three most common categories, otherwise the rest.

henceforth called common goods. The data was thus split into common and non-common goods based on this definition. Therefore the tests were done on four different sub-samples. Common goods and the rest in the complete sample, and then sub-sampled only for the extended items. The results are in table 4.6. Similarly to the tests in table 4.5, not many could reject the null. Over the whole sample, there were no signs of significant differences in the distributions in the two groups. The only sign of a difference comes from test three, looking at the sample with only extended items where the three most common goods were excluded. Here a negative z-statistic is found, and the left-side p-value is 0.0586. Implying that the distribution is larger for the treatment group than the control.

4.4 Effects on Bidding Behaviour

After looking at the more revenue-focused part of the experiment, the next test was more focused on the buyers' bidding behaviour due to the treatment after an extension had been triggered. Therefore, the focus was only on the items that have been extended, i.e., when at least one bid came in the last 30 seconds of an auction. Therefore, a new variable was created. This dummy variable showed whether at least one more bid was submitted after the first extension. Due to the low number of observations where an extension was triggered in the first place, the possibilities of sub-sampling and heterogeneity tests were limited. Two sample tests of proportions were then conducted on the proportions of items where at least a second bid was submitted in the treatment and control groups. As before, the sample was also split into gold and non-gold items.

Table 4.7 shows the results of the three tests. For all extended items, the proportion was 42.6 percent for the control. At the same time, it was 60 percent in the treatment, implying a 40.8 percent higher probability of a second bid being submitted than in the control group. This difference in means was significant at a 10 percent significance level with a p-value of 0.086. That the proportion in the treatment was larger than the control was significant at a five percent significance level. The same shift in proportions was found for the extended gold items. Here, the two-sided p-value was 0.037, i.e., it was significant at a five percent significance level. The estimated difference was even more prominent here. The increase in the probability of a second extension was 87.6 percent. Test three could not reject the null of a difference in proportions; the point estimate of the proportion in the treatment group was slightly higher than the control. Still, it was only 2.7 percentage points, less than a five percent probability difference. The tests of the proportions of another bid being submitted indicated an increase in the probability that at least one more bid would come in due to the longer extension. This increase was true for the whole sample and when looking at only gold items, but no statistical conclusion can be drawn on a change in bidding behaviour regarding non-gold items. This change in probabilities of another bid due to the longer extension in the whole sample seems to be driven by the gold items. This aligns with the earlier mentioned theory that late bidding is more frequent with common value goods where bidders want to hide their private information from competitors.

Table 4.7: Two-Sample Test of Proportions: Second extension

	Test 1	Test 2	Test 3
Mean (Treatment)	0.600	0.625	0.577
Std. Err. (Treatment)	0.069	0.099	0.097
Obs (Treatment)	50	24	26
Mean (Control)	0.426	0.333	0.550
Std. Err. (Control)	0.072	0.091	0.111
Obs (Control)	47	27	20
Estimated Diff. between Means	0.174	0.292	0.027
Total Obs	97	51	46
z-Statistic	-1.718	-2.083	-0.183
p-Value	0.086	0.037	0.855
Whole sample of extended	Yes	No	No
Gold	No	Yes	No

Note: Test 2 and 3 restrict the sample to gold and non-gold items.

4.5 Discussion

The first thing to mention about this experiment is that the balanced table 4.1 showed no signs of improper randomisation. The characteristics of the items up for auction in the two groups did not differ between them. The two groups were, on average, the same for characteristics that should not differ before treatment happened, i.e., the groups were balanced.

A discussion of the effects of changing the length of extensions needed to be added to the literature and to which this experiment contributes. As mentioned before by Boatwright et al. (2010), it is described that bidder awareness of auctions is random. The number of bidders partaking in auctions that were not extended was outside the scope of this thesis, but examining the distribution of bidders in those that received actual treatment was of interest. To see whether new bidders partook due to the longer extension or left, for that matter, due to higher monitoring costs. As shown in figure 4.1, some items had eight and nine unique bidders in the treatment group. This number was not observed in the control group. The Epps-Singleton test for these discrete distributions could not find any significant difference. I.e., there was statistically significant proof of the longer extension bringing in new bidders. Therefore, the longer extension gave already participating bidders a bigger chance to respond to developments in the auctions. It should be noted that this does not necessarily mean that new bidders do not join the auction due to longer extensions. It could simply be that the extra 60 seconds was not enough to make a statistical difference in the awareness of new bidders.

When focusing on how extensions affect revenue, the effects of the longer extensions are ambiguous. As discussed in section 2.1, the differences between hard-close and soft-close auctions have been widely studied. Late bidding for common value items is a dominant strategy not to reveal information about the item or not to get into bidding wars with incremental bidders (Roth & Ockenfels, 2002). Ariely et al. (2005) and Glover & Raviv (2012) find that extensions lead to higher revenue and efficiency instead of fixed deadlines. While Cao et al. (2019) argues that extensions do not mitigate late bidding due to the monitoring costs.

These effects of a longer extension on the revenue could not be statistically proven by the data in any direction. For the whole sample and when looking at only the extended items, no difference in means could be proven by the t-tests. Neither could any statistical difference between the two groups be shown when looking at the non-gold items. For the sample of every sold gold item, there was a weak statistical significance, using the right-sided p-value, implying an average treatment effect of the longer extension that lowered the revenue. However, this effect was not observed when looking at only the extended gold items. Although there was a lack of significance, the point estimates of all three tests on the extended items indicated a positive effect on the revenue due to longer extensions. For the aggregate test, the point estimate of the increase in revenue was 7.4 percent.

When using the non-parametric Wilcoxon Rank-Sum Test to look for different distributions between the two groups, the two-sided p-values could not show any differences in distributions for the whole sample. However, it should be noted that when looking at the right-sided p-value, at a ten percent significance level, there was an indication that the distribution in control was higher than in the treatment. Nevertheless, when looking at the sample of all extended items, there was a sign of weak significance that the distribution was higher for the treatment—implying a positive revenue effect of the longer extensions. Similar significant results as in the t-tests came from the sample of every gold item, which was significant at the five percent level. The observed rank-sum was higher for the control, suggesting that the rank of the ratios was higher and from a different distribution. Similarly to the results of the t-tests, the same effect was not shown when examining the distribution of only the extended gold items.

The results indicate that gold items, on average, were sold at higher prices in the control than in the treatment group when evaluating both the mean ratio and distribution. If the longer extension brought a lower price for gold items, one would expect that the treatment effect on the treated, i.e., the sample of only extended gold items, would be larger than the average treatment effect due to attenuation bias. In other words, it would be expected that the negative effect on revenue would be even more prominent for the extended items. Since this was not the case, on the contrary, the point estimates showed a higher mean ratio and a higher distribution for treated gold items, although insignificant. Therefore, random variation in the data likely explains the negative effect shown for the whole sample of gold items. Since the bidders did not know of the changed extension, it is not likely that the longer extensions lowered the prices without extensions taking place.

Heterogeneity in the effects of how the extensions affected the distributions in the two groups was also tested, where few signs of the distributions being different were shown. Only when excluding the three most common items up for auction (rings, pendants and bracelets) could the left-sided p-value show a weak sign of significance, but only at the ten percent level. These tests could indicate that the categorisation done so far has been naive and that there was too much heterogeneity between the items for any apparent effects of the extension to be found. However, at the same time, it should be noted that this test was only done over 50 observations, so much statistical power was lost. Therefore, this is a topic for future research to examine with more observations.

Although this paper cannot find clear evidence of increased revenue due to longer extensions, the point estimates of the extended items were positive for all items. Additionally, the 95 percent confidence intervals of the difference in ratios between the treatment and control are open for further research. When looking at the whole treatment and control sample, the estimated confidence interval ranges from a -10.1 to 13.8 percent difference in ratios between closing and starting price. If this experiment is repeated many times, it can be said with 95 percent confidence that the true population difference would lie in these intervals. For example, this does not rule out the possibility of a ten percent increase in revenue due to the longer extension, although that is not statistically proven. The possibility of such a difference could still be economically meaningful, making it commercially attractive for firms to explore further. In the sub-sample of only extended items, this confidence interval was wider due to fewer observations but still ranging from -23.3 to 48.3 percent, which means that the difference for the extended items could be quite large. Therefore, an auction firm which observes frequent late bidding and extensions to their auctions, possibly increasing the length of the extensions, could be highly relevant. In this sample, only 7.8 percent of the sold items were extended, but the average starting price was approximately 2300 SEK. Suppose the true population difference in the means is the 7.4 percent increase in the price ratio due to a 60-second longer extension. Which is the point estimate of the difference when looking at all extended items. Then, on average, Pantbanken could earn roughly 170 SEK extra on each extended item. At the same time, this argumentation should be taken cautiously since the confidence intervals do not rule out the possibility of a loss in revenue. As argued by Cao et al. (2019), extensions leads to monitoring costs to bidders, therefore increasing the length of the extensions could deter bidders with higher alternative costs. Thus resulting in a revenue loss if those bidders have the highest valuations.

The effects of longer extensions on bidding behaviour in the extended auctions have been examined in addition to the potential revenue gains. The two-sample test of proportions for the whole sample of extended items was significant at a ten percent significance level. When looking at the left-sided p-value, it was significant at a five percent level. This estimated proportion difference showed a 40.8 percent increase in the probability of at least one more bid coming in after the first extension was triggered. No significant difference could be found when the same test was used on the sub-sample of all non-gold items. In contrast, the estimated effect was strongly significant for the gold items with a p-value of 0.037. The treatment group's probability of at least an additional bid was almost twice as large. The proportion was 87.6 percent larger for the gold items in the treatment group. Suppose gold items are appropriately classified as common value goods. As argued by Roth & Ockenfels (2002), Dang et al. (2015) and Ely & Hossain (2009), it is a dominant strategy for experienced bidders to bid late on common value goods not to reveal information to competitors. Therefore, these results indicate that prolonging the extension could mitigate potential revenue losses of late bidding since it can show that the extra 60 seconds increases the probability of further bids after the first extension. Bidders had more time to respond to developments in the auctions they were partaking in.

The analysis results focusing on the revenue and efficiency aspects of the auctions, although positive, were statistically insignificant. However, when examining the effects on further bids after the first extension, there were strong indications that the longer extension brought more bids to the auction. This increased probability of additional bids, combined with the confidence intervals of the differences in ratios of closing and starting prices, does not rule out economically meaningful effects of longer extensions, especially for gold items. This should encourage future research, not only for the academia but also out of commercial interest for *Pantbanken* and actors selling the same type of goods with similar auctions. As mentioned, a limitation of this experiment was the number of observations. The attenuation bias could be why no revenue results have been found for the whole sample since less than ten percent of the observations were extended. The low number of treated observations also affects the statistical power of the positive point estimates. It limits the possibility of diving deeper into the data to examine if there are heterogeneous effects across the auction data. The experiment should be run over a more extended period for further studies. To have a larger sample of extended items to analyse further the consequences of a longer extension in a soft-close auction. Additionally, a similar experiment over an extended period would also be helpful to investigate if there are equilibrium effects due to changes in the extensions. Supposedly, bidders could change their behaviour as they learn the game's rules, as opposed to this experiment where they did not know that there were changes to the extensions.

Chapter 5

Conclusion

This study evaluated the effect on revenue and bidding behaviour of changing the extension length in online auctions through a randomised controlled experiment. This is a topic that has not been previously explored or studied within formal economic theory, particularly through experimental research. The study was conducted in collaboration with *Pantbanken Sverige*. This Swedish pawn shop sells defaulted loan securities on its auction platform. The experiment randomised items for sale into a treatment and control group. In the case of a bid in the last 30 seconds on items in the treatment group, the auction deadline was increased by 90 seconds compared to the control with an extension of only 30 seconds. The point estimates for extended items indicated an increase in revenue due to the longer extensions when examining changes in the closing price of the auctions, although they were not statistically significant. That directional effects were not statistically proven could be because a change in the extension in a soft-close auction does not affect the bidders' behaviour. However, these results cannot rule out the possibility of economically meaningful impacts of longer extensions. The confidence intervals of the ratios between the closing and starting price do not rule out the possibility of an increase in prices of ten percent. The potential increase in revenue, combined with the increased probabilities of further bids after the first extension, indicates potential profitability in increasing the extension length, especially for gold items. The longer extension increased the probability of at least one more bid by 40.8 percent. The increase was even more prominent for gold items, with a probability increase of 87.6 percent. Gold items are considered to have a stronger common value component, where longer extensions more strongly mitigate late bidding. Potential gains in revenue and consumer surplus should therefore be further examined, especially out of commercial interest. This study leaves room for future research on the role of extensions in online auctions, where more data could bring new insights into further heterogeneity, equilibrium effects and the magnitude of the revenue effects of changing the length of the extensions.

Bibliography

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Ariely, D., Ockenfels, A., & Roth, A. E. (2005). An experimental analysis of ending rules in internet auctions. *The Rand journal of economics*, 36(4), 890–907.

Boatwright, P., Borle, S., & Kadane, J. B. (2010). Common value vs. private value categories in online auctions: A distinction without a difference? *Decision Analysis*, 7(1), 86–98.

Cao, W., Sha, Q., Yao, Z., Gu, D., & Shao, X. (2019). Sniping in soft-close online auctions: empirical evidence from overstock. *Marketing Letters*, 30(2), 179–191.

Cortinhas, C. & Black, K. (2012). *Statistics for Business and Economics*. Wiley.

Dang, C., Hu, Q., & Liu, J. (2015). Bidding strategies in online auctions with different ending rules and value assumptions. *Electronic Commerce Research and Applications*, 14(2), 104–111.

Ely, J. C. & Hossain, T. (2009). Sniping and squatting in auction markets. *American economic journal. Microeconomics*, 1(2), 68–94.

Glover, B. & Raviv, Y. (2012). Revenue non-equivalence between auctions with soft and hard closing mechanisms: New evidence from yahoo! *Journal of Economic Behavior & Organization*, 81(1), 129–136.

Goerg, S. J. & Kaiser, J. (2009). Nonparametric testing of distributions—the epps–singleton two-sample test using the empirical characteristic function. *The Stata Journal*, 9(3), 454–465.

Krishna, V. (2009). *Auction Theory*. San Diego: Elsevier Science, 2nd edition.

Mann, H. B. & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50 – 60.

Ockenfels, A. & Roth, A. E. (2006). Late and multiple bidding in second price internet auctions: Theory and evidence concerning different rules for ending an auction. *Games and Economic behavior*, 55(2), 297–320.

Pantbanken Sverige (2023a). *Guldpriser*. https://www.pantbanken.se/lana/guldpriser/. Accessed: 2023-03-26.

Pantbanken Sverige (2023b). *Villkor*. https://www.pantbanken.se/om-oss/villkor/. Accessed: 2023-03-22.

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

Welkowitz, J. (2012). *Introductory statistics for the behavioral sciences*. John Wiley and Sons, Inc., 7th edition.

Appendix A

Appendix

A.1 Randomisation

Table A.1: Complete Balance table

	Control		Treatment			t-test	
Variable	Mean	SD/SE	Mean	SD/SE	Diff.	t-stat/z-stat	p-value
Starting price	2240	4170	2400	4480	160	(-0.92)	0.35
Sold	0.537	0.015	0.547	0.015	0.009	(-0.451)	0.652
Extended	0.041	0.006	0.043	0.006	0.002	(-0.284)	0.776
Gold	0.691	0.014	0.675	0.014	-0.016	(0.813)	0.416
Silver	0.056	0.007	0.064	0.007	0.008	(-0.761)	0.447
Ring	0.262	0.013	0.257	0.013	-0.006	(0.312)	0.755
Pendant	0.131	0.010	0.129	0.010	-0.002	(0.173)	0.863
Necklace	0.151	0.011	0.145	0.010	-0.006	(0.405)	0.686
Bracelet	0.118	0.010	0.113	0.010	-0.005	(0.372)	0.710
Karlstad	0.087	0.008	0.085	0.008	-0.001	(0.113)	0.910
Skanstull	0.088	0.008	0.091	0.008	0.002	(-0.180)	0.857
Farsta	0.094	0.009	0.095	0.009	0.000	(-0.031)	0.975
Solna	0.090	0.008	0.090	0.008	0.000	(0.039)	0.969
Observations	1143		1149		2292		

Complete balance table, the content is discussed in 4.1.

A.2 Excluded outliers and Sensitivity Analysis of t-tests

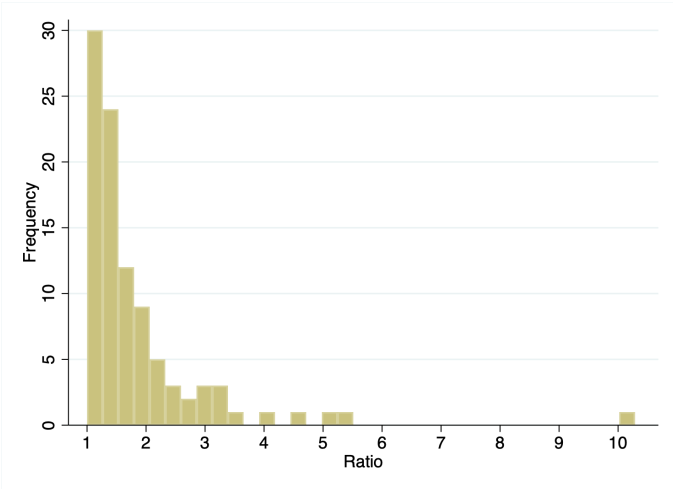

Figure A.1: Frequency Distribution of Ratios

Figure A.1 shows the distribution of ratios for all extended items in the original sample. The histogram shows the observations for both the treatment and control groups. The single observation with a ratio above ten seems like an anomaly by visual inspection. This observation was a collection of wine glasses, with a starting price of 350 SEK sold for 3600 SEK. As can be seen in the graph, the second largest ratio takes the value 5.5. When excluding all observations above ten, for both extended and non-extended, only two observations are that large in the whole sample. Meaning that the statistical power lost due to this is very low. Both of these observations belonged to the non-gold category. It should be mentioned that it is unclear where to draw the line of what is an unreasonable value for the ratio between starting and closing price. Still, in conversation with *Pantbanken*, they agreed it was a very unusual value. Therefore, it was decided to drop observations where the ratio was above ten. A sensitivity analysis can be found in table A.2, where these otherwise dropped observations were included.

In table A.2 the same tests as in table 4.3 but with all observations where the ratio was above ten were included. As can be seen in the table, only two observations in the whole sample had such a high ratio. These were non-gold items, out of which only one was treated, i.e., it was actually extended. Therefore, the only results that

Table A.2: t-Tests Results: Testing the variable Ratio

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Mean (Treatment)	1.500	1.114	1.963	1.818	1.488	2.122
Std. Dev. (Treatment)	1.185	0.260	1.619	0.840	0.439	1.002
Obs (Treatment)	625	341	284	50	24	26
Mean (Control)	1.478	1.148	1.909	1.876	1.449	2.451
Std. Dev. (Control)	1.108	0.308	1.544	1.552	0.547	2.194
Obs (Control)	614	348	266	47	27	20
Estimated Diff. between Means	0.022	-0.034	0.054	-0.058	0.039	-0.329
Total Obs	1239	689	550	97	51	46
t-Statistic	-0.343	1.555	-0.403	0.230	-0.280	0.680
p-Value	0.731	0.121	0.687	0.819	0.781	0.500
Sample	Whole	Gold	Non-gold	Whole	Gold	Non-gold
Only extended	No	No	No	Yes	Yes	Yes

were affected due to the excluded observations were test one, three, four and six. It did not have any affect on the significance levels of the test, i.e., no results that were insignificant before were significant due to the exclusion or vice versa. What should be noted though is that for tests four and six, the point estimates changed sign as compared to the results in table 4.3. Here those results indicate that extending the extension length lowers the revenue. For all extended items, this effect was 3.2 percent, and for the extended non-gold items, this was 15.5 percent. However, it is likely, that these negative effects should be disregarded, since out of approximately 1200 observations, only two observations had such a high ratio between the closing and starting price. It is simply these outliers skewing the point estimates.