HOUSING LAB WORKING PAPER SERIES

2021 | 2

Is the housing market an inequality generator?

Terje Eggum and Erling Røed Larsen

OSLOMET

HOUSING LAB
Nasjonalt senter for boligmarkedsforskning

Is the housing market an inequality generator?*

Terje Eggum† and Erling Røed Larsen‡

Abstract

We study inequality generated by capital gains in the housing market 2007-2019 by exploiting a combination of two countrywide data sources in Norway: a registry of housing units and a database of transactions. The combination allows us to construct a panel dataset of owners. We identify and follow all individuals in Norway, belonging to six birth cohorts, who were owners in the period 1 January 2007 – 1 January 2019, and estimate the sum of their actual and potential capital gains from all of their owned and sold properties. We document substantial increase in capital gains inequality over the period; both across and within geographical strata and across and within birth cohorts. The house price index Granger-causes capital gains inequality and market downturns are associated with reductions in inequality. Using variation across municipalities, we demonstrate that capital gains inequality is positively associated with changes in income and house price levels.

Keywords: Capital gains, Granger causation, housing inequality, owner-occupiers

JEL Codes: C20, R21, R31

*The authors are grateful to André Kallåk Anundsen, Andreas Eriksen, Jeanette Fjære-Lindkjenn, Chao-Hsi Huang, Lorien Rice, and Dag Einar Sommervoll for comments and suggestions. The authors also thank Eiendomsverdi for access to transaction and registry data and participants at the 94th Western Economic Association International conference; in addition to participants at seminars and work-shops.

†Eiendomsverdi

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

1 Introduction

Inequality is a topic that sits atop many contemporary discussions on economic challenges and a need to document its patterns and a desire to understand its sources have emerged in these discussions. In fact, the interest among economists into inequality and sources of inequality seems to be fast increasing. In economics, however, studies of inequality typically examine aspects such as wealth (Heathcote et al. (2010)), income (Furceri et al. (2018)), and consumption (Krueger and Perri (2006), Blundell and Etheridge (2010), Attanasio and Pistaferri (2014, 2016)). Few, if any, studies follow individual homeowners over a period of time; across multiple years and multiple transactions; and investigate the time development of the inequality created by individual accumulation of capital gains in the housing market. One reason for this paucity of analysis is the lack of data on individual owners across multiple years. Analysts of housing capital gains inequality need unique identifiers of individuals, houses, and transactions to be able to estimate capital gains by following individual owners over years. Moreover, to create a panel of capital gains accumulation over longer observation periods for owners who do not sell, but continue to own the unit, requires an accurate valuation method based on a sufficient number of observed attributes. This article combines data sources to overcome these data challenges and asks one simple question: Is the housing market an inequality generator?

Our study answers in the affirmative and our contribution comprises accessing data, mapping of capital gains inequality, and econometric estimation of relationships. It consists of four key empirical findings on capital gains inequality. First, capital gains inequality increases over the time-period 1 January 2007-1 January 2019. Second, house prices Granger-cause capital gains inequality. Third, capital gains inequality displays substantial variability along spatial and age dimensions. Fourth, capital gains inequality is tied to income development. The findings on Granger causality and that capital gains are unevenly distributed and linked to income may not be surprising, but, nevertheless, this article presents potentially useful patterns that document these links and offers estimate on magnitudes. In fact, we consider

the magnitudes of the estimated effects as the cornerstones of our contribution. For example, we show that the Gini index of house values 1 January 2007 was 0.26. On 1 January 2019 it was 0.29. We find that in the segment of owners who have owned a unit in Oslo, the 90th percentile of capital gains 2007-2019 is NOK 3.35 million. For comparison, the average monthly wage in Norway in 2019 across all sectors was NOK 45,610, thus the 90th percentile of capital gains is almost 80 times larger than the average monthly wage before tax. At the same time, the 90th percentile of 2007-2019 capital gains among owners who owned outside of Oslo is found to be NOK 1.67 million. The difference between the 90th percentiles in the Oslo segment and the non-Oslo segment indicates substantial dispersions in housing capital gains.

In fact, when we partition owners into 20 groups of owned values on 1 January 2007, and sort by magnitudes, we find that the capital gains over the period 2007-2019 for each of these 20 groups, not only was a curve with a positive slope, but a curve with an increasing slope. The group with the highest top five percentiles of owned values in 2007, i.e. group 20, experienced a capital gain over the next 12 years of NOK 2,862,503 while the group with the second highest owned values in 2007, i.e. group 19, experienced a capital gain of NOK 1,778,632. This means that group 20 had 61 percent larger capital gains than group 19. Group 19 in turn had 84 percent larger capital gains than group 10, which had capital gains of NOK 966,636.[1]

In order to obtain these empirical findings, we combine data sources that allow us to follow individual owners over the twelve-year period and to estimate the value of a non-transacted unit: a registry of all housing units and their owners and a registry of transaction data. We first identify every individual, among six birth cohorts spanning the period 1965-1990 in five year intervals, i.e. 1965, 1970, 1975, 1980, 1985, and 1990[2] who owned a home on 1 January 2007 and in every subsequent quarter. We then follow these individual owners through each

[1]These, and more, statistics can be seen illustrated in our Figures 7 and 8.

[2]There are only 72 individuals from 1990 in our panel. They are included in order shed light on the development among very young owners. For these under age owners, there are special laws.

quarter for 12 years. In our finest granularity, we are able to estimate capital gains on a quarterly basis for each owner-occupier. We combine these micro data with aggregated tax data for each Norwegian municipality for the period 2007-2017 in order to find out how income is linked to capital gains inequality in the housing market across municipalities.

We limit our study to individual owners and do not study the inequality that arises between owners and tenants. We also exclude firms. Our focus of attention is on individual, private owners, but we do include owners who hold more than one unit. We also include individuals who own shares of a unit by co-owning with a spouse, a partner, or a friend. We compute the capital gains each owner accumulates for each ownership period and each share, and we estimate the capital gains both for units they have owned throughout, units they have owned and sold, and units they have purchased and still own on 1 January 2019. Our panel consists of 77,554 owners who owned at least one housing unit on the start of each quarter in the period 1 January 2007-1 January 2019.

There are three categories of capital gains. The simplest category comprises individuals that have purchased and sold within the period. Then we have observations on both transaction prices, and we denote the difference realized capital gains. Such capital gains are readily observable from transaction data. However, an individual may have owned a unit at the start of the period, i.e. on 1 January 2007, and then sold it during the period or purchased a unit during the period and owned it to the end of the period, i.e. 1 January 2019. Then we have one actual transaction price observation and need to estimate the value of the unit at the other point in time. We may denote the difference as semi-realized capital gains. The third category comprises individuals that have owned the unit throughout the whole period. For these holders, we need to estimate the value of the unit both at the beginning and at the end of the 2007-2019 period. We denote the difference between the two estimates potential capital gains. Total capital gains for an individual is defined as the sum of realized, semi-realized, and potential capital gains.

In order to compute the potential gains gains for each owner in each quarter, also when

they hold without selling, we employ Eiendomsverdi's AVM (Automated Valuation Method). In the appendix, we demonstrate the validity of this approach by showing the high precision of the AVM.[3]

We study only the capital gains individuals have enjoyed in the housing market, and not gains from changes in labor income, payments on principal, inheritance, stock market returns, or any other source of wealth accumulation. The idea is to zoom in on the housing market only and the gains made therein. In fact, our aim is to isolate the housing market from all other sources of income and wealth, and only study this source of wealth accumulation.

We believe our exercise is useful because the results may be relevant when policymakers think about the sources of inequality and whether or not they can or want to do something about it. While multiple authors have mapped sources and effects of inequality arising from differences in income, ability, consumption, and financial wealth, fewer have been able to map the differences in purchasing power that arises with differences in housing capital gains. We want to examine whether the housing market is a systematic source of inequality, and if so, the magnitude of the wealth accumulation it offers. Put differently, we seek to find the inequality that potentially arises between individual choices made on where and how to live. Since many households in modern economies partake in housing capital gains and since the gains appear to be large compared to e.g. labor income, the outcomes can influence decisions and perceptions in the whole economy. For perspective, in Norway there are between about 2.6 million housing units[4], depending among other things how one counts secondary homes. The mean transaction value in early 2019 was NOK 3.7 million per unit. A back-of-the-envelope estimate of the market value of the stock of houses in Norway would be 9.6 trillion, about 2.7 times the GDP of about NOK 3.5 trillion (including off-shore oil activity). Thus, inequality arising from capital gains in the housing market may be a major

[3]For this exercise, we use the 23,374 transacted units among the 77,554 owners in the six birth cohorts. We compute the spread between estimated value and observed transaction price as percentage of observed transaction price. The median spread is -1.3 percent. The 10th and 90th percentiles are, respectively, -12 percent and 11 percent.

[4]See Statistics Norway for current estimates; available online https://www.ssb.no/en/bygg-bolig-og-eiendom/bolig-og-boforhold/statistikk/boliger.

source of inequality and it could even be a determinant of the economic path of the nation. Thus, mapping this inequality should provide relevant information source for policymakers and economists. It should also be relevant outside Norway's borders since many countries have experienced similar booms in the housing market.

This article is organized in the following way. The next section presents a brief literature review. The third section describes the data, the institutional framework, and the principles behind the automatic valuation method (AVM) we employ. In the fourth section, we go through which empirical techniques we use and in the subsequent section we present a few motivating, basic patterns seen in the data. Section six comprises our main empirical results. Section seven discusses the geographical dimension and the relationships between income developments and capital gains developments on a municipality level. The last section concludes and offers policy implications.

2 Literature

There exists a large research literature on the housing market and there is also a substantial literature on inequality. However, the literature on the intersection of the housing market and inequality is smaller. This article's brief literature review is meant as a brief overview exercise, in which we place relevant articles in a circle of interest surrounding a topic center of capital gains inequality. The underlying idea of this arrangement is that capital gains inequality is related to many other studies by common themes such as determinants, units of study, and empirical techniques, but that these studies come from an array of branches of economics.

For example, Krueger and Perri (2006) ask whether income inequality leads to consumption inequality. They find that the increase in income inequality does not spill over into a corresponding rise in consumption inequality. They measure consumption as a flow of goods and services, and the consumption of housing services is captured by paid rent (tenants)

and self-reported hypothetical rent (owner-occupiers). Aguiar and Bils (2015) attempt to correct for systematic measurement errors in the Consumer Expenditure Survey and find that consumption inequality tracks income inequality more closely than expenditure-based evidence. Albouy and Zabek (2016) study inequality in house prices and rents and while pointing out the paucity of studies on housing outcomes.

Attanasio and Pistaferri (2014) explain that one of the limitations of consumption inequality studies has been that the only source for estimation has been the Consumer Expenditure Survey (CES). They use a new measure sourced from the redesigned Panel Study of Income and Dynamics (PSID) data, and they emphasize that observations on rent is an important ingredient even if there is no information on rent equivalents for non-homeowners. For rent, they use an imputed rent measure equal to six percent of self-reported home value. In their broader review of the inequality literature, Attanasio and Pistaferri (2016) use only survey information on imputed services such as self-reported rent for homeowners. They study food, certain durables, and leisure, but do not discuss the role housing has in utility production.

In an attempt to broaden the scope from wage observations and consumption measures based on expenditures, Heathcote et al. (2010) examine wealth inequality using the Survey of Consumer Finances (SCF) and demonstrate that the net worth Gini coefficient increases by 5 points from 1983 to 2007. Piketty and Zucman (2014) study wealth-to-income ratios in the long-run and find that wealth-income ratios has risen strongly from 1970 to 2010. Blundell and Etheridge (2010) show that, for Britain, inequality growth has been episodic. It rises in the early 1980s, then stabilizes, then rises in the late 90s. They write: "The transmission from wages and income through to consumption is of considerable interest in understanding the workings of the economy at both the macro and micro levels." They demonstrate a difference in the development of income and consumption inequality in Britain since the two series break apart in the late 80s. They mention that "especially the value of real estate" as a possible explanation, but cannot offer empirical support. Benhabib and Bisin (2018)

survey the literature on the mechanisms underlying wealth distributions. Benhabib, Bisin, and Luo (2017) say: "The literature has largely emphasized the role of earnings inequality in explaining wealth inequality." They show, however, that the relationship is dubious, at best, by demonstrating that across the world, earnings Gini indices have little correlation with wealth Gini indices. They do not consider the housing market.

Effects from policy or the business cycle on inequality is a topic on which there exists several contributions. For example, Furceri et al. (2018) find that contractionary monetary policy appears to increase income inequality. Karahan and Ozkan (2013) ask whether an income shock persists through the life cycle. They answer that for young workers, shocks to earnings are only moderately persistent. For middle-age workers, shocks are persistent. Barlevy and Tsiddon (2006) find supporting evidence for a model that implies that recessions exacerbate trends. Their findings are consistent with recessions having an adverse effect on inequality. We shall see below that the opposite appears to be the case for capital gains inequality in the housing market.

The effect from the business cycle on inequality raises the deeper question of locating the sources of inequality. Van Nieuwerburgh and Weill (2010) ask why house price dispersion has gone up and construct a model in which households with heterogeneous abilities exit and enter areas and since the housing supply cannot respond rapidly, house prices respond instead. Hugget et al. (2011) study sources of lifetime inequality. They find that differences in starting conditions at age 23 account for more of variations in lifetime outcomes than shocks after 23. Di Nardi and Fella (2017) ask why some people are wealthy and others are poor and say the answers require that we understand why people save. They study inter-generational mechanisms, human capital, preferences, earnings, medical risks, random shocks, and entrepreneurship, but say little about the role played by houses and choices of residential purchases.

The above-mentioned questions of inequality trends, sources of inequality, life-cycle stages, and effects from the business cycle on inequality share key words with our study. These stud-

ies indicate that there is a paucity of knowledge about the inequality of housing capital gains. This article seeks to address that paucity.

3 Data and institutional background

3.1 Data on house transactions and ownership

We combine the countrywide registry of units and transactions and construct a dataset of 40,926,589 records. We use the registry to classify an owner as a person who has bought a property, but not yet sold it. A person can buy a share of a property and can change this share over the period we study. The number of records is large since we count as a record an individual's ownership status for each quarter in our period, which is a necessary requirement for constructing our panel. The starting point for ordering data is the sub-set of registered, unique individuals belonging to birth year cohorts the period 1965-1985 who were owners on 1 January 2007. The resulting dataset is intractably large. Thus, we limit our study to cohorts from 1965, 1970, 1975, 1980, 1985, and 1990. For more information on Norwegian data sources, see Fagereng et al. (2020).

It is possible to suggest that the 1960 cohort is more useful than the 1990 cohort. The 1960 cohort, however, is very large and presumably quite similar to the 1965 cohort. In contrast, there are only 72 individual owners in our panel from 1990, thus they do not pose a data tractability challenge. They could, however, potentially offer us a glimpse into a youth dimension, i.e. that the 1990 cohort is relatively different from the 1985 cohort.[5]

In our procedure, we examine the owners in each quarter, i.e. on 1 January 2007, 1 April 2007, 1 July 2007, ..., 1 January 2019, and identify a list of individuals who were owners in all quarters. For these 77,554 always-owners, we compute the cumulative capital gains for each quarter. The resulting dataset is a balanced panel of individuals who were owners on

[5]In Norway, an individual becomes a legal adult at 18. Under-age individuals may be the legal owner of a unit, i.e. from inheritance. However, under-age owners cannot fully dispose of the property, but have to consult with legal guardians.

the first day of each quarter during our 12 year period.

Transaction data and data on housing unit characteristics is the basis for Eiendomsverdi's Automatic Valuation Method (AVM). We access this AVM in order to obtain estimates of the market value for all housing units owned by our 77,554 always-owners. The AVM estimates were computed on a given date, 18 May 2019. From this date, we back-estimate values using Eiendomsverdi's house price index, which is constructed with a high spatial resolution. This index is based on the same methodology as the more parsimonious, official index of Real Estate Norway. In the appendix, we include a validation exercise that demonstrates the accuracy of the AVM by comparing differences between AVM estimates and transaction prices.

Capital gains belong to one of several categories: realized, semi-realized, or potential. If an owner has bought and sold a unit, the capital gains are realized and directly observable. For this category, we subtract the observed purchase price from the observed sell price. Negative capitals gains are possible. If an owner sells the unit she owned on 1 January 2007, and had purchased before 1 January 2007, the capital gains are semi-realized. We then compute the difference between the observed sell price on the sell date and the AVM value on 1 January 2007. This difference is not equal to the actual capital gains this seller experiences, since the purchase pre-dates 1 January 2007. Conversely, if a seller bought a unit during 1 January 2007 and 1 January 2019, and owned it for the remainder of the period, these capital gains are also semi-realized. For this type of capital gains, we compute the estimated semi-realized capital gains by taking the difference between the AVM value on 1 January 2019 and the observed purchase price. If a seller owns a unit throughout the period, from 1 January 2007 to 1 January 2019, all capital gains are potential. We compute these capital gains by taking the difference between the AVM value on 1 January 2019 and the AVM value on 1 January 2007. If an individual purchased and sold multiple times, we sum these realized capital gains. If an individual has an ownership share below one, we apply this ownership share.

We remove individuals who have undertaken non-market operations between 1 January 2007 – 1 January 2019 (inheritance, within-family transfer, divorce settlement etc.) by requiring that a transaction started with an advertisement on the online sale platform Finn.no. We also the trim data based on transaction information. We define an uncertain observation as an observation that satisfies one or several conditions: i) not observed sell date within 2007-2019, ii) absolute value of ask price less sell price on sell price larger than 0.7[6], and iii) other tags (multiple properties connected to transaction, bankruptcy transaction). The implication is that one given individual owner may see one unit in his portfolio removed given the tag of "uncertain observation ". In computing the owner's capital gains, we employ the ownership share. In Norway, it is possible to own shares ranging from zero to unity.

Table 1 summarizes the distribution of a few selected variables. We tabulate statistics on three variables: house values, owner values, and capital gains. We observe that the Gini index of house values increase from 0.26 in 2007 to 0.29 in 2019. The number of houses owned by owners in these six cohorts increases from 75,592 to 77,591 over the period. Since the number of individual owners is constant at 77,554 in this panel, the implication is that some owners own more units at the end of the period than they did at the start of the period, which is intuitive given the age cohorts. For all distributions, the mean is larger than the median, indicating a thick right tail. The last row presents statistics on capital gains. We do not attempt to compute the Gini index as some capital gains are negative. We observe that the 90th percentile of capital gains on 1 January 2019 is NOK 1,986,294. Such capital gains constitute more than half the value of the median house at the same time, which is NOK 3,536,648; evidence that supports the claim that capital gains are sizeable.

[6]The rationale is that a very small sell price might be a non-market sale while a very large sell price sometimes happen in active auctions.

Table 1. Selected summary statistics

Type	Unit (N)	Gini	P10	Median	Mean	P90
House values Jan 1 07	House (75,592)	0.263	1,109,981	1,964,056	2,230,537	3,628,694
House values Jan 1 19	House (77,591)	0.289	1,908,621	3,536,648	4,117,303	6,866,956
Owner values Jan 1 07	Owner (77,554)	0.277	1,131,606	1,996,998	2,321,336	3,810,775
Owner values Jan 1 19	Owner (77,554)	0.305	1,976,528	3,667,659	4,396,820	7,564,927
Capital gains Jan 1 19	Owner (77,554)	NA	421,334	837,137	1,094,006	1,986,294
			Birth year cohort			
	1965	1970	1975	1980	1985	1990
No. of owners	26,697	24,948	16,793	7,638	1,406	72

Notes: The four right-most columns in the upper panel are measured in NOK. The notation '1.1.07' refers to 1 January 2007. 'House' is short notation for houses and apartments, i.e. housing units. After trimming and matching for 5 cohorts: No. records 18,150,169. NA for Gini index is due to negative capital gains. Active trimming filters: We remove uncertain observations, which are observations that satisfy at least one of several conditions: i) Not observed sell date within 2007-2019; ii) absolute value of ask price less sell price on sell price larger than 0.7, iii) other tags (multiple properties connected to transaction, bankruptcy transaction).

3.2 Income data

The income data are acquired from official statistics[7] and are aggregate income levels for each municipality. In Norway, employers, financial institutions, and firms register wages paid (and non-monetary benefits) with the tax authorities. The tax register covers all income, both taxable and non-taxable. Statistics Norway defines households and classifies individuals into households. A household is defined as all individuals who live in a house and share common housekeeping. Income is defined as all income from employment, returns to financial capital, transfers, support, and stipends. From Statistics Norway we use data on all income within households and information on distribution statistics within municipalities (percentiles).

[7]See Statistics Norway at ssb.no. We use information from tables 06944 and 09114.

3.3 The Automated Valuation Method (AVM) and the house price index

We employ estimated market values from Eiendomsverdi, a bank-owned firm that specializes in estimating market values for banks and realtors and is member of the European AVM Alliance, which is a non-profit organization consisting of member firms that provide AVMs in their respective European countries.

This allows our value estimate to attain a higher accuracy than a simple hedonic model would have allowed since this firm employs muliple estimators. In particular, the company's AVM is based on an algorithm that assigns weights to separate value estimators. The weights of these estimators are also characteristics-dependent and functions of time. In the appendix, we have included the results from a validation exercise in which we inspect the accuracy of the predictions for each of the 23,374 units that were involved in transactions among the 77,554 owners in our birth cohorts. The median spread, i.e the difference between the estimated value and the observed transaction price as percentage of transaction price, is -1.3 percent. The 10th and 90th percentiles are, respectively, -12 percent and 11 percent.

The firm Eiendomsverdi constructs both the house price index used in their own AVM and the official house price statistics published by Real Estate Norway[8] every month in a press conference. They follow the same index methodology in both index constructions, with only minor differences. The key element is the combination of observed sell prices and estimated values using the SPAR-methodology (Bourassa et al. (2006) and de Haan et al. (2009)). In this set-up, one studies the ratio, for each transacted unit, of the observed sell price on the estimated market value for a base period. For each month one uses the median sell-price-on-estimated-value ratio, i.e. the SPAR, as the index level. For intuition, it is possible to view the SPAR set-up as an advanced square meter price index in which one adjusts the observed sell price not only by size, but all other relevant attributes as well.

[8]Available online: https://eiendomnorge.no. The statistics offered by Real Estate Norway is commonly viewed as the official house price statistics of Norway, and the data are used by banks, media, analysts, researchers, and the government.

3.4 The capital gains panel

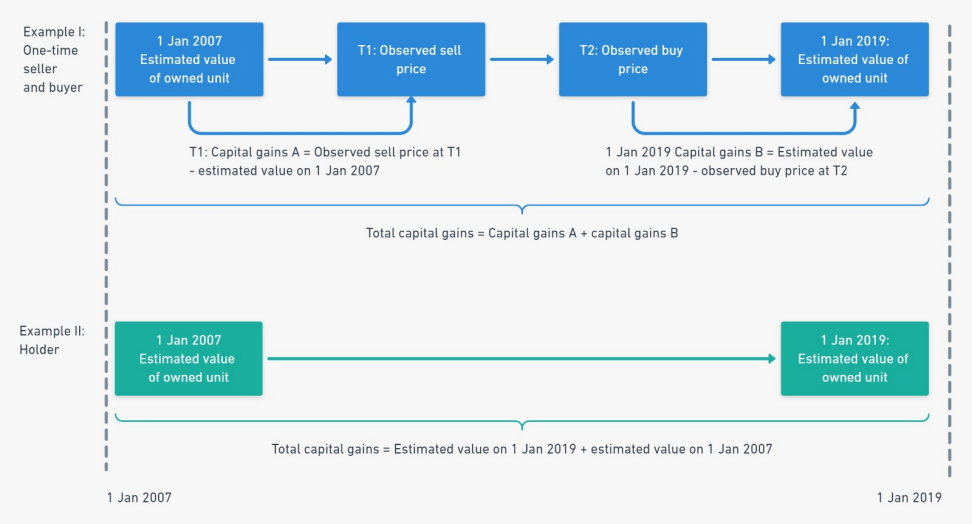
Figure 1 shows how we compute capital gains in two examples:

(I) An owner-occupier who sells the unit she was observed owning on 1 January 2007 on the date T1 and buys another unit on the date T2. T2 lies between 1 January 2007 and 1 January 2019. She owns the second unit on 1 January 2019. The capital gains of both the first and the second unit are semi-realized capital gains since one of the two values are estimated using the AVM.

(II) An owner-occupier who holds the unit during the whole period 1 January 2007-1 January 2019. This is potential capital gains since the values are estimated using the AVM, both at the start of the period and at the end of the period. These two cases are not exhaustive since there are other types. For example, a person may buy and sell within the period. Moreover, some individuals may do multiple sales.

Other owners may do combinations of the above-mentioned types. However, Figure 1 illustrates the thrust of our thinking. At any given point in time, we are able to trace an individual's holding of, and sale of, a given property. At any given point in time, we are also able to estimate the value of an individual's property. Since we can both observe a sell price or estimate a house value, we can compare a sell price or a house value with both initial and final house values computed at given dates. Such differences are either realized, semi-realized, or potential capital gains. For each individual who was registered as an owner on 1 January 2007, and who has been an owner in each quarter since, we observe what the owner has gained in a unit she holds or what she gained when she sold it. If there are repeated transactions, we compute the capital gains for each transaction and sum them.

Figure 1. Examples of the computation of capital gains for sellers and holders



3.5 Institutional background

Transactions in the Norwegian housing market are organized as ascending bid (English) auctions. The typical transaction starts with an owner-occupier who decides to move house. First, he decides whether to buy or sell first. In Norway, we see a mix of buy-first and sell-first strategies. Often, owner-occupiers are involved in both processes simultaneously, and there is usually no problem for moving owner-occupiers to obtain interim funding if the household needs to hold two houses, and have two mortgages, for a short period of time.

In order to sell, a moving owner-occupier contacts a realtor with whom she discusses a sales strategy before setting an ask price. Then, the realtor announces, in an online advertisement, a date for the open house (public showing). After the open house, typically the day after, the auction commences. In this auction, bidders submit bids to the realtor, most often using digital platforms. Bids may be conditional and may have expiration time and date. Bids and acceptances of bids are legally binding. Since bids and acceptances are legally binding, we are able to pin down the exact date on which the transfer of ownership took place. This fine temporal granularity allows us to establish an accurate time line of individual capital gains.

An alternative route is to buy before selling, and one often observes that the moving owner-occupier attempts to make arrangements with both the buyer of the old house and the seller of the new home to coordinate move-out and move-in dates.

Around four fifths of Norwegian households are owner-occupiers (Røed Larsen and Sommervoll (2009)). Most households finance purchases using variable interest rate mortgages. In the Norwegian capital, Oslo, the time-on-market is typically short, often only a few weeks. The wider metropolitan area surrounding and including Oslo accounts for about one fifth of the Norwegian population of 5.3 million citizens.

4 Inequality framework and empirical techniques

4.1 Lorenz curves, Gini-indices, and the P90-P10 measure

Due to the cyclicality of capital gains, we do observe negative values. In plotting Lorenz curves and estimating Gini-indices, we impute zero for negative values of capital gains. However, since this imputation skews the impression of inequality, we make use of, and prefer to use, the difference between the 90th and the 10th percentile of estimated capitl gains as our inequality measure. We denote this measure P90-P10. The P90-P10 has several advantageous features, and one advantage is that it does not require conversion of negative capital gains. A second advantage of this dispersion measure is that it directly measures the difference in purchasing power between individuals at two percentiles. Thus, we prefer the difference between P90 and P10 to the ratio of P90 to P10 since the latter does not capture the purchasing power inherent in the difference between two monetary values.

To see this, consider a thought example. In one segment, spatial or temporal, P90 is NOK 4 million and P10 is NOK 2 million. In another segment, P90 and P10 are NOK 12 million and NOK 6 million. While the ratio P90/P10 is 2 in both segments, the difference P90-P10 is NOK 2 million in the first segment and NOK 6 million in the second segment. Since distance is not a feature captured by a ratio the ratio may not always reflect the

purchasing power differences within and between segments. Since households can purchase goods with capital gains, we prefer the difference measure P90-P10 to the ratio measure P90/P10. While the Gini index is unitless, the P90-P10 measure is measured in monetary units, i.e. Norwegian krone (NOK).

4.2 Granger causality

We test for a time series association between the house price index and our inequality measure by testing for Granger causality. Let the house price index and the inequality measure at time t be denoted HPt and It. We test:

$$HP_t = \alpha + \beta_1 L(HP_t) + \beta_2 L(I_t) + e_t, \qquad HP_t = \alpha + \beta_1 L(HP_t) + e_t, \tag{1}$$

$$I_t = \theta_0 + \theta_1 L(I_t) + \theta_2 L(HP_t) + u_t, \qquad I_t = \theta_0 + \theta_1 L(I_t) + u_t, \tag{2}$$

in which L is short-notation for the lag-operator of an unspecified number of lags and e and u are error terms assumed to be zero-mean. The left-hand side equation represents the unrestricted model and the right-hand side equation represents the restricted model in which $\beta_2, \theta_2 = 0$. If the restricted model in (1) is rejected, we say that the inequality measure, I, Granger-causes the house price index, HP. If the restricted model in (2) is rejected, we say that the house price index Granger-causes the inequality measure.

4.3 The Monte Carlo bootstrap simulation of the distribution of inequality measures

We perform Monte Carlo non-parametric bootstrap simulations of the inequality measure in order to estimate its distribution without relying on parametric assumptions about its shape.

We do this because there exist factors in the auction arrangement that imply a long right tail. For example, when two bidders with high match-utility enter into a bidding contest, the outcome is a sell price that exceeds the expected value; see Anundsen and Røed Larsen (2018) for documentation of reversion in Norwegian housing auctions.

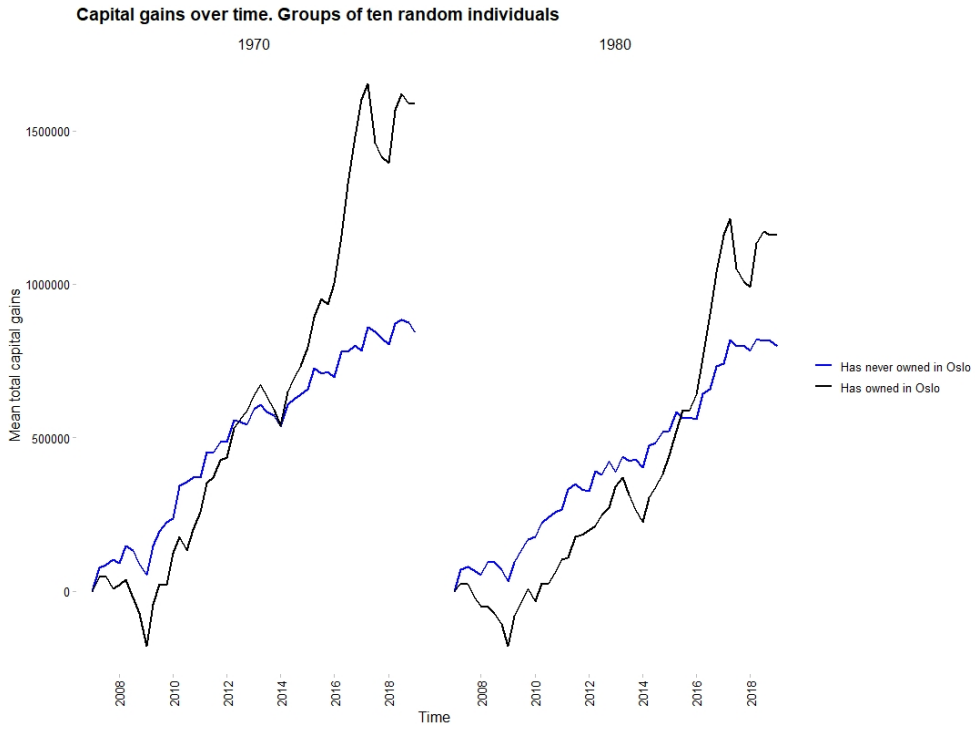
To perform the non-parametric Monte Carlo simulation, we construct 1,000 samples of the same size as the original sample by drawing from the original sample with replacement. For each bootstrapped sample, we compute the inequality measure. The result is a simulated distribution of the inequality measure upon which we make statistical inference.

5 Motivating patterns

5.1 Example capital gains

In Figure 2, we have drawn four groups of 10 random individual owners using two partitions: i) born in 1970 or in 1980 and ii) have owned in Oslo during the period or not. We then compute the mean capital gains for each of the four groups at each quarter in the period 1 January 2007 and 1 January 2019. For illustration, we plot the time development in two graphs, one for owners who were born in 1970 (left panel) and one for owners born in 1980 (right panel).

Figure 2. The development of capital gains from 2007 to 2019 for a few random owners. Oslo vs non-Oslo. Birth years 1970 vs 1980



Note: 'Has owned in Oslo' means that the owner has been registered as having owned a unit, or share of a unit, in Oslo at least once during the period 1 January 2007-1 January 2019.

First, we observe a time development since capital gains in all four groups increase. However, we do see a reduction in capital gains during the financial crisis. During that crisis, Norwegian house prices decreased for about 18 months, from May 2007 until December 2008.

Second, we observe a birth cohort effect since the capital gains in the 1970 cohort is larger than in the 1980 cohort. Most likely, owners in the older 1970 cohort had been able to purchase a more valuable home, which increased more in value.

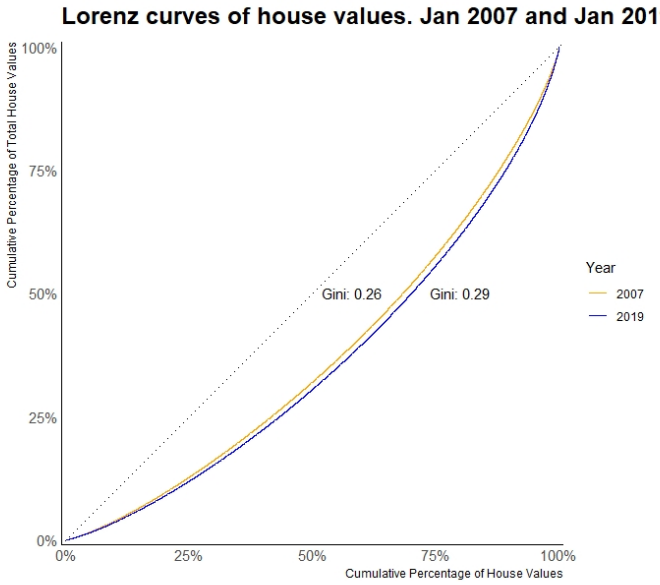
Third, we find a spatial component since owners who had been owners of a unit in Oslo experience larger capital gains than non-Oslo owners. This illustrates three of the dimensions we explore in more detail below: time development, birth cohort effects, and geographical

differences. While Figure 2 plots capital gains, we also scrutinize differences between capital gains.

5.2 Lorenz curve and Gini indices of estimated AVMs of housing units

In Figure 3, we plot the Lorenz curve for the AVM values of the housing units. We observe that the Lorenz curves take the conventional shape known from Lorenz curves of wages, earnings, income, wealth, but consumption, and that the curve for 2019 differs from the curve for 2007.

Figure 3. Lorenz curves and Gini index of house values 2007 and 2019. Norway, birth cohorts 1965, 1970, . . . , 1990



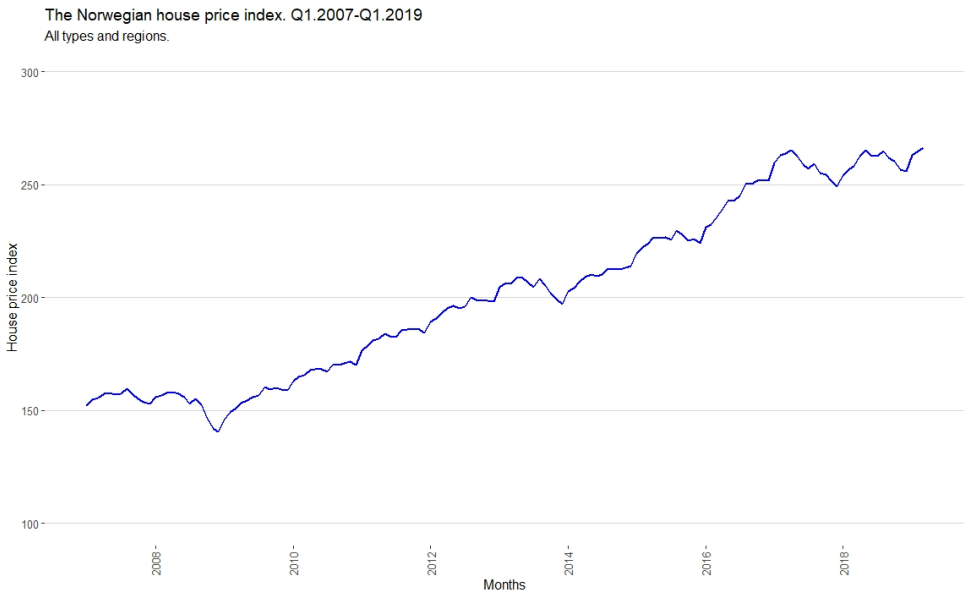
Notes: Value of units owned by individuals in birth year cohorts 1965, 1970, . . . , 1990 as estimated by Eiendomsverdi's AVM on 1 January 2007 and 1 January 2019. See Table 1 for summary statistics.

We compute Gini indices of house values on 1 January 2007 and on 1 January 2019, and notice that while the Gini index for 2007 is 0.26 it increases to 0.29 for 2019, indicating a substantial increase in the dispersion of owned values. This increase is statistically significant as is demonstrated by the non-parametric Monte Carlo simulations of the distributions in Figure A1 in the appendix.

In Figure 4, we plot the time development of the house price index and in Figure 5 we plot the time development of the Gini index of house values. The idea behind comparing these two figures is that we seek to motivate the idea that there is an association between the time development of house prices and the time development of housing inequality.

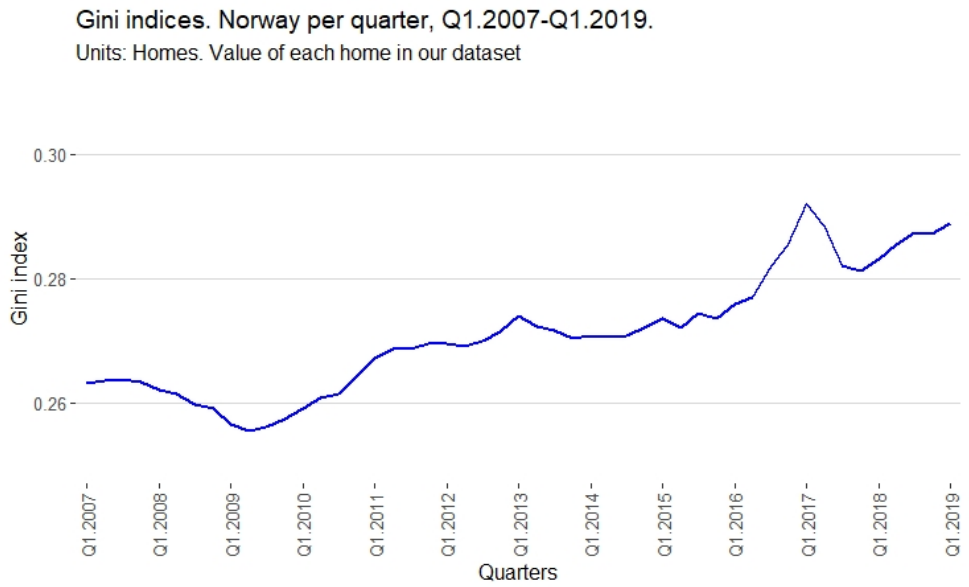
We make two observations. First, both the house price index and the Gini index have a rising trend. Second, for both the price index and the Gini index there are deviations from trend and they seem to occur at roughly the same time. For example, the minimum Gini index occurred in 2009, just a short time after the housing market's downturn during the financial crisis, in which Norwegian house prices reached their minimum in December 2008 (Røed Larsen (2018)). Moreover, the maximum of the Gini index series was reached on 1 January 2017. This is the same time as the end of a period with high growth in the house price index. Figure 4 and Figure 5 are exhibits that support a notion of co-movements in the house price index and the Gini index. Below, we test the hypothesis that house prices Granger-cause housing inequality.

Figure 4. The house price index. Norway, 2007-2019



Note: The index is sourced from Eiendomsverdi/Eiendom Norge; see above and eiendomnorge.no for a description of methodology.

Figure 5. Lorenz curves and Gini index of house values 2007 and 2019. Norway, birth cohorts 1965, 1970, . . . , 1990



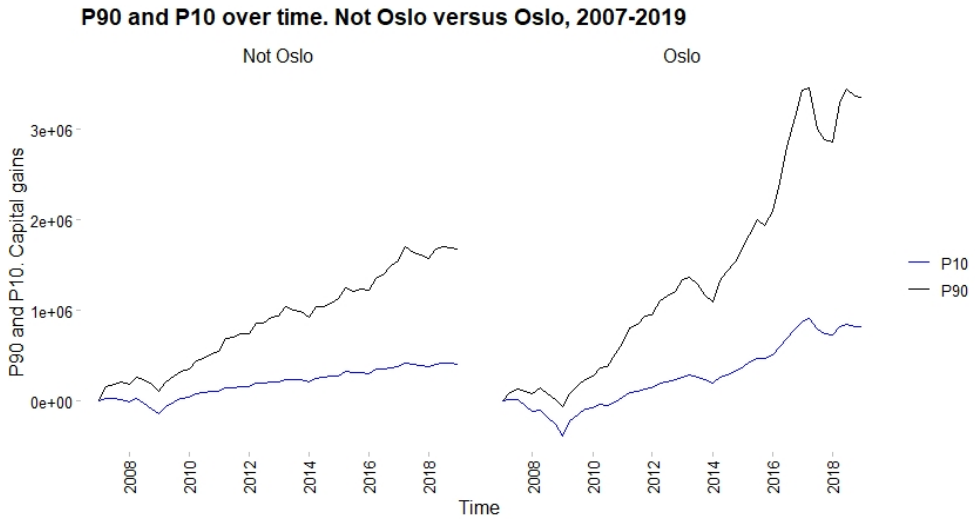
6 Empirical results

6.1 Time development of capital gains inequality

In Figure 6, we plot P90 and P10 of total capital gains for the six birth year cohorts over the time period 1 January 2007-1 January 2019. Visual inspection leads us to two observations: First, there is a rising time trend of capital gains inequality during the period since the 90th percentile increases more strongly over time than does the 10th percentile. Second, this pattern is accentuated in Oslo. In Oslo, the capital gains of the 90th percentile on 1 January 2019 is NOK 3.4 million. The 90th percentile outside of Oslo increases much less than within Oslo.

Figure 6 is our main exhibit, and the difference between the 90th percentile and the 10th percentile in capital gains our preferred inequality measure. The results demonstrate increasing inequality in housing capital gains, both within-markets and between-markets.

Figure 6. 90th percentile and 10th percentile of capital gains per individual owner. Not Oslo and Oslo, 1 January 2007-1 January 2019. Birth cohorts 1965, 1970, . . . , 1990



Notes: 'Oslo' is short notation for having owned a unit in Oslo. This graph was generated by constructing

a list of owners in each quarter 1 January 2007-1 January 2019 among birth cohorts 1965, 1970, 1975, 1980, 1985, and 1990. For each point in time t (quarter), we compute the capital gains between 1 January 2007 and t. Oslo is defined as owners who have owned in Oslo at one point. Among these estimated capital gains, we identify the 10th and the 90th percentile.

Visual inspection tells us that the difference between the 90th percentile and the 10th percentile increases over time, and Table 2 tabulates results that support this notion as it contains the estimates of fitting a linear trend to the P90-P10 measure over the period. The difference P90-P10 in Oslo increases at a substantial rate over time; the estimated slope coefficient is 54,914. The interpretation is that for each quarter the difference in capital gains between the 90th percentile and the 10th percentile increases by NOK 54,914, which is more than the average monthly pre-tax wage.

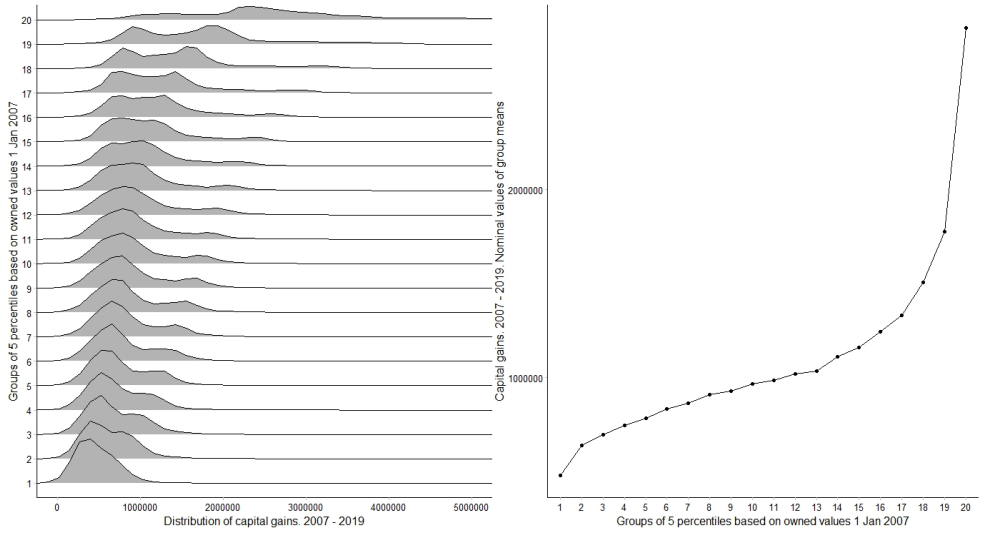
Figure 7 allows us to explore this result in more detail. It plots results from a grouping of individual owners into 20 value groups based on the estimated value of their ownership on 1 January 2007. Group 1 comprises owners with the smallest owned values, i.e. the least valuable units. Group 20 comprises owners with the largest owned values, i.e. the most valuable units. The left-hand side panel plots the distribution of capital gains 2007-2019 for each of the 20 groups. The right-hand side panel plots the within-group mean capital gains 2007-2019 for each group in increasing order of 2007 owned values. We observe from the right-hand side panel that groups with the highest-value ownership also experience the largest capital gains, in nominal monetary terms, consistent with the proposition that the housing market is an inequality generator. We also see from the left-hand side panel that there is substantial heterogeneity within the 20 groups. In particular, the distributions display large right tails.

Table 2. Regression of P90-P10 on linear time. Not Oslo and Oslo, 2007-2019

	Not Oslo		Oslo	
Intercept	34,219	(14,580)	-261,396	(62,176)
Slope	26,195	(508)	54,914	(2,165)
Adj. R2	0.982		0.931	
No. owners	67,254		10,300	

Notes: The data used in the regression was generated by constructing a list of owners in each quarter 1 January 2007-1 January 2019 among birth cohorts 1965, 1970, 1975, 1980, 1985, and 1990. For each point in time t (quarter), we compute the capital gains between 1 January 2007 and t. Oslo is defined as owners who have owned in Oslo at one point.

Figure 7. Capital gains by groups of 5 percentiles of owned values in 2007, Norway



Notes: We sort individual owners by their owned value on 1 January 2007 and group them in 20 groups. Group 1 comprises individuals with the smallest owned values on 1 January 2007. Group 20 comprises individuals with the highest owned values on 1 January 2007. The left-hand panel plots the distribution of

capital gains across the period 2007-2019 within these 20 groups. The right-hand panel plots the within-group mean capital gains across the period 2007-2019. We trimmed the distributions in the left panel on 0 and NOK 5 million.

6.2 The Granger causality of the house price index and capital gains inequality

Table 3 reports the results from our Granger causality test of the house price index (HP) and the inequality measure P90-P10 (I). In the first part of the test, which is reported in the upper part of the table, we smooth both the house price, HP, and the inequality metric, I (P90-P10), using a loess-function with a span parameter of 0.2 before performing the Granger-causality tests. In the second part of the test, we do not smooth the measures HP and I. We see from the reported p-values in the table that the null of no Granger-causality is rejected at the 1 percent level for the case of HP Granger-causes I with one lag both for the non-smoothed and smoothed test. The p-values are 0.0041 and 0.0036, respectively.

With two lags and three lags, the null is rejected at a p-levels 0.04 and 0.02, respectively, for no smoothing in the same direction, i.e. house prices Granger-cause inequality. there is no evidence supporting the hypothesis that the inequality measure, I, Granger-causes the house price index, HP. In summary, the evidence suggests that in Norway, house prices Granger-cause capital gains inequality with one lag.

Table 3. Testing for Granger causality between the house price index (HP) and the capital gains P90-P10 inequality measure (I). Birth years 1965, 1970, ..., 1990. Norway, 1 January 2007-1 January 2019

Tests of Granger causality	
$HP_t = \alpha + \beta_1 L(HP_t) + \beta_2 L(I_t) + e_t,$	$HP_t = \alpha + \beta_1 L(HP_t) + e_t,$
$I_t = \theta_0 + \theta_1 L(I_t) + \theta_2 L(HP_t) + u_t,$	$I_t = \theta_0 + \theta_1 L(I_t) + u_t$

Smoothing	Number of lags, F-statistic (p-value)		
	L1	L2	L3
HP Granger-causes I	9.1 (0.0041)	2.7 (0.079)	1.7 (0.19)
I Granger-causes HP	1.1 (0.29)	1.1 (0.35)	0.32 (0.81)

No smoothing	Number of lags, F-statistic (p-value)		
	L1	L2	L3
HP Granger-causes I	9.5 (0.0036)	3.5 (0.040)	3.6 (0.023)
I Granger-causes HP	0.93 (0.34)	1.7 (0.20)	1.5 (0.23)

Notes: The data used in the regression was generated by constructing a list of owners in each quarter 1 January 2007-1 January 2019 among birth cohorts 1965, 1970, 1975, 1980, 1985, and 1990. For each point in time t (quarter), we compute the capital gains between 1 January 2007 and t and then identify percentiles. Smoothing involves using the loess-function in R with a span parameter of 0.2. In the smoothing part of the test, both the HP and I measure are smoothed before the Granger-causality tests are performed.

6.3 Controlling for cohort effects

We segment individual owners into birth year cohorts from 1965 to 1990 and plot the inequality measure P90-P10 in Figure 8. All cohorts display an increasing time trend of the inequality measure and the deviations from trend occur at the same time. However, both the slope of the time trend and the magnitudes of deviations from trend differ between birth year cohorts. Inequality is largest among the oldest, consistent with the notion that heterogeneity

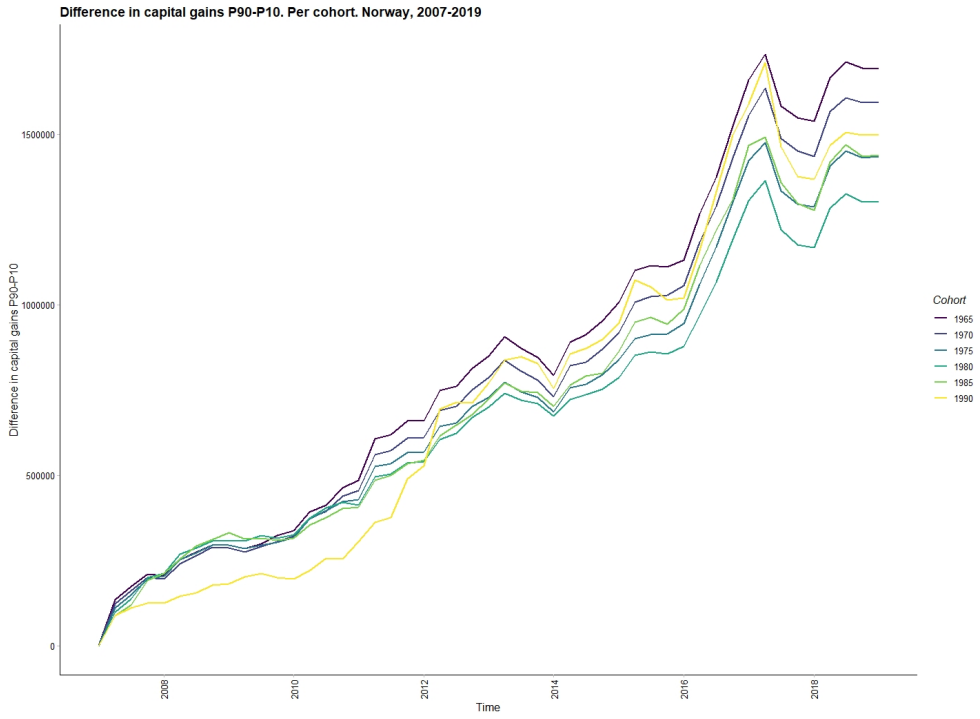
in owned values increases with age as life outcomes tend to diverge with time. Inequality tends to be smaller among younger, but the 1990 cohort deviates from the pattern. This heterogeneity among the very young indicates that this cohort might be selected differently than the others. Since individuals born in 1990 turned 17 in 2007, inheritance played a role for some of these owners. The notion that inheritance displays large heterogeneity is consistent with the development seen in Figure 8.

The difference between the inequality measure developments for cohorts 1965 and 1985, however, is statistically significant.[9] To demonstrate this, we ran a Monte Carlo bootstrap simulation in which we constructed simulated same-size-samples through a sampling with replacement algorithm and computed the inequality measure on the simulated samples. Table 4 reports the simulation results. We see that while the 99.5th percentile of the 1985 cohort is 1,572,632, the 0.5th percentile of 1965 cohort is 1,653,505. Thus, the two distributions do not overlap, and we reject the null of no difference.

The empirical evidence is consistent with the notion that capital gains inequality in the period 2007-2019 for the 1965 birth year cohort is larger than the capital gains inequality for the 1985 birth year cohort. In summary, capital gains inequality increases both within and between birth year cohorts.

[9]We use the birth year cohort 1985, not 1990, due to sample size and selection issues.

Figure 8. Difference between the 90th percentile and the 10th percentile of estimated capital gains. Norway, 1 January 2007-1 January 2019. For each birth year 1965, 1970, . . . , 1990



Note: We identify the 90th and 10th percentiles of capital gains in each of the birth year cohorts and compute the difference, P90-P10.

Table 4. Bootstrap simulation of the P90-P10 distribution on 1 January 2019 for cohorts born in 1965 and 1985

	Percentiles of P90-P10 in year 2019				
Birth year cohort	0.5	2.5	50	97.5	99.5
1965	1,653,505	1,659,679	1,691,446	1,719,913	1,724,673
1985	1,313,681	1,355,333	1,438,024	1,553,688	1,572,632

Notes: The data used in the regression was generated by constructing a list of owners in each quarter 1 January 2007-1 January 2019 among birth cohorts 1965, 1970, 1975, 1980, 1985, and 1990. For each point in

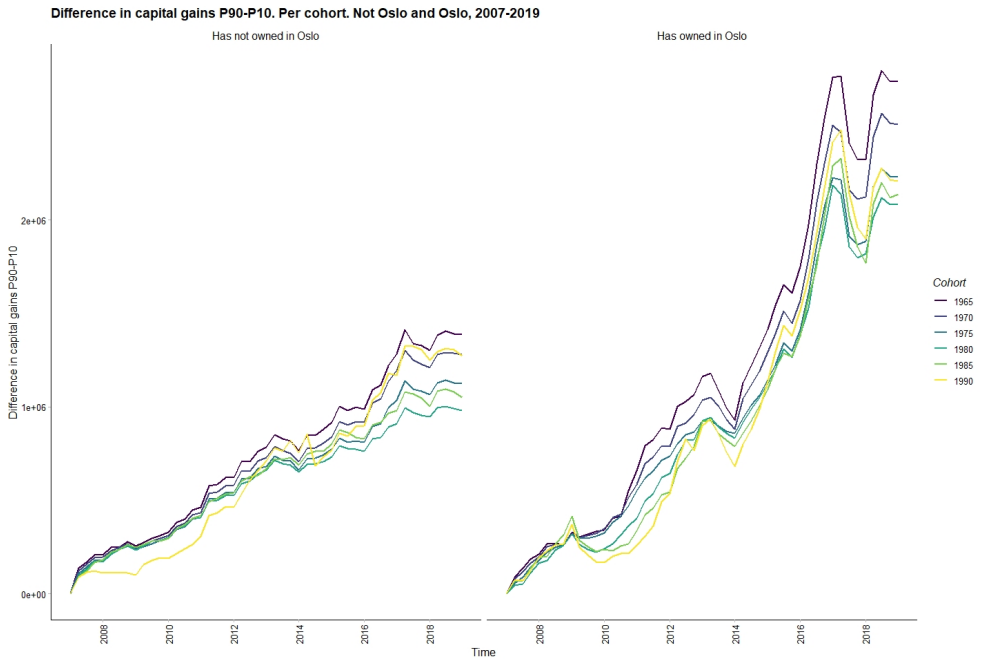
time t (quarter), we compute the capital gains between 1 January 2007 and t. We then identified the 90th and the 10th percentile and computed the difference, P90-P10, in each quarter for each birth year cohorts. For the bootstrap simulation, we randomly drew with replacement same-size samples 1,000 times for both the 1965 cohort and the 1985 cohort on 1 January 2019 and computed the P90-P10 for each of the 1,000 samples.

7 Discussion

7.1 Spatial dimensions of inequality

To further explore the geographical dimension of capital gains inequality, we segment owners into two segments, one segment for individuals who have owned a unit in Oslo during out time period and one segment for everyone else. We also segment into the birth year cohorts. Figure 9 plots the development of our capital gains inequality measure, P90-P10, for each birth year cohort. We observe that the time trend in Oslo is steeper than the rest of Norway for all cohorts. While P90-P10 for the rest of the country on 1 January 2019 is just around NOK 1 million, the P90-P10 measure in Oslo is between NOK 2 and 3 million, depending on birth year cohort. The interpretation is that the spatial component is a factor in understanding capital gains inequality.

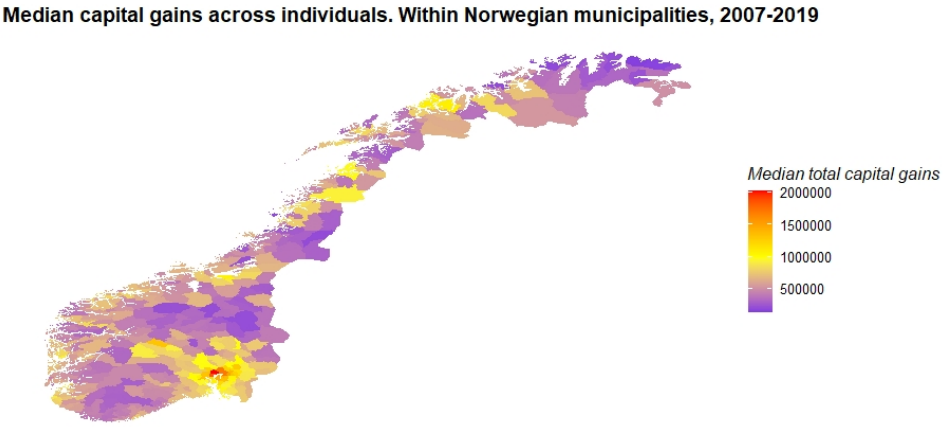
Figure 9. Difference 90th percentile and 10th percentile. Oslo and Not-Oslo, 1 January 2007-1 January 2019. For each birth year 1965, 1970, . . . , 1990



Notes: Individuals belonging to the not-Oslo segment is plotted to the left and individuals belonging to the Oslo-segment is plotted to the right. An individual belongs to the Oslo segment if this owner owned a unit in Oslo during the time period 1 January 2007-1 January 2019.

To further emphasize the spatial dimension, we plot in Figure 10 the median across-period capital gains for each municipality in Norway. Peripheral, rural, and mountainous areas of Norway have seen smaller median capital gains over the 12-year period. In these municipalities, median capital gains are around NOK 500,000. In contrast, the median in the southeast area that includes Oslo is much higher, which is indicated by the yellow color. The large capital gains in Oslo is seen by the red color. This graph illustrates the strong presence of spatial components in the development of capital gains inequality.

Figure 10. Median capital gains across individuals within a municipality. Norway, 2007-2019



Notes: For each individual owner we compute total capital gains 1 Jan 2007-1 Jan 2019. For each municipality we find the median total capital gains across individuals. For each individual owner we ascribe a municipality. If their ownership is sequential, we ascribe to each individual the municipality in which they owned a unit on 1 January 2007. If they owned several units at that point in time, we use the municipality in which they owned the most valuable unit. The map shows Norwegian municipalities and the Oslo region is the one in red color in the middle of the low-right part (south-east) of the country. The borders of Norwegian municipalities dates from 2018, and these borders are changing as there is an ongoing restructuring project of Norwegian municipalities.

7.2 House prices, income, and capital gains inequality across Norwegian municipalities

We cannot here map determinants of the capital gains inequality, but in this discussion section we seek to present some evidence of the role played by income development. In Norway, there were 428 municipalities until recently[10], and there considerable variation across

[10]The borders are in the process of being redrawn based on political decisions, thus the number of municipalities decreases.

these municipalities in terms of size, population, infrastructure, and economic activity. We seek to exploit this variation to investigate whether there is evidence to support a claim of an association between income development and capital gains inequality development.

Table 5 reports the result of municipality fixed effect regressions. The evidence suggests that there is an association between capital gains inequality and income and house prices. Model III shows that an increase in median income of NOK 100,000 is associated with an increase in the inequality measure P90-P10, across municipalities, of NOK 160,000. Moreover, model III shows that an increase in median house price of NOK 100,000 is associated with, across municipalities, of an increase in P90-P10 of 17,000.

Table 5. Municipality fixed effect regressions of capital gains inequality on income and lagged income

	Dependent variable		
	P90-P10 capital gains		
Independent variable	I	II	III
Median income	1.91 (0.16)		1.60 (0.16)
Median house price		0.174 (0.0071)	0.165 (0.0071)
Year FE	YES	YES	YES
Municipality FE	YES	YES	YES
Period	2008-17	2008-17	2008-17
Adj. R2	0.678	0.724	0.732
No. municipalities		369	
N	3,690	3,510	3,510

Notes: Regressions are run using yearly observations for each municipality. Municipalities without all observations are removed in order to keep a panel structure. The robust standard errors in parentheses are computed using the function vcovHC.plm in the lmtest-package in R, employing the "arellano" method, which handles autocorrelation and heteroscedasticity. The arellano method clusters on the fixed effects in

the regression, here municipality and year. Year FE is short notation for using a collection of year dummies. Income is acquired from Statistics Norway (ssb.no) and is obtained using gross household income data from the Norwegian IRS (tax records) from Statistics Norway table no. 06944. Individuals are classified based on the municipality of the ownership of largest value.

7.3 Capital gains taxes and debt

Our study does not compute after tax capital gains, but instead concentrates on pre-tax capital gains as defined as the difference between a market value at one point in time and a market value at another point in time. There are three reasons for this. First, in Norway housing capital gains are tax exempt if the seller has resided in the unit at least 12 months during the last 24 months before the sale. Most sellers are owner-occupiers and thus enjoy this tax exemption. The implication is that most of the realized capital gains we have computed are not subject to taxation and the potential capital gains we have computed are also would not be subject to taxation once it is realized.

Second, if we had attempted to apply the tax rate onto realized capital gains and semi-realized capital gains from within-period selling, but not on potential capital gains or semi-realized capital gains from within-period buying, we would need to be able to differentiate between sellers for which the tax exemption was applicable. This would require a mapping of sellers' historical home addresses, which in practice is infeasible.

Third, our study aims at documenting patterns in the inequality of capital gains of housing, and our choice of concentrating efforts on pre-tax capital gains does not impact the patterns we find since potential capital gains would become realized capital gains at one point in the future. After-tax capital gains would be somewhat smaller, and thus after-tax capital gains inequality could be somewhat smaller. Further research might be able establish a difference between pre-tax capital gains and after tax capital gains.

We do not study debt. The implication is that that we do not study returns to equity. Thus, if households A and B both purchase a unit of NOK 4 million, and the units appreciate

to NOK 6 million, we focus attention on the capital gains of NOK 2 million. However, if household A financed its purchase using NOK 4 million of equity while household B financed its purchase using NOK 2 million of equity, household A has experienced a 50 percent increase on their equity while household B has experienced 100 percent increase on their equity. It is interesting to know whether leverage plays a role in inequality and whether leverage displays heterogeneity across spatial and age dimensions. However, it is left for future research. Here, we choose to study the purchasing capacity of the capital gains, i.e. NOK 2 million in the thought example, not the financial acumen or access to financing among households, and we believe this is the more relevant method when one wants to answer whether the housing market is an inequality generator.

8 Concluding remarks and policy implications

While many inequality studies have studied the development of inequality in wages, income, wealth, and consumption, few, if any, studies have mapped the time development in the inequality of capital gains in the housing market. This paucity of capital gains studies may be due to a challenge of data access. In order to study the time development of capital gains inequality analysts need access to data that include transactions, owners, and units. Capital gains studies require the computation of realized, semi-realized, and potential capital gains and the ability to follow owners and units over time.

This study constructs a panel dataset of 77,554 owners with housing unit identification. We follow these owners throughout the period and we are able to compute the capital gains these owners made in the housing market in Norway during the period 1 January 2007 and 1 January 2019. At the start of the period, these 77,554 owners owned 75,592 units.[11] At the end of the period they owned 77,591 units. During the period they sold 23,374 units. We compute the total capital gains for an individual owner by summing realized, semi-realized, and potential capital gains. The classification depends upon how the value of the housing

[11] Owners may own shares in the same house.

unit is observed. Observed purchases and sales allow the computation of realized capital gains. When either a purchase or a sale is not observed, because they happened outside of the time period, we estimate semi-realized capital gains by employing an AVM for the value of the unit instead of the market transaction. When neither the purchase nor the sale is observed, and the owner holds the unit throughout the period, we estimate both the entry and exit value using the AVM. We denote the estimated difference potential capital gains.

Our preferred measure of capital gains inequality is the difference between the 90th and 10th percentile of the capital gains across individual owners. It is measured in monetary units and thus reflects a difference in actual purchasing capacity between individual owners. Alternative measures comprise the Gini index, the coefficient of variation, and the P90/P10. These measures have the disadvantages that they must be re-scaled and transformed to deal with negative values and/or are unitless and thus less intuitive to interpret.

For data tractability reasons, we study individual owners belonging to birth year cohorts from 1965, 1970, 1975, 1980, 1985, and 1990. Our results show that the housing market is an inequality generator. During the time period of 12 years the capital gains accumulated in the housing market varied substantially.

We find that while the 90th percentile of capital gains 2007-2019 for owners who have owned in Oslo is NOK 3.35 million, it is NOK 1.67 million for owners outside of Oslo. Moreover, the top five percentiles with the highest owned values in 2007 experienced a capital gain over the next 12 years of NOK 2,862,503 while the next five percentiles in terms of owned values in 2007 experienced a capital gain of NOK 1,778,632. This means that the top five percentiles had 61 percent larger capital gains than the next five.

We find considerable inequality trend differences between and within birth year cohorts and across geographical strata in Norway. Median capital gains in some municipalities were at least four times larger than median capital gains in other municipalities.

What causes these increases in capital gains inequality? It is only natural to look to house prices since capital gains are defined as increases in home values. We document that

the house price index Granger-causes the P90-P10 inequality measure, but not the other way round. Moreover, we also find an association across municipalities between the P90-P10 and income using municipality fixed effect regressions.

The increase in capital gains inequality represents an increase in the difference of economic opportunity and it may affect households in ways that cannot easily be escaped. For example, when a household needs to move house because of labor market events, it matters not only where it happens or what year it happens, but also even what month it happens. This follows from the spatiality, cyclicality, and seasonality of the housing market. For example, Nenov et al. (2016) show that transaction seasonality in the housing market is associated with thick market effects. An implication of their findings may be that if individuals seek to solve the dual search-and-matching problem of finding both a job and a house, or if they simply seek a better match between their preferences and house attributes, they may discover that their future economic opportunities are affected by the month in which they are searching, selling, or buying. Anundsen and Røed Larsen (2018) show that when a sell price is higher than an appraisal value, the sell price tends to revert towards an appraisal value in the next sale. The implication is that if an individual buys at a high price in a thick market, the capital gains will tend to be lower compared to an individual who bought at a low price in a thin market.

An individual's timing of choices in both labor and housing markets may be deliberate and due to skill, but timing may also be due to unforeseen events and thus due to bad luck. This means that not only is the inequality that arises with capital gains caused by year cohort effects and geographical effects, but also by the season in which individuals chose to or had to move.

The development of capital gains inequality shows that there are large differences between individuals. Thus, these findings open up the question of policy intervention. Norwegian tax rules, as do rules around the world, allow house owners several tax advantages. Such rules and polices make housing an attractive asset in Norway – again, as in many other countries.

Given a political aim of limiting the increases of inequality, and leveling the playing field, the findings in this study invite a renewed look at whether and how to tax capital gains in the housing market.

References

Aguiar, M. and M. Bils (2015): Has consumption inequality mirrored income inequality? *American Economic Review*, **105**: 9, pp. 2725-2756.

Albouy, d. and M. Zabek (2016): Housing inequality, *NBER Working Paper* No. 21916.

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

Attanasio, O. P. and L. Pistaferri (2016): Consumption inequality, *Journal of Economic Perspectives*,**30**: 2, pp. 3-28.

Attanasio, O. P. and L. Pistaferri (2014): Consumption inequality over the last half century: Some evidence using the new PSID consumption measure, *American Economic Review*, **104**: 5, pp. 122-126.

Benhabib, J. and A. Bisin (2018): Skewed wealth distributions: Theory and empirics, *Journal of Economic Literature*, **56**: 4, pp. 1261-1291.

Benhabib, J, A. Bisin, M. Luo (2017): Earnings Inequality and Other Determinants of Wealth Inequality, *American Economic Review*, **107**:5, pp. 593 - 397.

Barlevy, G. and D. Tsiddon (2006): Earnings inequality and the business cycle, *European Economic Review*, **50**, pp. 55-89.

Blundell, R. and B. Etheridge (2010): Consumption, income, and earnings inequality in Britain, *Review of Economic Dynamics*, **13**, pp. 76-102.

Bourassa, S. C., M. Hoesli, and J. Sun (2006): A simple alternative house price method, *Journal of Housing Economics*,**15**, pp. 80 – 97.

De Haan, J., E. van der Wal, and P. de Vries (2009): The measurement of house prices: A review of the sale price appraisal ratio method, *Journal of Economic and Social Measurement*, **34**, pp. 51 – 86.

Di Nardi, M. and G. Fella (2017): Saving and wealth inequality, *Review of Economic Dynamics*, **26**, pp. 280-300.

Fagereng, A., M. Blomhoff Holm, and K. Næss Torstensen (2020): Housing wealth in Norway, 1993-2015, *Journal of Economic and Social Measurement*,**45**, pp. 65-81.

Furceri, D., P. Loungani, and A. Zdzienicka (2018): The effects of monetary policy shocks on inequality, *Journal of International Money and Finance*,**85**, pp. 168-186.

Heathcote, J., F. Perri, and G. L. Violante (2010): Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006, *Review of Economic Dynamics*,**13**, pp. 15-51.

Hugget, M., G. Ventura, and A. Yaron (2011): Sources of lifetime inequality, *American Economic Review*,**101**, pp. 2923-2954.

Karahan, F. and Serdar Ozkan (2013): On the persistence of income shocks over the life cycle: Evidence, theory, and implications, *Review of Economic Dynamics*,**16**, pp. 452-276.

Krueger, D. and F. Perri (2006): Does income inequality lead to consumption inequality? Evidence and theory, *Review of Economic Studies*, **73**, pp. 163-193.

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.

Piketty, T. and G. Zucman (2014): Capital is back: wealth-income ratios in rich countries 1700-2010, *Quarterly Journal of Economics*,**129**: 3, pp. 1255-1310.

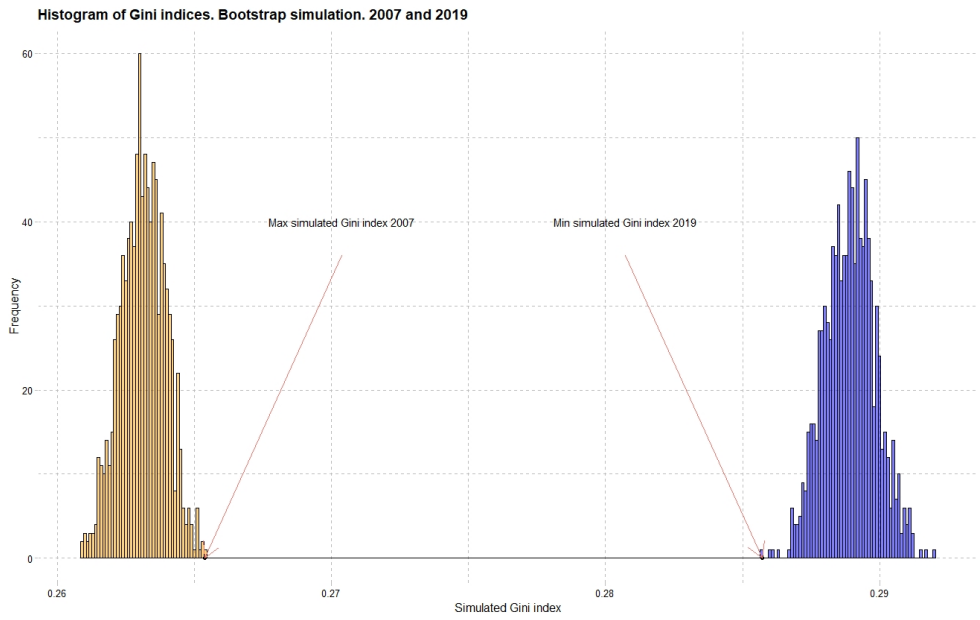
Røed Larsen, E. (2018): Can monetary policy revive the housing market in a crisis? Evidence from high-resolution data on Norwegian transactions, *Journal of Housing Economics*, **42**, pp. 69 – 83.

Røed Larsen, E. and D. E. Sommervoll (2009): The impact on rent from tenant and landlord characteristics and interaction, *Regional Science and Urban Economics*,**39**:3, pp. 316–322.

Van Nieuweerburgh, S. and P.-O. Weill (2010): Why has house price dispersion gone up? *Review of Economic Studies*,**77**, pp. 1567-1606.

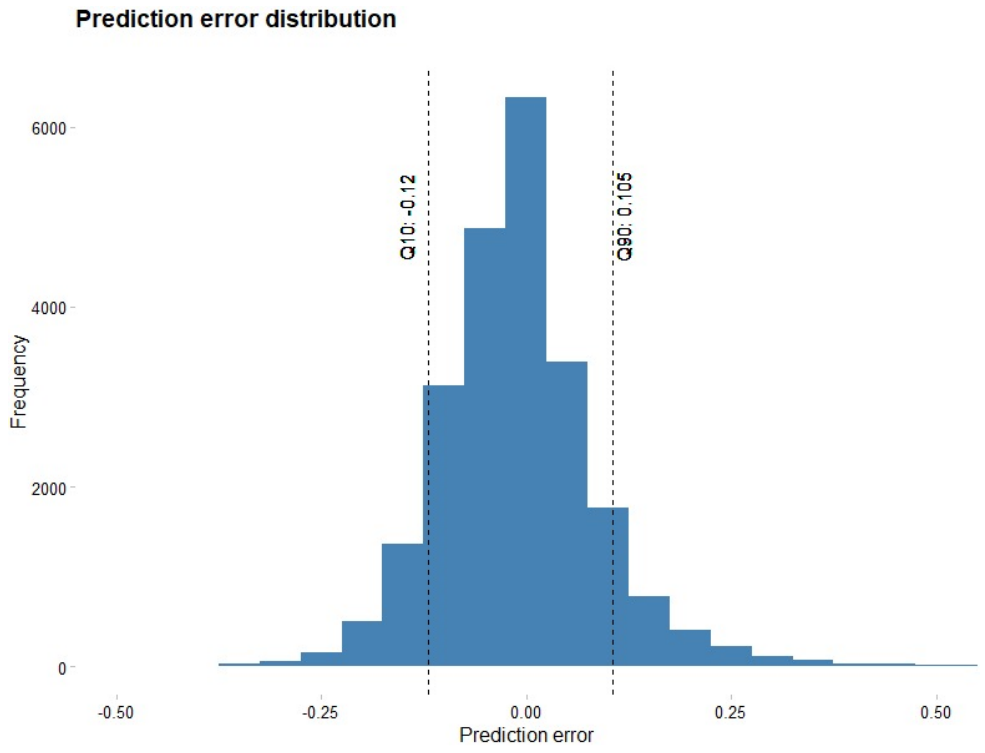
9 Appendix

Figure A1. The Monte Carlo simulated distribution of the 2007 Gini index and the 2019 Gini index of house values



Notes: We use two samples, 1 January 2007 and 1 January 2019. From these two samples, we draw with replacement 1,000 equal-sized samples and compute the Gini index for each sample. The graph shows the two histograms for these 1,000 simulated Gini indices.

Figure A2. The distribution of the difference between the estimated AVM value and the observed transation price as percentage of transaction price. Norway, 2007-2019



Notes: We compute the spread between estimated value and observed transaction price as percentage of observed transaction price. We plot the 10th and the 90th percentile of these 23,374 spreads. During the period 1 January 2007 and 1 January 2019 the owners in the birth year cohorts were involved in 23,374 transactions.

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 Selvaag, Sparebank1-Gruppen and OBOS. We are greatful for the financial support. All views expressed in our papers are the sole responsibility of Housing Lab, and do not necessarily represent the views of our sponsors. The authors are grateful to DNB Eiendom and Eiendomsverdi for auction and transaction data.

Authors:
Terje Eggum,
Eiendomsverdi AS, Norway;
email: te@eiendomsverdi.no
Erling Røed Larsen,
Housing Lab, Oslo Metropolitan University, Norway;
email: erlingro@oslomet.no

ISSN: 2703-786X (Online)

OSLOMET

HOUSING LAB
Nasjonalt senter for boligmarkedsforskning