HOUSING LAB WORKING PAPER SERIES

2020 | 5

House price seasonality, market activity, and the December discount

Erling Røed Larsen

HOUSING LAB
Nasjonalt senter for boligmarkedsforskning

House price seasonality, market activity, and the December discount*

Erling Røed Larsen†

Abstract

In Norway, house prices tend to drop in December. This regularity is persistent across cities and over time. I exploit a transaction data set with high temporal granularity to document and estimate the size of the December discount. To show its existence, I control for a composition effect using a hedonic model. I account for unobserved unit and seller heterogeneity by looking at repeat sales and controlling for unit fixed effects. By segmenting Norway into sub-markets, I investigate determinants of price seasonality. Evidence suggests that the December effect is linked to time-on-market for each unit and transaction volumes within each sub-market. The findings show clear co-movements of prices, time-on-market, and sales.

Keywords: Auctions, December discount, house price seasonality, search-and-matching, transaction volume

JEL Codes: C21, D12, R31

*The author is grateful to André Kallåk Anundsen, Terje Eggum, Andreas Eriksen, Jeanette Fjære-Lindkjenn, William Larson, Shawn Rohlin, and Dag Einar Sommervoll for comments, suggestions, and code assistance. The author also thanks Eiendomsverdi for transaction data and participants at the 92nd Western Economic Association International conference, the 13th meeting of the Urban Economic Association, and the AREUEA-ASSA 2020-meeting in San Diego; in addition to participants at multiple seminars and work-shops. Housing Lab is grateful for financial support from SpareBank1 and Selvaag Bolig.

†Housing Lab, Oslo Metropolitan University and BI Norwegian Business School.

1 Introduction

In Norway, mean house prices fall in December. At first blush, this price pattern should encourage people to buy in December and discourage people from selling in December. One would think that this December discount could be exploited or avoided, so the existence of a predictable price seasonality begs scrutiny. The trivial explanation would be that units sold in December simply are different from units sold in September. If such a composition effect accounts for the whole difference, the price drop is no discount nor is there any special buying opportunity or discouragement from selling. This article deals with this and other possibilities, and asks two questions: Is there a December effect and if there is, what generates it?

The answer is that there is a December effect, and its magnitude is about 1.5 percent, when sell prices are compared with September. It is linked to a non-repeatable time-on-market (TOM) effect that is unrelated to the the unit and a market activity effect. I deal with three econometric challenges in order to ensure that my findings are valid: the composition effect, unobserved unit heterogeneity, and unobserved seller heterogeneity. I control for these effects by using a hedonic model, by controlling for unit fixed effects in repeat-sales models based on ask prices and appraisal values, by exploiting the information in ask prices and appraisal values, and by utilizing an instrumental variable set-up.

I use data on the ask price of each unit that is transacted, the appraisal value of about half the units, the date on which a bid was accepted by a seller, and a unique unit identifier that allows me to follow repeat-sales of the same unit. These variables contain information that lets us control for effects that would otherwise bias the estimates. The ask price, for example, reflects the market value set by the person who knows the most about the unit, namely the seller. The appraisal value is another example. It is exogenous to the seller and is a valuable source of information if we worry that the ask price is set strategically (see e.g. Anglin et al. (2003); Anglin and Wiebe (2013); Anundsen et al. (2020b)). The date on which a bid is accepted is useful since its presence implies that we can differentiate sharply between bids accepted on 31 December and 1 January. The unique unit identifier lets us set up repeat-sales models, not only of the sell price, but also for sell-ask spreads and sell-appraisal spreads. In fact, combining a repeat-sales model with spreads makes it possible to control for time-invariant latent attributes and time-varying unobserved quality of the unit. In addition, using the unique unit identifier, we may explore whether a long time-on-market (TOM) for a given unit in the first sale is linked to a long TOM for the same unit in a second sale. It turns out that it is not. TOM is not connected to the unit, and they tend to be mean-reverting.

The novelty of the study lies in exploiting the temporal granularity of the sell date and for-sale advertisement date, combined with information contained in ask prices and appraisal values. My contribution is purely empirical. The findings, however, are relevant in a broader context. First, the main thrust of my argument is aimed at showing the importance of the level of market activity in understanding how well markets match buyers and sellers. Thus, this is a paper about markets and when they work at their best. The paper best can be thought of as a case study of search and matching activity, an activity meticulously studied in labor economics, even if it is less studied in the housing market. Moving house involves solving a dual search problem in that it requires the moving owner-occupier to find both a seller and a buyer. Information on how market activity affects the sell and buy processes is helpful to policymakers and market participants, especially if prices display seasonality patterns. Potentially, market coordinators could seek to flatten the seasonality pattern or to nudge market participants to move their activity to other months. The latter would be the temporal equivalent to establishing spatial concentration, i.e. markets, and from temporal concentration there could be coordination benefits to be reaped.

Second, the housing market is sufficiently influential to the macroeconomy (Leamer (2015)) that any price regularity, be it weekly (Røed Larsen (2020)), yearly, or cyclically (Leamer (2007)) should be of interest. After all, what moves the housing market has the potential to have real economic consequences, which is one of the reasons why forecastability of house prices has been of long-standing interest to economists and policymakers (Case and Shiller (1989), Røed Larsen and Weum (2008)).

Third, the estimate of a 1.5 percent December discount allows policymakers to entertain some ideas on the societal value of arranging markets that match households with houses accurately, transparently, and quickly. Keep in mind that the 1.5 percent December discount is measured relative to September. If we compare the fourth quarter to the second quarter, the effect could be larger (but not as precisely estimated). The December discount is consistent with a notion that few high-quality matches[1] is the result of low-activity markets. A non-trivial portion of all sales takes place in December and the overall value of the housing stock

[1] Hsieh and Moretti (2019) have shown substantial mis-allocation losses result from spatial mis-matches caused by inoptimalities on the supply side of housing markets. Potentially, mis-matches sourced in temporal factors, such as seasonality, may also matter.

in Norway is estimated to be 2.5 times GDP.[2] In other words, even one percent matters if it is caused by mis-matches between preferences and attributes.

The empirical strategy is straightforward and let me offer a few details here. I use a hedonic model, including a range of housing attributes, to account for composition effects. To control for a trend in house prices, I use a linear time trend. The December discount survives this initial control. However, there are differences in what types of units are sold in December compared to other months. In particular, units sold in December are smaller, both ask prices and appraisal values are lower, and the share of units transacted in Oslo is lower. At the same time, we observe that market conditions are different in December. The transaction volume in December is a third of the volume in September. Furthermore, while TOM in September is 39.1 days, it is 55.2 days in December. These statistics hint at what we should look for when we seek to map out the co-variates of price seasonality.

In order to account for permanent unobserved unit heterogeneity, I study units that have been sold twice and control for unit fixed effects. Controlling for unit fixed effects does not deal with time-varying quality changes. To control for this, I study repeat sell-ask and sell-appraisal spreads instead of sell prices. After all, ask prices reflect what a seller believes the market value should be, taking into account potential unobserved quality changes (e.g. renovation or lack thereof). However, ask prices are endogenous to the seller, and could include a strategic element (Anundsen et al. (2020b)). Thus, I also consider specifications in which

[2]This is a back-of-the-envelope computation using the following numbers: The Norwegian GDP for the year 2017 in market value is 3,300 billion NOK; see http://www.ssb.no. The firm Eiendomsverdi computed the market value of the whole Norwegian housing stock in September 2017 using their automtaic valuatio method (AVM) on the registry of all units. Their estimate was 8,000 billion NOK (contact eiendomsverdi.no).

I control for the appraisal value. The December discount survives these controls. It is, however, fathomable that there is unobserved heterogeneity among sellers if one type of seller is prone to sell in December, while other types are not. I examine what an instrumental variable approach can do to mitigate this possibility. The December discount survives all these specification and robustness checks.

To study determinants of the December discount, I use two approaches. First, in a micro-based approach, I link individual TOM to the sell-appraisal spread of the same transaction. When segment the data based on TOM, the estimated December coefficient becomes similar to coefficients of September and October within the segments. TOM itself is not linked to the unit, since a regression of the second TOM on the first TOM shows that TOM is highly mean-reverting and there is little persistence. Instead, it appears that the long-TOM transactions in December, which are associated with discounts, result from random processes in which some units simply are not sold quickly. Second, in a macro-based approach, I estimate market characteristics of sub-markets of houses and apartments in municipalities and show that there is an association between low market activity and the magnitude of the December discount.

Related literature

This article examines the empirical traces of the idea that seasonality in house prices is related to seasonality in market activity. The basis for the idea can be found in search theory (see e.g. Diaz and Jerez (2013), Genesove and Han (2012), Kashiwagi (2014), Maury and Tripier (2014), and Krainer (2001)). The key assumption is that the number of market participants influences the probabilities of different levels of match-quality between buyer preferences and house attributes, which in turn affects realized prices (see Ngai and Tenreyro (2014), Kaplanski and Levy (2012), Nenov et al. (2016), Novy-Marx (2009), and Anundsen and Røed Larsen (2018)). When seasonality implies fewer sellers and fewer buyers in December, it also implies longer TOM and lower probabilities of high quality matches.

Essentially, seasonality is as a special case of cyclicality, and so the underlying thinking for understanding the pattern of a December discount shares features with Diaz and Jerez (2013). They note how there is a joint cyclical behavior of house prices, sales, and TOM. In principle, the 31 days of December is a miniature version of a cyclical downturn, and so their observation of a co-movement of prices, sales, and TOM should be found in this paper's seasonality measures as well. It is.

The results in this article shed light on, and is consistent with, the model Albrecht et al. (2007) construct in which agents enter the market relaxed and becomes increasingly desperate to sell as TOM grows, and so the expected price falls with TOM. Sales with long TOM in December might have been the result of a one-on-one negotiation rather than an auction (Coles and Muthoo (1998)) even if the mean-reversion of TOM implies that the long TOM is most likely the result of a random process. My results are less compatible with the idea proposed by Taylor (1999), in which TOM is a sign of quality, in that I show that the second TOM of a given unit is almost orthogonal to the first TOM of the same unit.

This article is structured in the following way. Section two presents the data sources and gives a few details on the institutional background. Section three goes through the empirical techniques I use throughout the paper. Section four contains empirical results. In section five, I explore what can explain the December discount and present a skeleton model that rationalizes my findings. The section also presents evidence that suggests TOM and market activity are linked to the December discount. Section six discusses a few ways to probe deeper into measures of market activity. Section seven concludes and offers a few potential policy implications.

2 Data and institutional background

2.1 Data source

From the collaboration with real estate agencies, the bank-owned firm *Eiendomsverdi* obtains information on units advertised for sale on the online platform *Finn.no* that covers more than 70 percent of the market. The data are combined, and cross-checked, with public registry transaction data. The transaction data used in this article have the same source as the data used in Anundsen and Røed Larsen (2018) and Røed Larsen (2020), but this article's data coverage is wider since I have three more years of transactions. The data include, but are not limited to, unit identifier, transaction price, common debt, ask price, date of acceptance of highest bid, date of online advertisement when unit was put up for sale, unit attributes (e.g. construction type, size, construction year, number of rooms, lot size), and geographical location. For about half the observations, the data set also includes the appraisal value set by an appraiser.

I trim the data in order to remove extreme observations or observations with missing values, duplicates, suspicious entries, and typos. Common debt[3] is in-

[3]Common debt is more common in co-ops, which are not included in the data set. Self-owned units sometimes have common debt, e.g. if row house owners own a shared parking lot and there is shared debt for the construction of it.

cluded in sell price, ask price, and appraisal value. I trim on 0.1 and 99.9 percentiles. I study self-owned units, and not co-ops[4], because the co-op ownership type does not have a unique unit identifier for all years in the data set. I use two versions of the data set: i) all transactions and ii) observations with appraisal value.

Table 1 presents summary statistics for the data. The total data set after trimming consists of 691,192 transactions during the period 1 January 2002 and 1 February 2017. Out of these observations, 373,373 have appraisal values. The appraisal data have units with a slightly lower mean size, owing in part to the much higher Oslo share. While the overall data set has a 16 percent Oslo share, the appraisal data have an Oslo share of 25 percent. Part of the reason for this is institutional. During the period I study, it was not common in all cities to make use of an appraiser, but it was common in Oslo.[5] For both the overall and the appraisal data, the median sell-ask spread (sell price less ask price on ask price) is zero. For the appraisal data set, the median appraisal spread (sell price less appraisal value on appraisal value) is zero.

The lower panel includes observations from Oslo transactions with appraisal values. We see that these transactions involve smaller units; the mean size is 73 square meters while the overall mean size is 111 square meters. While the overall data have an apartment share of 40 percent, Oslo appraisal data have a share of 80 percent. Neither the median sell-ask spread nor the appraisal spread for Oslo transactions are zero. This could potentially be due to a rising price trend

[4] Cooperatives are organized such that the occupier buys a right to live within the compound. All occupiers share financial and other responsibilities.

[5] Today, it is no longer common to obtain an appraisal value. The realtor handles the valuation.

in Oslo, in which ask prices and appraisal values might have been lagging sell prices. Anundsen et al. (2020b) argue that part of this non-zero spread is due to strategies among realtors. For this article, since I compare spread with spread, which essentially is taking first differences, it does not affect the results when I compare December coefficients with September coefficients.

Table 1. Summary statistics. Transaction data. Norway, 2002-2017

	Min	25th percentile	Median	Mean	75th percentile	Max
			Data with all observations			
			N = 691,192			
Date	1 Jan 2002	19 Oct 2006	22 Sep 2010		16 Dec 2013	1 Feb 2017
Sell	326,000	1,600,000	2,275,000	2,640,965	3,225,000	14,750,000
Ask	250,000	1,590,000	2,225,000	2,598,211	3,190,000	18,900,000
Sell-ask spread	-0.305	-0.0303	0.000	0.0185	0.0571	0.573
Size	21	75	111	119	153	378
Sell/size	2,874	14,535	21,765	25,297	32,500	100,303
			Share Oslo = 0.16			
			Share apartments = 0.36			
			Data containing observations with appraisal value			
			N = 373,373			
Date	2 Jan 2002	31 Jan 2007	6 Dec 2010		11 Nov 2013	1 Feb 2017
Sell	330,000	1,680,000	2,375,000	2,772,720	3,392,124	14,750,000
Ask	250,000	1,650,000	2,304,580	2,717,162	3,300,000	18,900,000
Appraisal	250,000	1,690,000	2,350,540	2,768,529	3,370,000	22,000,000
Size	21	72	108	117	150	378
Sell/size	2,874	15,625	23,664	27,082	35,278	100,271
Sell-ask spread	-0.305	-0.0269	0.00253	0.0230	0.0643	0.566
Appraisal spread	-0.375	-0.0476	0.000	0.00556	0.0536	0.543
			Share Oslo = 0.25			
			Share apartments = 0.40			
			Data with only observations with appraisal value. Oslo			
			N = 93,716			
Date	2 Jan 2002	23 Aug 2006	21 Jun 2010		9 Aug 2013	31 Jan 2017
Sell	540,000	2,050,000	2,900,000	3,480,289	4,300,000	14,750,000
Ask	520,000	1,956,875	2,791,499	3,361,810	4,100,000	18,900,000
Appraisal	520,000	2,000,000	2,850,000	3,432,444	4,200,000	22,000,000
Size	21	54	73	89	110	378
Sell/size	5,379	30,000	40,411	41,987	52,083	100,271
Sell-ask spread	-0.304	-0.0106	0.0270	0.0430	0.0899	0.555
Appraisal spread	-0.364	-0.0333	0.00885	0.0220	0.0738	0.543
			Share apartments = 0.80			

Notes: Prices are in NOK. Size in square meters; rounded to square meters. Date is date of acceptance of bid.

2.2 The construction of repeat-sales data

Since the data include units that are transacted multiple times, there is an element of panel structure in the data. These repeat-sales units are, however, not transacted at the same time, thus the structure is that of a repeated cross-section. In order to construct the repeat-sales data, I retain only units that are sold exactly twice. The reason why I leave out units that are sold multiple times is that some of these units can be buy-to-let units or otherwise different from units sold exactly twice. In the overall data set, there is 213,394 number of observations of units sold exactly twice. These observations encompass 111,244 observations of transactions involving units that have transacted exactly twice and, in addition, have appraisal values, i.e. there are 55,622 such repeat-sale units in the appraisal data set.

2.3 The construction of sub-market aggregate data

In the analysis of sub-markets, I study patterns of geographical variation. To construct the appropriate data set, I first require that municipalities have at least 800 transactions over the period out of a total of 428 municipalities.[6]

I then partition into two types of units: apartments and non-apartments.[7] I left out one sub-market with no December sales, and was left with 247 sub-markets. These 247 sub-markets represent 627,405 transactions.

[6]There is an an element of art to the choice of a cut-off of 800, and I did experiment with several cut-offs. A higher cut-off leaves us with fewer, but thicker sub-markets. A lower cut-off leaves us with more, but thinner sub-markets.

[7]Again, there is an an element of art into partitioning. Houses could potentially be further partitioned into detached, semi-detached, and row houses. Alternatively, one could partition into segments below and above the median size. These partitions would have left us with more, but thinner sub-markets. I chose not to.

2.4 Institutional background

In Norway, about four out of five are owner-occupiers. The proportion depends upon whether one measures households, individuals, or housing units. The typical housing career involves renting while in school, then buying the first unit upon entering the first job. As an individual grows older, she typically buys a larger unit. After retirement, many households sell their house and move into easier-to-maintain and centrally located apartments.

Geographically, Norway is a relatively large country with few inhabitants. There is a substantial difference between rural areas and urban centers, and there is considerable heterogeneity at the local sub-market level (Røed Larsen (2020)). This geographical variation allows us to study the co-movement of prices and market activity and test the hypothesis that the December discount is more pronounced in sub-markets with more pronounced transaction seasonality.

2.4.1 Financing a purchase

When an individual or a household has decided to buy a new home, the buyer first visits a bank and obtains proof-of-financing. This is a financial certificate that serves as evidence of financial capability and there exists a regulatory framework that stipulates what banks can do. In general[8], they cannot grant loans when the total household debt exceeds five times household income. Moreover, there is a requirement on ability to sustain a five percentage point interest rate increase, computed on the basis of estimated budgets[9] There is also a requirement of at

[8]There is a macroprudential regulatory framework that governs the credit extension behavior of banks.

[9]This budget is estimated by research; e.g. the institute SIFO estimates what individuals and households need in order to sustain a minimum level of acceptable material standard of

least a 15 percent down payment.[10]

2.4.2 Auctions

The Norwegian housing market is organized around a transaction process that involves an ascending-bid (English) auction. This auction commences the day after the last open-house (public showing). The realtor leads the auction and informs participants, both active bidders and interested parties, about the bidding activity. Bids may be extended digitally, which involves a country-wide digital identification system. Bids may also be extended verbally, or, mostly relevant to the early period in my data set, through fax and sms. All bids are legally binding, but conditional bidding is allowed.[11]

Acceptances of bids are legally binding. The implication is that once the highest bid is accepted transfer of ownership between seller and buyer is locked in. In contrast to the situation in other Nordic countries, there is no grace period in which agents may walk away from the agreement. This means that, it is possible in Norway to construct a daily price ticker of house prices (see Anundsen et al. (2020a) for an application).

2.4.3 The buying and selling process

A seller first contacts a realtor, who assists her through the selling process. It is legally possible for sellers to do most of the process themselves, but few choose this option. By law, the realtor is obliged to be the caretaker of the interests of

living.

[10] In this setting, the market value of the unit is defined as the purchase price.

[11] Conditions may include contingencies upon financing or take-over dates. Often, conditions include expiration times, for example a statement that the bid lasts for three hours.

both the buyer and the seller, but since the realtor is paid by the seller, the realtor consults mostly with the seller. The realtor, however, must make sure that all relevant information about the state of the unit is given to the buyer.

In order to become a realtor, one is required to obtain a licence. A realtor, typically, has gone through a Bachelor-program in realty, which is a specially designed program for prospective realtors at business schools.[12]

Until 2016, it was common practice in most cities that the realtor contacted an appraiser who would inspect the unit, issue a technical report, and announce an appraisal value. After 2016, the appraiser typically concentrates his effort on the inspection and the report and does not announce any appraisal values. Part of the reason for this change, is that the typical background of the appraiser is in engineering and it was considered more appropriate to let realtors handle the valuation side. The appraisal value that was commonly issued before 2016 was an independent value assessment, neither related to taxation nor the financial situation of the buyers. A buyer's ability to obtain a loan is connected to his income and home equity.[13] While the mortgage in today's regulatory framework cannot exceed 0.85 of the final price of the unit to which the mortgage is tied, rules were different throughout the period I study. The appraisal value was issued before the auction and did not in itself impose any constraints on the bidding.

Before the online advertisement, the realtor and the seller discuss what ask price to announce. The seller has some room for manoeuver in setting the ask price, but there is regulation requiring that the realtor must ensure that the ask

[12]More information on real estate agencies can be found at eiendomnorge.no. More information on realtors can be found at nef.no.

[13]For more information on Norwegian appraisers, consult with: norsktakst.no.

price is realistic and reflects the seller's reservation price to a high degree. [14]

A seller puts his unit up for sale on the online platform Finn.no. In the advertisement, he announces both the ask price and one or several dates for the open-house (public showings) on which any interested buyer may come and inspect the unit. Typically, a seller in Oslo puts her advertisement up on a Friday and announces the open-house for the Sunday or Monday nine and ten days, respectively, later. The advertisement includes all relevant information about the unit and typically has a large number of photographs. This information makes it possible for prospective buyers to obtain a sense of the match between their preferences and the attributes of the unit, and thus to make informed decisions on which open-houses to visit. Due to the time it takes to visit an open-house, no buyer can visit more than a small fraction of the open-houses on any given Sunday.

Buyers study these advertisements and form plans on which open-houses to visit. The realtor is the host of the open-house, but the owner is sometimes present. Participants in the ensuing auction are mostly recruited from visitors to the open-house, but some bidders may have chosen not to visit the open-house because they obtained a sufficient amount of information from the advertisement. The Norwegian registry of houses is public information so a prospective bidder is also able to find relevant information from the registry.

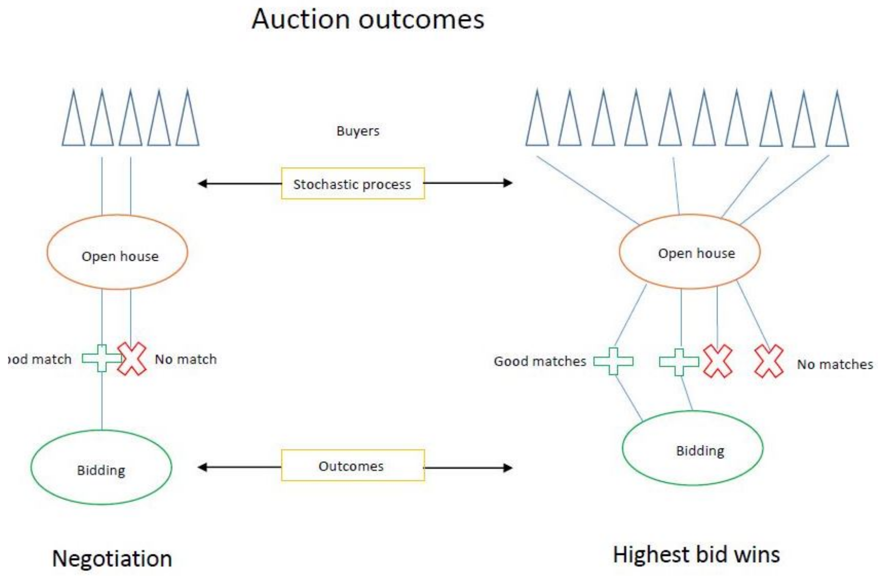
Figure 1 presents an image of the process and depicts the stochastic element in the arrival of bidders (Anundsen and Røed Larsen (2018)). Bidders are drawn randomly from the pool of buyers in an unknown stochastic process.[15] Bidders with a high-quality match between their preferences and the attributes of the unit

[14] If the authorities discover that a realtor agency systematically is associated with multiple transactions in which the ask price is set artificially low, fines may be issued.

[15] Often, one thinks of this process as Poisson process.

would have a higher willingness-to-pay as long as income is equal. However, higher income bidders with a medium match-quality may have higher willingness-to-pay than medium income bidders with a high match-quality and vice versa.

Figure 1. The open-house and auction process



3 Empirical methodology

3.1 The hedonic model

I use a hedonic model to control for composition effects. The model is similar in spirit to, and is based on the same data source (although wider in coverage) as, Anundsen and Røed Larsen (2018) and Røed Larsen (2020). The hedonic model is a regression of the logarithm of sell price[16] (including common debt) onto a space spanned by determinants:

[16]Notice that I use the logarithm of sell price as dependent variable when the aim is to inspect the estimated coefficients of the December dummy. If the aim is to predict the sell price, it may

$$log(P_h) = \alpha + \beta_1 log(Size_h) + \beta_2 Sq(log(Size_h)) + \sum_k \gamma_k A_{h,k} + \eta m_h + \sum_T \theta_T M_h + \epsilon_h, \quad (1)$$

in which the subscript h refers to unit h, $A_{h,k}$ denotes attribute k in a collection of K attributes that characterize unit h, m_h is a counting variable that takes on the running number (across the period) of the month in which the transaction of unit h took place,[17] and M_h is a calendar month dummy which takes on the value 1 if the transaction of unit h took place in that calendar month and 0 otherwise.[18] The error term ϵ_h is assumed mean-zero, constant variance. The collection of attributes includes dummies of type, interaction variables of type and the size polynomial, interaction variables of Oslo and the size polynomial, a collection of construction periods which is unity if the unit was constructed in that period,[19] city dummies for the 18 largest cities in Norway, regional dummies for the 19 administrative regions (counties) Norway consisted of during the period,[20], and a weekday-city interaction variable for the two largest cities, Oslo and Bergen.[21]

Note that I use the subscript h as short notation for the transaction of house h. Transaction of unit h is defined as having taken place on the date on which

be preferable to use the sell price as dependent variable as taking logarithms is a non-linear transformation and so Jensen's inequality implies a prediction bias.

[17]The month number (counting variable) runs from the first to the last month in the sample period, January 2002 - February 2017.

[18]The collection of calendar month dummies consists of 11 dummy variables from January to December excluding July (which is default).

[19]The collection of construction year dummies consists of three periods, 1950-1979, 1980-1999, and after 2000. The period before 1950 is default.

[20]The exceptions are Oslo, which is also a city (the capital), and Troms and Finnmark, which constitute the default region.

[21]Røed Larsen (2020) demonstrates that there are intra-week price patterns in the Norwegian housing market driven by the mode of the distribution of the day of the open-house.

the highest bid was accepted since acceptance of bids (as well as bids themselves) are legally binding in Norway. Sell price, ask price, and appraisal value include common debt.

3.2 Controlling for unobserved heterogeneity

The hedonic model is fully specified and the coefficients contian no bias as long as there are no omitted variables that are not orthogonal to the determinants yet influences the sell prices. To see the potential challenges, let us think of the sell price of house h, P_h, as a function of observed attributes A_h for house h, unobserved quality ξ_h, a December discount D_h, and a residual component e_h:

$$P_h = a + bA_h + c\xi_h + dD_h + e_h, \tag{2}$$

in which the error term e_h contains both a match-utility component that is sourced from the unique matching between attributes of unit h and buyer preferences, potential non-observable seller-specific components, and a mean-zero, constant variance white noise element. I shall now describe how I deal with potential quality bias caused by unobserved unit and seller heterogeneity.

The hypothesis that a December discount exists is the hypothesis that the coefficient d is statistically significantly different from zero and negative. If c is non-zero, i.e. ξ_h not orthogonal to e_h, OLS suffers from omitted variable bias. The intuition is that in an OLS-regression of sell price onto attributes and a December-dummy, we could risk that a selection mechanism would make units with certain

ξs be associated with December sales. For example, low-quality units could be sold at higher frequencies in December.

A repeat-sales set-up is a way of controlling for unobserved unit heterogeneity in the quality xi_h. A repeat-sales set-up allows us to follow the same unit over time, and observe same-unit transactions in December and non-December. In such a set-up, one collects observable attributes and unobserved quality in a unit-specific intercept, $\omega_h = a + bA_h + c\xi_h$:

$$P_{h,s} = \omega_h + dD_{h,s} + e_{h,s}, \tag{3}$$

in which $s = 1, 2$ denotes transaction number of unit h. Regressing the sell price P in a unit fixed model onto a space that includes a dummy for a December-sale would allow us to obtain an estimate of d.

However, while the repeat-sales set-up controls for time-invariant unobserved heterogeneity, it does not control for time-varying unobserved heterogeneity. It is possible that quality varies over time[22] so that $\omega_{h,t} = a + bA_h + c\xi_{h,t}$.

The seller would know this. She is the agent with the most knowledge about the unit. Thus, the ask price reflects both observed attributes and unobserved quality, i. e. the ask price can be written as $AP_{h,t} = a + bA_h + c\xi_{h,t}$. We may combine the repeat-sales set-up with ask prices in order to control for both time-invariant and time-varying unobserved unit heterogeneity. To this end, I use the sell-ask spread, i.e. the sell price, $P_{h,t}$, less the ask price, $AP_{h,t}$ as a fraction of the ask price. The numerator becomes $P_{h,t} - AP_{h,t} = a + bA_h + c\xi_{h,t} + dD_{h,t} - a - bA_h - c\xi_{h,t} + u_{h,t} =$

[22] E.g if the unit is renovated or not.

$dD_{h,t} + u_{h,t}$, in which $u_{h,t}$ is an error term. Thus, using the sell-ask spread of repeat-sales in a unit fixed effect regression is a remedy for solving the challenge from time-invariant and time-varying unobserved unit heterogeneity.

However, the ask price may contain a strategic element (Anglin et al. (2003); Anglin and Wiebe (2013); Anundsen et al. (2020b)). To control for this possibility, I use appraisal values when they are available by studying sell-appraisal spreads in stead of sell-ask spreads. Again, the appraisal value of unit h, $APP_{h,t}$, reflects both observed attributes and unobserved quality, $APP_{h,t} = a + bA_h + c\xi_{h,t}$.[23]

It is fathomable that the December discount is related to unobserved seller heterogeneity. To see this, let us consider a thought experiment in which there are two seller types, A and B. Say type B sells in December and this type tends to set ask prices in relation to sell prices, or accept bids in relation to ask prices, in a different fashion than type A who sells in other months. It could cause omitted variable bias in our estimated December dummy coefficient.

In order to control for this possibility, I re-estimate the hedonic model with a two-stage set-up. In the first stage, I regress the logarithm of ask price onto the logarithm of appraisal value, which is exogenous. In the second stage, I use repeat-sales in a unit fixed-effect set-up in which the logarithm of the sell price is dependent variable and the instrumented logarithm of the ask price from the first stage is one of the independent variables.

[23]see Anundsen and Røed Larsen (2018) Tables A3-A5 for an assessment of the level of information in appraisal values.

3.3 Price seasonality across sub-segments

For each of the 247 sub-markets, I construct a measure of price seasonality. I compute the mean sell-ask spread in December and the mean sell-ask spread in the other eleven months (excluding the year 2017 for which we only have observations in January and February). I use the spread, not the sell price, since the ask price both reflects a time trend and accounts for attributes and quality of the unit. Thus, the spread captures the extent to which the sell price exceeds an expected value. My price seasonality measure is the difference between the two spreads, i.e. the difference between the December spread and the non-December spread.

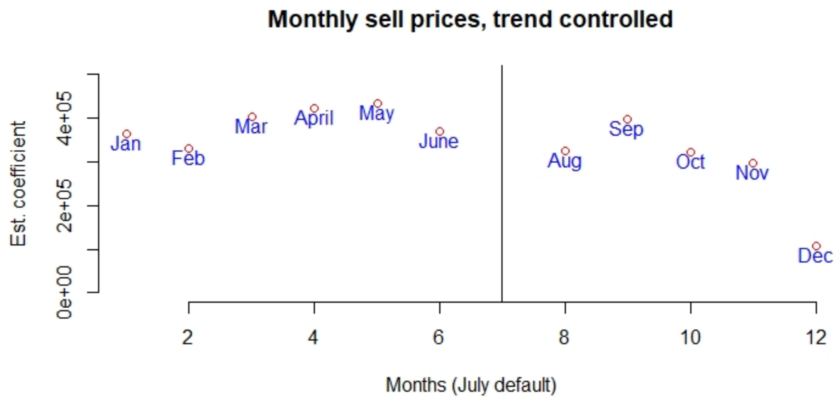
I use two measures of market activity, one based on transactions and one based on listings. The transaction based measure is the ratio of the mean number of transactions in a given sub-market in December relative to the mean number of transactions in the same sub-market in the other eleven months. The listing-based measure is the ratio of the mean number of new listings in a given sub-market in December relative to the mean number of new listings in the same sub-market in the other eleven months.

4 Empirical results

4.1 The regularity of a December price drop

Figure 2 displays the estimated coefficients of month dummies in a regression of sell price onto a space spanned by a linear trend (a counter in month number since January 2002) and calendar month dummies. We observe that the estimated coefficient for December is substantially lower than other months, which reflects the lower trend-controlled sell prices in December.

Figure 2. House price seasonality. Norway, 2002-2017



Notes: Coefficients are estimated using a regression model in which sell prices are regressed onto a month counter running over months within 2002-2017 and calendar month dummies (July is default). We do not depict deviations from trend; we plot estimated coefficients for month dummies, i.e. in addition to the linear trend. The data cover the period described in the upper panel of Table 1. N = 691,192. Adj. R2 = 0.149. All monthly coefficients were statistically significant at the 0.001 level.

4.2 Composition effects

I examine the possibility that units sold in December are different from units sold earlier in the Fall. I first perform a check for balance of key attributes of the unit (size, type, location) and key attributes of the transaction (sell price, TOM). Table 2 tabulates the results of the check for balance. We observe that transacted units tend to be somewhat smaller in December (116 square meters) than in September (122 square meters) and that apartments tend to have a larger share in December

(37 percent) compared to September (34 percent). We also see that the sell price is ten percent lower (NOK 2.4 mill) in December compared to September (NOK 2.7 mill). Time-on-market (TOM) is 55 days in December, which is much longer than the 39 days in September. The transaction volume in December is one third of the volume in September, 25,245 sales versus 75,853 sales. Table 2 indicates that both the composition of transacted units and the market conditions are different in December.

Table 2. Check for balance. Transaction data. Norway, 2002-2017

	Aug	Sep	Oct	Nov	Dec
No. obs.	64,149	75,853	67,660	56,884	25,245
Size	116.6	121.5	120.1	118.7	116.2
Sell	2,647,516	2,729,536	2,649,470	2,631,532	2,412,158
Ask	2,588,463	2,686,929	2,619,987	2,610,616	2,425,791
TOM	43.9	39.1	39.2	40.5	55.2
Share apartments	0.37	0.34	0.35	0.36	0.37
Share Oslo	0.17	0.16	0.16	0.16	0.14

Notes: Sell and ask prices are in NOK. TOM is in days; size in square meters. The dates used to compute TOM is the date of acceptance of bid less the date of advertisement posted online (on Finn.no).

4.3 Controlling for composition bias, attributes, and time trend

In order to disentangle the December effect from a composition bias, I start out by controlling for observed attributes through the construction of a hedonic model. The hedonic model is estimated on two data sets, the whole transaction data set and the subset that contains only observations with appraisal values. The estimated dummy coefficients for December are -0.0117 and -0.00714, while the coefficients for September are 0.0389 and 0.0488. Having controlled for composition, attributes, and a time trend, the results show a substantial December discount. These estimates indicate a large December discount of five percent.

Table 3. Hedonic model of log sell prices on determinants. Norway, 2002-2017

	I	II
	All data	Appraisal data
Intercept	12.31 (7.0e-2)	11.93 (0.10)
Logsize	-9.80e-2 (2.9e-2)	5.36e-2 (4.1e-2)
Sqlogsize	8.41e-2 (2.9e-3)	7.11e-2(4.2e-3)
Type FE	YES	YES
Interaction	YES	YES
Constr. year	YES	YES
Large lot FE	YES	YES
City FE	YES	YES
Region FE	YES	YES
Weekday*City	YES	YES
Linear trend	4.70e-3 (7.1e-6)	4.63e-3 (9.9e-6)
Jan-June FE	YES	YES
Sep	3.89e-2 (1.9e-3)	4.88e-2(2.6e-3)
Oct	2.44e-2 (1.9e-3)	3.20e-2 (2.7e-3)
Nov	1.80e-2 (2.0e-3)	2.27e-2(2.7e-3)
Dec	-1.17e-2 (2.4e-3)	-7.14e-3 (3.4e-3)
No. obs.	691,192	373,383
(Deleted due to missingness)	(6,432)	(2,295)
Adj. R2	0.711	0.721
F-statistic (p-value)	2.44e4 (2.2e-16)	1.39e4 (2.2e-16)

Notes: White heteroskedasticity-consistent standard errors are computed using the R-function vcovHC. Interaction variables comprise products of (Oslo,logsize), (Oslo, sqlogsize), (apartment, logsize), and (apartment, Sqlogsize). The specification also includes dummies for construction year periods; see the Data section. City FE city refers to the inclusion of dummies for the 18 largest cities in Norway. Region FE denotes dummies for all Norwegian counties, except Oslo, which is also a city, and Troms and Finnmark, which are default. Weekday*City involves five dummies for each of the days in the work-week, Monday-Friday, multiplied by dummies for Oslo and Bergen. Linear trend denotes a counting variable that counts month number since January 2002, which is default.

4.4 Unobserved heterogeneity

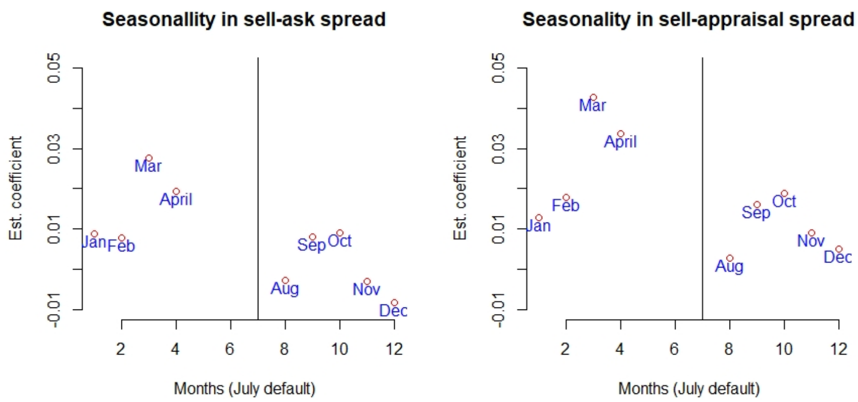
Having controlled for composition, attributes, and time trend, we are left with the challenge from unobserved heterogeneity. If negative unobserved qualities are associated with units transacted in December or if unobserved seller types are involved in December transactions, the December dummy estimate would contain both a season effect and a quality, or seller type, effect.

4.4.1 Unobserved heterogeneity in units

I deal with unobserved heterogeneity in units by combining two remedies, a repeat-sales model in which the same unit is sold twice and information on the ask price and the appraisal value. As a preliminary exercise, Figure 3 plots the seasonality in sell-ask spreads and sell-appraisal spreads.

We see from Figure 3 that when we control for unobserved time-invariant attributes inherent in the ask price and appraisal value, in addition to a linear time trend, the December discount pattern remains. The December spreads are lower than spreads in September, October, and November.

Figure 3. Seasonality in sell-ask and sell-appraisal spreads. Norway, 2002-2017



Notes: The coefficients are estimated using a regression model in which sell-ask and sell-appraisal spreads are regressed onto a space spanned by year dummies for 2002-2016 (2002 default) and calendar month dummies (July is default). 2017 data are included, but the regression specification does not include a dummy for 2017 since 2017-data only cover January and 1 February. We do not depict deviations from trend; we plot estimated coefficients for month dummies in addition to a linear trend. The graph in the left-hand panel was generated using all transaction data, described in the upper panel of Table 1. N = 691,192. The graph in the right-hand panel is generated using appraisal data, described in the middle panel of Table 1. N = 373,373. The Adj. R2 were 0.0304 for the left-hand side regression and 0.0433 for the right-hand side regression.

Table 4 tabulates the results from regressions on spreads from repeat-sales data. Models I and III are ordinary least-square models and they are included for comparison. Models II and IV are unit fixed effect models. Model II is estimated on all data for which we use the sell-ask spread as dependent variable. Model IV is estimated on appraisal data for which we use the sell-appraisal spread. For both types of dependent variable and both type of data, we see that the estimated December coefficient is substantially lower than the coefficients for September-November. The December discount pattern is intact when we control for unobserved unit heterogeneity.

I view Table 4 as my main exhibit. The repeat-sales structure takes care of time-invariant unobserved unit heterogeneity and the ask price in the sell-ask spread and the appraisal value in the sell-appraisal spread takes care of time-varying elements. We observe that the December coefficient is substantially smaller than the coefficients of September and October and that the effect, compared to September, is about 1.5 percent (-0.0072 - 0.0075 = -0.0147) for model II and 1.7 percent for model IV.[24]

[24] I have run similar regressions for after segmenting on major cities and the pattern is intact. I do not report these exercises, but results are available upon request.

Table 4. Spreads in repeat sales. Norway, 2002-2017

	All data		Appraisal data	
	Sell-ask spread		Sell-appraisal spread	
	I	II	III	IV
	OLS	Unit FE	OLS	Unit FE
Const.	2.62e-3 (1.2e-3)		-0.0182 (1.8e-3)	
Year FE	YES	YES	YES	YES
Jan-Aug FE	YES	YES	YES	YES
Unit FE	NO	YES	NO	YES
September	0.0118 (9.0e-4)	7.5e-3 (1.2e-3)	0.0113 (1.4e-3)	6.61e-3 (1.9e-3)
October	7.50e-3 (9.1e-4)	3.83e-3 (1.2e-3)	4.86e-3 (1.4e-3)	3.19e-3 (1.9e-3)
November	3.73e-3 (9.3e-4)	-1.21e-3 (1.3e-3)	-6.05e-4 (1.5e-3)	-4.37e-3 (1.9e-3)
December	-6.51e-3 (1.1e-3)	-7.20e-3 (1.5e-3)	-0.0106 (1.8e-3)	-0.0105 (2.4e-3)
	No. sales = 2		No. sales = 2	
	N = 213,394		N = 111,244	
R2	0.0370	0.0460	0.0569	0.0756

Notes: Repeat sales data include units that are transacted exactly two times in the 2002-2017 period. FE is short notation for a fixed effect regression run using the plm-function in R and the within-model. Year FE denotes a collection of year dummies (2002 default). 2017 data are included, but the regression specification does not include a dummy for 2017 since 2017-data only cover January. Jan-Aug FE denotes a collection of seven month dummies from January to August, excluding July (default). I report the R2, not Adj. R2, due to the high number of constants in FE models. The sell-ask and the sell-appraisal are spreads computed by taking the difference and dividing by the ask price and the appraisal value, respectively. These spreads are on a percentage basis, so I do not take the logarithms of the dependent variables. The reported standard errors in model II and IV are White heteroskedasticity controlled errors computed using the vcovHC-function in R while standard errors for models I and III are controlled for autocorrelation (due to the presence of two sales per unit for models I and III) and heteroskedasticity computed using the vcovHAC-function in R.

4.4.2 Unobserved heterogeneity in sellers

In order to control for the possibility of unobserved seller heterogeneity, I re-estimate the hedonic model in Table 3 using an instrumented logarithm of ask price from a first stage as an independent variable. I employ the logarithm of appraisal value as an instrument for the logarithm of ask price. The underlying idea is that when I regress the log(ask) onto log(appraisal), potential strategic elements of the ask price are removed in the projection onto the exogenous space that appraisal values represent. At the same time, since both the ask price and the appraisal value reflect the unit's market value, unobserved unit heterogeneity is controlled for.

Table 5 tabulates the results from regressing the logarithm of the sell price onto a space spanned by year fixed effects, month fixed effects, the instrumented logarithm of ask price, and a December dummy plus an interaction term with December and instrumented log(ask).[25] The idea behind the regression is that the last term, the interaction of the December dummy and the instrumented log(ask), would indicate the magnitude of possible disturbance from unobserved seller heterogeneity in the December effect.

We see that the estimated December coefficient in model ii is negative and sta-

[25] The reason why I do not simply include the hedonic model attributes here, or a predicted value based on attributes alone in a base period, is that the attributes and a base period prediction, are constant across the two sales so this would invalidate the unit fixed effect model.

tistically significant. Since the coefficient of the last term in model iii, which is the interaction term, is estimated with a t-value of -1.93 and is statistically significant, there is (some) evidence that hints at the presence of a seller heterogeneity effect. I do not emphasize this result, however, since it is only circumstantial evidence and since there is no increase in R-square between model ii and model iii. The December discount pattern is intact.

Table 5. Log(sell) on instrumented log(ask). Repeat sales. Norway, 2002-2017

		Appraisal data log(sell) on	
	i	ii	iii
	OLS	Unit FE	Unit FE
Year FE	YES	YES	YES
Month FE	YES	YES	YES
September	0.0141 (1.4e-3)	8.04e-3 (1.8e-3)	8.00e-3 (1.8e-3)
$\hat{log}$(ask)	0.985 (6.4e-4)	0.808 (3.9e-3)	0.808 (3.9e-3)
December	-9.52e-3 (1.8e-3)	-7.00e-3 (2.3e-3)	0.110 (6.1e-2)
December×$\hat{log}$(ask)			-8.04e-3 (4.2e-3)
		No. sales = 2	
		No. obs.: 111,244	
R2	0.973	0.928	0.928
P-value	2.2e16	2.2e-16	2.2e-16

Notes: I report the R2 not Adj. R2, because of the high number of intercepts in the FE regressions. $\hat{log}(ask)$ is the predicted value from a two-stage set-up in which I first regress log(ask) on log(app) then using the estimated coefficients to predict log(ask) on the basis of log(app). The White heteroskedasticity-consistent standard errors are estimated using the vcovHC-function in R.

5 Exploration of explanatory mechanisms

5.1 Duration of sale

In the appendix, I outline a skeleton model that I use as an interpretative framework for understanding how market activity affects seasonality in prices. In essence, the model uses the assumption that the arrival of bidders is a stochastic variable, outside of the control of the seller. When there is seasonality in market activity, this spills over into seasonality in sell prices because there are fewer high-quality matches that lead to high prices. One implication is that units are randomly sorted into short and long TOM. Some units that are put on the market in August and September randomly end up as unsold in December. These units receive fewer bidders and bids and, if sold, obtain lower prices. This is a testable implication using segmentations based on TOM.

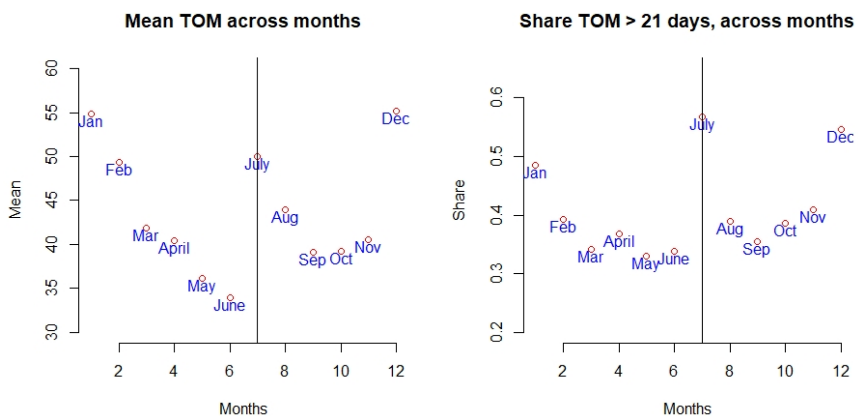
I re-estimate the regressions from above while controlling for TOM. In order to handle the approach to the outcome variable TOM, I first segment into two segments based on TOM, then run the regression without TOM. The results are tabulated in column a and b in Table 6. We observe that for the long TOM segment, there is little difference between the estimated coefficients of September, October, November and December, indicating that when we look at only sales of units that had a long duration on the market (as inventory), the differences in spreads are minimal. The interpretation is that if a unit in September had been sold with a TOM typical of December, then the resulting sell-appraisal spread would tend to be as low as they typically are in December. The converse also holds, if a unit had been sold with a long TOM in December and the TOM was equally long as the one in September, the difference in spreads would be minimal. Figure 4 plots mean TOM across months and shows that there are substantial differences in TOM across months. In the spring and early summer months of May and June, TOM is at its lowest. These months also have the lowest share of long TOM transactions, as defined by the share of TOMs above 21 days. We notice that mean TOM is almost as long in July as it is in December, and we observe that the share of long TOM transactions is at least as large in July as in December. Thus, the long TOM effect does not appear to be merely a winter effect. Rather, it appears to be a market activity effect.

Table 6. Segmentation on TOM. Appraisal data, repeat sales. Norway, 2002-2017

	Appraisal data	
	Sell-appraisal spread	
TOM segment	FE 0-21 days, both sales a	FE 22- days, both sales b
Year FE	YES	YES
Jan-Aug FE	YES	YES
Unit FE	YES	YES
September	-0.00534 (2.7e-3)	-0.0108 (2.8e-3)
October	-0.00737 (2.8e-3)	-0.0138 (2.8e-3)
November	-0.0123 (2.8e-3)	-0.0186 (2.9e-3)
December	-0.0133 (3.5e-3)	-0.0138 (3.4e-3)
	No. sales = 2	
No. obs.	61,466	25,356
R2	0.0637	0.0535

Notes: In column a, I first segment appraisal data on TOM equal to or below 21 days (238,308 obs.), then retain only units that are sold exactly twice. (Some observations are lost because one TOM is above 21 days.) In column b, I first segment appraisal data on TOM equal to or larger 22 days (135,065 obs.), then retain only units that are sold exactly twice. FE is short notation for a fixed effect regression run using the plm-function in R and the within-model. Year FE denotes a collection of year dummies (2002 default). 2017 data are included, but the regression specification does not include a dummy for 2017 since 2017-data only cover January. Jan-Aug FE denotes a collection of seven month dummies from January to August, excluding July (default). I report the R2, not Adj. R2, due to the high number of constants in FE models. The sell-appraisal spreads are computed by taking the difference and dividing by the appraisal value. The spreads are on a percentage basis, so I do not take the logarithms of the dependent variable. The reported standard errors are White heteroskedasticity controlled errors computed using the vcovHC-function in R.

Figure 4. Mean TOM and share long TOM across months. Norway, 2002-2017



Notes: The graphs were generated using all transaction data, described in the upper panel of Table 1. N = 691,192.

5.2 The repeatability of TOM

Since the TOM in an individual transaction is related to the eventual sell price of a given unit, it is possible that TOM is an outcome variable that is associated with unobserved unit heterogeneity. If so, some units should tend to have long TOMs;

other short TOMs. The implication is that a unit's TOM should be forecastable on the basis of that unit's TOM in an earlier transaction. In Table 7, I present evidence to the contrary. The table tabulates results from a regression in which the second TOM for each unit is regressed onto the first TOM. We make two observations. First, the adjusted R-square is 0.000651, only a little more than half of one percent of the variation in the second TOM is explained by the variation in the first TOM. Second, when a given first-sale TOM is 10 days longer than the intercept of 36 days, the second-sale TOM for the same unit tends to be less than 1 day longer. There is substantial reversion to the mean. This evidence suggests that TOM is not linked to units; it is determined by other processes.

Table 7. Repeatability of time-non-market for each units. Norway, 2002-2016

	Transaction data Time-on-market second sale	
Intercept	36.4	(0.26)
Time-on-market first sale	0.0838	(4.2e-3)
	T = 2 No. of obs.: 106,697 (units) Adj. R2 = 0.00651 p-value = 2.2e-16	

Notes: The regression was run on a data set in which units have been observed sold exactly twice. Time-on-market (TOM) is computed as the difference in days between the date on which the unit was announced for sale on the online platform Finn.no and the date on which the highest bid was accepted. The standard errors were computed using the vovHC-function that controls for heteroskedasticity.

5.3 Market activity

In Table 8, I present results from a regression of market spreads on market activity. First, I identify municipalities with a sufficient amount of transactions, then I compute the sell-ask spread for December and January-November for each sub-market. I regress the 247 mean sub-market sell-ask spreads onto relative December vs. non-December market activity. Here, market activity is either transaction volume or number of new advertisements of for-sale units.[26] We observe that in all four regressions the sign is positive, thus there is an association between market activity and the sell-ask spread. Using the sell-ask spread is convenient because the ask price reflects unobserved unit heterogeneity and it also encompasses the price trend. Regressions a, c, and d use transaction volumes as a metric of market activity. Since there could be an endogeneity issue in that both the spread and the transaction volume are based on the acceptance of the highest bid, I also include regression b, which uses the number of new advertisements of for-sale units as a metric of market activity. Results from regressions a, b, and d are clearly statistically significant. Regressions c and d are run on segments of markets. In column c, I use markets with long TOM (above or equal to the median over the within-market means). This is the regression with the coefficient with the lowest t-value. The t-value is higher for the short TOM segment. However, I urge caution in interpreting these numbers since t-values are stochastic variables

[26] Note that these for-sale advertisements represent units that were eventually sold since my data set is a transaction data set. At the time, there were, presumably, some advertisements for units that were never sold and thus were never included in the transaction data set. I have no information on such never-sold units.

that are functions of sample size and since both the dependent and independent variables are constructed variables from aggregate statistics. Thus, there is an element of data mining in these exercises. That said, the evidence appears to indicate an association between market activity in December and sell-ask spread. More activity is associated with larger spread.

Table 8. Regressions of market observations. Market sell-ask spread on market activity in December. Norway, 2002-2016

	Transaction data			
	Sell-ask spread. Dec less not December			
	All markets		Long TOM	Short TOM
Market activity measure:	sales	ads	sales	sales
	a	b	c	d
Intercept	-0.0336 (5.2e-3)	-0.0250 (2.4e-3)	-0.0382 (9.6e-3)	-0.0332 (5.9e-3)
Rel. Dec. vol.	0.0285 (0.011)		0.0315 (0.019)	0.0348 (0.013)
Rel. Dec. for-sale		2.0e-3 (8.1e-4)		
No. of markets	247	247	124	123
No. obs. 247 markets	627,405			
Adj. R2	0.0462	0.0381	0.0405	0.0852

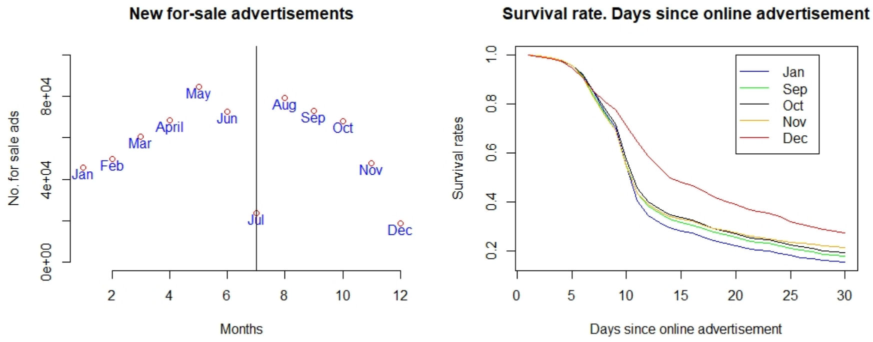
Notes: The regressions were run after partitioning Norway into different markets. First, I removed municipalities with less than 800 transactions in the period. This leaves us with 627,405 transactions. Then, I partitioned each municipality into apartments and non-apartments (detached houses, semi-detached houses, and row houses). I also removed transactions from the year 2017 since my records are not complete for this year. Markets with no December transactions were removed. For each market, I compute the mean sell-ask spread in December and non-December and the transaction volume in December and non-December. I also compute number of units that were put up for sale in each market (online advertisements) for December and non-December. The for-sale registration date is the date on which the unit was announced for sale on the online platform Finn.no. The variables Relative December transaction volume and Relative December for-sale registration are the number of transactions and registrations in December compared to the other 11 months. Long TOM and short TOM markets were defined as markets with intra-market mean TOM above or equal to across-markets median TOM and intra-market mean TOM below across-markets median TOM, e.g. a market belongs to the short TOM segment if the mean TOM in that market is below the median TOM among the within-market means across markets. The standard errors were computed using the vovHAC-function that controls for both heteroskedasticity and autocorrelation.

6 Discussion

To explore market activity in more detail, one avenue is to study the arrival of new for-sale advertisement. After all, the frequency with which new sellers arrive in the market is per definition market activity. In Figure 5, the left-hand side panel plots the number of new advertisements for each month during the period 2002-2017. The month with the highest number of new advertisements is May. The month that has the lowest number of new advertisements is December. This low rate of arrival of new supply translates into fewer matching opportunities, which in turn imply lower rate of bids, as long as inventory from earlier months does not obtain off-setting higher rates of bids.

If buyers observe that fewer units are put on the market, their estimate of their probabilities of good matches change (see the search-and-matching framework in the Appendix). The implication is that the new inventory is associated with longer TOM since the new units stay unsold longer. The right-hand side panel of Figure 5 plots the number of units that have not been sold within 30 days, i.e. the survival rate, for each month of registration. The red line represents December. We see that the rate with which units are sold is lower for each day on the market for units put on the market and advertised in December. This pattern supports the notion that both new sellers and new (acceptable) bids arrive in smaller numbers in December, offering supporting evidence of the idea of co-movement of prices, TOM and market activity.

Figure 5. New ads and survival rates. Norway, 2002-2017



Notes: The left-hand side panel plots the mean number of new for-sale units advertised on the online platform for each month in the period 2002-2016. The right-hand side panel graphs segments on month of online registration and retains units that were sold within 30 days. It shows the rate of unsold units (survival rate) for each day 1-30 since the advertisement was posted on the online platform Finn.no

7 Conclusion and policy implications

There have been reports for some time in Norway that house prices tend to fall in December. I document that this is indeed the case, and rule out that a composition effect can explain all of the December discount. A fully specified hedonic model with a time trend yields statistically significant and economically substantial estimates of a December coefficient.

I control for unobserved unit and seller heterogeneity using a battery of techniques. I use repeat-sales to control for time-invariant unit-specific effects, and ask prices to account for time-varying unit-specific effects. To control for a potential strategy-effect in ask prices, I complement the analysis using appraisal values instead of ask prices. In addition, I combine ask prices and appraisal values in an instrumental variable approach. To some extent, this last approach can mitigate unobserved heterogeneity effects among sellers because a potential difference in using the ask price in relation to accepted bids (reservation price) is removed when I regress ask prices onto the exogenous plane consisting of appraisal values. The December discount is intact across specifications and robust to data set changes.

My results indicate that the December discount is 1.5 percent compared to September price levels. The December discount is associated with long TOMs, since segmentation of sales into TOM segments imply that the difference between the estimated September coefficient and the estimated December coefficient vanishes for long TOMs. TOM is not related to the unit, since a repeat-sales set-up reveals that there is little persistence between first sale TOM and second sale TOM of the same unit.

The December discount is linked to market activity. Segmenting Norway into 247 sub-markets, we observe that the December discount is larger in sub-markets with lower market activity in December.

Since the December discount is at least 1.5 percent of the price of a house, and since part of this appears to be linked to market activity and sub-optimal matching, the December discount might be indicative of welfare losses. The evidence is consistent with a search-and-matching model in which the low prices result from a low number of high-quality matches between buyer preferences and unit attributes. Thus, there could be welfare gains to be made if one arranges housing markets to ensure optimal matching between buyers and units. Potentially, better matching would be achieved either by nudging more sales in low-activity periods or inducing sales in low-activity periods to be moved to high-activity periods.

Sellers and buyers may also take notice. The December discount appears to be associated with long TOMs, so one advice to sellers could be to remove an unsold unit from the market if it has been on the market for a long time and wait for periods of highre market activity. Sellers would also be advised to keep in mind the high survival rates in December. Among all months, the survival rates are the highest in December, which means that for new advertisements December is the month with fewest quick sales. There is, however, potentially bargains to be made for buyers. Repeat-sale analysis shows that a unit that previously was sold in a non-December month at no discount could sell in December with a discount. The most likely reason is the few market participants. The implication is that a buyer would have fewer competitors in bidding rounds in December. Thus, if a buyer knows what she wants and finds a unit with high-match quality, and that unit has been on the market for a long time, it could be possible to negotiate a discount. It appears that December is a month in which such negotiations lead to a sale. The caveat for buyers is that the probability of a good match between preferences and attributes is lower in low-activity markets.

References

Albrecht, J., A. Anderson, E. Smith, and S. Vroman (2007): Opportunistic matching in the housing market, *International Economic Review*, **48**: 2, pp. 641 - 664.

Anglin, P. M., R. Rutherford, and T. M. Springer (2003): The trade-off between the selling price of residential properties and time-on-market: The impact of price setting, *Journal of real estate finance and economics*, **26**: 1, pp. 95 - 111.

Anglin, P. M. and R. Wiebe (2013): Pricing in an illiquid real estate market, *Journal of real estate research*, **35**: 1, pp. 83 - 102.

Anundsen, A. K. and E. Røed Larsen (2018): Testing for micro efficiency in the housing market, *International Economic Review*, **59**: 4, pp. 2133 - 2162.

Anundsen, A. K., B. K. Kivedal, E. Røed Larsen, and L. A. Thorsrud (2020a): Behavioral changes and policy effects during Covid-19: Evidence from day-by-day sales and bid-by-bid auction logs in the housing market, Housing Lab Working Paper Series 2020-3.

Anundsen, A. K., E. Røed Larsen, and D. E. Sommervoll (2020b): Strategic price-setting and incentives in the housing market, submitted manuscript.

Bulow, J. and P. Klemperer (1996): Auctions versus negotiations, *American Economic Review*, **86**: 1, pp. 180 - 194.

Case, K. E. and R. J. Shiller (1989): The efficiency of the market for single-family homes, *American Economic Review*, **79**: 1, pp. 125 - 137.

Coles, M. G. and A. Muthoo (1998): Strategic bargaining and competitive bidding in a dynamic market equilibrium, *Review of Economic Studies*, **65**: 2, pp. 235 – 260.

Diaz, A. and B. Jerez (2013): House prices, sales, and time on the market: A search-theoretic framework, *International Economic Review*, **54**: 3, pp. 837 - 872.

Genesove, D. and L. Han (2012): Search and matching in the housing market, *Journal of Urban Economics*, **72**: 1, pp. 31 - 45.

Hsieh, C.-T. and E. Moretti (2019): Housing constraints and spatial misallocation, *American Economic Journal: Macroeconomics*, **11**: 2, pp. 1 - 39.

Kaplanski, G. and H. Levy (2012): Real estate prices: An international study of seasonality's sentiment effect, *Journal of Empirical Finance*, **19**: 1, pp. 123 - 146.

Kashiwagi, M. (2014): A search-teoretic model of the rental and homeownership markets, *Journal of Housing Economics*, **26**, pp. 33 - 47.

Krainer, J. (2001): A theory of liquidity in residential real estate markets, *Journal of Urban Economics*, **49**: 1, pp. 32 - 53.

Leamer, E. E. (2007): Housing is the business cycle, NBER Working Paper No. 13428.

Leamer, E. E. (2015): Housing really is the business cycle: What survives the lessons of 2008-09? *Journal of Money, Credit, and Banking*, **47**: 1, pp. 43 - 50.

Nenov, P., E. Røed Larsen, and D. E. Sommervoll (2016): Thick Market Effects, Housing Heterogeneity, and the Determinants of Transaction Seasonality, *Economic Journal*, **126**: 598, pp. 2402 - 2423.

Ngai, L. R. and S. Tenreyro (2014): Hot and cold seasons in the housing market, *American Economic Review*, **104**: 12, pp. 3991 - 4026.

Novy-Marx, R. (2009): Hot and cold markets, *Real Estate Economics*, **37**: 1, pp. 1 - 22.

Maury, T.-P. and F. Tripier (2014): Search strategies on the housing market and their implications on price dispersion, *Journal of Housing Economics*, **26**, pp. 55-80.

Røed Larsen E. (2020): Intra-week price patterns in the housing market, forthcoming, *Scandinavian Journal of Economics.* DOI: https://doi.org/10.1111/sjoe.12403.

Røed Larsen, E. and S. Weum (2008): Testing the efficiency of the Norwegian housing market, *Journal of Urban Economics*, **64**, pp. 510 - 517.

Taylor, C. (1999): Time-on-the-market as a sign of quality, *Review of Economic Studies*, **66**, pp. 555 - 578.

Appendix

A skeleton model of search-and-match when there is seasonality

The December discount implies that when a unit is sold in December it sells for a price that is lower than the price had been if, counter-factually, the same seller had sold the same unit in September the same year. The skeleton model in Anundsen and Røed Larsen (2018), which is a simpler framework than the models of Albrech et al. (2007) and Diaz and Jerez (2013), is a guide that structures the search and match thought process I follow when I consider the mechanism that could generate the co-movement in prices, transactions, and TOMs. Let the number of bidders $N_{h,t}$ for a housing auction of unit h at time t be Poisson distributed, $N_{h,t} \sim \text{Poisson}(\mu_{h,t})$, in which $\mu_{h,t}$ is the expected number of bidders for unit h at time t. Let $\mu_{h,t} = \mu$ be common to all auctions for units of the similar type across time. The probability that an auction of unit h draws k bidders is a stochastic variable.

Let bidder b be among the k bidders. Bidder b has preferences F_b and considers the housing consumption utility $u(M_{b,h})$ he may extract from a match $M_{b,h} = m(F_b, AT_h)$ with house h with attributes A_h. For simplicity, I classify match-quality into three types, high, medium, or low:

$$M_{b,h} = \begin{cases} H, & m(F_b, AT_h) \geq m_H \\ M, & m_L < m(F_b, AT_h) \leq m_H, \\ L, & \text{otherwise,} \end{cases} \tag{4}$$

in which $m(F_b, AT_h)$ is the outcome when preferences and attributes are in-putted into the match-quality function.

Bidder b forms his willingness-to-pay (WTP) for a unit h on the basis of a utility-maximization over the utility stream from house h and other consumption. In this utility-maximization, bidder b optimizes what match-quality he may derive from unit h, $M_{b,h}$, by comparing the possible match-quality of other units and the utility from the consumption other goods while heeding his financial budget constraint imposed by his income I_b and equity E_b: $WTP_{b,h} = w(M_{b,h}, I_b, E_b)$. The function $w(.)$ incorporates access to credit and bank LTV-regulations.[27] The function $w(.)$ is increasing in all three arguments.

Given the supply of houses for sale at time t, S_t, and the number of buyers on the market, B_t, the probabilities of good, medium, and low matches are $\rho_{G,t}(S_t, B_t), \rho_{M,t}(S_t, B_t)$ and $\rho_{L,t}(S_t, B_t)$ such that $\sum_j \rho_{j,t}(S_t, B_t) = 1$, $j = G, M$, and L. Assume that the probability functions are time-invariant, $\rho_{j,t}(S_t, B_t) = \rho_j(S_t, B_t)$, $j = G, M$, and L, even if the inputs, i.e. the numbers of sellers S_t and buyers B_t are not. The number of bidders $N_{h,t}$ for unit h consists of three types, $N_{G,h,t} + N_{M,h,t} + N_{L,h,t} = N_{h,t}$. The expected number of good matches is $E(N_{G,h,t}(S_t, B_t)) = \rho_G(S_t, B_t)E(N_{h,t}(S_t, B_t))$. Thus, the number of expected good matches $E(N_{G,h,t})$ is increasing in the expected number of bidders $E(N_{h,t})$: $\frac{\partial E(N_{G,h,t}(S_t,B_t))}{\partial E(N_{h,t})} = \rho_G(S_t, B_t) \geq 0$, which is a result consistent with Bulow and Klemperer (1996, p. 185) who compare auctions with N bidders to auctions with $N + 1$ bidders and find that "the auction with the extra bidder yields a higher expected revenue". Thus, in low-activity markets we expect a reduced number of

[27] In Norway, the financial authorities ask that banks limit credit to five times income and require 15 percent equity; with the possibility of waivers in certain cases after a "speed-limit" enforced by the authorities.

bids that classify as acceptable to the seller.

Let the seller's reservation price of unit h be $R_{h,t}$. Assume that the reservation price is time-invariant so $R_{h,t} = R_h$. The sell price for unit h becomes equal to the second highest WTP across $WTP_{b,h}$ for the number of bidders $N_{h,t}$ when $N_{h,t} \leq 2$ as long as the second highest WTP is above the reservation price R_h. When only one willingness-to-pay among bidders b, $WTP_{b,h}$, is above the reservation price R_h, the sell price $P_{h,t}$ becomes the reservation price:

$$P_{h,t} = \begin{cases} \pi_h = max_{-1,B}(WTP_{B,h}), & N_{h,t} \geq 2, max_{-1,b}(WTP_{b,h}) \geq R_h \\ R_h, & max_b(WTP_{b,h}) \geq R_h, max_{-1,b}(WTP_{b,h}) < R_h \\ \text{no transaction}, & \text{otherwise.} \end{cases} \tag{5}$$

The notation $max_{-1,B}(WTP_{B,h})$ denotes the second highest WTP for unit h among all bidders in the set of bidders B when the number of bidders is at least two. Since π_h is at least as high as the reservation price, the transaction price $P_{h,t}$ is non-decreasing in number of bidders $N_{h,t}$. To see this, keep in mind that an increase in $N_{h,t}$ in expectation is associated with an increase in the number of good matches, $N_{G,h,t}$.

As in Anundsen and Røed Larsen (2018), the probability that the number of good matches, given $N_{h,t}$, is equal to n follows a binomial distribution. Since the highest price π_h requires at least two good matches, the probability that that the sell price is equal to the highest price, $Prob(P_{h,t} = \pi_h)$, is the sum of probabilities that the number of good matches is equal to two, three, or more. Following Anundsen and Røed Larsen, we then see that the probability of a high price $Prob(P_{h,t} = \pi_h)$ is increasing in the number of bidders $N_{h,t}$, $\frac{\partial Prob(P_{h,t}=\pi_h)}{\partial N_{h,t}} > 0$. Thus, if there are few bidders in December, there is a decrease in the expected frequency of auctions that result in high prices.

Nenov, Røed Larsen, and Sommervoll (2016) show that December is, indeed, a month of low market activity in Norway. Here, Table 2 tabulates that the number of transactions in December is only one third of the number in September. The implication is that sell prices should be lower in December, everything else being the same. Moreover, since search takes time, fewer buyers implies that time-on-market increases in December.

Acknowledgements:

Housing Lab is partly funded by the Norwegian Ministry of Finance and the Norwegian Ministry of Modernisation and Municipalities. Housing Lab also receives financial support from Krogsveen and Sparebank1-Gruppen. We are greatful for the financial support. All views expressed in our papers are the sole responsibility of Housing Lab, and does not necessarily represent the views of our sponsors. The authors are grateful to DNB Eiendom and Eiendomsverdi for auction and transaction data.

Authors:

Erling Røed Larsen, Housing Lab, Oslo Metropolitan University, Norway;

email: erlingro@oslomet.no

ISSN: 2703-786X (Online)

OSLOMET

HOUSING LAB
Nasjonalt senter for boligmarkedsforskning