Norwegian School of Economics

Bergen, Fall 2021

Adverse selection in iBuyer business models

A study of adverse selection in the use of automated valuation models for iBuyers

Arne Johan Pollestad & Eirik Helgaker

Supervisor: Are Oust

Master thesis, Economics and Business Administration

Major: Business Analytics and Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Preface

This thesis was written as a part of our Master of Science in Economics and Business Administration degree at Norwegian School of Economics, during the autumn semester of 2021. The thesis has a scope of 30 ECTS, and was written by Business Analytics major Arne Johan Pollestad and Financial Economics major Eirik Helgaker. All work is independent and original.

We want to express our gratitude to our supervisor, Associate Professor Are Oust, for essential guidance and constructive feedback during the entire research process. Furthermore, we would like to thank Ulf Jakob Flø Aarsnes, and the rest of the team in Solgt.no, for providing us with data and expertise within the field of AVMs, and wish them the best of luck with future development. We would also like to thank Assistant Professor Kyrre Kjellevold for inspiration and help with choosing a research field of interest. Lastly, we want to thank family and friends for their help and support along the way.

Norwegian School of Economics

Bergen, December 2021

Arne Johan Pollestad Eirik Helgaker

Abstract

The purpose of this thesis is to examine how adverse selection can affect the average resale profits for iBuyers, and how simple strategic purchasing rules can help limit this potential problem. The rise of instant buyer (iBuyer) businesses in the past years has made automated valuation models (AVMs) an important part of the property market. Acting as an intermediary between sellers and buyers, the iBuyers provide liquidity and convenience to the market. Although iBuyer services are in demand, large actors within the segment have reported dissatisfying profits over time.

In this thesis, hedonic sales, extreme gradient boosting, and support vector machine AVMs are first trained to predict apartment prices in Oslo, Norway. The dataset consists of 84,905 apartment transactions in Oslo, where 80% of the data were used in training. Next, the predictive accuracies of the AVMs are analyzed for different sub-groups of apartments, before purchasing rules are formulated to prevent automated bidding in apartment groups that are hard to price. At last, using the remaining 20% of the data, the average expected resale profits per apartment are examined for a hypothetical iBuyer operating in the Norwegian capital, with and without adverse selection and purchasing rules.

We find that adverse selection has a large negative impact on average profits for the hypothetical iBuyer, causing a reduction from 6.29-7.96% to 0.19-1.21% per apartment, across different models and scenarios. Furthermore, the simple purchasing rules are able to limit this reduction with around 1 percentage point per apartment when adverse selection is present. The findings are robust when altering the initial market assumptions, leading to a conclusion that adverse selection poses a noticeable threat to the iBuyer business model. In addition, we conclude that simple purchasing rules can help improve the average profits.

Contents

List of Figures

List of Tables

1. Introduction

Buying and selling homes are the largest transactions most people make during their lifetime. Getting a good price and selling to the right buyer is thus important. Traditionally, the process of selling a house is through a broker, with a listing and an auction. Instant buyers, or iBuyers, challenge this process. The iBuyer business model involves using automated valuation models (AVMs) to predict the market value of a home, before using this prediction to give a fast bid on the dwelling. This rapid and convenient process, however, does not come without challenges.

AVMs are statistical prediction models that try to predict the value of an object. Although progress is continuously made to develop AVMs with as high accuracy as possible, they are not able to fully reflect reality and capture all the factors that affect the price of a home. Adverse selection is a situation in which the seller knows more about an object and its value, than the buyer. Because of this asymmetric relationship concerning information about the home's value, in an iBuyer setting, the seller will often have an idea of whether the iBuyer offer is based on a correct, a too high, or a too low, price prediction. Intuitively, rational homeowners are more likely to accept a high bid rather than correct bid, suggesting that iBuyers may suffer from purchasing overpriced dwellings. Buchak, Matvos, Piskorski & Seru (2020) argue that iBuyers must focus on the most liquid dwellings to limit adverse selection. Little, or no, research beyond Buchak et al. have studied adverse selection in the context of iBuyer businesses. The purpose of this paper is to examine how adverse selection affects the profits of an iBuyer operating in Oslo, Norway, and how simple purchasing rules can be applied to limit this effect.

During the fall of 2021, at the time of writing, one of the biggest companies competing in the iBuyer market in the United States, Zillow, is pulling the plug on its iBuyer operations (Financial Times, 2021). Overvaluing and buying dwellings with low liquidity were named as causes for the failure. The company bought around 10,000 dwellings but managed to sell only 3000 of them. As a consequence, Zillow had to take a writedown on inventory of 300 million USD. This event underlines some of the challenges in the iBuyer business model and motivates our research and thesis.

The dataset used in the thesis consists of 84,905 apartment transactions in Oslo. Firstly, three AVMs are trained to predict apartment prices, using 80% of the data. The valuation models are based on linear regression, gradient boosting, and a support vector machine. After training the models, SHAP values and standardised coefficients are used to determine the most important predictor variables. These predictors are consequently used as dimensions for dividing the data into sub-groups. We analyse the predictive performances of the AVMs for the different sub-groups, before formulating purchasing rules to avoid bids in groups with bad performance. Lastly, the average expected resale profit per apartment is examined for a hypothetical iBuyer using a test set with the remaining 20% of the dwellings. The profit calculations are done with, and without, adverse selection and purchasing rules, to address their financial impacts. Adverse selection is implemented through accept probability distributions, indicating how likely a seller is to accept an offer when the bid is a certain percentage lower/higher than the seller's perceived valuation.

The thesis finds that adverse selection leads to a large reduction in expected profit per dwelling for the hypothetical iBuyer, from 6.29-7.96% without adverse selection, to 0.19-1.21% when adverse selection is included. As anticipated, the results imply that adverse selection poses a noticeable threat for iBuyer businesses. On the other hand, the paper also finds that implementing simple purchasing rules increases the average expected profits per apartment with between 1.03-1.57 percentage points. These results are later shown to be robust to changes in market assumptions, one of which is using both repeat sales and list price as proxies for sellers' perception of dwelling value.

In section 2, Literature review, we present the relevant literature for this thesis. Section 3, Data, describes our dataset, the cleaning process, and the different dependent and independent variables. Section 4, Methodology, introduces the hedonic price, extreme gradient boosting, and support vector machine models, in addition to the SHAP framework. Section 5, Model outputs and purchasing rules, examines the model performances and the creation of purchasing rules. Section 6, Results, presents the average profits for a hypothetical iBuyer. Section 7, Discussion, looks at the results in light of previous literature and discusses the impact for iBuyer businesses, before section 8 concludes and gives our final remarks.

2. Literature review

This chapter introduces the relevant literature for the study. It covers the iBuyer business model, use of AVMs in real estate, adverse selection, the lemon problem, and gives an overview of the property market in Oslo, Norway.

2.1 iBuyers and AVMs in Real Estate

iBuyers are prop-tech companies buying and re-selling dwellings, profiting on advanced automated valuation models (AVMs) to accurately assess the value of the dwellings (Gores, 2019). As the name suggests, iBuyers have the advantage of reducing the time spent in the traditional property sales process, as an offer can be received almost "instantly". The traditional sales process requires getting in touch with a real estate agent, who performs research, communicates with potential buyers, creates listings, before arranging a bidding process. Selling to an iBuyer removes most of these steps. The seller provides information on variables such as size, location, and condition, before this data is used by the iBuyer to create a price prediction. The user then receives an offer equal to this prediction minus a margin captured by the iBuyer, illustrated in equation 2.1.

$$Offer = Predicted\ market\ value - iBuyer\ margin \tag{2.1}$$

Opendoor, one of the leading iBuyers in the United States, began its operations in 2014. The company has grown in recent years, after making big steps in 2019 and 2021, even after suffering from a setback following the Covid-19 pandemic (Marquand, 2021). Other important iBuyers in the US are Offerpad, Keller Williams, and Redfin. Zillow Offers used to be one of the largest actors, before shutting down its iBuyer operations (Financial Times, 2021). A report by Mike Delprete of the University of Colorado (2020) suggests that the average profit margin of an iBuyer is 3.7% per dwelling, while Zillow Offers reported a negative per-unit profit of -2% in the 4th quarter of 2019. Delprete (2020) further finds that bids from Zillow and Opendoor generally corresponded to 98.6% of the AVM predicted value.

The use of AVMs in real estate is a much-discussed subject. Kok et al. (2017) finds strong evidence of the superiority of automated valuation models over traditional appraisals in terms of lower absolute error, as well as being more time-efficient and less costly. Furthermore, Mooya (2011) finds "no theoretical or practical reasons why AVMs should not completely replace traditional valuers". Others have, in contrast, suggested that AVMs should be used as a supplement to enhance, rather than an alternative to replace, manual appraisals (e.g., Reed, 2008; Waller et al., 2001).

For iBuyers, there are no manual appraisals or physical inspections involved in the housing transactions. Whereas manual appraisals are subject to human bias and subjectiveness, the advantages of AVMs are quick and consistent valuations (Jahanshiri, Buyong, & Shariff, 2011), that are non-biased (Fortelny & Reed, 2005). However, due to this lack of physical inspection in the iBuyer business model, there may be aspects affecting the price of a dwelling that the AVMs do not capture fully.

2.2 Adverse Selection and the "Lemon Problem"

Adverse selection is a well-known phenomenon within agency-contract theory, and a consequence of asymmetric information between two parties in a contractual agreement (Wilson, 1989). The research area has received much attention and was pioneered by the 2001 Nobel laureates in economic sciences George A. Akerlof, Michael Spence, and Joseph E. Stiglitz who won the prize "for their analysis of markets with asymmetric information" (Nobel Prize Outreach AB, 2001). The subject of adverse selection in the case of iBuyers is less explored.

Adverse selection generally occurs when a seller has more information about a product than a buyer (Wilson, 1989). The buyer cannot to a full extent observe the quality of the product, only the distribution of "good" and "bad" products sold in the past, and the seller thus has the incentive to market a "bad" unit as a "good" one (Akerlof, 1970).

Akerlof goes further to describe the "Lemon problem". The buyer cannot know or observe whether an item is a lemon (bad) or not, and the risk of purchasing a lemon reduces the average reservation price of buyers. This reduced reservation price makes non-lemon sellers less interested in selling,

increasing the proportion of lemons further. Genesove (1993) suggests four criteria that must be met for a market with adverse selection suffering from a lemon problem. There must be (I) asymmetric information regarding the quality of the good between the seller and the buyer at the time of the purchase, both (II) the seller and the buyer must value quality, the (III) price must be determined by the party with less information, and there must be (IV) no institutions completely removing uncertainty related to the quality of the good (Genesove, 1993).

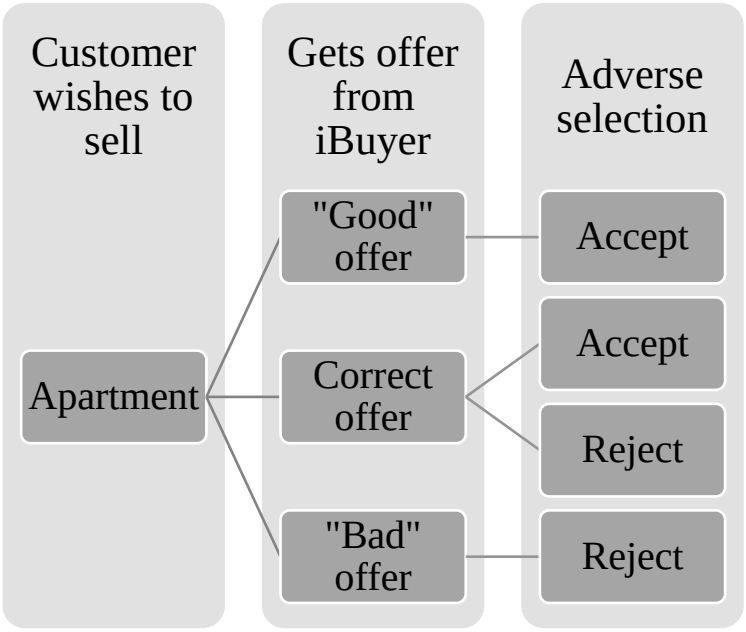

Figure 1: Adverse selection for iBuyers, derived from Wilson (1989), Akerlof (1970), and Buchak et al. (2020). The table illustrates the likely scenario for how sellers of dwellings are more likely to accept an offer when it is "good", rather than correct or "bad".

Buchak et al. (2020) points out the problems of adverse selection for iBuyers. iBuyers generally operate with higher fees than conventional realtors, indicating that the homeowners selling to these businesses trade profit for a quick transaction. However, the implementation of such quick transactions usually comes at a cost of information loss. There are aspects affecting the value of a house that are not easily quantifiable, such as the view or the condition of the neighboring housing units (Buchak et al., 2020). Furthermore, the lack of internal inspection of the property may lead to information loss related to structural and conditional attributes (Tretton, 2007; Fortelny & Reed, 2005). While the algorithmic valuation models of the iBuyers may not be able to fully capture this

information, it is usually known by the homeowner. This asymmetry in information between the buyer and the seller gives rise to problems related to adverse selection, comparable with the way Akerlof describes it. Sellers receiving too high offers compared with their perceived valuation are more likely to accept the offer than sellers who receive a correct offer, as illustrated in Figure 1 (Akerlof, 1970).

In the iBuyer case, the Genesove (1993) criteria hold, and adverse selection can result in a lemon problem. This happens if the iBuyer increases the premium to stay profitable, after taking the increased risk of buying apartments valued too high (i.e., lemons) into account. These increased premiums can again result in more "correct" valuations being turned down, thus increasing adverse selection further. The cycle follows the results of studies such as Palm (2015), and Emons & Sheldon (2007), in the real estate office market and the used car markets respectively. A visualization of the cycle is shown in Figure 2.

Spence (1974) suggests adverse selection can be dealt with through market signaling, where sellers undertake efforts to inform buyers about the quality of the product to change the initial asymmetric information structure of the market. Hence actions are taken by the party with the most information. Another way of preventing problems related to adverse selection is through screening, to bridge the information gap between the two parties (Stiglitz, 1975). Here, actions are taken by the party with the least information. Homebuyers will normally attempt to learn as much as possible about the quality of an apartment before purchasing, to reduce the risk of paying too much. This may be done through viewings and questioning the realtor. On the other hand, the purpose of selling to an iBuyer is to reduce time, and thus the screening process will also need to be shortened (Buchak et al., 2020).

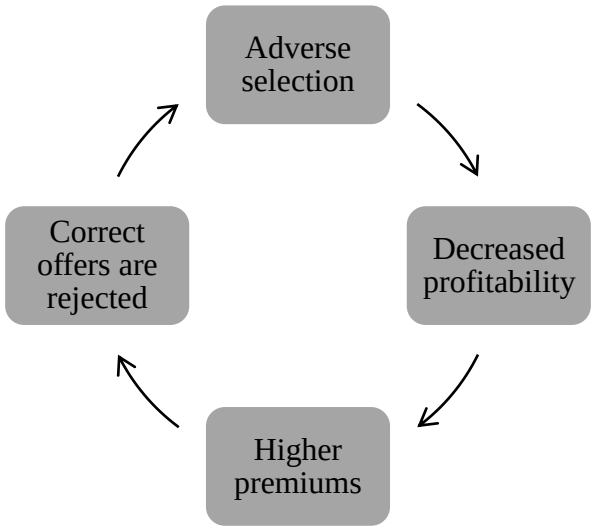

Figure 2: The lemon problem in the context of iBuyers, derived from Akerlof (1970). Decreased profitability from adverse selection issues may trigger higher premiums from the iBuyer, thus increasing the amount of offers that are rejected, which again increases adverse selection.

Previous research thereby suggests that adverse selection might be a problem in the iBuyer business model. However, the most widely accepted actions to reduce adverse selection in general, signaling and screening, are difficult to implement efficiently within this segment. To reduce problems with adverse selection, the aim is to avoid the homeowners receiving too "good" or too "bad" offers, from Figure 1. On the other hand, the question of what defines a good, bad, and correct offer is not definite. Too good is, in this case, referring to an offer that is noticeably higher than the real market value, and the perceived value of the homeowner. Too bad, in contrast, means that the offer is well below the actual value. Both scenarios are damaging for the business model. A too good offer results in negative profits, while too bad damages the credibility of the business. Furthermore, the real market value of an apartment is not known until it has been sold on the open housing market. The iBuyer relies on the AVM for settling the price, and the seller usually has his/her own perceived opinion on the value. This perception of price introduces further complications, as it is not always a reflection of the true value.

Tversky and Kahneman (1974) introduced the theory of anchoring and adjustment. It states that when people make value predictions, they usually start with a single estimate and subsequently adjust this estimate with new information. Furthermore, people are generally shown to be too

optimistic in this valuation prediction process and will often settle with an estimate of their liking rather than a correct one (Lovallo & Kahneman, 2003). Starting with an optimistic anchor and adjusting this subject to new information, will create a biased, overvalued, result. In the case of homeowners, an upwards biased estimate may influence which offers from iBuyers are accepted and which are refused, even if they are "correct" following the real market value.

Previous literature covers the area of creating well-performing AVMs, as examined further in chapter 4. An AVM with good performance will, all other things be equal, result in profits for the iBuyer, if all offers are accepted. However, when introducing the aspect of adverse selection, all offers being accepted does not seem like a realistic assumption for the iBuyer businesses in practice. Buchak et al. (2020) suggest that iBuyers only purchase the most liquid, and easy to value, houses. This is one way of dealing with potential adverse selection.

2.3 Property market in Norway and Oslo

To utilize pricing models properly, it is helpful to understand the dynamics of the relevant market. A high degree of openness, open listings, and open auctions are general attributes of the Norwegian property market, as well as high market participation. 77% of all households in Norway own the dwelling they live in, and housing is the largest asset class (Statistics Norway, 2020). The property market thus affects most Norwegian citizens noticeably. Buying and selling homes in the Norwegian property market is, in general, done through an open auction, where 90 percent are sold via English auctions (Olaussen, Oust, & Sønstebø, 2018).

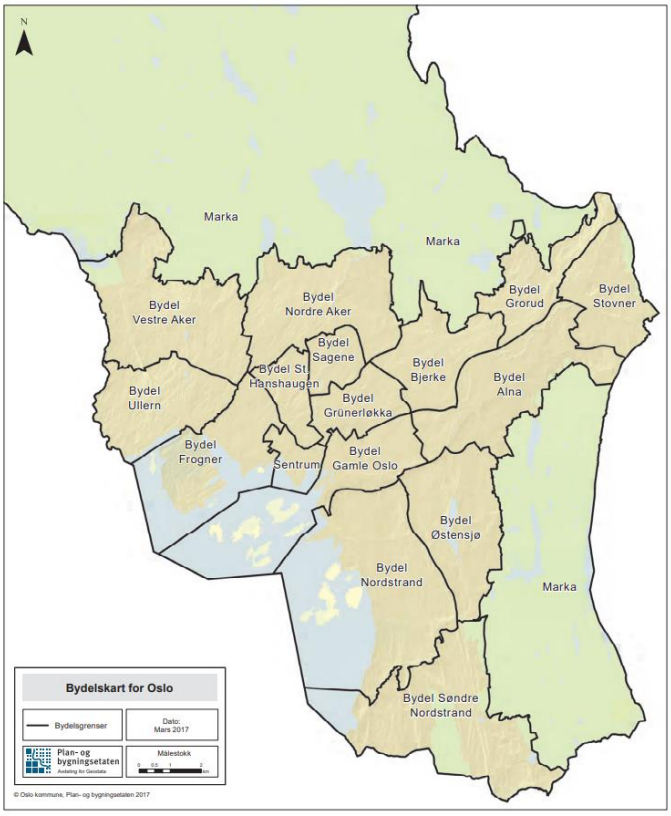

Figure 3: Map over administrative districts in Oslo (Oslo Kommune, 2017).

The property market in Oslo can roughly be divided in two parts: the west and the east. Square meter prices are on average lower in the east part of Oslo than in the west, divided by the river Akerselva (Sørgjerd, Murray, & Hager-Thoresen, 2020). Price differences started as early as in the late 19th century when the richer part of the city moved west of the factories built alongside Akerselva (Oust, 2012). Around 80% of all people with background from Asia, Africa, and Latin America live on the east side of Oslo (Sloan & Aarbakke, 2016). As elaborated more extensively in part 3.3.3, Oslo has 15 administrative districts, in addition to Marka and Sentrum (Oslo Kommune, 2021). Each district has a district committee that organizes and provides services. A map of the districts is shown in Figure 3. Between 2010 and 2020, apartments in Norway increased in price by 78.9% (Statistics Norway, 2021), and during the same decade, the prices have more than doubled in the capital, as seen in Figure 4.

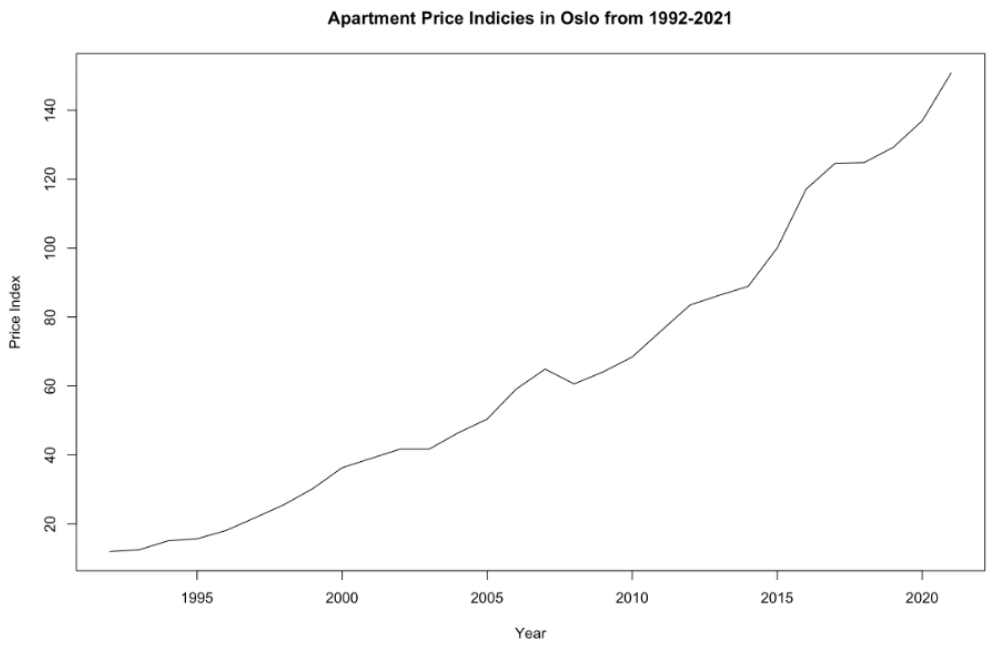

Figure 4: Price development for apartments in Oslo between 1992 and 2021, based on yearly indices published by Statistics Norway. During these nearly three decades, the overall apartment prices have thus increased by 1156 % nominally.

3. Data

The primary data used in the study was provided by Solgt.no, a Norwegian prop-tech startup operating in Oslo. The dataset contains information on housing transactions listed on Finn.no, the largest online marketplace for private property in Norway, combined with public data from the Norwegian Mapping Authority (NMA). The relevant transactions took place between 2007 and 2021 and consist of apartments in the Norwegian capital. This section will describe the relevant variables in the data material, and how the data was filtered and cleaned for usage and interpretation purposes.

3.1 Housing type stratification

Before examining the variables in the dataset, the different housing types are assessed. A stratification process of organizing the data based on housing type shows that 85.2% of all the transactions were apartments, as displayed in Table 1. Oslo generally stands out in Norway as a region with a high proportion of apartments. To observe a large share of the listed transactions in the dataset being apartments is thus not unexpected, albeit slightly higher than the 73% of Oslo households living in apartment blocks found by SSB in 2018 (Statistics Norway, 2018).

Only using one AVM applied on multiple housing types could yield unsatisfying predictions, as the residencies differ noticeably related to predictor variables such as total living area, location, floor, etc. The apartment segment alone, however, constitutes a more homogenous sub-market within private residencies, thus providing more appropriate terms for comparison and valuation modelling.

When taking the large proportion of apartments into account, as well as the homogenous qualities of the housing type as a separate sub-market, this study focuses on apartments exclusively. This decision further corresponds with the business model of prop-tech companies such as Solgt.no, as the company only purchases apartments. For this reason, only apartments observations, registered as primary residences and approved living units, are included in the study.

	Number of transactions
Apartment	148,249
House	11,114
Row house	7,439
Semi-detached house	5,991
Other	1,211

Table 1: Housing type frequency in the data set. Most observations are apartments. For this reason, and apartments having homogenous characteristics, the remaining housing types from the data set are not used in the study.

3.2 Data pre-processing

Before providing a descriptive examination of the final dataset, we elaborate on the process of removing missing values and erroneous observations. Missing values occur as many of the observations do not contain values for all the relevant variables. Furthermore, the data was extracted from ads on Finn.no and is thus subject to errors if realtors include obvious wrong information in the housing ads. Therefore, the data cleaning process does not only consider the missing values, but also examines whether the existing values appear realistic.

Therefore, to ensure a suitable contribution to the prediction models, several modifications are done to the variables, see Table 2. The original data contains information on both list price and actual sales price. Observations with sales price more than twice as high, or half as low, as the list price, are removed. This is done to prevent the inclusion of observations where the sales price is wrong, as this variable is an essential part of the models. Furthermore, observations with a living area of less than 9 m^2 or over 300 m^2 are removed. These cut-offs are chosen because observations outside this range are generally not seen as realistic, as prices, number of bedrooms, and ad-titles do not correspond with such small or large values of living area. Observations with a building year before 1600 are removed, as these contain wrong values. Similarly, we remove observations with more than ten bedrooms or located on a floor higher than 20.

Some of the latitude and longitude coordinates in the original data have also been mixed up. Latitude values are registered for the longitude variable and longitude values for the latitude variable. This is fixed, in addition to removing the few observations with coordinates far away from Oslo. The geographical location of apartments is not included directly in the model in the form of coordinates, but the coordinates are used to place properties into the correct district. This will be explained further in section 3.3.3 about the district variable. Lastly, four incomplete months with less than 100 observations left are removed.

In total, the cleaning process reduces the number of observations from number 178,001 to 85,087, displayed further in Table 2. Even after a strict cleaning process, this new dataset is still deemed sufficiently large for creating prediction models (Ogundimu, Altman, & Collins, 2016). The data is then split into a training and a test set, corresponding to sizes of respectively 80% and 20% of the total dataset. This subsampling is done by randomly assigning 80% of the observations to a training set and using the rest as a test set. The focus of the study is to identify and account for groups of apartments that the AVMs fail to predict correctly, thus giving rise to adverse selection problems. These groups are assumed to be relatively constant through time. The training/test split is thus done randomly, instead of separating the data based on a cut-off in time. The test set is then put aside until the models are built and ready to be evaluated. In the following sections, where the data is explored further, only the training data observations are examined. The reason for not examining the test set is to prevent these observations from impacting the modelling. However, the test dataset was cleaned equally to the training set to avoid faulty values.

Data pre-processing	Observations left
Original data	178,001
Only apartments (primary residency)	143,039
Keep observations with sales price more than 0.5*list price, and less than 2*list price	142,510
Keep observations with living area larger than 9 square meters and less than 300 square meters	138,173
Keep observations with build year newer than 1600	137,405
Keep observations with less than 10 bedrooms	129,380
Removing observations with faulty coordinates	128,247
Keep observations with floor larger than 0 and less than 21	113,748
Keep observations with at least one bathroom	85,271
Remove dwellings sold in months with less than 100 observations	84,905

Table 2: Steps taken in data pre-processing. In addition to the steps, the right column shows how many observations are left in the data after implementing the relevant step. The raw data set included 178,001 observations, while the cleaned data set included 84,905 observations. Thus, almost half of the observations were rejected in the cleaning process, due to containing faulty information.

3.3 Descriptive statistics

3.3.1 Dependent variable

The aim of AVMs is generally to predict the price of a good. The total transaction price, including both price and debt associated, for the different apartments in the dataset is therefore the dependent variable. There are large differences in apartment prices, as shown in the histogram in Figure 5. The visualization does not represent the whole dataset, as it is limited to a maximum price of NOK 15,000,000. Apartments exceeding this price constitute a small minority of the observations, and a graphical representation of sales volume among the most expensive apartments is therefore not

deemed suitable. In contrast, the descriptive statistics in Table 3 highlights the wide range of apartment prices found in the data material. The most expensive apartment cost NOK 60,000,000, while the median value was NOK 3,500,000.

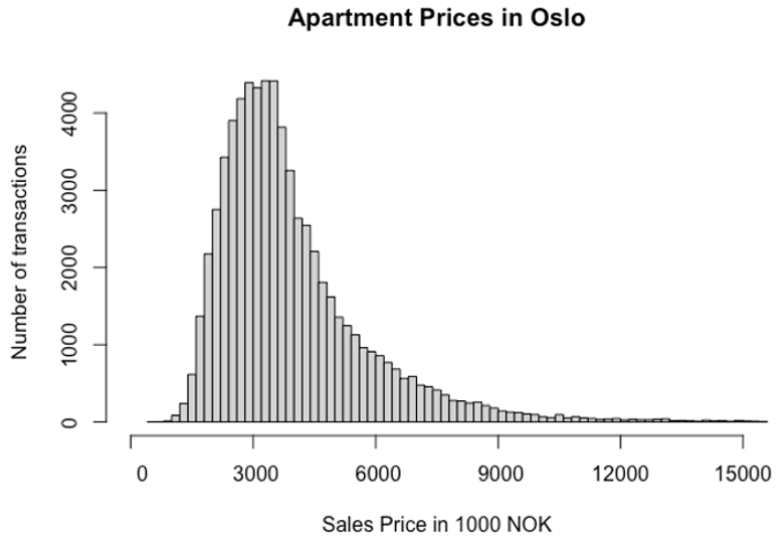

Figure 5: Histogram showing the number of transactions for sales prices. The distribution follows a right-skewed distribution, with a mean of 3.9 million NOK.

Sales price (in thousands NOK)

Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
594	2,724	3,500	3,986	4,640	60,000

Table 3: Summary statistics for the sales price variable (in thousands NOK).

3.3.2 Physical variables

After having examined the apartment price, it is relevant to have a further assessment of the independent variables relevant for the AVM predictions. The first group of explanatory variables is related to the size and the physical dimensioning of the apartments. These variables are related to the total interior livable area (living area), which floor the apartment is located on, the number of bedrooms, and the number of bathrooms. The living area was chosen as the measure for size instead of usable square meters, as it is deemed more precise to value an apartment. Usable square

meters contain all space in an apartment, while primary rooms are the square meters other than storage space, such as kitchen, bedrooms, and bathrooms.

As mentioned in the pre-processing section, apartments with less than 1 registered bathroom were removed. Although there may be apartments in the data that do not have a bathroom, it is not plausible, based on comparable data from Statistics Norway (2021), that this is the case for such a large proportion of the data. Furthermore, several of the observations registered without a bathroom were relatively expensive, and it did not seem realistic for them to have no bathroom.

Table 4 gives an overview of the size and dimensionality variables. As the table shows, the median apartment in the dataset is 64 m^2, has one bathroom, two bedrooms, and is located on the third floor. A large proportion of the observations only has a single bathroom, indicated by a mean value close to 1 as well as the third quartile also being 1. There are, however, observations in the data with up to five bathrooms. The number of bedrooms varies from 0 to 8, while the floor varies from 1 to 20. The smallest apartment among the observations in the training data is 10 m^2, while the largest is almost 30 times larger.

Physical variables

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
Bedrooms	0	1	2	1.73	2	8
Bathrooms	1	1	1	1.06	1	5
Floor	1	2	3	3.2	4	20
Living area	10	50	64	66.58	78	299

Table 4: Summary statistics for size and physical variables.

After data cleaning, the oldest apartment in the data set is from 1667, and the newest ones are from 2021. The median build year is 1968. The number of dwellings built grew steadily from 1900 to 1980. Furthermore, there was a large increase in the number of dwellings built between 2000 and 2021.

The mean living area in square meters is 66.58m^2, the median apartment size is 64m^2, and the distribution is right skewed. 58% of all observations are equal to, or smaller than, the average of 67m^2. Figure 6 displays the number of apartments in different sizes.

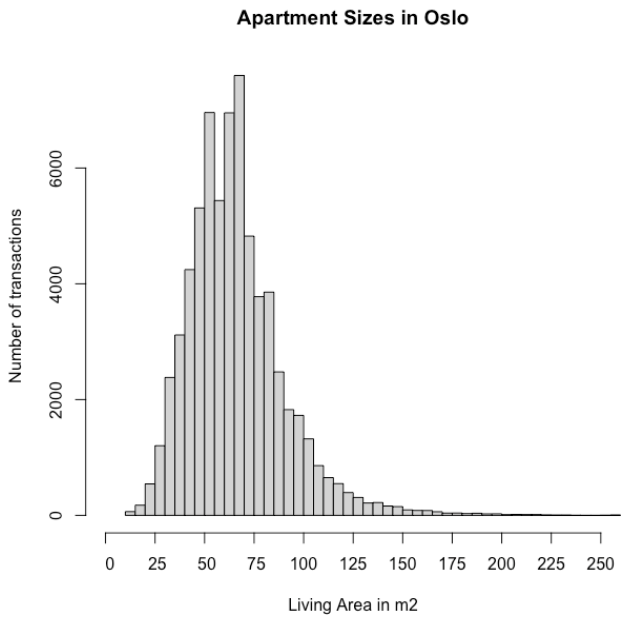

Figure 6: Number of apartments in different sizes, measured as total interior livable area in square meters. The distribution is right skewed with a mean of 66m^2 and resembles the histogram for prices in Figure 5.

3.3.3 District variable

Geographical location is a central factor in determining the price of an apartment. The location could have been included in the dataset in several ways, as the original dataset contains information on both coordinates and full address with postal code. In this study, the apartments are mainly sorted into geographical groups based on the administrative district the property belongs to. Finding a suitable trade-off between low and high spatial aggregation is influential (Sommervoll & Sommervoll, 2018). Low spatial aggregation captures more of the systematic spatial variation but also reduces the number of observations in each region. Meanwhile, a high aggregation has the opposite effect. Administrative districts were found to have satisfyingly similar intra-regional location premiums, thus making them a good candidate for capturing spatial effects.

There are 17 different administrative districts in the Norwegian capital, differing in size, population, and average square meter prices. There are some inter-regional variations, and apartments close to the border areas may be affected by neighboring regions, but the districts are still deemed appropriate spatial groupings. As the smallest of the districts, Sentrum, has a limited number of transactions, it is included in the district of St. Hanshaugen. This is done to reduce the possible impact of potential outliers. Consequently, in the rest of the study, we operate with 16 districts.

The apartments are sorted into their relevant districts based on the postal codes in the address. Some of the postcodes can, however, cover properties across multiple districts. The apartments with these post numbers, therefore, need to be labeled in another way. The districts were hence separated over two steps consecutive; (I) giving a district value to each property with a postal code that could only belong to a single district, and (II) using a K-nearest neighbor-approach (KNN) to label the remaining apartments in the same district as the majority of its 10 nearest neighbors. "Nearest" is, in this setting, defined as the observations having the smallest Euclidian distance from the apartment that needs labeling, in terms of standardized longitude and latitude. The KNN algorithm is described further in appendix A.1.3 District Labelling with K-Nearest-Neighbors.

After sorting each observation into a region, information on the different districts can be found in Table 5. As seen, there are noticeable differences in both mean apartment prices, the number of transactions, and internal variation (standard deviation) across the different districts. Frogner and Ullern stand out as the districts with the highest average apartment price in the dataset, both having a mean greater than NOK 5,000,000. Grorud, Stovner, and Søndre Nordstrand stand out on the opposite end of the pricing scale, averaging apartment prices of less than NOK 2,800,000. There are variations between the districts concerning average sizes, which further affect the average prices. Naturally, regions with smaller apartments will tend to have lower prices, and a study of the total prices between regions may therefore not give the full picture. The total price is, however, the variable we seek to explain. Furthermore, the districts are included in the models as explanatory variables and examining the relationship between the total price and the district is thus relevant.

District prices (in thousands NOK)

	Mean	Standard deviation	Minimum	Maximum	Number of transactions
Alna	2,910	731	1,060	9,500	4,344
Bjerke	3,382	1,229	1,048	10,450	2,647
Frogner	5,380	3,170	950	50,020	8,174
Gamle Oslo	3,782	1,716	950	42,000	9,537
Grorud	2,629	568	1,058	5,501	1,819
Grünerløkka	3,986	1,514	873	14,115	8,878
Nordre Aker	4,617	2,330	1,001	19,900	1,726
Nordstrand	3,709	1,710	985	23,000	3,360
Østensjø	3,283	1,184	1,123	12,700	2,861
Sagene	4,094	1,639	594	13,300	6,534
Søndre Nordstrand	2,784	814	620	7,500	2,265
St. Hanshaugen	4,094	1,722	975	22,000	7,594
Stovner	2,684	656	844	6,400	2,348
Ullern	5,138	2,230	1,165	20,500	3,124
Vestre Aker	4,940	2,681	1,166	60,000	2,713

Table 5: Regions in the dataset with mean, standard deviation, minimum and maximum price in millions NOK, and the number of transactions in each area. There are noticeable differences between areas in the mean price and number of transactions, which will be important to consider later when creating purchasing rules.

3.3.4 Sales time variable

Sales time is a categorical variable providing information on what year and month the property was sold. Property valuations will vary over time, affected by factors such as interest rates and inflation. Two identical apartments would not necessarily be valued at the same price in 2010 as in 2020, even though the other explanatory variables remain equal in the two cases. For this reason, it is necessary to include sales periods in the models to account for temporal heterogeneity (Helbich, Brunauer, Hagenauer, & Leitner, 2013).

The observations are divided into monthly categories based on the date for judicial registration of the new ownership. This is the date from which the buyer officially owns the purchased apartment, derived from information from the NMA. The time aspect is included monthly instead of yearly, to capture seasonal effects within the individual years. However, note that the time between the actual sale and the judicial registration may vary slightly. The dataset consists of transaction data

from 2007 and up to the time of writing (2021). The first and the last months in the data each have less than 100 observations, and these periods are thus excluded.

3.3.5 Seller valuation variables

The next variables to describe are not used as predictors in the AVMs but play central roles in the computation of expected profits for iBuyer companies. To incorporate how likely a homeowner is to accept an offer from the iBuyer, a measure for what the seller believes the apartment is worth, is necessary. In this paper, two different variables for homeowner valuation are used. The first one is the list price found in the sales ad. This represents the price that a professional real estate broker chose to value the apartment at. Such an appraisal is available to homeowners considering selling their apartment and is thus deemed appropriate for representing the seller's valuation of the apartment.

The second measure for the perceived apartment valuation of the homeowner is based on a repeat sales calculation. The repeat sales valuation uses the price of the previous transaction of an apartment before adding the general price development in the relevant market, from the previous sales time to the current one. As seen in equation 3.1, the price development is included through apartment price indices for the Oslo property market, provided by Statistics Norway. These indices are shown graphically in Figure 4, with data going back to 1992. As repeat sales valuations incorporate the previous purchase price of the apartment, the repeat sales may seem like a reasonable estimate for what sellers think their home will be sold later. However, a proportion of the apartments in the dataset only had a single owner since the building time. These apartments, without a previous transaction price, cannot be given a new valuation through repeat sales calculation. For this reason, the proportion of the dataset with repeat sales value estimates consists of 52,667 apartments, compared to the complete training set of 67,924.

$$Value_t = \frac{Sales\ Price_{t-1} \times Price\ Index_t}{Price\ Index_{t-1}} \qquad 3.1$$

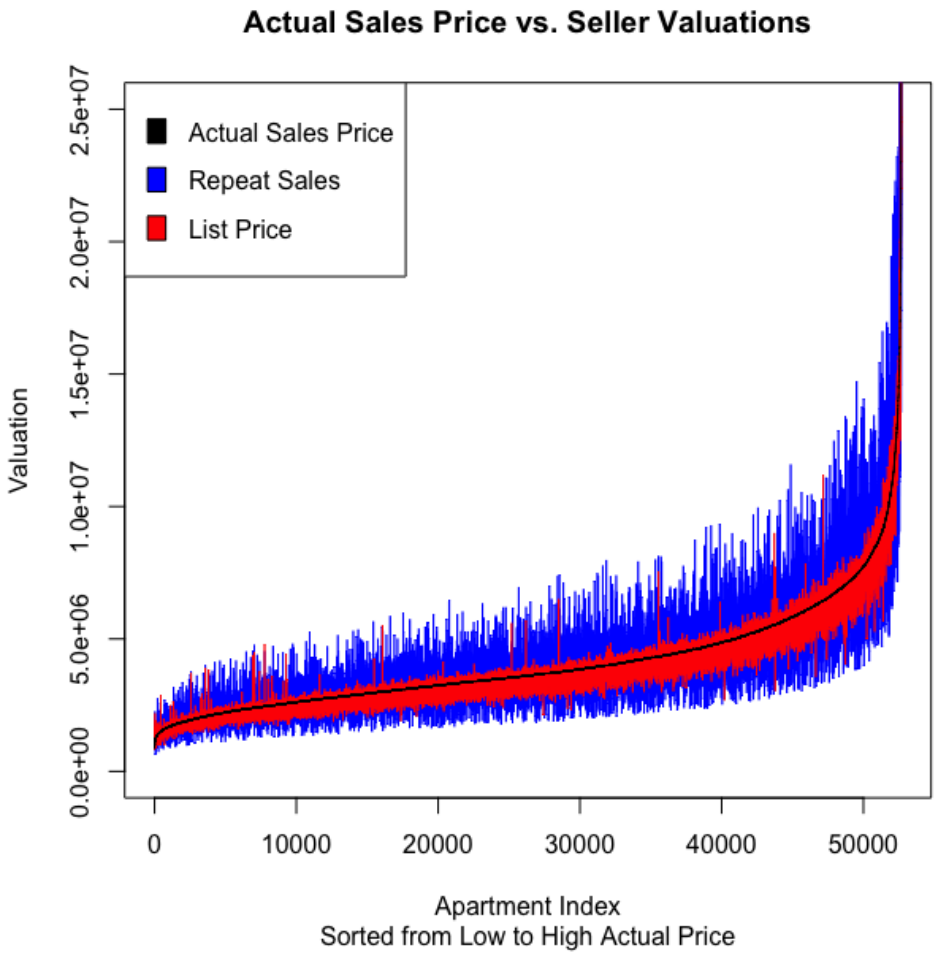

Figure 7: Graphical visualization of sales price and seller valuations. The apartments are sorted from low to high price, and plotted with sales price (black), repeat sales value (blue), and list price value (red). As seen, the red list price line is generally closer to the black sales price line, indicating that this is a more accurate valuation method than the repeat sales. However, the purpose of these valuations is not to be as accurate as possible, but rather to represent the seller's perceived valuation of the apartment. Repeat sales is included as it incorporates the price the homeowner paid to acquire the home in the first place, which intuitively is important for the homeowner when later considering selling the same apartment.

The list price on average deviates 5.8% from the actual sales price. Table 6 indicates that the list price most often is lower than the sales price, as both the mean and the quantiles are lower. This remark also corresponds with the plot in Figure 7, where the list price generally appears to be lower

than the sales price more often than higher. The repeat sales seller valuation, on the other hand, on average deviates 11.5% from the actual sales price. The deviations appear to be more evenly distributed above and below the black line, but the spread is noticeably wider than for the list price.

Repeat sales against list price and actual price

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
Repeat sales	625	2,735	3,558	4,064	4,773	45,573
List price	832	2,690	3,491	3,960	4,650	55,000
Actual price	844	2,824	3,609	4,100	4,800	50,000

Table 6: Summary statistics for the repeat sales valuations, list price valuations, and the actual sales prices. Only the apartments that has been sold more than once since 1992, and thus has a repeat sales value, are included.

Table 6 and Figure 7 above compare the two seller valuation variables with the actual price, for the 52,667 apartments sold more than once since 1992. For the complete training dataset, there are slight changes in summary statistics. While the dependent variable was already described in section 3.3.1, Table 7 gives the summary statistics for the list price when considering the complete training dataset. The full dataset, including apartments with only one transaction between 1992 and the date of the sale, has both lower list prices and lower quartiles. This implies that the apartments not included in Table 6 generally had a slightly lower list price than the rest.

List price on total data set

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
List price	560	2,572	3,390	3,846	4,500	59,000

Table 7: Summary statistics for the list price valuations. The full training dataset is included.

3.3.6 Renovation variable

A binary indicator showing whether the relevant apartment is a renovation project or not was also included in the data set. Renovation projects require additional investments after the initial purchase and are therefore valued lower than the same apartment would be if it did not need these extra costs. The renovation variable thereby says something about the condition of the apartment. This variable is derived from the information found in the title of the ad. If the title includes one or

several keywords related to need for renovation, the variable takes the value 1, otherwise 0.[1] There is a degree of uncertainty concerning this estimate. However, it seems reasonable to assume that most realtors will include a need for renovation in the title. Yet, the ad titles of some apartments were changed after the sale, to inform potential buyers that they were no longer for sale. Whether the apartment initially needed renovation or not, would therefore not have been extracted. On the other hand, it seems reasonable that the variable captures most of the apartments that have an extensive need for renovation in the dataset. In total, there are 1147 renovation projects out of the 67,924 transactions in the training dataset.

3.3.7 Facility variables

The last group of variables concerns facilities of the dwelling. To provide information on the condition of the apartment, in addition to the physical measures, several binary variables on facilities were included, as seen in Table 8. There are a total of 22 different facilities to choose from in the Finn.no advertisement interface, and the inclusion of any of these is completely optional.

It seems likely that several of the possible facilities do not contain significant information for predicting the property price. The facility variables were carefully considered, as the process of writing a Finn.no ad may differ between real estate brokers. One broker may consider a dwelling to be "modern" while another may not. In addition to this problem, related to subjective opinions, it is also likely that many ad creators did not include a facility even if the apartment had the relevant attribute. A reason for doing this will be that the ad seems more structured if only the most important facilities are included, in contrast to adding close to 20 different. One of the variables that might suffer from not being prioritized in this way, is parquet. There are very few observations with "parquet" listed as a facility, much less than common sense would assume.

[1] The keywords searched for in the ad titles are "oppussing", "renovering", "renovasjon" and "utbedring"

With the aforementioned problems in mind, it is clear that there exists a varying degree of uncertainty related to the facilities. By examining the total number of apartments having different facilities, we can assess whether the relevant number seems plausible. In combination with an evaluation of whether the variable seems interesting from a predictive point of view, we have chosen to include the variables that seem most relevant. These are elevator, balcony, child-friendly, garage, fireplace, quiet, and view. The included dummy variables are considered reasonable but are subject to uncertainty, which could potentially bias our results.[2]

Number of facility variables

	Elevator	**Balcony**	**Child friendly**	**Garage**	**View**	**Fireplace**	**Quiet**
Yes	28,757	47,114	39,747	27,553	15,273	24,298	26,168
No	39,167	20,810	28,177	40,371	52,651	43,626	41,756

Table 8: Facility variables. The table shows the number of observations in the training data that have the different attributes.

[2] *Elevator facility is supplied with data from Norwegian Ministry of Trade and Industry-owned company Ambita, and thus more reliable*

4. Methodology

With the aim of examining how adverse selection can be considered in an iBuyer business model, the first step is to create the AVMs. We will study three different models: a hedonic regression model, a Support Vector Machine, and a gradient boosting model known as XGBoost (eXtreme Gradient Boosting). The reason for creating different AVMs is to underline that the methodology and results are applicable and generalizable for a wide range of models. The hedonic linear regression model is chosen for its interpretability in addition to the fact that it is widely used in the field of real estate valuations, and this model will serve as a baseline for the others. The two machine learning approaches are chosen based on predictive accuracy found in previous literature, see Table 20 in the appendix. The XGBoost algorithm is still a relatively recent addition to the machine learning world at the time of writing, but studies of the approach have already shown encouraging results within the area of real estate valuation.

Note that the aim of the paper is not to identify the best-performing AVMs in terms of predictive accuracy, but rather to examine how the different models may be used and modified to reduce problems related to adverse selection for iBuyers. There are several ways of measuring the performance of a prediction model. Section A.2.2 Performance evaluation in the appendix contains a brief discussion on the most relevant performance metrics in this context. In the first parts of the methodology chapter, however, there will be sections providing algorithmic descriptions of the three modeling approaches. In the end, a framework for interpreting the outcome of machine learning models, known as SHAP, is described.

4.1 Hedonic regression model

Many objects and products can be broken down into a set of separable factors. These internal and external factors all have market values, referred to as implicit or hedonic prices. The price of the product can be explained by the sum of its characteristics' hedonic prices (Rosen, 1974). In the property market, an apartment price can consequently be explained as the value of the given

apartment's factors, such as living area, number of bedrooms, location, etc. Hedonic regression models are thus popular methods for predicting the sales prices of apartments.

The relationship between the dependent variable (sales price) and the independent variables is generally non-linear. The marginal price effect of a one-unit increase in a predictor value is different from cheap to expensive apartments. Figure 10 in appendix A.1.2 Price versus space plot gives an example of the non-linear relationship between the sales price and the living area of apartments in the dataset. Therefore, the dependent variable in the hedonic regression model is the natural logarithm of the sales price.

$$ln(P_i) = \beta_0 + \sum_k \beta_k X_{ki} + \varepsilon_i, \qquad (\varepsilon \sim i.i.d.) \tag{4.1}$$

Using the natural logarithm of an apartment's price, $ln(P_i)$, as the dependent variable, the regression function is given by equation 4.1. Here, β_0 is the intercept, β_k is the coefficient of predictor k, and X_{ki} is the feature value of predictor k for a given apartment i. The aim is to find the coefficients β_k of the different predictors that minimizes a given loss function for all the apartments in the dataset. ε_i is the error term of apartment i, with an expected value of zero. The error terms are identically and independently drawn.

Property markets data tend to suffer from outliers, thus applying robust regression models is important (Janssen, Söderberg, & Zhou, 2001). A common approach to create a linear regression model would be to apply the ordinary least squares method (OLS). Earlier literature using hedonic models on property data, however, indicate that the least absolute deviation (LAD) method is preferred, due to being more robust towards outliers (Yoo, 1999). LAD was first introduced by Koenker and Basset (1978). The difference between LAD and OLS is the loss function the algorithm seeks to minimize. Whereas OLS finds the feature coefficients, β_k, that minimizes the squared prediction errors, LAD minimizes the absolute value of these errors. As the errors are not squared, the LAD loss function is less sensitive to outliers (Stock & Watson, 2019). The LAD function will work to minimize 4.2.

$$\underset{\beta}{argmin} \sum_{i=1}^{n} |\ln(P_i) - \beta X_i| \qquad 4.2$$

In other words, the algorithm finds the vector of β_k values from equation 4.1, that minimizes the absolute value of the deviations between the price and the prediction for all apartments i=1, 2, 3, …, n in the dataset. ln (P_i) is the natural logarithm of the sales price of apartment i, while βX_i is the predicted value of the same logarithm. The prediction is computed by multiplying the vector of coefficients, β, with the vector of predictor values, X_i.

4.2 Support Vector Machine

While originally created for classification problems, support vector machines (SVM) have been developed to handle regressions as well. Support vector regression (SVR), which builds on the foundation of the SVM algorithm, was introduced by Drucker, Burges, Kaufman, Smola and Vapnik (1997).

Support vector regression utilizes an ε-tube and slack variables to find the regression line. ε represents the allowed error within which all errors are disregarded. For observations outside the tube, errors are measured as the deviation between the actual response value and the ε-tube itself, rather than the regression line. This separates SVR from methods like the hedonic price model, where the goal is to minimize the residual errors of all observations. Observations inside the ε-tube are referred to as support vectors.

With regression optimization in mind, the support vector machine aims to solve the following primal function:

$$\underset{\xi^*,\xi,w,b}{min} \; C(\sum_{i=1}^{n} \xi_i^* + \sum_{i=1}^{n} \xi_i) + \frac{1}{2}(w^t w) \qquad 4.3$$

Subject to:

$$y_i - w^T\phi(x_i) - b \leq \varepsilon + \xi_i \quad (4.4)$$

$$w^T\phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \quad (4.5)$$

$$\xi_i^*, \xi_i \geq 0, i = 1, \dots, n \quad (4.6)$$

As a solution to the optimization problem with no observations outside the ε-tube is not always feasible, the slack variables ξ_i and ξ_i^* are introduced for apartments $i = 1,2, \dots, n$ in the training set. The slack variables represent the errors that exceed the value of ε, and thus impose what is referred to as a soft margin. Furthermore, b is a bias term, y_i represents the response value, $\phi(x_i)$ is the mapping of observation i with predictor values x_i in feature space, and w the weights of the regression line. $w^T\phi(x_i)$ is thereby the prediction from the regression line.

C is the regularization parameter in the model, determining how much to penalize the slack variables of response values outside the ε-tube. The C-parameter thus controls the trade-off between bias and variance, which is a common issue for all machine learning models. Increasing C will punish errors more, reducing the number of errors, giving higher bias, and lower variance. Decreasing C will increase the number of allowed errors, resulting in lower bias, and higher variance. A too high C will risk underfitting the model, and a too low C could potentially give a very accurate model on the training data, but risk overfitting such that it does not have a satisfactory prediction power.

One of SVMs greatest strengths is being able to map input's attribute space into higher dimensions. The primal optimization problem can be rewritten in a less computationally expensive Lagrange dual form, as done in equation 4.7. Kernel functions make this transformation feasible and reduces the computational effort needed, known as the "Kernel trick". In equation 4.7 the Kernel function is noted as $K(x_i, x_j)$. When satisfying the Mercer condition, the Kernel function can, in the dual function, compute the inner products without transforming the input features.

The dual formula for non-linear SVM regression seeks to find the coefficients that solves the following optimization problem:

$$\min_{\alpha,\alpha^*} \frac{1}{2}(\alpha - \alpha^*)^T(\alpha - \alpha^*)K(x_i, x_j) + \varepsilon \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) \quad 4.7$$

Subject to:

$$\sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0, \quad 4.8$$

$$\alpha_i \geq 0, \alpha_i^* \leq C, i = 1, \dots, n \quad 4.9$$

Where α_i and α_i^* are the dual coefficient vectors. Solving the dual formula allows new observations to be predicted using the following formula:

$$f(x) = \sum_{i \epsilon SV}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b \quad 4.10$$

Subject to the calculation of b:

$$b = y_i - \sum_{i=1}^{n} \alpha_i^* \; K(x_i, x_j) - \varepsilon, \forall i \; such \; that \; 0 < \alpha_i^* < C \quad 4.11$$

A random search method is used to obtain optimal hyperparameters for C and ε, as suggested by Villalobos-Arias et al. (2020). The tuning process is explained in appendix A.3.1 ML models tuning and hyperparameters.

4.3 eXtreme Gradient Boosting

XGBoost, or eXtreme Gradient Boosting, is an implementation of gradient boosting introduced by Tianqi Chen in 2014. Since then, the algorithm has received attention for contributing to winning

several ML competitions.[3] The former winner of the Avito Context Ad Clicks competition, Owen Zhang, said in an interview "when in doubt, use XGBoost" (Zhang O. , 2015). To understand XGBoost, it is relevant to have a brief review of boosting methods as described by Freund & Schapire (1996), and more specifically the gradient boosting machines of Friedman (2000).

In boosting methods, the first step is to fit a simple regression tree on the training data, to explain the dependent variable. In the next step, a new tree, referred to as a weak learner, is fitted on the residuals of the model from step one, instead of the response variable itself. The weak learner is then added to the full ensemble model, referred to as the strong learner, and the residuals are updated. New weak learners are sequentially fitted on the residuals of the strong learner, and the process is repeated over a fixed set of iterations. By adding new weak learners aiming to explain the residuals of the strong learner repeatedly, the boosting algorithm will slowly improve the strong learner in the areas where it does not perform well by seeking to find the relationship between the predictors and the current residuals.

XGBoost algorithm		Equation nr.
Data:	Training data and hyperparameters	
Initialize $f_0(x)$;		
for k = 1,2,…, M **do**		
	Compute the gradients g_i;	4.16
	Compute the hessians h_i;	4.17
	Determine the structure of the tree by choosing the splits that maximize gain;	4.21
	Determine the optimal leaf weights w*;	4.20
	Determine the weak learner;	4.18
	Add the weak learner to the ensemble model;	4.19
end		
Result:	Ensemble model from sum of weak learners	4.12

Algorithm 1: XGBoost, an implementation of the ensemble method known as gradient boosting.

[3] For examples, see https://github.com/dmlc/xgboost/tree/master/demo#machine-learning-challenge-winning-solutions

The gradient boosting machine (GBM) is a different formulation of the boosting method, although the basic principles are the same as the ones described above. In gradient boosting, the weak learners are sequentially fitted to correlate maximally with the negative gradient of an arbitrary differential loss function related to the strong learner. There are several possible loss functions to choose from in this setting, e.g., MSE for regression problems and log-loss for classification problems. This ability to choose the loss function allows for greater flexibility.

GBM's may, on the other hand, suffer from problems related to overfitting, and the number of iterations (e.g., number of weak learners) must be chosen carefully. The XGBoost approach, however, is an optimized implementation of gradient boosting where a variety of regularization options help avoid problems with overfitting. Furthermore, the XGBoost approach also allows for parallel processing to improve computation speed, tree pruning, and handling of missing data.

The XGBoost algorithm is explained in the steps above in Algorithm 1, following the methodology of Chen & Guestrin (2016) and Choi (2019). To perform these steps, however, there are a few mathematical equations that needs elaborating. Firstly, the final ensemble model is equal to the sum of all the weak learners.

$$f(x) = \sum_{k=0}^{M} b_k(x) \tag{4.12}$$

Where f(x) is the ensemble model, $b_k(x)$ are the weak learners, and k=1, 2, …, M is the number of weak learners. The aim is for the predictions f(x) to be as close to the real values y as possible. This is done by determining the weak learners that minimizes the following formula, where l is an arbitrary loss function:

$$\hat{b}_k(x) = argmin_b \sum_{i=1}^{n} l\left(y_i, f_i^{(k-1)}(x) + b_k(x_i)\right) + \Omega(b_k) \tag{4.13}$$

Where:

$$\Omega(b) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \quad 4.14$$

i = 1,2, …, n is the number of observations in the training dataset. $\Omega(b)$ is the regularization term (e.g., complexity of the tree), where γ is a pruning factor, j = 1, 2, …, T is the number of leaves in the tree structure, λ is the regularization term for the weights, and w_j are the weights of each leaf. The loss function can further be approximated by a second order Taylor expansion, and after removing the constants, the tree to be added to the ensemble in iteration t is the one that minimizes:

$$\hat{b}_k(x) \approx argmin_b \sum_{i=1}^{n} \left[g_i b_k(x_i) + \frac{1}{2} h_i b_k^2(x_i) \right] + \Omega(b_k) \quad 4.15$$

g_i and h_i are the gradients (first derivative) and the hessians (second derivative) of the loss function respectively:

$$g_i = \partial_{f_i^{(k-1)}} l\left(y_i, f_i^{(k-1)}\right) \quad 4.16$$

$$h_i = \partial^2_{f_i^{(k-1)}} l\left(y_i, f_i^{(k-1)}\right) \quad 4.17$$

I_k is a set of indices of all the observations assigned to leaf j for j=1, 2, …, T. Each of the weak learners can be determined by multiplying the leaf weights with the indices:

$$\hat{b}(x) = \sum_{j=1}^{T} w_j I[x \in R_j] \quad 4.18$$

After each iteration when a weak learner is determined, the additive strong learner is updated with the "boost" from the weak learner:

$$f_k(x) = f_{k-1}(x) + \hat{b}(x) \quad 4.19$$

R_j is the region of the datapoints in the relevant leaf j, and w_{jk} is constant in R_j. The optimal weights w_j^* for each leaf can be found by:

$$w_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad 4.20$$

Lastly, the binary splits in the weak learners are chosen by maximizing the gain. Gain is given by the equation 4.21, where L and R indicate the left and right branches of the split. The third term, which is subtracted, represents the score of the original leaf before the split. If the gain from performing the split is smaller than the pruning factor γ, it will not be added.

$$Gain = \frac{1}{2}\left[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda}\right] - \gamma \quad 4.21$$

In the AVM produced in this study, we used the "xgboost" package in R. The implementation requires choosing a value for several hyperparameters. The tuning process and chosen hyperparameters are described further in appendix A.3.1 ML models tuning and hyperparameters.

4.4 SHAP

The tradeoff between complexity and interpretability is a central aspect to consider in prediction modeling. Whereas the hedonic model may not allow for the most accurate predictions compared to the more complex machine learning techniques, the readability and general interpretability is higher. More precisely, the linear model allows for an intuitive understanding of how each of the predictors affect the prediction, simply by multiplying the coefficients with the predictor values. For ML models such as SVM and XGBoost, examining how the different predictors affects the final predictions is impractical. SHAP, or SHapley Additive exPlanations, is a useful framework for interpreting these individual predictor impacts for more complex models, based on Shapley values. Shapley values were first introduced by Shapley (1953) in the field of coalitional game theory. In 2010, Štrumbelj & Kononenko introduced Shapley values for ML models, based on the weighted average marginal effect the relevant variable implies when entering all the possible coalitions of predictors (Štrumbelj & Kononenko, 2010).

Mathematically, the Shapley value ϕ_i of a predictor i on a prediction $f_{S\cup\{i\}}(X_{S\cup\{i\}})$ can be found by using equation 4.22. F is a set of all the predictors, while S is any subset of F not including predictor i. The second part of the equation is the marginal contribution that predictor i has on the relevant prediction. The first part of the equation is the weight. The weight is given by dividing the number (i.e., $|S|!\,(|F|-|S|-1)!$) of different coalitions giving the same marginal contribution, by the total number of possible coalitions $|F|!$. ϕ_i is the local importance of the predictor on a single prediction, and to find the global importance ϕ_i is aggregated for all the observations in the training data.[4]

$$\phi_i = \sum_{S\subseteq F\setminus\{i\}} \frac{|S|!\,(|F|-|S|-1)!}{|F|!} \times \left[f_{S\cup\{i\}}(X_{S\cup\{i\}}) - f_s(X_S)\right] \qquad 4.22$$

There are several methods for measuring feature importance, and the SHAP framework is a relatively recent addition to this area (Lundberg & Lee, 2017). However, whereas Lundberg et al. (2018) prove other methods to be prone to inconsistency, the SHAP framework will always increase the measured importance of a predictor when its true impact increases. The main problem of utilizing Shapley values is, on the other hand, the computing time. In a model with k different predictions, there are 2^k possible coalitions. To overcome these computing challenges, the SHAP framework either estimates the Shapley values based on subsampling of observations and predictors for non-tree-based models (Lundberg & Lee, 2017), or uses the more effective TreeSHAP implementation for tree-based models such as the XGBoost (Lundberg, Erion, & Lee, 2018). We use the SHAP functions included in the "xgboost" package in R in this study.

[4] In a hypothetical model with the five predictors A, B, C, D, and E, the marginal contribution of i = C is equal for coalitions ABCDE, BACDE, ABCED and BACED. Instead of calculating the marginal contribution of C entering these coalitions four times, and each assigning them a weight 1/(5!), equation 4.22 only requires computing the marginal contribution once, before assigning it a weight of (2!*(5-2-1)!)/(5!) = 4/(5!).

5. Model outputs and purchasing rules

In the following chapter, model outputs are examined. The AVMs were trained as described in the methodology chapter, and more information on the tuning process is enclosed in A.3.1 ML models tuning and hyperparameters in the appendix. Firstly, the predictive accuracies of the different AVMs are assessed. Following an examination of feature importance, we then divide the apartments into groups. The predictive accuracies are then examined for the different groups before creating simple rules to prevent bidding on apartments from groups that are hard to price.

5.1 Model performance

We start by looking at the predictive performance of the three AVMs on the test data set, using the metrics described in appendix A.2.2 Performance evaluation. The linear model is, as expected, performing worse than the two ML models. There is approximately a two-percentage point difference between the ML AVMs and the hedonic price model in terms of Mean Average Percentage Error (MAPE), as seen left in Table 9. The XGBoost model performs slightly better than SVM, yielding better results for MAPE and PPE_{10}. However, the SVM has somewhat lower Root Mean Squared Error (RMSE). Generally, both ML models have satisfying accuracy, albeit lower than what would be expected from commercially used models. [5]

Model performance metrics on the training data are shown to the right in Table 9. As these observations are used to train the models, it is natural that the accuracy is somewhat higher than for the test observations. However, as the ML models were tuned to avoid overfitting, the differences are not too noticeable. The hedonic model performs relatively similar on the training observations, while the two ML models predict apartment prices slightly more accurately for these

[5] Previous literature suggests PE_{10} should be at least 65 %, and PE_{15} at least 85 % (Rossini & Kershaw, 2008; AVMetrics, 2018). Furthermore, Veros (2018), suggests that high quality AVMs has a PE_{10} as high as 80-90 %. Ecker, Isakson, & Kennedy (2020) argues that a test MAPE below 10 % is strong.

data points. In the rest of chapter, the training data is still in focus, to prevent information in the test data from influencing methodological and strategical choices.

Model performance test data	Mean absolute percentage error (MAPE)	Root mean squared error (RMSE)	Percentage Predicted Error (PPE), 10% (PPE)
Hedonic price model	11.25 %	867,824.0	56.17 %
Support vector machine (SVM)	9.22 %	659,194.2	65.13 %
eXtreme Gradiant Boosting	9.14 %	673,696.3	65.73 %

Model performance train data	Mean absolute percentage error (MAPE)	Root mean squared error (RMSE)	Percentage Predicted Error (PPE), 10% (PPE)
Hedonic price model	11.11 %	909,438.1	56.37 %
Support vector machine (SVM)	8.30 %	631,047.0	71.18 %
eXtreme Gradiant Boosting	7.31 %	425,739.8	74.87 %

Table 9: Predictive performance metrics for the three AVMs. Left panel gives metrics for the test set, while the right panel gives metrics for the training data. The evaluation metrics are described in greater detail in appendix section A.2.2 Performance evaluation. The two machine learning models perform somewhat better on the training data than the test set.

5.2 Feature importance

In the next section, the aim is to examine which of the independent variables hold most explanatory power. Feature importance can be evaluated in several ways, we chose to assess the standardized regression coefficients of the hedonic LAD model and the mean absolute SHAP value of the XGBoost AVM.

With interpretability being one of the main advantages of the LAD AVM, the variable coefficients provide an interpretation of how the individual variable impacts the dependent variable through a one-unit change. However, the coefficients are subject to scaling issues, as the one-unit impact on the natural logarithm of total price will be larger for features measured in big units. For this reason, a direct comparison of the absolute value of the regression coefficients does not provide an accurate evaluation of the importance of each variable. To overcome this challenge, we standardize the coefficients by multiplying each variable coefficient with the standard deviation of the relevant variable and dividing by the standard deviation of the intercept. The absolute values of some standardized coefficients are shown in Table 10.

Independent variable	Square meters primary rooms	Bedrooms	Build year	Fireplace	Floor	Garage
Standardized coefficient	0.555	0.095	0.076	0.051	0.048	0.038

Table 10: Standardized coefficients for different independent variables from the LAD AVM. Higher values indicate higher importance in the model predictions, implying that living area is the variable that explains most of the variation in apartment prices.

Note that in the table above, district and sales time were not included for readability purposes. These are categorical variables, and thus require dummy variable representations for interpreting the coefficients, which would give a total of 183 variables. Table 10 shows the six predictors with the highest absolute standardized coefficient values, after excluding district and sales time. The two excluded predictors did, however, both have multiple classes with a standardized coefficient higher than bedrooms[6], but none higher than square meter living area. Living area, district, time of sale, number of bedrooms, and build year are therefore seen as the most important independent variables in the hedonic prediction model.

Independent variable	Total interior living area	Build year	Bedrooms	Floor	Fireplace	Quiet
Mean absolute SHAP value	813,479	318,302	179,433	65,093	64,527	62,618

Table 11: Mean absolute SHAP values for different independent variables in the XGBoost AVM. High values indicate that the inclusion of the relevant variable in the AVM has a large contribution to the final price prediction. Total interior living area is thus the most important predictor variable for the XGBoost model, followed by build year.

As described in the methodology chapter, machine learning trades interpretability for complexity. For this reason, we introduced the framework of SHAP. Table 11 shows the variables with the highest mean absolute SHAP for the XGBoost model. Like the previous table, district and sales

[6] Grünerløkka, Bjerke and Søndre Nordstrand, for instance, had standardized coefficients of 0.2614, 0.2316 and 1.812 respectively.

time are excluded for readability purposes. Some districts did, however, have a noticeably higher mean absolute SHAP than the bedrooms variable, while no sales time category had a value exceeding 46,000.

Both the standardized coefficients and the SHAP values indicate that living area is the most important variable for describing variation in apartment prices by a clear margin. District and build year are other variables that stand out for both models. These variables will be examined further in the next section.

5.3 Creating subgroups

5.1 gave a summary of the predictive performance of the different AVMs. In this section, the performance is examined further on different subgroups of the data. The purpose of dividing the training set into different subgroups is to highlight potential systematic differences in predictive accuracy for certain types of apartments. Throughout the section, MAPE for the different groups is presented in tables Table 12 to Table 15. A group marked with red color indicates that the MAPE of these apartments is more than 5% higher than the average MAPE across all apartments. Likewise, yellow groups are within +/- 5% of the average, while green groups are more than 5% lower. Intuitively, the color codes imply that groups with red color are predicted with less accuracy than the average, and green with higher accuracy.

A natural place to start is with the districts, as these groups already exist in the dataset. In Table 12, MAPE is shown for the three AVMs for each district. Nordre Aker is among the districts with the worst MAPE. Nordstrand, Søndre Nordstrand, and Vestra Aker are other districts in which all three models generally struggle to provide accurate apartment price predictions. A common trait for all these underperforming districts is that they contain less than 4,000 observations. This suggests a plausible correlation between having few observations and worse predictive performance, although other small districts such as Grorud and Stovner prove that this is not always the case. There are also some differences between the three AVMs when it comes to how accurately apartments in the individual region are priced.

District	Number of observations	MAPE Hedonic model	MAPE XGBoost model	MAPE SVM model
Alna	4,344	0.114	0.067	0.075
Bjerke	2,647	0.140	0.074	0.096
Frogner	8,174	0.121	0.075	0.089
Gamle Oslo	9,537	0.094	0.073	0.078
Grorud	1,819	0.131	0.058	0.072
Grünerløkka	8,878	0.093	0.070	0.074
Nordre Aker	1,726	0.144	0.083	0.101
Nordstrand	3,360	0.122	0.083	0.097
Østensjø	2,861	0.098	0.073	0.086
Sagene	6,534	0.092	0.065	0.070
Søndre Nordstrand	2,265	0.131	0.089	0.100
St.Hanshaugen	7,594	0.104	0.077	0.086
Stovner	2,348	0.129	0.071	0.077
Ullern	3,124	0.136	0.076	0.091
Vestre Aker	2,713	0.138	0.077	0.093

Table 12: Predictive performance for the AVMs in different districts. Red cells indicate that the group is underperforming, yellow are around average, and green are overperforming. Nordre Aker, Vestre Aker, Nordstrand and Søndre Nordstrand are districts in which all three models struggle to predict apartment prices accurately.

As shown in section 5.2, living area and build year are other important variables for explaining the differences in apartment prices in the training data. These are, in contrast to the districts, not pre-divided in groups appropriate for examining the predictive performance. [7] For this reason, the variables are divided into groups based on the feature values. Firstly, the build year is divided into subgroups based on decades, starting in 1880. The apartments built before this year are placed in a single group, due to being few, as well as less affected by the build year[8]. There does, however, not appear to be any systematic patterns indicating that buildings from a certain year are

[7] The individual build years could be used as separate groups, although this would yield small-sized groups, highly impacted by certain outliers that are difficult to value.

[8] A building from 1820 does not necessarily imply worse apartments than a building from 1850, as both buildings will more than likely have gone through multiple restorations over the past decades.

significantly harder to predict for the AVMs, see Table 13. The difference in MAPE between the most accurate decades and the least accurate decades is smaller than the corresponding numbers found for the districts.

	Build year	Number of observations	MAPE Hedonic model	MAPE XGBoost model	MAPE SVM model
1	< 1880	898	0.109	0.074	0.084
2	1880-1889	1,132	0.095	0.070	0.076
3	1890-1899	7,038	0.101	0.070	0.076
4	1900-1909	1,862	0.103	0.074	0.079
5	1910-1919	1,192	0.117	0.073	0.092
6	1920-1929	1,981	0.117	0.073	0.088
7	1930-1939	4,659	0.107	0.072	0.081
8	1940-1949	1,656	0.091	0.066	0.076
9	1950-1959	8,280	0.108	0.071	0.081
10	1960-1969	6,080	0.107	0.066	0.076
11	1970-1979	5,250	0.143	0.072	0.084
12	1980-1989	5,560	0.126	0.084	0.097
13	1990-1999	3,723	0.107	0.075	0.082
14	2000-2009	12,319	0.105	0.075	0.084
15	2010 <=	6,294	0.117	0.077	0.090

Table 13: Predictive performance for the AVMs on groups of apartments built in different years. Red cells indicate the group is underperforming, yellow are around average and green are outperforming.

The training set observations are also stratified based on the size of the living area, as displayed in Table 14. These groups are made based on distance to the mean apartment size, measured in standard deviations. The reason for dividing the data into groups based on standard deviations from the mean, rather than specifical pre-defined values, is to make the study more generalizable for cities where average apartment size differs noticeably from this study on Oslo. The size of each group is 0.5 standard deviations. There appears to be systematic differences in training MAPE for apartments of different sizes, more specifically for the smallest and the largest apartments. Row 1, in addition to 5-11, generally has higher MAPE across the models, although the differences are noticeably smaller from row to row for the XGBoost model.

	Living area, standard deviation	Living area, square meters	Number of observations	MAPE Hedonic model	MAPE XGBoost model	MAPE SVM model
1	< -1.5	< 29.6	1,668	0.158	0.075	0.088
2	- 1.5 to -1.0	29.6 to 42	6,543	0.093	0.064	0.070
3	-1.0 to -0.5	42 to 54.2	14,384	0.088	0.065	0.070
4	-0.5 to 0	54.2 to 66.6	15,593	0.094	0.072	0.077
5	0 to 0.5	66.6 to 78.9	12,789	0.107	0.078	0.087
6	0.5 to 1.0	78.9 to 91.2	8,268	0.134	0.079	0.094
7	1.0 to 1.5	91.2 to 103.6	4,034	0.147	0.083	0.098
8	1.5 to 2.0	103.6 to 115.9	2,026	0.156	0.085	0.106
9	2.0 to 2.5	115.9 to 128.2	1,134	0.177	0.093	0.115
10	2.5 to 3.0	128.2 to 140.6	550	0.183	0.088	0.123
11	3.0 <	140.6 <	935	0.284	0.071	0.129

Table 14: Predictive performance for the AVMs in different sizes in living area. Red cells indicate the group is underperforming, yellow are around average and green are outperforming.

	Price prediction, standard deviation	Price prediction, price in thousands NOK	Number of observations	MAPE Hedonic model	MAPE XGBoost model	MAPE SVM model
1	< -1.0	< 1,954	4,168	0.130	0.082	0.095
2	-1.0 to -0.5	1,954 to 2,970	18,055	0.102	0.071	0.077
3	-0.5 to 0.0	2,970 to 3,985	20,489	0.100	0.069	0.076
4	0.0 to 0.5	3,985 to 5,001	10,945	0.105	0.076	0.083
5	0.5 to 1.0	5,001 to 6,017	6,035	0.119	0.079	0.091
6	1.0 to 1.5	6,017 to 7,033	3,523	0.128	0.077	0.093
7	1.5 to 2.0	7,033 to 8,048	1,954	0.138	0.075	0.097
8	2.0 to 2.5	8,048 to 9.065	1,123	0.164	0.078	0.109
9	2.5 to 3.0	9,065 to 10,080	593	0.174	0.077	0.113
10	3.0 <	10,080 <	1,039	0.234	0.071	0.133

Table 15: Predictive performance for the AVMs in different XGBoost predicted price groups. Red cells indicate the group is underperforming, yellow are around average and green are outperforming.

Lastly, the predictive performance is examined based on price groups, see Table 15. As price is not a value known to the iBuyer, the predicted price must be used instead. For illustrative purposes, the XGBoost price predictions are used to divide the observations into different groups. These groups are, as for the living area variable, also separated based on distance from the mean price measured in standard deviations. Like the living area, there is a higher MAPE for the smallest and

the largest apartments. Groups 2-4 have the smallest errors. The predicted price groups and the living area groups show similar systematic patterns, as expected since living area is the most important variable for determining the price prediction. On the other hand, the two are not perfectly correlated as there are some differences.

5.4 Purchasing rules

As described in the literature review, adverse selection problems in the iBuyer business model occur when the AVMs perform unsatisfyingly. In the next section, the aim is to create simple purchasing rules deciding which apartments not to bid on. The groups from the previous section provide a useful foundation for creating such rules, for the districts, build years, living area sizes, and predicted prices.

Table 14 indicated that the models generally perform poorly for groups 1 and 5-11. To improve the predictive performance of the models, the first rule should therefore exclude these groups. Such a rule involves only bidding on houses within -1.5 and 0 standard deviations from the average living area. For the dataset of Oslo housing transactions, this corresponds to a purchasing range of apartments with a living area between $30m^2$ and $67m^2$.

For the predicted price, Table 15 shows a similar pattern. Groups 1 and 5-10 on average have higher absolute percentage errors, and the second rule is therefore to only bid on houses within -1 and +0,5 standard deviations from the mean predicted price. This corresponds to a range between NOK 2,036,853 and 4,961,239 in Oslo. For apartments outside these two values, the models perform noticeably worse, and bids should therefore not be submitted.

For build year and district, some values also stand out negatively. Especially the apartments built in the years from 1980 to 1989, as well as apartments located in the districts of Søndre Nordstrand, Nordstrand, Nordre Aker, and Vestre Aker. Before making purchasing rules related to these variables, it is relevant to examine these hard-to-price groups further. An examination of the training data shows that 22% of all the apartments with a build year in the 1980s were in Søndre Nordstrand. In contrast, the same district only constitutes about 3% of the apartments in the

complete dataset. Furthermore, more than half of the buildings in Søndre Nordstrand were built in this decade. By the looks of it, the models' bad performances in both Søndre Nordstrand and the 1980s are thus related.

Purchasing rule		
1. District: No bids in Nordre Aker, Vestre Aker Søndre Nordstrand, and Nordstrand	**2.** Living Area: Only bid on houses within -1.5 and 0 standard deviations from mean. This corresponds to a range between 30 and 67m^2	**3.** Price prediction: Only bid on houses within -1 and +0.5 standard deviations from the mean. This corresponds to a range between NOK 2,036,853 and 4,961,269

Table 16: Three individual purchasing rules were made subject to systematic worse predictive performance for the AVMs for these apartments. The use of the rules implies not bidding on groups of apartments in which adverse selection is believed to be a greater problem.

The three other districts, on the other hand, have no clear connection to the decade of 1980. Apartments from these districts also appear to be similarly distributed across different living area and predicted price groups as the full dataset. However, the regions are among the ones with the least observations. Furthermore, the regions are aggregated into groups that still cover relatively large areas and may contain significant intra-regional differences. The low predictive performance may indicate that the combination of few observations and intra-regional differences makes the apartments difficult to price. The same goes for Søndre Nordstrand, which is the third smallest district in the dataset.

As the errors in the 1980s seem to be a consequence of the relevant decade having a large proportion of apartments located in Søndre Nordstrand, it is not deemed necessary to create an individual purchasing rule for excluding the decade. For districts, however, excluding the regions of Søndre Nordstrand, Nordstrand, Nordre Aker, and Vestre Aker might help improve the predictive performance and reduce problems with adverse selection, and thus constitutes the third and final purchasing rule. The rules are summarized in Table 16.

6. Results

In this chapter, the aim is to examine the financial impact of adverse selection and the purchasing rules for a hypothetical iBuyer. We begin disclosing our assumptions, before reporting the results from calculating profits for a hypothetical iBuyer with, and without, adverse selection and purchasing rules. The results are computed for apartments in the separated test set, which did not impact the training of the models nor the formulation of the purchasing rules.

6.1 Assumptions

Before diving into the results of our study, we lay out the assumptions our findings rest upon. Firstly, it is assumed that (I) the hypothetical iBuyer does not face competition from other corresponding companies. The iBuyer can therefore decide internally how large margins to take. The margin is the difference between the AVM predicted price and the actual bid, as shown in 6.2, and (II) is initially assumed to be 6%. Results with margins 3% and 9% can be found in appendix A.4.1 iBuyer Margins.

Average expected profit per apartment is calculated as a percentage of the bid, illustrated in equation 6.1, and we assume that (III) this is the KPI that the iBuyer will want to improve. We use percentage profits instead of absolute profits because the iBuyer cannot purchase all apartments in the market. Costs of financing, employee salaries, administration, and additional costs are not included in the equation. The aim of the iBuyer in this simplified illustration is to purchase apartments for less money than they are sold for, to generate a profit.

$$Profit = \frac{1}{n}\sum_{i=1}^{n}\frac{(Total\ price_i - Bid_i) * P(accept)_i}{Bid_i} \quad 6.1$$

Where:

$$Bid_i = Prediction_i * (1 - iBuyer\ margin) \quad 6.2$$

The probability of a bid being accepted, P(accept), depends on the bid from the iBuyer, as well as both the homeowner's perceived valuation and a probability distribution for how likely a seller is to accept a bid that is X% higher, or lower, than the perceived valuation. The iBuyer works as a substitute for selling through a traditional real estate agency assumed to have a provision of 2% of the sales price.[9] With this in mind, (IV) we use 98% of the list price in the sales ads as an initial proxy for what the seller believes (s)he will receive by selling through a broker, and thus the seller's perceived valuation of the property.

The final factor needed in the equation is the probability distribution. In section 2.2, adverse selection for iBuyers was introduced from a theoretical point of view. We presented the probable issue of homeowners being more likely to accept a bid based on a too-high predicted value, than a bid based on a correct value. Sellers are biased upwards when it comes to their own valuations of the properties, as given by the theories of Lovallo & Kahneman (2003), further implying that an overvalued dwelling is more likely to be bought than an undervalued one. The purpose of the accept-probability distribution, P(accept), is to include adverse selection in the profit equation above.

The probability function is assumed normally distributed around a mean replicating a convenience factor. The convenience factor determines where the centre of the probability distribution is located. This implies half of the homeowners that get a bid corresponding to their perceived valuation, after subtracting brokering commissions, minus the convenience factor will accept the bid. Initially, (V) the convenience factor is assumed to be 4% for the hypothetical iBuyer.

The width of the distribution is determined by its standard deviation. A narrow distribution means few sellers accept bids that are too low, and almost all accept too high bids. In contrast, a wider distribution implies more people accept low bids, and fewer accept high bids. As the accept probability distribution is a limited reflection of reality, three different scenarios of probabilities are created. The three scenarios are referred to as "pessimistic", "neutral", and "optimistic". (VI)

[9] Provision is normally between 1-3.6% for Norwegian real estate brokers, according to André Øren in DNB Eiendom (2021).

The neutral distribution has a standard deviation of 6%, while the pessimistic scenario has a narrower distribution with a standard deviation of 4% and the optimistic one wider with 8%.

The neutral scenario works as a benchmark, and this is the distribution of probabilities that is assumed most likely to reflect reality. As Figure 8 shows, this implies that about 1 in 6 homeowners will accept a bid that is 10% lower than their perceived market price after broker provision, while 3 in 4 will sell when they receive a bid equal to their valuation. 93% will accept an offer that is 5% higher than the seller's perceived valuation after broker provision, and nearly all bids more than 10% higher than the seller's valuation will be accepted.

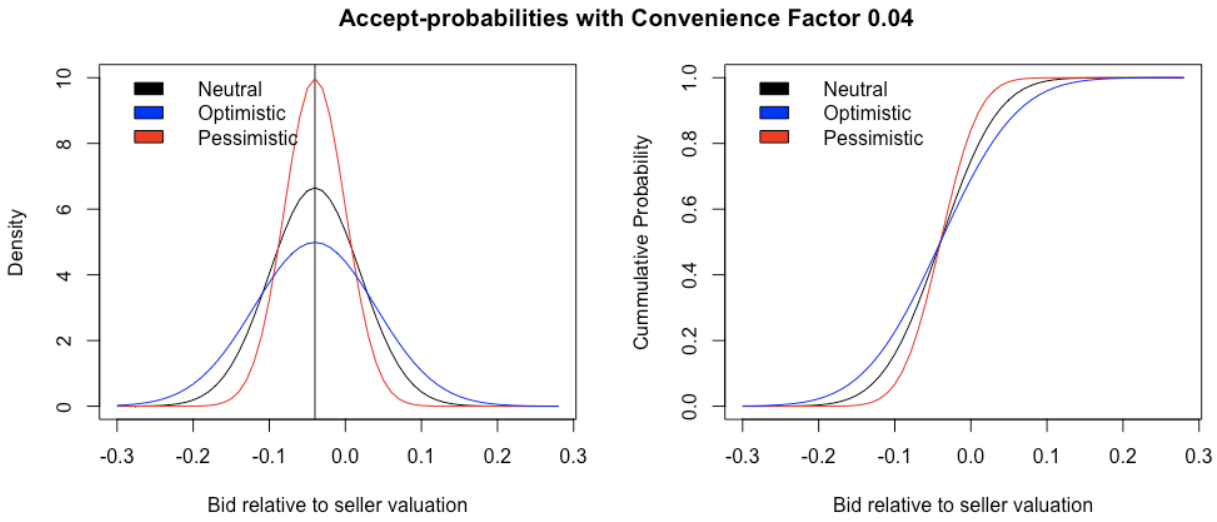

Figure 8. Left panel: Density distribution curves for the normally distributed bid acceptance probabilities. The x-axis indicates in percentage how much higher/lower the bid from the iBuyer is than the seller's perceived value. The curves are all centered around a convenience factor of 4%, implying that half of the sellers in the market would accept a bid that is 4% lower than their own valuation, in return for a quick sale. The three scenarios have curves that differ in width, with the neutral probability distribution having a standard deviation of 6%, and the pessimistic and optimistic distributions having standard deviations of 4% and 8% respectively. Right panel: The cumulative probabilities of a seller accepting a bid that is a certain percentage higher or lower than the seller's perceived valuation, for the same scenarios as the left panel. The three curves meet for bids 4% below the seller's valuation, where half of the sellers accept the bid for all three scenarios.

After having described the different factors in the profit equation, the next assumption is that (VII) the iBuyer will be able to sell the apartment for the same price as the apartment bought for in the dataset. Furthermore, the iBuyer can resell the apartments it purchases within a short enough time frame to avoid general price changes in the market. If the AVM models were able to completely predict this selling price, i.e., the predictions were 100% accurate, the average profit per apartment in an initial market without adverse selection would equal the bid margin of 6%.

To sum up, we assume (I) a hypothetical market with no competition, that (II) the iBuyer gives bids that are 6% lower than the AVM price prediction, and that (III) the company wants to improve average expected resale profit per apartment as a percentage of the bid. We assume that (IV) homeowners believe they will receive 98% of the list price after realtor provision, and by extension that this is what people value their home. Furthermore, (V) a convenience factor of 4% is assumed, implying that half of all offers that are 4% lower than the sellers' perceived valuations are accepted, and that (VI) the market in a neutral state is reflected by a normally distributed acceptance probability distribution with a standard deviation of 6%. Corresponding distributions in the case of pessimistic or optimistic scenarios, have standard deviations of 4% and 8% respectively. Lastly, it is assumed that (VII) the iBuyer can sell the apartments for the same price as a broker.

6.2 Profit calculations

The next step is to examine profits. In a scenario without adverse selection, we assume that all bids are accepted, regardless of how high or low the bid is compared to the seller's opinion of a fair price. Both undervalued and overvalued dwellings are bought, and the average expected resale profit is determined with P(accept) from equation 6.1 equal to 1 for all apartments. The top left panel in Table 17 displays the average expected profits per apartment in the scenario without adverse selection with an iBuyer margin of 6% and a convenience factor of 4%. As seen in the first row, where none of the purchasing rules are applied, the profits for all models exceed the iBuyer margin of 6%, with SVM and LAD giving a profit of 7.29% and 7.96% respectively, and XGBoost 6.29%.

iBuyer average profits

6% margin between predicted and bid price and 4% convenience factor, with list price as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	7.96%	6.29%	7.29%
Primary rooms	7.87%	6.27%	7.02%
Price	6.80%	5.89%	6.54%
District	7.47%	6.20%	7.14%
Primary rooms and price	7.76%	6.33%	6.97%
All	7.91%	6.35%	6.92%
Difference between All and None	-0.05%	0.06%	-0.37%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.19%	0.97%	1.21%
Primary rooms	1.59%	1.92%	2.07%
Price	0.43%	1.36%	1.47%
District	0.26%	1.04%	1.31%
Primary rooms and price	1.68%	2.02%	2.17%
All applied at once	1.82%	2.06%	2.24%
Difference between All and None	1.63%	1.09%	1.03%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.02%	0.86%	1.08%
Primary rooms	1.45%	1.86%	1.98%
Price	0.03%	1.28%	1.37%
District	0.10%	0.94%	1.19%
Primary rooms and price	1.54%	1.96%	2.08%
All applied at once	1.68%	2.00%	2.16%
Difference between All and None	1.66%	1.14%	1.08%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.38%	1.09%	1.35%
Primary rooms	1.75%	1.99%	2.16%
Price	0.59%	1.46%	1.59%
District	0.45%	1.16%	1.44%
Primary rooms and price	1.82%	2.09%	2.26%
All applied at once	1.97%	2.13%	2.33%
Difference between All and None	1.59%	1.04%	0.98%

Table 17: Average expected resale profits with six percent iBuyer margin and four percent convenience factor. The iBuyer margin represents the difference between the AVM predictions and the bids, while the convenience factor represents the loss at which half of all bids are accepted. The top left panel is a market with no adverse selection, where all bids from the iBuyer is accepted. The remaining three panels show profits in markets with different assumed probabilities for accepting bids. As seen, the introduction of adverse selection reduces profits noticeably, from 6.29-7.96% to 0.19-1.21% in the neutral scenario with no rules. Furthermore, the rules increase profits with 0.98-1.66 percentage points in the markets with adverse selection, clearly contributing to limiting the adverse selection loss.

Before implementing any purchase rules, the test set contains 16,981 observations. These observations represent the available apartments the iBuyer can bid on. When implementing all the purchasing rules, specified under 5.3, the number of apartments to bid on is decreased to 7142. In practice, this means that approximately 59% of all inquiries from homeowners wanting to receive an offer on their apartment get rejected, creating a more homogeneous data set.

Applying all the purchasing rules, without adverse selection, increases the profits for the XGBoost AVM iBuyer by 0.06 percentage points, but slightly decreases the profits for the other models. The rules are intended to remove groups of apartments the models struggle to predict accurately,

whether the apartments are over- or underpriced. Underpriced apartments are, however, profitable for the iBuyer when assuming that all bids are accepted, as homeowners may sell their houses far below market value. The slight decrease in profits from the rules for the LAD and the SVM AVMs indicates that the apartments removed from the purchasing rules generally are underpriced by these models. The purchasing rules do thus not appear to improve average expected resale profits per apartment in a market without adverse selection.

Having examined the profits of the iBuyer in a market where all offers are accepted regardless of price, the next step is to introduce adverse selection. It is therefore no longer assumed that all offers are accepted, but rather that the chance of an offer being accepted follows the probability distributions introduced in 6.1. This inclusion of adverse selection in the hypothetical iBuyer market further presents two interesting questions: (I) How does adverse selection affect the profits of the iBuyer, and (II) how the predefined purchasing rules influence this effect. We begin looking at the profits after implementing adverse selection.

The top right panel in Table 17 gives the new average expected resale profits after introducing adverse selection with the neutral probabilities. The profits are 0.19%, 0.97%, and 1.21% for the LAD, the XGBoost, and the SVM models respectively. This corresponds to reductions of 7.77, 5.32, and 6.08 percentage points compared to the scenario with no adverse selection. This sharp reduction implies that adverse selection poses a threat to the hypothetical iBuyer, as a large proportion of the iBuyer's profit margin is lost.

Considering the reduced profits after introducing adverse selection, it is also interesting to assess whether implementing the purchasing rules can help limit the loss. The change in profits from none to all purchasing rules applied is displayed in the bottom row in the top right panel of Table 17 with neutral accept probabilities. The lower panels show the same effect in the pessimistic and optimistic probability scenarios. The increase in profits surpasses 1 percentage point for practically all models and scenarios, suggesting that implementing the purchasing rules will increase the profits of the hypothetical iBuyer with a 6% iBuyer margin and a 4% customer convenience factor. Furthermore, the purchasing rule based on living area gives the highest isolated increase in profits

out of all the individual rules. The price rule gives the second-highest increase, followed by the district rule. However, a combination of the rules improves the average profit per apartment further.

To sum up, the sharp reduction in average expected resale profits per apartment thereby suggests that (I) adverse selection is problematic in the iBuyer business model. Furthermore, (II) the introduction of simple purchasing rules can help improve the average profit by around 1 percentage point. This is a noticeable increase, considering the low initial profits with no rules.

The results discussed until now are based on several assumptions, and the hypothetical iBuyer market can only be seen as a limited reflection of the real property market in Oslo. To assess whether the findings are robust, or a result of random chance, we continue by altering the initial assumptions to examine if the effects vary. One strength related to robustness was already shown in the previous section, as the effect of the purchasing rules on expected resale profits was displayed for three different AVMs and three different accept probabilities. Increased average profits from the purchasing rules were therefore not a model- or distribution-specific result.

A similar assessment can be made for different iBuyer bid margins. Table 23 and Table 24 in appendix A.4.1 iBuyer Margins display the corresponding results with iBuyer margins of 3% and 9%. The 9% iBuyer margin overall generates higher profits than the lower margins, although such a high margin leads to more bids being rejected. The 3% iBuyer margin has the opposite effect, with more bids being accepted and lower average resale profits. However, despite differing profits, the tables show that both the higher and the lower iBuyer margins give similar results as the initial assumptions in Table 17. Profits drop noticeably when adverse selection is implemented, and the purchasing rules improve the profits by around 1%.

Different customer convenience factor values were also tried, in case the initially assumed 4% does not represent reality. Table 25 and Table 26 in appendix A.4.2 Convenience factors show the results from calculating profits with 2% and 6% convenience factors respectively. Both additional convenience factor values give similar results as the benchmark model with 4%. The findings are thus robust to changes in how much customers value the services of the iBuyer.

A final robustness check was implemented related to the perceived valuation of the seller. In case the list price does not serve as a realistic proxy for the actual seller valuations, the profits were computed with a repeat sales proxy as an alternative. As explained in 3.3.5, repeat sales prediction of dwelling value utilizes the previous dwelling price and the general market indices to estimate the present value. Table 18 displays the results from using the repeat sales. Introducing adverse selection and purchasing rules have similar effects as under the list price assumption. However, with profits decreasing from 7.94% down to 1.38% on average across the AVMs, the iBuyer is slightly more profitable with this seller estimate. The effects of the purchasing rules are also lower.

iBuyer average profits

6% margin between predicted and bid price and 4% convenience factor, with repeat sales as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	8.94%	7.03%	7.94%
Primary rooms	8.31%	6.92%	7.58%
Price	7.77%	6.73%	7.25%
District	8.42%	6.92%	7.75%
Primary rooms and price	8.18%	7.00%	7.51%
All	8.29%	6.96%	7.42%
Difference between All and None	-0.65%	-0.07%	-0.52%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	1.03%	1.42%	1.70%
Primary rooms	1.72%	1.95%	2.07%
Price	1.11%	1.66%	1.79%
District	1.04%	1.44%	1.73%
Primary rooms and price	1.85%	2.10%	2.21%
All applied at once	1.94%	2.09%	2.24%
Difference between All and None	0.91%	0.67%	0.55%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.89%	1.31%	1.59%
Primary rooms	1.59%	1.86%	1.96%
Price	0.98%	1.57%	1.69%
District	0.90%	1.34%	1.62%
Primary rooms and price	1.73%	2.02%	2.11%
All applied at once	1.82%	2.00%	2.14%
Difference between All and None	0.93%	0.69%	0.55%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	1.19%	1.54%	1.82%
Primary rooms	1.87%	2.06%	2.19%
Price	1.25%	1.78%	1.91%
District	1.20%	1.56%	1.86%
Primary rooms and price	1.99%	2.20%	2.32%
All applied at once	2.08%	2.18%	2.35%
Difference between All and None	0.90%	0.65%	0.53%

Table 18: Average expected resale profits with a six percent iBuyer margin and a four percent convenience factor. Under these altered market assumptions, a repeat sales estimate is used as a proxy for the sellers' own valuations of the apartments. The top left panel is a market with no adverse selection, where all bids from the iBuyer is accepted. The remaining three panels show profits in markets with different assumed probabilities for accepting bids. As seen, the introduction of adverse selection reduces profits noticeably, from 7.03-8.94% to 1.03-1.70% in the neutral scenario with no rules. Furthermore, the rules increase profits with 0.53-0.93% in the markets with adverse selection, clearly contributing to limiting the adverse selection loss.

One potential reason for being more profitable in the repeat sales case is that around ¼ of the data is removed, due to not having any previous transactions. An examination of the removed data points shows that these are apartments that the AVMs generally struggle to price accurately. Furthermore, as seen in Figure 7, the repeat sales estimate has a larger error compared to the actual sales price. Too low seller valuations are profitable for the company, while too high ones do not affect the returns noticeably as iBuyer bids are likely rejected. Despite these differences in profitability, the general effects are similar as adverse selection largely reduces the profits while the purchasing rules help limit this loss.

Interestingly, the results of the hypothetical iBuyer case study do not seem to be a consequence of random chance, nor only being significant under specific circumstances. Adverse selection largely reduces the profits of the iBuyer, while the purchasing rules help limit the loss. These results are found across three different AVMs, three different accept-probability distribution widths, three different convenience factors, and three different iBuyer margins. The results thereby seem robust, although several assumptions are made to simplify the reality of an iBuyer market.

7. Discussion

After having presented the assumptions and average resale profits of a hypothetical iBuyer operating in Oslo, the results and corresponding implications can be discussed within both a theoretical and practical context. How consistent are the results with previous research and what impact could the findings have for iBuyers? These are questions that follow the results from the hypothetical iBuyer case and will be discussed further in this section.

As underlined in section 2.2, the research of adverse selection in the iBuyer business model is limited. From an objective point of view, considering the general well-established theories of adverse selection, the four criteria of Genesove (1993) hold. iBuyers facing adverse selection problems does therefore seem both reasonable and likely from a theoretical point of view. A rational homeowner will, according to Akerlof's original paper on adverse selection (1970), be incentivized by a high sales price, and by extension is more likely to accept an offer based on a too-high AVM prediction than a correct one. Furthermore, the iBuyer cannot know, at the time of the transaction, whether the apartment is a lemon (i.e., is overpriced by the AVM) or not. General papers on adverse selection do therefore imply that the business model of iBuyers may be prone to such problems. One of the few iBuyer market-specific studies done on adverse selection, by Buchak et al. (2020), strengthens this hypothesis further.

Whereas previous research thereby suggests that the introduction of adverse selection may prove problematic for iBuyers, it is also possible to examine this effect from a mathematical point of view. There are three value components in the profit equation: the seller's valuation, the AVM prediction, and the actual sales value that the iBuyer will get for the apartment once it is resold. The combination and correlation of these three value components are eventually deciding the impact of introducing adverse selection on the average resale profits. The magnitude of the impact on profits is determined by how well sellers know the values of their own apartments. In a case where sellers know this value relatively accurately, high bids will be accepted and low bids will be rejected, thus contributing to a reduction in profits for the iBuyer. In contrast, if sellers have inaccurate value estimates for their apartments, underpriced bids may still receive a high P(accept) weight if the seller values the apartment even lower than the AVM. Likewise, overpriced bids may

still be weighted with a low P(accept) if the owner values the apartment even higher than the AVM. Low bids being accepted, and high bids being rejected, is beneficial for the iBuyer.

From a mathematical point of view, the magnitude of the impact of introducing adverse selection on average resale profits is therefore dependent on how well both the iBuyer and the sellers can predict the actual sales price. The list price does, however, seem to reflect the actual sales price well. This implies that adverse selection will have a noticeable negative effect on the average resale profits, as was also expected from previous literature.

The results from Table 17 show that both previous literature and mathematical intuition were correct for the hypothetical iBuyer. When implementing adverse selection the profit margins decrease dramatically, by around 5-7% for the various models in the neutral scenario. The reduction in average profits varies somewhat across the different scenarios, iBuyer margins, convenience factors, and choices of seller valuation proxy, but the effect of introducing adverse selection is always noticeably negative.

When it comes to the purchasing rules, previous literature by Buchak et al. (2020) suggests that iBuyers can reduce adverse selection problems by avoiding the most heterogeneous and hard-to-predict objects. In essence, the purchasing rules from this study are simple solutions on how iBuyers can filter out some of these hard-to-price apartments. Furthermore, the approach involves increased screening. The iBuyer, as the party with the least information in the asymmetric information transactions, acts to identify and avoid groups that are hard to predict. Spence (1974) can thereby underline the hypothesis of Buchak et al. (2020), that avoiding hard to predict apartments may reduce the problems related to adverse selection for iBuyers.

Once again, the findings in the hypothetical iBuyer study are supported by previous research, as simple purchasing rules contribute to reducing adverse selection problems. In the neutral scenario from Table 17, implementing all the purchasing rules improves the average resale profits by between 1.03 and 1.63 percentage points across the different AVMs.

On the other hand, financial results are not only a consequence of higher expected average percentage profits. Imposing purchasing rules will imply that the iBuyer does not bid on potentially

profitable apartments, because they are too large, too expensive, or in the wrong district. How many of the rules should be used, will depend on the liquidity and the market share of the relevant iBuyer. If the company gets the opportunity to bid on more apartments than it can purchase due to limited liquidity, the iBuyer is best served by implementing strict rules for only buying the most profitable apartments. In such a case, combining the primary rooms, price, and district rules is beneficial. However, for an iBuyer receiving few inquiries, with available funds, a strategy of not bidding on many apartments might be counterproductive. Such an iBuyer should focus on purchasing enough profitable apartments, rather than only bidding on the most profitable ones. In this case, a softer rule strategy might be beneficial, for instance only implementing the rule related to square meters of living area.

The approach of identifying groups of apartments the AVM struggles to price accurately and creating rules for not bidding on these apartments, therefore underlines a trade-off between transaction volume and average expected resale profitability. Higher average profitability can be achieved by ruling out a large proportion of the apartments in the market, while purchasing more apartments can be achieved by accepting lower average profitability. Individual iBuyers will need to consider this trade-off, identify the relevant needs of the company, and find the optimal balance between the proportion of the market to submit bids for and required average profits.

Lastly, the findings from the hypothetical iBuyer case do not necessarily have to imply that the company must avoid the apartments restricted by the purchasing rules completely. Although these apartments are hard to predict for the AVMs, and thus should not receive automated offers, human expertise can be used as a supplement to improve the predictions. In addition to the trade-off between transaction volume and average resale profits, the individual iBuyers should therefore consider whether the apartments avoided by the purchasing rules never should receive a bid, or whether to supplement the AVM prediction with a human, physical appraisal before giving a bid. The iBuyer could also operate with a higher iBuyer margin for apartments that are hard to predict, to compensate for increased risk or potentially the cost of human appraisals.

8. Conclusion

The use of AVMs in real estate has grown increasingly important in recent years. Suddenly, homeowners could sell apartments in a matter of days rather than weeks and months. In a market full of rich data, AVMs constantly improve to make the bids as correct as possible. However, during the same period, several iBuyers reported disappointing financial returns. One of the large actors, Zillow, pulled out of the automated bid segment. A question that follows is how iBuyers, with advanced prediction models and services that are greatly in demand, still are not able to produce satisfying returns.

Buchak et al. (2020) suggested adverse selection to be a problem in the iBuyer business model, a hypothesis that may also be derived from the well-known theories of Akerlof (1970) and Genesove (1993). Furthermore, Buchak et al. (2020) points out that iBuyers, in order to limit these problems, might tend to purchase the most liquid and easy-to-price homes. Based on previous literature, the purpose of this paper has been to examine the effects of adverse selection for iBuyers, and investigate whether simple strategic changes in the use of AVMs can help reduce the potential threat.

The study was conducted by creating three different AVMs, before examining the predictive performance of each of these models for different groups of apartments. These groups were made based on the most important predictor variables. Consequently, a set of simple purchasing rules were made to prevent iBuyers from purchasing apartments in groups with bad performance. Thereafter, a hypothetical iBuyer case was examined using the aforementioned AVMs, where average expected resale profits were computed both with and without adverse selection and purchasing rules.

As expected by the previous literature, the paper finds that the average expected resale profits per apartment drop dramatically when introducing adverse selection in the hypothetical iBuyer market. In the baseline scenario, the inclusion of adverse selection makes the average iBuyer profit decrease from between 6.29-7.96% per apartment, to a new profit of 0.19-1.21%. All profit calculations are done on test set apartments not used in the training of the models. Furthermore, similar effects were

found when altering the assumptions of the hypothetical market. By using the different purchasing rules, average profits increased with between 1.03 and 1.63 percentage points per apartment, in the neutral scenario. This effect was also robust to alterations in initial market assumptions and using repeat sales as an additional proxy. The paper thereby does not only find that average resale profits for iBuyers are highly affected by adverse selection, but also that simple purchasing rules can help reduce these problems.

Whereas the topic of creating optimal property market AVMs is well-explored in current literature, the subject of adverse selection for iBuyers is less studied. The findings of this paper thus highlight several new possibilities for future research. The study of the hypothetical iBuyer is prone to multiple assumptions. Although the results are robust when altering these assumptions, further analysis of accept probabilities in iBuyer property markets would be an interesting addition to the field of research. Techniques for developing more sophisticated purchasing rules could also be examined further. As suggested by the third criteria of Genesove (1993), adverse selection follows a situation where the party with the least information determines the price. Another possible way of dealing with adverse selection could be to alter the iBuyer business model and let the party with the most information determine the price. In practice, this could imply that the seller suggests a price, and the iBuyer chooses whether to accept. Further examining of profits in such a scenario could be of interest in future research.

References

Akerlof, G. A. (1970). The market for «Lemons»: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84 (3)*, pp. 488–500.

Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2020). Why is intermediating houses so difficult? Evidence from iBuyers. *NBER Working paper series*.

Chai, T., & Draxler, R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development. 7.* , pp. 1247-1250.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. pp. 785-794.

Choi, D.-K. (2019). Data-Driven Materials Modeling with XGBoost Algorithm and Statistical Inference Analysis for Prediction of Fatigue Strength of Steels. *International Journal of Precision Engineering and Manufacturing, 20*.

CoreLogic. (2011). *Automated valuation model testing.* Retrieved from CoreLogic: www.corelogic.com/downloadable-docs/automated-valuation-model-testing.pd

Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory, IT-13(1)*, pp. 21–27.

Delprete, M. (2020, May 5). *The 2020 iBuyer Report.* Retrieved from Mikedp.com: https://www.mikedp.com/ibuyer-report

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support Vector Regression Machines. *Advances in Neural Information Processing Systems 9*, 155-151.

Ecker, M., Isakson, H., & Kennedy, L. (2020). An Exposition of AVM Performance Metrics. *Journal of Real Estate Practice and Education. 22.*, pp. 22-39.

Edvardsen, K. (2021, February 12). *DNB Eiendom*. Retrieved from Hva koster det å selge en bolig?: https://dnbeiendom.no/altombolig/kjop-og-salg/tips-til-selgere/hva-koster-det-a-selge-bolig

Emons, W., & Sheldon, G. (2007, June 14). The Market for Used Cars: New Evidence of the Lemons Phenomenon. University of Bern, Department of Economics.

Financial Times. (2021, November 3). *Zillow: the models underneath a housing hedge fund did not hold*. Retrieved from Financial Times: https://www.ft.com/content/cb3f4448-990f-4052-b4a9-0938b147e402

Fix, E., & Hodges, J. L. (1951). Discriminatory analysis, nonparametric discrimination: Consistency properties. *Technical Report 4, USAF School of Aviation Medicine*.

Fortelny, A., & Reed, D. R. (2005). The increasing use of Automated Valuation Models in the Australian mortgage market. *Australian property journal, 36(6),* pp. 681-685.

Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm.

Friedman, J. H. (2000). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*, pp. 1189–1232.

Genesove, D. (1993). Adverse Selection in the Wholesale Used Car Market. *Journal of Political Economy*, *vol. 101*, *no. 4*, pp. 644–65.

Gores, P. (2019, October 25). *iBuyers use technology to take the time and hassle out of home selling. And they could be in Milwaukee soon.* Retrieved from Milwaukee Journal Sentinel: https://eu.jsonline.com/story/money/2019/10/25/ibuyers-who-buy-homes-less-hassle-may-coming-milwaukee/2455392001/

Helbich, M., Brunauer, W., Hagenauer, J., & Leitner, M. (2013). Data-Driven Regionalization of Housing Markets. *Annals of the Association of American Geographers*, 871-889.

Ho, W. K., Tang, B.-S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research 38.1*, pp. 48–70.

Huang, Y. (2019). Predicting home value in California, United States via machine learning modeling. *Statistics, optimization and information computing, Vol. 7 No. 1*, pp. 66-74.

Jahanshiri, E., Buyong, T., & Shariff, A. R. (2011). A Review of Property Mass Valuation Models. *Pertanika Journal of Science & Technology, 19(1)*, pp. 23-30.

Janssen, C., Söderberg, B., & Zhou, J. (2001). Robust estimation of hedonic models of price and income for investment property. *Journal of Property Investment & Finance*, 342-360.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting, Elsevier, vol. 32(3)*, pp. 669-679.

Koenker, R., & Basset, G. (1978). Regression Quantiles. *Econometrica*, 33-50.

Kok, N., Koponen, E.-L., & Martínez-Barbosa, C. A. (2017). Big Data in Real Estate? From Manual Appraisal to Automated Valuation. *The Journal of Portfolio Management Special Real Estate Issue 2017, 43 (6)*, pp. 202-211.

Kontrimas, V., & Verikas, A. (2011). The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing, Vol. 11 No. 1*, pp. 443-448.

Lam, K. C., Yu, C. Y., & Lam, C. K. (2009). Support vector machine and entropy based decision support system for property valuation". *Journal of Property Research, Vol. 26 No. 3*, pp. 213-233.

Lovallo, D., & Kahneman, D. (2003, July). *Delusions of Success: How Optimism Undermines Executives' Decisions.* Retrieved from Harvard Business Review: https://hbr.org/2003/07/delusions-of-success-how-optimism-undermines-executives-decisions

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems 30*, pp. 4765–4774.

Lundberg, S. M., Erion, G. G., & Lee, S.-I. (2018). Consistent Individualized Feature Attribution for Tree Ensembles.

Marquand, B. (2021, September 3). *What Is an iBuyer?* Retrieved from Nerdwallet.com: https://www.nerdwallet.com/article/mortgages/understanding-ibuyers

Mayer, M., Bourassa, S. C., Hoesli, M., & Scognamiglio, D. (2019). Estimation and updating methods for hedonic valuation. *Journal of European Real Estate Research, Vol. 12 No. 1*, pp. 134-150.

Mooya, M. (2011). Of Mice and Men: Automated Valuation Models and the Valuation Profession. *Urban Studies, 48(11)*, pp. 2265–2281.

Mu, J., Wu, F., & Zhang, A. (2014). Housing value forecasting based on machine learning methods. *Abstract and Applied Analysis, Vol. 4*, pp. 1-7.

Nobel Prize Outreach AB. (2001, October 10). *Press release*. Retrieved from NobelPrize.org: https://www.nobelprize.org/prizes/economic-sciences/2001/press-release/

Ogundimu, E. O., Altman, D. G., & Collins, G. S. (2016, August). Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of Clinical Epidemiology*, pp. 175-182.

Olaussen, J. O., Oust, A., & Sønstebø, O. J. (2018). Bidding Behavior in the Housing Market under Different Market Regimes. *Journal of Risk and Financial Management*.

Oslo Kommune. (2017). *Oslo Bydelskart.* Retrieved from Oslo Kommune: https://www.oslo.kommune.no/getfile.php/13206469-1490274697/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Statistikk/Geografiske%20inndelinger/Oslo_Bydelskart_20170221_A3.pdf

Oslo Kommune. (2021, October 4). *Geografiske inndelinger*. Retrieved from Oslo Kommune: https://www.oslo.kommune.no/statistikk/geografiske-inndelinger/#gref

Oust, A. (2012). House price indices Oslo 1970-2011.

Palm, P. (2015). "The office market: a lemon market? A study of the Malmö CBD office market". *Journal of Property Investment & Finance, Vol. 33 No. 2*, pp. 140-155.

Reed, R. (2008). The use and misuse of AVMs. *Australian and New Zealand property journal, vol. 1, no. 8*, pp. 651-656.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 34-55.

Rossini, P., & Kershaw, P. J. (2008). Automated valuation model accuracy: some empirical testing.

Shapley, L. S. (1953). A Value for n-Person Games. *Contributions to the Theory of Games*, pp. 307–317.

Sloan, L., & Aarbakke, M. (2016). *En trygg voksen.* Oslo: Oslo Kommune, Bydel Alna.

Sommervoll, Å., & Sommervoll, D. E. (2018). Learning from man or machine: Spatial fixed effects in urban econometrics. *Regional Science and Urban Economics*, 239-252.

Spence, M. (1974). Market Signalling: Information Transfer in Hiring and Related Screening Processes. Cambridge: Harvard University Press.

Statistics Norway. (2018, May 30). 7 av 10 Oslo-husholdninger bor i blokk.

Statistics Norway. (2020). *Dette er Norge 2020.* Statistics Norway.

Statistics Norway. (2021, October 1). *Prisstigning for brukte boliger siste ti år*. Retrieved from Statistics Norway: https://www.ssb.no/bygg-bolig-og-eiendom/faktaside/bolig

Statistics Norway. (2021, April 07). *Statistikkbanken.* Hentet fra Dwellings, by type of building and number of toilets (M) 2007 - 2021: https://www.ssb.no/en/statbank/table/06516

Stiglitz, J. E. (1975). The theory of «screening», education, and the distribution of income. *American Economic Review 65 (3)*, pp. 283–300.

Stock, J. H., & Watson, M. W. (2019). *Introduction to Econometrics.* New York: Pearson.

Štrumbelj, E., & Kononenko, I. (2010). An Efficient Explanation of Individual Classifications Using Game Theory. *Journal of Machine Learning Research 11*, pp. 1–18.

Sørgjerd, C., Murray, S. M., & Hager-Thoresen, F. (2020, November 28). *Se oversikten: Så store er prisforskjellene innad i Oslo*. Retrieved from Aftenposten.no: https://www.aftenposten.no/oslo/i/aPOM4A/se-oversikten-saa-store-er-prisforskjellene-innad-i-oslo

Tretton, D. (2007). Where is the world of property valuation for taxation purposes going? *Journal of Property Investment & Finance, 25(5)*, pp. 482-514.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *American Association for the Advancement of Science*, 1124-1131.

Villalobos-Arias, L., Quesada-López, C., Guevara-Coto, J., Martínez, A., & Jenkins, M. (2020). Evaluating hyper-parameter tuning using random search in support vector machines for software effort estimation. *PROMISE 2020: Proceedings of the 16th ACM International Conference on Predictive Models and Data Analytics in Software Engineering* (pp. 31-40). New York: Association for Computing Machinery.

Waller, B. D., Riley, N. F., & Greer, T. H. (2001). An appraisal tool for the 21st Century: Automated Valuation Models. *Australian Property Journal, 36(7)*, pp. 636–641.

Wang, X., Wen, J., Zhang, Y., & Wang, Y. (2014). Real estate price forecasting based on SVM optimized by PSO. *Optik, Vol. 125 No. 3*, pp. 1439-1443.

Wilson, C. (1989). Adverse Selection. In J. Eatwell, M. Milgate, & P. Newman, *Allocation, Information and Markets*. London: Palgrave Macmillan.

Yoo, S.-H. (1999). A robust estimation of hedonic price. *Applied Economics Letters*, 55-58.

Zhang, O. (2015). Avito Winner's Interview: 1st place, Owen Zhang.

Zhang, S. H. (2012). Application of Support Vector Machine in determination of real estate price. *Advanced Materials Research, Vol. 461*, pp. 818-821.

A Appendix

A.1 Data

A.1.1 Correlation between variables

Studies were undertaken to discover potential multicollinearity in the data set. Between the numerical variables, some variables are highly correlated. Number of bedrooms and living area in square meters are highly correlated, at 75%. Bathrooms and living area are correlated at 39%, and bathrooms and bedrooms are correlated at 26%, as showed in Figure 9. Of the binary variables, child friendly and Balcony seems to be moderately correlated, at 45%. Quiet and child friendly are also moderately correlated, at 46%.

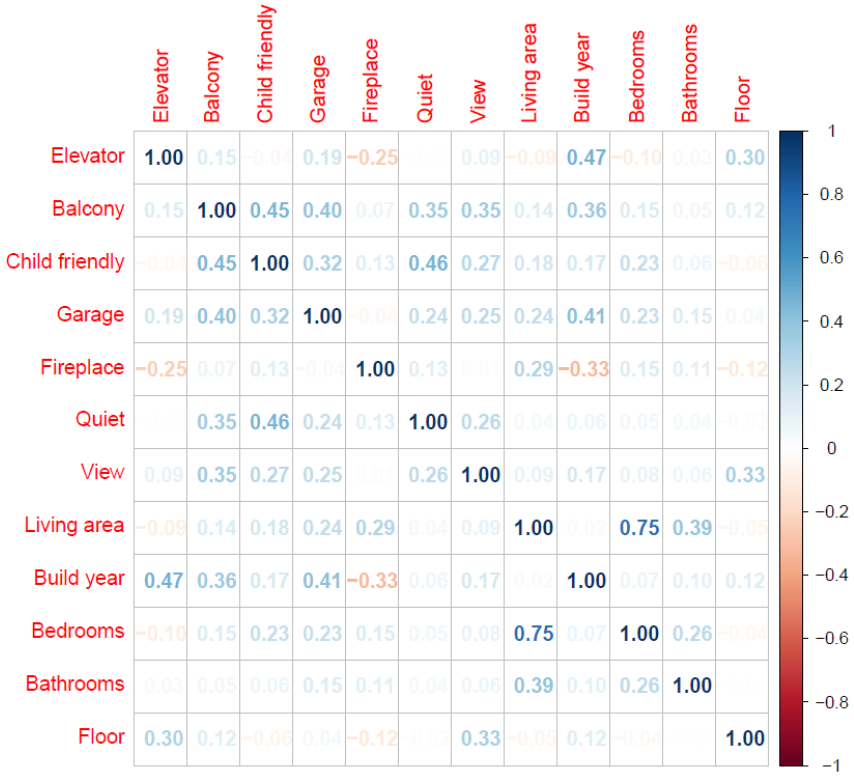

Figure 9: Correlation matrix. A number close to 1 or -1 indicate a high correlation, and numbers closer to 0 indicate low or no correlation.

a

In contrast, a Variance Inflation-test (VIF) on the variables indicates that no serious problems related to multicollinearity can be found between the variables. The output from the VIF-test is displayed in Table 19.

Variable	GVIF	Degrees of freedom	GVIF^(1/(2* Df))
Living area	2.95	1.00	1.72
Bathrooms	2.46	1.00	1.57
Floor	1.24	1.00	1.11
Buildyear	2.06	1.00	1.15
Elevator	1.49	1.00	1.43
Balcony	1.63	1.00	1.22
Child friendly	1.64	1.00	1.28
Garage	1.65	1.00	1.28
Fireplace	1.51	1.00	1.28
View	1.40	1.00	1.23
Renovation	1.01	1.00	1.18
District	1.91	14.00	1.02
Sales time	1.28	158.00	1.00

Table 19 Variance Inflation Test (VIF) on the training data, indicating no threat of multicollinearity for the Least Absolute Deviation model (LAD).

A.1.2 Price versus space plot

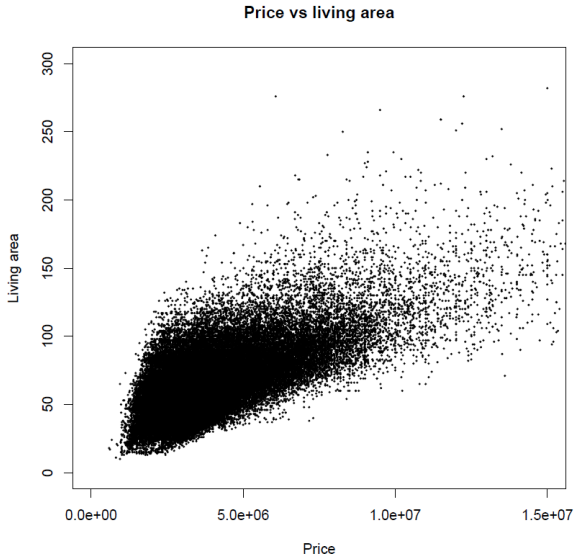

Figure 10 Plot of price on living area, indicating a non-linear relationship between higher prices and square meters of living area.

As part of the descriptive study of the data set, price versus living area indicated a non-linear relationship, arguing for a log transformation of the hedonic model.

A.1.3 District Labelling with K-Nearest-Neighbors

In section 3.3.3, where the apartments are placed in their relevant administrative districts, we mentioned how postal codes were used for this spatial labelling. However, some of the postal codes belonged to two or more districts, and did therefore have to be labelled in another way. This constitutes a classification problem where the apartments with only one possible district value was used as a training dataset, D, to fit a model to label the remaining target apartments. In a classification problem as such, the test error is minimized by labelling each observation into the class by which the probability of it belonging is highest, given the relevant predictor values. In other words, a target observation with predictor values x_0 should be placed into the class j for which the probability in equation A.1 is largest. Equation 1 is known as the Bayes classifier.

$$\Pr(Y = j | X = x_0) \qquad \text{A.1}$$

However, the probabilities in the Bayes rule are not previously known in the case of apartments and corresponding districts, and thus need to be determined before classification is possible. For this purpose, the KNN (K nearest neighbors) method was chosen. KNN was first introduced by Fix & Hodges (1951), before getting expanded by Cover & Hart (1967).

KNN algorithm		Equation nr.
Data:	Training data, D, a target object, x_0, being a vector of predictor values for the object that should be classified, a set of possible labels, L, and the number of the nearest neighbors to consider, K.	
for each object $x_i \in D$ do		
	Compute the Euclidian distance between the training observation, x_i, and the target x_0;	A.2
end		
Select $N_0 \subseteq D$, the set (neighborhood) of K nearest training points to x_0;		
Compute the probabilities of x_0 belonging to the class j;		A.3
y_{x0} is is the class j that x_0 has highest probability of belong		A.4
Result:	A label, $y_{x_0} \in D$, for the target observation with predictor values x_0.	

Algorithm 2: Class labelling with the K-Nearest-Neighbor algorithm

The KNN algorithm uses the K nearest points in the dataset, measured in Euclidian distance, to label an observation with the predictor values x_0. The first step is thus to compute the distances $d(x_0, x_i)$ between the observation that needs labelling, x_0, and the apartments in the training data, $x_i \in D$. In equation A.2, V refers to the independent variables to consider when computing the distances, and k is a variable in V. In the case of apartments and administrative districts, latitude and longitude of the apartments are the two independent variables, V.

$$d(x_0, x_i) = \sqrt{\sum_{k \in V} (x_{0_k} - x_{i_k})^2} \qquad \text{A.2}$$

d

After having computed the distances, the next step is to determine a dataset consisting of the K datapoints with closest predictor values to x_0, the set denoted as N_0. Consequently, the algorithm computes the probability of the target observation, x_0, belonging to class j. This is done by dividing the sum of apartments in the neighborhood, N_0, belonging to class j, by the total number of apartments in the neighborhood, K. The mathematical formulation for finding the probabilities is given by equation A.3, where I is a binary indicator taking the value 1 if the district of apartment i, y_i, is equal to j, and 0 if not.

$$\Pr(Y = j | X = x_0) = \frac{1}{K} \sum_{i \in N_0} I(y_i = j) \qquad \text{A.3}$$

When the probabilities are computed, the Bayes rule is applied, and the target observation is labelled, y_{x_0}, in the class it has highest probability of belonging to. In equation A.4, y_{x_0} is the class of the target apartment with predictor values x_0, j is a label in L, i is an observation in N_0, and y_i is the class of apartment i.

$$y_{x_0} = \underset{j \in L}{argmax} \ \frac{1}{K} \sum_{i \in N_0} I(y_i = j) \qquad \text{A.4}$$

Different values of K were tried, and eventually K=10 was chosen. In practice, this implies that each of the non-labeled apartments were placed in the district of which most of its 8 closest geographical training set neighbors were located in.

A.2 Methodology

A.2.1 Litterature on AVMs

Literature	*Findings*
Zhang S.H., 2012	SVM predicts the value of houses in China with greater accuracy than ANN in a case with small training data

Mu et al., 2014	SVM outperforms PLS for dealing with non-linearity in forecasting the value of Boston suburb houses
Wang et al., 2014	Proposed SVM model shows good forecasting performance for real estate in Chongqing, China
Huang, 2019	SVM provides a "dramatic improvement" compared to linear regression and other ML-methods[10] for predicting the value of houses in California, US
Lam et al., 2009	SVM provides better property valuations than MRA and ANN in case studies of Hong Kong real estate
Kontrimas et al., 2011	SVM clearly outperformed OLS- and MLP-based models for predicting real estate values in Lithuania
Kok et al., 2017	XGBoost is superior to OLS and RF for predicting the value of multifamily homes in California, Florida and Texas, US
Mayer et al., 2019	Gradient Boosting outperforms five other estimation-methods[11] in terms of accuracy, for predicting the value of single-family houses in Switzerland
Ho et al., 2021	Gradient Boosting achieves better accuracy than SVM for appraisal of Hong Kong property[12]

Table 20: Previous literature on the use of SVM and XGBoost for predicting housing prices

A.2.2 Performance evaluation

To measure the performance of the AVMs, we here introduce the metrics used in the study. There are several ways of quantifying model performance, and the choice of metrics reflects different sides of the problem we want to solve. The first of the performance measures we will use in this study, is the Mean Absolute Percentage Error (MAPE), as shown in equation A.5. MAPE is chosen

[10] Decision trees, boosting, and random forest

[11] Linear least squares, robust regression, mixed-effects regression, random forests, and neural networks

[12] However, Ho et al. (2021) still acknowledge SVM as a useful algorithm due to for instance its capability to produce relatively accurate predictions within a tight time constraint

as it is widely used in model performance evaluation and has an intuitive understanding (Kim & Kim, 2016). MAPE gives a measure of how large the error between the AVM valuation and the real selling value is on average, as a percentage of selling price. In other words, a 5% MAPE means that the AVM predictions on average are 5% different from the actual sales value, in absolute terms.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \quad \text{A.5}$$

The next performance metric chosen is the Root Mean Squared Error (RMSE), as shown in A.6. Although it is less intuitive to understand whether a given RMSE value is satisfying or not, the measure serves well for comparison of different models. Furthermore, an interesting aspect of the RMSE is that the squared error term penalizes large errors more (Chai & Draxler, 2014). In other words, will a prediction with an error of 10% have more than twice as big impact on the RMSE than a prediction that misses by 5%. In terms of avoiding adverse selection in the iBuyer business model, it is reasonable to assume that this is the case, as the largest errors are substantially more problematic than the smaller ones.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad \text{A.6}$$

The last metrics that are used to evaluate the performance of the models are called Percentage Predicted Error (PPE) buckets. These buckets indicate how large proportion of the predictions fall within a certain percentage difference from the actual value. As an example, PPE_{10} indicates how large the proportion of predictions that are within +/- 10% of the sales value is. The PPE buckets are symmetric, and thus account for negative and positive errors outside the pre-defined percentage threshold equally. Furthermore, the PPE buckets only measure how many of the observations are within the relevant bucket, and not the magnitude of the errors with bad predictions (Ecker, Isakson, & Kennedy, 2020).

In a study of predictive performance of AVMs with focus on preventing adverse selection problems, however, it can be argued that the proportion of predictions having errors larger than a

g

given threshold, and thus giving rise to adverse selection, is more interesting than the magnitude of the errors in themselves. This is because good strategical use of the AVM will involve identifying and choosing to not purchase the types of apartments that the model fails to predict accurately. PPE buckets are therefore central in the discussion and results of this study, and PPE_{10}, PPE_{15}, and PPE_{20} are chosen as these metrics are most used in AVM performance evaluation (CoreLogic, 2011).

A.3 Models

A.3.1 ML models tuning and hyperparameters

Both machine learning models in the paper have hyperparameters that affects the training process and predictive performance. To ensure the use of appropriate hyperparameter values, a 5-fold cross validation method was used to perform a random search in pre-defined ranges of possible values. In contrast to performing a grid search, where all possible combinations of hyperparameter values within a grid are tested, the random search method tried 100 different random combinations. In the end, the combination that gave the lowest cross validation test RMSE was returned and used as estimates for the theoretical optimal hyperparameter values. For the SVM model, a radial kernel was used. The final optimal values are to be seen in Table 21 and Table 22.

XGBoost model

Hyperparameter	Description	Value
Eta	Learning rate	0.1
Max depth	Maximum depth of each tree	10
nround	Number of boosting iterations	560
min_child_weight	Minimum number of observations needed in each node	10
colsample_bytree	Size of subsample of columns when training a new tree	0.95
subsample	Size of subsample of rows when training a new tree	0.60

Table 21: Optimal values of hyperparameters after tuning XGBoost.

SVM model

Hyperparameter	Description	Value
cost	Penalty factor	148
epsilon	Width of insensitivity zone where classification errors are not penalized	0.089

Table 22: Optimal values of hyperparameters after tuning SVM.

A.4 Results

A.4.1 iBuyer Margins

To ensure robust results, profits are also computed using two additional values for the iBuyer margin. Table 23 displays results with a 3% margin and Table 24 with a 9% margin.

iBuyer average profits

3% margin between predicted and bid price and 4% convenience factor, with list price as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	4.61%	3.00%	3.98%
Primary rooms	4.53%	2.98%	3.71%
Price	3.50%	2.61%	3.24%
District	4.15%	2.92%	3.83%
Primary rooms and price	4.43%	3.05%	3.66%
All	4.57%	3.06%	3.62%
Difference between All and None	-0.04%	0.07%	-0.36%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-1.25%	-0.66%	-0.33%
Primary rooms	0.14%	0.25%	0.49%
Price	-1.11%	-0.32%	-0.14%
District	-1.18%	-0.61%	-0.25%
Primary rooms and price	0.20%	0.36%	0.59%
All applied at once	0.37%	0.40%	0.65%
Difference between All and None	1.62%	1.07%	0.98%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-1.25%	-0.66%	-0.33%
Primary rooms	0.14%	0.25%	0.49%
Price	-1.11%	-0.32%	-0.14%
District	-1.18%	-0.61%	-0.25%
Primary rooms and price	0.20%	0.36%	0.59%
All applied at once	0.37%	0.40%	0.65%
Difference between All and None	1.62%	1.07%	0.99%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-1.01%	-0.52%	-0.16%
Primary rooms	0.31%	0.30%	0.58%
Price	-0.92%	-0.22%	-0.02%
District	-0.95%	-0.47%	-0.09%
Primary rooms and price	0.36%	0.41%	0.66%
All applied at once	0.53%	0.45%	0.72%
Difference between All and None	1.54%	0.97%	0.88%

Table 23 iBuyer average profits with 3% margin between predicted and bid price. Run as part of a robustness check on our results. The table indicate that a lower margin yields similar results as with 6% margin, thus strengthening our findings.

iBuyer average profits

9% margin between predicted and bid price and 4% convenience factor, with list price as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	11.51%	9.79%	10.83%
Primary rooms	11.42%	9.77%	10.55%
Price	10.32%	9.38%	10.05%
District	11.01%	9.70%	10.68%
Primary rooms and price	11.30%	9.85%	10.49%
All	11.46%	9.86%	10.44%
Difference between All and None	-0.05%	0.07%	-0.39%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	1.16%	2.10%	2.23%
Primary rooms	2.53%	3.06%	3.09%
Price	1.50%	2.53%	2.56%
District	1.24%	2.18%	2.34%
Primary rooms and price	2.64%	3.14%	3.20%
All applied at once	2.74%	3.18%	3.27%
Difference between All and None	1.58%	1.08%	1.04%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.92%	1.91%	2.02%
Primary rooms	2.30%	2.90%	2.91%
Price	1.29%	2.35%	2.37%
District	1.00%	2.00%	2.13%
Primary rooms and price	2.41%	2.98%	3.01%
All applied at once	2.51%	3.02%	3.08%
Difference between All and None	1.58%	1.11%	1.06%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	1.43%	2.30%	2.45%
Primary rooms	2.78%	3.21%	3.28%
Price	1.74%	2.71%	2.76%
District	1.50%	2.38%	2.56%
Primary rooms and price	2.87%	3.30%	3.38%
All applied at once	2.99%	3.33%	3.46%
Difference between All and None	1.56%	1.03%	1.00%

Table 24 iBuyer average profits with 9% margin between predicted and bid price. Run as part of a robustness check on our results. The table indicate that a higher margin yields similar results as with 6% margin, thus strengthening our findings.

A.4.2 Convenience factors

In the accept probability distribution scenarios, the mean corresponds to -1 times the convenience factor. The value indicates how much lower a bid can be than the perceived valuation of the seller, for 50% of sellers to accept. In the benchmark model a 4% convenience factor was used. Table 25 and Table 26 display the profits with 2% and 6% convenience factors respectively. In the first, homeowners value the services of the iBuyer less, while in the latter they value it more.

iBuyer average profits

6% margin between predicted and bid price and 2% convenience factor, with list price as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	7.96%	6.29%	7.29%
Primary rooms	7.87%	6.27%	7.02%
Price	6.80%	5.89%	6.54%
District	7.47%	6.20%	7.14%
Primary rooms and price	7.76%	6.33%	6.97%
All	7.91%	6.35%	6.92%
Difference between All and None	-0.05%	0.06%	-0.37%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-0.32%	0.48%	0.69%
Primary rooms	1.02%	1.39%	1.51%
Price	-0.08%	0.85%	0.95%
District	-0.25%	0.55%	0.79%
Primary rooms and price	1.11%	1.48%	1.60%
All applied at once	1.24%	1.52%	1.68%
Difference between All and None	1.55%	1.05%	0.98%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-0.52%	0.32%	0.51%
Primary rooms	0.83%	1.26%	1.36%
Price	-0.26%	0.71%	0.79%
District	-0.45%	0.40%	0.61%
Primary rooms and price	0.92%	1.35%	1.45%
All applied at once	1.05%	1.40%	1.52%
Difference between All and None	1.56%	1.08%	1.01%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	-0.09%	0.65%	0.89%
Primary rooms	1.23%	1.52%	1.67%
Price	0.13%	1.01%	1.12%
District	-0.02%	0.72%	0.98%
Primary rooms and price	1.31%	1.62%	1.76%
All applied at once	1.45%	1.65%	1.83%
Difference between All and None	1.53%	1.00%	0.94%

Table 25 iBuyer average profits with 2% convenience factor. Run as part of a robustness check on our results. The table indicate that a lower mean convenience factor yields similar results as with 4% margin, thus strengthening our findings.

iBuyer average profits

6% margin between predicted and bid price and 6% convenience factor, with list price as proxy for seller valuation

Without adverse selection

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	7.96%	6.29%	7.29%
Primary rooms	7.87%	6.27%	7.02%
Price	6.80%	5.89%	6.54%
District	7.47%	6.20%	7.14%
Primary rooms and price	7.76%	6.33%	6.97%
All	7.91%	6.35%	6.92%
Difference between All and None	-0.05%	0.06%	-0.37%

With adverse selection, neutral probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.75%	1.50%	1.77%
Primary rooms	2.21%	2.47%	2.65%
Price	0.99%	1.90%	2.04%
District	0.84%	1.57%	1.87%
Primary rooms and price	2.29%	2.58%	2.76%
All applied at once	2.45%	2.62%	2.84%
Difference between All and None	1.70%	1.12%	1.07%

With adverse selection, pessimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.63%	1.46%	1.71%
Primary rooms	2.14%	2.48%	2.64%
Price	0.91%	1.89%	2.00%
District	0.73%	1.53%	1.81%
Primary rooms and price	2.22%	2.59%	2.75%
All applied at once	2.38%	2.64%	2.83%
Difference between All and None	1.75%	1.18%	1.12%

With adverse selection, optimistic probability

Purchasing rules applied	*LAD*	*XGB*	*SVM*
None	0.88%	1.56%	1.85%
Primary rooms	2.30%	2.47%	2.68%
Price	1.09%	1.93%	2.08%
District	0.97%	1.63%	1.94%
Primary rooms and price	2.37%	2.58%	2.78%
All applied at once	2.53%	2.62%	2.85%
Difference between All and None	1.64%	1.05%	1.00%

Table 26 iBuyer average profits with 6% convenience factor. Run as part of a robustness check on our results. The table indicate that a higher mean convenience factor yields similar results as with 4% margin, thus strengthening our findings.