Norwegian School of Economics
Bergen, Fall 2021

Measuring The Heat of Oslo's Housing Market

A Composite Indicator to Improve the Informational Efficiency in the Residential Real Estate Market

Eivind Vinsand and Håvard Hauger Sjong

Supervisor: Jonas Andersson

Master thesis, Economics and Business Administration

Major: Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

First of all, we would like to thank our excellent supervisor Jonas Andersson from the Department of Business and Management Science at the Norwegian School of Economics, for valuable feedback and writing tips.

Furthermore, we extend our sincere gratitude to Aslak Bergersen and VIRDI for providing essential data material and constructive input throughout the process.

We also gratefully recognize the help of William Becker. His knowledge and contributions within the field of composite indicator construction have been paramount for this project.

Finally, we want to thank our friends, families, and significant others who have helped us along the way.

Norwegian School of Economics

Bergen, December 2021

Eivind Vinsand

Håvard Hauger Sjong

Abstract

This paper develops a composite indicator that estimates the relative bargaining power between buyers and sellers in Oslo's residential real estate market. It constitutes a geometric mean of three min-max scaled indicators that measure (1) how long it takes to sell a home, (2) the disparity between sale and listing prices, and (3) the relative housing supply. The paper's objective is to develop a robust measurement of market temperature that improves the informational efficiency in the Norwegian real estate market. We find that information on bargaining power can guide market participants with prospecting, bid, and sales strategies. We also find that it can support decision-makers in monitoring the impacts of policies, assessing market dynamics, and benchmarking regional differences. Uncertainty analysis suggests that the index is generally unbiased. Variance-based sensitivity analysis reveals normalization to be the only significant uncertainty factor. We show that index trends coincide with Oslo's home appreciation rates and the media's perception of market heat.

Keywords – Housing Market, Composite Indicator, Bargaining Power

Contents

List of Figures

List of Tables

1 Introduction

The housing market is an essential part of the Norwegian economy. Each year, a magnitude of one in twenty homes sell on the Norwegian market, with a total value of more than 500 billion NOK (Regjeringen, 2021). In 2019, investments in residential housing accounted for one quarter of gross fixed capital investments and over six percent of mainland gross domestic product (SSB, 2021b). Housing is also the most common form of saving and the single most important financial decision a typical household makes (Regjeringen, 2021). Despite its importance, there is evidence that inefficiencies exist in the housing market because of barriers to information (Oslo Economics, 2020). For example, purchasing a home can induce substantial transaction costs in the form of expenses for prospecting, appraisals, attorneys, and real estate agents (Herath & Maier, 2015). These costs imply that information on prices and market conditions are not available to all or in an appropriate form. In other words, the market is not informationally efficient.

The Government and other agencies devote considerable effort to ensure that decision-makers and consumers are well informed to mitigate market inefficiencies. Statistics Norway (SSB), guided by Real Estate Norway, is a principal channel through which the Government provides information (SSB, 2021a). The agency publishes reports covering national and regional home appreciation rates, housing inventories, and transaction volumes. The statistics form a basis for policy-making , monitoring, and informed decision-making in the general public. Although individual statistics offer valuable details on market conditions, interpreting innumerable indicators can be challenging and time-consuming. Thus, it is surprising that no agency produces an index that combines them to measure market performance systematically. Composite indicators, also known as indices, can provide a big picture, are easy to interpret, and facilitate effective communication (OECD, 2008).

This paper aims to fill the void by developing a composite indicator that estimates the relative bargaining power between buyers and sellers in Oslo's residential real estate market. We compute the index monthly from January 2017 to May 2021 using second-hand residential property transactions from FINN. The index constitutes a geometric mean of three min-max scaled indicators that measure (1) the relative housing supply, (2) how long it takes to sell a home, and (3) the disparity between sale and listing prices.

These metrics denote domains of supply, demand, and price negotiation, which collectively delimit and quantify the elusive phenomenon of "market temperature". The composite lies on the inclusive interval of 1 to 100. A value of 1 indicates a *cold market*, where buyers have relatively more bargaining power. In contrast, a value of 100 indicates a *hot market* where sellers have more bargaining power.

The purpose of the index is to be a robust aggregate that improves the informational efficiency in the Norwegian real estate market. Creating a robust and useful index is, however, not a trivial exercise. It poses conceptual and methodological challenges that stem from data that form them and the multifaceted development process. While this thesis generally follows the *Handbook on Constructing Composite Indicators* developed by OECD (2008), the construction depends more on sound decisions than on universally acknowledged rules for encoding. Thus, we develop the index following three core principles: *simplicity*, *robustness*, and *transparency*. Simplicity implies that it should be easy to understand for a non-technical audience. Yet, simplicity does not prevail over technical robustness. Thus, our thesis extensively discusses plausible alternatives in each development step. The alternatives constitute sources of uncertainty, which we translate into a set of input factors for a Monte Carlo simulation. The simulation involves estimating the index multiple times, randomly varying the factors. Uncertainty and sensitivity analysis is then conducted by means of the simulation to evaluate the robustness.

The robustness evaluation reveals an insignificant uncertainty in the index scores and find normalization to be the only influential input uncertainty. Moreover, to assess the explanatory power of the composite, we correlate the index with a news-based sentiment of "market heat", and home appreciation rates in Oslo. We find that index trends coincide with both measures. These results suggest that our index generally provides an unbiased estimate of market temperature. Subsequently, the index has numerous expected applications. We suggest that it can improve market efficiency through the benefit of decision support for policy-makers and market participants. For example, we show that understanding contemporary bargaining power provides insights to guide prospecting, bid, and sales strategies. We also indicate that the composite can help decision-makers monitor the impacts of particular policies, with examples from new mortgage regulations and Covid-19 restrictions.

This thesis proceeds as follows. Section 2 defines the concept of market temperature and offers a brief review of recent indices that measures the phenomenon. Section 3 outlines a sequence of methodological steps to construct a composite indicator. Section 4 describes the implementation of methods leading to the development of the index. Section 5 presents the index and results from the uncertainty and sensitivity analysis. Section 6 evaluates the composite's explanatory power and implications before discussing limitations and topics of further research. Finally, Section 7 provides concluding remarks.

2 Background

This section is twofold: first, we discuss the concept of market temperature before exploring contemporary efforts of measuring it. We reason that market temperature represents the relative bargaining power between market participants. What influences the bargaining power is housing market conditions. Attempts to measure market temperature empirically are limited, but we find examples in the academic work of Carrillo (2013), who estimates sellers' bargaining power through a structural search model. Moreover, three prominent American real estate companies, Zillow, Realtor, and Redfin, measure the phenomenon by summarizing diverse housing market conditions.

2.1 Defining Market Temperature

People use a variety of measures to classify the housing market, from isolated indicators, such as home appreciation rates, to subjective experiences of bidding rounds. Despite many appraisal methods, participants generally aim to evaluate the competitiveness between buyers and sellers. More specifically, it is an assessment of relative bargaining power.

Bargaining power affects market participants' ability to exert influence over each other (Wilhelmsson, 2008). Buyers search across sellers until they find an appealing dwelling in the market. If the buyer's reservation price exceeds the seller's listing price, the price difference defines a surplus that can be bargained over. Thus, when sellers have strong bargaining power, the buyers' cost of dispute is high. That is, they are unable to secure an agreement with the sellers on beneficial terms. For example, buyers might have to pay their reservation price, or if multiple parties are interested, they might be unable to purchase at all. Either way, the price of the property increases. Hence, we define a market where sellers have strong bargaining power as *hot* and vice versa as *cold*. Equivalently, a *neutral* market describes a balance in bargaining power. Over time, the market transitions between hot, neutral, and cold periods, driven by the underlying market conditions.

Housing market conditions have a significant influence on the housing temperature (Novy-Marx, 2003). Sellers do well in specific periods, not because of a systematic improvement in skills but because market conditions support their bargaining position. For example, a market with inadequate housing supply indicates decreasing options for buyers. When demand exceeds supply, we generally observe that homes sell quickly and prices increase. Since market conditions play an important role in market temperature, it is also the foundation for empirically measuring it.

2.2 Measuring Market Temperature

As we illustrate in this section, attempts to measure market temperature mainly involves creating composite indicators. Thus it is necessary to begin with a discussion on the topic. Formally, a composite indicator is a mathematical aggregation of individual indicators applied to measure a multidimensional concept (OECD, 2008). We can further define an indicator as a quantitative measure of a phenomenon derived from a series of observed facts (European Commission, n.d.-a). The definitions reveal the practical value of an index. It facilitates the summary of complex concepts without significantly reducing the underlying information base. Thus, indicators are in many ways a powerful way of conveying information. They allow comparisons over time and between units which can aid with evidence-based decision making (OECD, 2008). Besides, a non-technical audience might find composite indicators easier to interpret than a bundle of separate indicators (OECD, 2008).

Due to the benefits, composite indicators have gained widespread adoption in many research areas and by global institutions. For example, Bandura (2011) recognizes more than 400 official indices that assess countries based on environmental, political, social, and economic achievements. Nonetheless, the fact that indicators only "indicate" reveals a trade-off. Since indicators describe a simplified reality, they heavily rely on the usefulness of data that constitute them and the development process. Moreover, since there is no universally accepted scientific approach, attempts at measuring the same phenomenon may differ significantly. For example, three prominent real estate companies, Zillow, Realtor, and Redfin, offer indices measuring market temperature in America. Despite aiming to provide users with insights on the same phenomenon, their methods differ. In the following, we briefly present each index.

The *Market Hotness Index* from Realtor (2021) shows how areas are undergoing changing supply and demand dynamics. The platform uses Listings Views to indicate demand and Median Days on Market to indicate supply. They rank and score metro areas, and counties relative to the rest of the country on a scale from 0 to 100.

The real estate marketplace, Zillow, provides the *Buyer-Seller Index (BSI)*, which constitutes the following three indicators, Days on Market, the Sale-to-List Price Ratio, and the Percentage of Units for Sale With Price Revisions (L. S. Becker, 2019). The index measures the temperature of all U.S regions. The composite ranges from 0 to 10, where 0 indicates a cold market and negotiating power in favor of buyers. Conversely, a value of 10 indicates a hot market and that sellers have more leverage.

Finally, Redfin (2021) computes the *Compete Score*, which indicates how competitive an area is on a scale from 0 to 100. The composite index comprises four inputs: Days on Market, the Sale-to-List Price Ratio, the Number of Competing Offers, and Waived Contingencies. Waiving contingencies is a strategy for a buyer to make their offer more appealing by giving up certain rights, such as the right to exit the transaction if the buyer cannot secure financing.

The heat of the housing market is evidently a topic of concern for real estate companies. However, academic literature offers limited work on the matter. Yet, we find a novel contribution by Carrillo (2013), who uses Sale-to-List Price Ratios and Time-on-the-Market indicators to estimate an index that measures sellers' and buyers' bargaining power in a structural search model. He estimates the index using aggregate transaction data from the Washington D.C. area by computing a structural parameter that measures market temperature yearly in 1998-2009. The index coincides with home appreciation rates and perceptions of heat in the area.

In an attempt to improve short-term price forecasting Miller and Sklarz (2012) employ market condition drivers which describe competitiveness. They find that indicators such as Home Inventory, Sales Volume, Expired Listings, Days on Market, Absorption Rates, and Sale-to-List Price Ratios improve short-run price forecasting. The authors argue that the metrics are valuable because they reflect supply and demand interactions that are leading indications of price trends. They propose that their application in price forecasting enables the model to occasionally catch turning points and price bubbles.

Guided by literature and contemporary indices, we have identified ten indicators that impact bargaining power and, correspondingly, market temperature. We summarize these in Table 2.1. Albeit interrelated, we group the indicators into three domains: price negotiation, demand, and supply. Seven of the identified indicators implicitly or explicitly focus on supply or demand conditions, while the remaining three emphasize price negotiation. The outcome of price negotiation indicates bargaining power between market participants. The domain comprises indicators influencing the participants' benefits. Increasing benefits in favor of sellers should imply a hot market and vice versa in cold markets. Supply and demand affect bargaining power through the relative availability of alternatives in the market. Thus, the domain comprises indicators that reflect options in the market. For example, the direction and pace at which housing supply changes indicate whether the opportunities for buyers are increasing or decreasing. When demand exceeds supply, we assume the market to heat up, and it begins to cool down when supply exceeds demand.

Table 2.1: Indicators of Market Temperature

Domain	**Indicators**	**Description**	**Relationship***
Price Negotiation	Listings With a Price Cut	The percentage of dwellings listed with a price cut	Inverse
	Sale-to-List Price Ratio	The ratio between sale and listing prices	Positive
	Waived Contingencies	The percentage of total offers with at least one waived contingency	Positive
Demand	Days on Market	The median number of days it takes to sell a dwelling	Inverse
	Listing Views	The average number of listing views	Positive
	Sales Volume	The total number of property sales	Positive
	Expired Listings	The percentage of expired listings	Inverse
	Competing Bids	The average number of competing bids on a dwellings	Positive
Supply	Absorption Rate	The ratio between dwellings sold and the total inventory	Positive
	Home Inventory	The total inventory of dwellings available for sale	Inverse

* Relationship describes the hypothesized association between the indicator and the phenomenon of market temperature

3 Theory

This chapter outlines a sequence of methodological steps to construct a composite indicator. The steps generally follow the *Handbook on Constructing Composite Indicators* developed by OECD (2008). A brief overview follows.

1. *Indicator selection* involves selecting indicators based on their relevance to the phenomenon being measured and statistical soundness.
2. *Normalization* involves making indicators comparable by adjusting indicators to a standard scale.
3. *Weighting* involves determining the trade-off between indicators before compiling them.
4. *Aggregation* involves compiling the indicators to a composite index and thus determining the degree of compensability among the indicators.
5. *Robustness* involves quantifying uncertainty in indicator scores and identifying sources of uncertainty through uncertainty and sensitivity analysis.

3.1 Indicator Selection

Indicator selection is the process of selecting variables to form the composite indicator. We summarize the importance of this step in two considerations. First, the indicators' quality directly affects how well we measure the phenomenon of market temperature. Second, the indicators' statistical structure influences how appropriate it is to compile them together. In the following, we present a quality framework to assess potential indicator candidates and the application of multivariate analysis to determine if it is reasonable to compile them to a composite.

3.1.1 Quality Framework

To ensure the quality of indicators, OECD (2008) employs a framework to guide indicator selection. The framework consists of six quality dimensions: (1) accessibility, (2) relevance, (3) accuracy, (4) timeliness, (5) interpretability, and (6) coherence. By assessing if each indicator conforms to these dimensions, we can ensure the integrity of the index:

1. Accessibility relates to how easily we can locate real estate specific data. It affects the cost of developing and maintaining the indicator over time. Moreover, accessibility influences the integrity of the index if it is difficult to replicate. Although it is essential to account for data quality, preference should also be given to accessible sources.

2. Relevance relates to the purpose of the indicators and whether they, in a balanced way, cover an adequate range of domains. That is, a meaningful index relies on indicators that explain conceptual areas of market temperature.

3. Accuracy relates to the credibility of the data source. The quality of the index depends on the objectivity of the data. Moreover, it should be produced by proper statistical standards.

4. Timeliness is essential to minimize the need for data imputation because sources often publish data at different times. Thus it is necessary to evaluate the consistency of when data can be retrieved.

5. Availability of metadata is vital to ensure the comparability of data over time. Metadata includes, for example, classifications and definitions used to produce the data.

6. Data must be coherent over time and across units (e.g., cities or boroughs), which implies being based on common definitions, concepts, and methodology to be comparable (OECD, 2008).

3.1.2 Multivariate Analysis

After assessing the quality of the indicators, it is useful to apply multivariate analysis to explore interrelationships between them. Multivariate analysis is a set of statistical procedures that allows for simultaneous observation of multiple variables (W. Becker, 2021d). It is a preliminary step to assess how suitable it is to aggregate the indicators to an index. Moreover, it guides later methodological steps, such as weighting. The first step is to evaluate pair-wise associations with Pearson's correlation coefficient. Then, estimate how consistent the indicators describe the same phenomenon with Cronbach's Coefficient

Alpha. Finally, indicator association is explored through Principal Component Analysis (PCA). In the following assume that Q is the number of indicators, and M the number of months in consideration, where:

x_{qm} : is the raw value of indicator q for month m, with $q = 1, ..., Q$ and $m = 1, ..., M$.

3.1.2.1 Pearson's Correlation Coefficient

Pearson's correlation coefficient measures the linear relationship between two indicators (Nickolas, 2021). Examining correlations is necessary because composite indicators require relationships between the variables (Hardeman, Van Roy, Vertesy, & Saisana, 2013). Yet, the coefficients should not be excessively high since it implies redundancy (Hardeman et al., 2013). The method involves computing the ratio between the indicator's covariance and the product of their standard deviations:

$$\rho = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}} \tag{3.1}$$

Hence, it is a normalized statistic of the covariance on a scale between -1 and 1. A coefficient of 1 implies that an increase in one indicator value returns a positive increase of a fixed proportion in the other indicator (Glen, 2021). Equivalently, -1 implies a negative decrease of a fixed proportion. Thus the absolute value of the coefficient describes how strong the bivariate relationship is.

3.1.2.2 Cronbach Coefficient Alpha

Cronbach Coefficient Alpha is a popular estimate of the internal consistency of data (Cronbach & Shavelson, 2004). In the context of indices, the coefficient relies on the correlation between individual indicators. That is, alpha estimates the proportion of variance that is systematic in the set of indicators. The formula to compute Cronbach's Alpha is:

$$\alpha = \left(\frac{Q}{Q-1}\right)\frac{\sum_{i \neq j} cov(x_i, x_j)}{var(x_0)} \qquad m = 1, ..., M\ ;\ i, j = 1, ..., Q \tag{3.2}$$

where $x_0 = \sum_{q=1}^{Q} x_j$ is the sum of all indicators (OECD, 2008). The statistic generally ranges from 0 to 1, and α increases with the covariance of each pair of indicators. If indicators are perfectly collinear, the alpha is 1, while if they are independent, the alpha equals 0. However, from the formula, we observe that increasing the number of indicators will increase alpha's size. For example, assume the reliability of four indicators to be 0.8 and the correlation among them constant; if we add another indicator, the reliability increases to 0.86.

Research suggests a coefficient value of 0.7-0.8 to be sufficiently reliable, depending on the context (Nunnally, 1978; Vaske, Beaman, and Sponarski, 2017). If the set of indicators exceeds the threshold, it is evidence that they individually measure the concept well. Nevertheless, the statistic is not a measure of unidimensionality (Ten Berge & Sočan, 2004). To better understand the dimensionality, we can conduct principal component analysis.

3.1.2.3 Principal Component Analysis

PCA is a statistical method that can provide valuable insight into the underlying structure of our data (OECD, 2008). The method involves finding linear combinations of Q indicators, $x_1, x_2, ..., x_Q$ to produce uncorrelated principal components $Z_1, Z_2, ..., Z_Q$ (Saisana & Tarantola, 2016), following:

$$Z_j = \sum_{i=1}^{Q} a_{ij} x_i \ , \ j = 1, 2, ..., Q \tag{3.3}$$

The weights a_{ij} are estimated to satisfy three conditions (Saisana & Tarantola, 2016). First, the principal components $Z_1, Z_2, ..., Z_Q$, are constrained to be uncorrelated, implying that they measure different dimensions in the data. Second, the first component, Z_1, contain the maximum proportion of the variance possible, while Z_2, explain most of the remaining variance, and so on. Third, the sum of squared loadings across all components equal one.

In brief, PCA involves computing the eigenvalues λ_j of the sample covariance matrix C,

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{1Q} \\ c_{21} & c_{22} & c_{2Q} \\ \ldots & & \\ c_{Q1} & c_{Q2} & c_{Q3} \end{bmatrix} \tag{3.4}$$

The diagonal elements of the matrix c_{ii} represents the variance of x_i, while c_{ij} are the correlation between pairs of x_i and x_j. The eigenvalues of the matrix constitute the variance in the principal components. The total variance explained by all components equals the total variance of the original indicators. To avoid having a single indicator significantly influencing the components, it is adviseable to standardize the Q indicators to have means of 0 and standard deviations of 1. In that case, the covariance matrix equals a correlation matrix, where the correlation between components Z and the indicators x are referred to as component or factor *loadings*. If the indicators are uncorrelated, the loadings are identical to the weights a_{ij}.

PCA is helpful because few components usually preserve a lot of variance (Saisana & Tarantola, 2016). A low-dimensional representation of the data makes it easier to observe trends graphically. By exploring how much the underlying indicators correlate with components, we better understand if indicators contain unique patterns or measure domains already reflected in other variables. Moreover, if PCA only yields one component with an eigenvalue above 1, we can consider the data unidimensional (Vaske and Sponarski, 2016).

3.2 Normalization

After selecting high-quality indicators that measures the same latent phenomenon, the third step involves making them comparable through normalization. Normalization is the process of adjusting the values of different scales to a common scale (European Commission, n.d.-b). Comparable indicators are necessary before summarizing them into a composite. There are numerous methods of normalizing indicators, each with different properties. In the context of measuring market temperature we can distinguish between two applicable methods; *linear methods* (min-max or z-scores) and *non-linear methods* (ranking).

The choice of a normalization method depends on the purpose of the composite indicator. If the purpose is to compare a unit's relative performance over time, rank is suitable (OECD, 2008). When the intention is to reward or penalize exceptional behavior, the absolute differences between indicator values are important, and linear methods are considered more applicable (European Commission, n.d.-b). Both linear and non-linear methods are useful in the context of measuring market temperature, and are described in further detail below, where we employ the notation:

I_{qm} : the normalised value of indicator q for month m.

3.2.1 Ranking

Ranking is the simplest normalization method and involves assigning ranks to units based on their indicator values (European Commission, n.d.-b). For instance, assigning rank 1 to the highest indicator value, rank 2 to the second-highest, and so on. We compute the rank as

$$I_{qm} = Rank(x_{qm}) \tag{3.5}$$

The main advantage with using rank is the simplicity and independence to outliers (European Commission, n.d.-b). In cases of time-dependent data, rank is usually conducted at each point in time which allows comparison of a month's relative performance (European Commission, n.d.-b). However, as information on absolute levels is lost, a unit's difference in absolute performance cannot be compared. For instance, a unit can improve from one month to the next, yet its ranking can decrease if other units improve more.

3.2.2 Min-max

Min-max normalization uses a linear function to rescale the indicators to a common scale, for example 0 to 100, by using the minimum and maximum values as reference points (European Commission, n.d.-b). The formula is as follows:

$$I_{qm} = \frac{x_{qm} - min(x_q)}{range(x_q)} \tag{3.6}$$

The advantage of min-max is that re-scaling can widen up the range of an indicator (OECD, 2008). This allows differentiation between months of similar levels of performance. Another advantage is that min-max re-scales indicators to the exact same scale (W. Becker, 2021e), which makes it easy to communicate and interpret. However, a disadvantage with min-max is that extreme values can distort the normalized indicators (European Commission, n.d.-b). This can happen if the maximum and minimum values are outliers. Such outliers can again introduce aggregation distortions when indicators have different means (European Commission, n.d.-b).

3.2.3 Z-Scores

Z-scores involve re-scaling indicators to a common scale, generally with a mean equal to zero and standard deviation of one (OECD, 2008). Normalizing using z-score is as follows:

$$I_{qm} = \frac{x_{qm} - \mu_q}{\sigma_q} \tag{3.7}$$

where μ_q and σ_q is the mean and standard deviation of the indicator. A benefit of z-scores is that aggregation distortions stemming from different indicator means are avoided (European Commission, n.d.-b). This allows indicators from different distributions to be directly compared. Using the standard deviation as a scaling factor results in extreme values having greater effect on the composite indicator (European Commission, n.d.-b). Hence, z-scores are suitable when the intention is to reward exceptional behavior. However, the drawback with using z-scores is that the indicators will have differing scales (European Commission, n.d.-b).

3.3 Weighting

Selecting a weighting scheme is the fourth step in constructing a composite indicator. We can interpret weights as a trade-off ratio between pairs of indicators (Munda & Nardo, 2009). Thus, the selection of weighting schemes might significantly affect the final index (Saisana, Saltelli, & Tarantola, 2005). Nonetheless, there is no uniformly agreed methodology for weighting (European Commission, n.d.-c). Weights can be assigned based on the quality of the indicators, statistical considerations, or participatory approaches (Greco, Ishizaka, Tasiou, & Torrisi, 2019). This subsection presents plausible weighing systems in the context of aggregating indicators to gauge market temperature. These include equal weighting, using principal component analysis, and neutralizing excessively high correlations.

3.3.1 Equal Weights

The most common weighting scheme for composite indicators is equal weights (Bandura, 2008). This method is often justified by (1) simplicity of construction, (2) lack of theoretical structure to justify unequal weights, (3) disagreement between decision-makers, (4) inadequate empirical or statistical knowledge, and (5) alleged objectivity (Freudenberg, 2003; OECD, 2008; Maggino and Ruviglioni, 2009; Decancq and Schokkaert, 2016). However, applying equal weight with no adequate justification, such as selecting weights based solely on simplicity, can bear a considerable oversimplification cost (Paruolo, Saisana, & Saltelli, 2013).

3.3.2 Principal Component Analysis

We can elicit weights using factor loadings of the first principal component. The procedure is desirable because of its objectivity and the transparency of the process. However, relying on the first component requires that it explains a large enough portion of the variance. Thus, more components are sometimes necessary. In general, components should (a) have eigenvalues greater than one, (b) explain more than 10 percent of the overall variance, and (c) cumulatively contribute to more than 60 percent of the total variance (OECD, 2008).

When developing indicators of product market regulation, Boylaud and Nicoletti (2000) show that we can construct weights by first rotating components using a varimax technique. The rotation maximizes the variance shared between the indicators and minimizes the number of indicators with high loadings on the same component. Weights are then computed based on their contribution to the overall variance in the associated factor. For example, assume that three indicators have equal loadings, $a = 0.6$. Then the sum of squared loadings equals the total variance explained by the component, $3 * 0,6^2 = 1,08$. Their respective weight is thus $0.6^2/1.08 \approx 0.33$.

Although PCA's statistical properties are advantageous, they can also be a drawback. Weights selected endogenously do not inevitably resemble the actual connection among the indicators (Saisana & Tarantola, 2002). Moreover, it does not necessarily ensure a sound theoretical foundation for the composite indicator (De Muro, Mazziotta, & Pareto, 2011).

3.3.3 Correlation Neutralization

We can also elicit indicator weights using correlation analysis. For example, in the instance of a strong correlation between a set of indicators, we might want to adjust weights to moderate the effect. The reason is that highly correlated indicators could introduce double counting. In other words, a specific phenomenon could implicitly receive more weight relative to other indicators if they were both included and collinear. The approach is applied in the index of the regional problems in the European Union (Saisana & Tarantola, 2002). The procedure is as follows. First compute the arithmetic mean of the coefficients of determination for each bivariate correlation that includes the given indicator:

$$u_i = \left(\frac{1}{Q-1}\right) \sum_{j \neq i}^{Q} \left(\frac{\operatorname{cov}(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}}\right)^2 \; j = 1, ..., Q, \;\; \forall\, i \epsilon \{1, ..., Q\} \tag{3.8}$$

Then compute the weight for a given indicator to be inversely proportional to the preceding mean:

$$w_i = \frac{1}{u_i} * \sum_{i=1}^{Q} u_i \;\; i = 1, ..., Q \tag{3.9}$$

Finally, the weights can be scaled to add up to 1.

3.4 Aggregation

Aggregation is the final step in construction since it involves compiling the individual indicators into a single index. As with selecting the weighting scheme, there is no ideal system of selecting aggregation schemes (Arrow, Raynaud, et al., 1986). The fundamental concern in selecting aggregation methods is to decide the degree of compensability between indicators (Greco et al., 2019).

Compensability refers to the option of compensating advantage on some variables by a disadvantage on another variable (Munda & Nardo, 2009). We divide aggregation methods into two distinctive categories: *compensatory* and *non-compensatory* approaches (Munda, 2016). Examples of compensatory aggregation methods are arithmetic, geometric and harmonic mean. Multi-criteria analysis is an example of non-compensatory aggregation.

With compensatory aggregation methods, a strong performance in an indicator can compensate for a weaker performance in another. In contrast, non-compensatory approaches allow no form of compensation between indicators. Because we consider weights to designate a trade-off ratio between pairs of indicators, non-compensatory aggregation is the only relevant approach (Greco et al., 2019). In the following, we employ the notations:

w_q : the weight of indicator q.

CI_m : the value of the composite indicator for month m.

We aggregate over the Q normalized indicators $I_{1m}, I_{2m}, ..., I_{Qm}$, such that $\sum_{q=1}^{Q} w_q = 1$ and $w_q > 0$.

3.4.1 Arithmetic Mean

Arithmetic mean is the most straightforward method of aggregating indicators. The weighted arithmetic mean is calculated by:

$$CI_m = \sum_{q=1}^{Q} w_q I_{qm} \tag{3.10}$$

Arithmetic aggregation implies perfect compensability, which means that a high score in an indicator can perfectly compensate for a low score in another indicator(Greco et al., 2019). Thus, the method is only applicable if a quantified substitution rate exists between the indicators (Munda & Nardo, 2009).

3.4.2 Geometric Mean

Weighted geometric mean is an alternative that can reduce the level of compensability between indicators. The geometric aggregation method uses the product of the indicators rather than the sum:

$$CI_m = \prod_{q=1}^{Q} I_{qm}^{w_q} \tag{3.11}$$

Geometric aggregation is appealing in instances where an indicator's high score should not fully compensate for low achievement in others (UNDP, n.d.). Moreover, if a units low score improves, its marginal utility will be higher compared to when high-achieving indicators improve (Greco et al., 2019).

3.4.3 Harmonic Mean

The weighted harmonic mean is the least compensatory aggregation method of the three methods (W. Becker, 2021b). It uses the mean of the reciprocals of the indicators, which we compute as follows:

$$CI_m = \frac{\sum_{q=1}^{Q} w_q}{\sum_{q=1}^{Q} \frac{w_q}{I_{qm}}} \tag{3.12}$$

Using the reciprocals aggravates the impact of small indicator values and mitigates the impact of extreme values. Thus, it is desirable to use it when very little compensation is required. The harmonic mean often provides an accurate estimate when aggregating indicators involving rates and ratios.

3.5 Robustness

Although aggregation is the final step in constructing the index, it is necessary to measure the robustness of the composite before concluding the work. Developing a composite indicator involves making decisions at multiple stages, which introduces uncertainties. Indicator selection, indicator construction, and methodological steps such as normalization and aggregation are all sources of input uncertainty. Even though it is impossible to measure uncertainties completely, factors with plausible alternatives warrant further analysis of their effect on the outcome. The combination of uncertainty analysis and sensitivity analysis can help to improve transparency and gauge the robustness of the composite indicator (OECD, 2008).

Sensitivity analysis (SA) is a method that allows us to quantify which and by how much input uncertainties are causing output uncertainty (OECD, 2008). The method focuses on the variance of model outputs and helps uncover the relative significance of factors. If the SA result indicates that specific factors do not vary the outcome, we might assign an arbitrary value or use methods that are easier to interpret. Sensitivity analysis is closely related to uncertainty analysis (UA). UA focuses on how input uncertainties affect the composite indicator values (OECD, 2008). This can help provide insights on how robust the index scores are to different methodology.

The main technique for UA and SA is to use Monte Carlo simulations (W. Becker, 2021c). In this paper we use the Monte Carlo design as presented by Saisana, Saltelli, and Tarantola (2005). It involves estimating the index multiple times, each time randomly varying the uncertain input alternatives. The benefit of the design is that it applies variance-based SA. When several uncertainty sources exist simultaneously, the model could become nonlinear and possibly non-additive, thus variance-based techniques for SA is the most appropriate (Saltelli, 2007). Additionally, variance-based techniques has several attractive features. First of all, they allow an analysis of the full range of variation of input factors (OECD, 2008). Further, they can detect interaction effects among input uncertainties and give measures of uncertainty that are easy to interpret (OECD, 2008). The following subsections describe the Monte Carlo design for UA and variance-based SA by Saisana, Saltelli, and Tarantola (2005).

3.5.1 The Monte Carlo Framework

The Monte Carlo simulation generates N replications of the index and computes two outputs for each replication. These two outputs are used to conduct UA and SA. The first output is the ranked index score for each month. Ranks are used because indicator scores can be very inconsistent depending on the methodology, while ranks are much more stable. The composite indicator CI for a given month m is assigned a rank defined by:

$$Rank(CI_m) \tag{3.13}$$

The second output is a single statistic capturing the relative shift in ranks across the whole period. This statistic is expressed as the mean absolute rank change between nominal ranks (ranks from the base model) and ranks from replication N. The mean absolute rank change is given by:

$$R_s = \frac{1}{M}\sum_{m=1}^{M} |Rank_{nom}(CI_m) - Rank(CI_m)| \tag{3.14}$$

where M represents total number of months, while $Rank_{nom}(CI_m)$ is the rank for month m in the nominal model, and $Rank(CI_m)$ the replicated model.

The sources of uncertainty is translated into a set of scalar input parameters, which can be sampled randomly according to their distributions for each replication. For example, assume the aggregation system causes uncertainty, and plausible options are linear and geometric aggregation. Then, the aggregation method would be a discrete input parameter X_i which can take integer values between 1 and 2 used to trigger the respective aggregation method. Similar triggers are created for the complete set of uncertainty inputs. An example of the complete procedure follows. Assume two uncertain input factors indicated by $X_i, i = 1, 2$, where i is normalization and aggregation schemes, then:

1. Assign a probability distribution function to each input parameter X_i. The first input parameter, X_1 will trigger a selection of normalization method; the second input parameter, X_2, will select the aggregation method. X_i are discrete random

variables and are produced by drawing a random integer ζ, uniformly distributed between [0,1]. Thereafter we use a russian roulette algorithm, e.g. for X_2, linear aggregation is selected if $\zeta\epsilon[0, 0.5)$ and harmonic aggregation is selected if $\zeta\epsilon[0.5, 1]$.

2. Generate a sample, l, of N random combinations of independent input parameters X^l , where $l = 1, 2, ..., N$.

3. For each sample l, select a method of normalization and aggregation based on X_1, X_2

4. Estimate the output Y^l, where Y^l is either $Rank(CI_m)$, the rank for each month, or R_s, the mean absolute rank change.

5. Close the loop over l, and analyse the output vector $Y^l, l = 1, 2, ..., U$

3.5.2 Uncertainty and Sensitivity Analysis

UA uses the ranks from the N simulations as output Y to assess the overall uncertainty. The mean, median and confidence intervals of the ranks across N replications are computed for each month. These summary statistics can then be visualized in comparison with nominal ranks. By comparing for instance the difference between the nominal ranks and the median rank, the uncertainty of the index values can be quantified. If there are considerable discrepancies between actual ranks and median ranks, the model can be considered biased.

In contrast, SA applies the mean absolute rank change R_s as output Y from all the N replications. The uncertainty in a single model output Y is encapsulated from its variance $V(Y)$ (W. Becker, 2021c). The higher the variance, the more uncertainty. This variance can then be decomposed into uncertainty caused by each input, and interaction effects between inputs. The total output variance $V(Y)$ of output Y can be decomposed as:

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} + ... + V_{12}...k \tag{3.15}$$

where:

$$V_i = V_{x_i}\{E_{x_{-i}}(Y|X_i)\} \tag{3.16}$$

$$V_{ij} = V_{X_iX_j}\{E_{X_{-ij}}(Y|X_iX_j)\} - V_{X_i}\{E_{X_{-i}}(Y|X_i)\} - V_{X_j}\{E_{X_{-j}}(Y|X_j)\} \tag{3.17}$$

and so on for higher terms. Here x_i denotes the i-th input parameter varied, k the number of uncertain input parameters. $E(.)$ denotes the expected value and $V(.)$ the variance operator. Equation 15 shows the first-order conditional variances, that is variance caused by factor X_i. The following example explain the intuition behind this equation.

Assume the fixed parameter X_i to a particular value $X_i{}^*$ in its range. This value can for instance be aggregation method, with a range of 2 possible alternatives (e.g linear and geometric aggregation). Then compute the mean of Y over all parameters except X_i: $E_{x-i}(Y|X_i = X_i{}^*)$. Thereafter take the variance of the function of $X_i{}^*$ over the all values in the range of the fixed parameter X_i. The results is the variance where the dependence on $X_i{}^*$ has been dropped. Consequently, V_i is a number between 0 and $V(Y)$ when all other parameters are non-influential at any order. If X_i does not contribute to Y at the first order, V_i is 0.

Equation 16 is the second-order term, that is, variance contribution caused by interactions between inputs (e.g.X_i and X_j). Interactions only exists if the first term in the equation, $V_{X_iX_j}\{E_{X_{-ij}}(Y|X_iX_j)\}$, is larger than the sum of the first order terms for X_i and X_j.

The first- and second-order terms can be used to describe the fractional contribution to the model's overall output variance, $V(Y)$, caused by the uncertainty in X_i. This is measured through *sensitivity indices.* Two sensitivity indices are created for each input uncertainty, *first-order sensitivity index* (S_i), and *total-order sensitivity index* (S_{Ti}). The *first-order sensitivity index* is the fraction of output variance caused by each uncertain input parameter alone. This can be computed as:

$$S_i = \frac{V_i}{V(Y)} \tag{3.18}$$

The first *first-order sensitivity index* can be defined as the "main effect" caused by input uncertainty i. Important input uncertainties are those that, when fixed individually, reduce the most variance in the output. Input uncertainties can be considered important if the S_i (main effect) is greater than $1/k$ of the total output variance (Saisana et al., 2005).

Terms above first-order is defined as "interaction effects". Models without interactions among input uncertainties are additive. Subsequently, the first-order conditional variance equation will suffice to decompose the output variance of the model. However, if interactions exist, higher-order sensitivity indices need to be estimated. The number of indices that require estimation is at max $2^k - 1$, and higher-order sensitivity indices are thus typically not estimated. The alternative is a compact sensitivity measure, the *total-order sensitivity index* (S_{Ti}), which expresses the total total contribution to the variance of Y caused by X_i alone and the interaction effects. With a model of three independent input uncertainties the total-order indices are:

$$S_{T1} = \frac{V(Y) - V_{X_2 X_3}\{E_{X_1}(Y|X_2, X_3)\}}{V(Y)} = S_1 + S_{12} + S_{13} + S_{123}. \tag{3.19}$$

Analogously:

$$\begin{aligned} S_{T2} &= S_2 + S_2 + S_{12} + S_{23} + S_{123} \\ S_{T3} &= S_3 + S_3 + S_{13} + S_{23} + S_{123} \end{aligned} \tag{3.20}$$

The *total-order sensitivity index* is defined in more general terms as:

$$S_{Ti} = \frac{V(Y) - V_{X_{-i}}\{E_{X_1}(Y|X_{-i})\}}{V(Y)} \tag{3.21}$$

where X_{-i} represents all input uncertainties except the ith. The interaction effect for input uncertainty i can be isolated by taking the difference between S_{Ti} and S_i. When there is a significant difference between S_{Ti} and S_i, the input uncertainty X_i has a noticeable interaction role in the output variance. The pair of sensitivity indices thus give a practical description of model sensitivities.

Computing the *total-order sensitivity index* can be computationally expensive and requires more replications than UA. High complexity and dimensionality can constraint the number of model runs available. Hence, selecting an efficient and robust first-order estimator is of high importance. Recent study by Puy, Becker, Piano, and Saltelli (2020), comparing total-order estimators for variance-based sensitivity analysis suggest that the estimator by Jansen (1999) is the most accurate in estimating "true" total-order indices. The study also conclude that Jansen's estimator is considered as most accurate when the aim is to measure inputs in terms of their contribution to output variance. The estimator requires $N(d+2)$ number of replications to compute the *total-order sensitivity index*, where N is the number of samples and d is the number of uncertain input parameters. The estimator is applied and recommended in the methodologies for construction and evaluation of composite indicators by the *The European Commission's Competence Center on Composite Indicators* (W. Becker, 2021f). Hence, the method can be considered applicable for developing composite indicators. Further explanation of the estimator will not be conducted. We refer to Jansen (1999) for more details .

4 Methodology

This section outlines the methodology we use to construct the composite indicator. It generally follows the sequence of steps outlined in Section 3. We begin by describing the data sample, which is the basis for selecting indicators that constitute the index. To remove erroneous data, we apply suppression rules, limiting dwelling sizes and square meter prices to reasonable figures. Then we proceed with a discussion of indicator selection. Potential indicators are those identified through the previous literature review and contemporary indices. By assessing their quality, their fitness of use, and statistical properties, we determine the indicators Days on Market, Sale-to-List Price Ratio, and Absorption Rate to be most suitable in explaining market temperature. After presenting the selection process, we discuss the rationale behind selecting min-max normalization, equal weights, and geometric aggregation, which are the methods forming the final composite model.

Decisions made in one step of the development have implications for the next. Thus, appropriate methodological choices also imply that we identify if they fit well together. Since coherence is essential and we aim to provide a transparent model, we emphasize examining the robustness of the index. Thus, we sequentially discuss plausible alternatives for each step of normalization, weighting, and aggregation. The alternatives constitute the sources of uncertainty, which we translate into a set of input factors for the uncertainty and sensitivity analysis described in the final part of this section.

Each step of the development process is carried out using the R programming language. In particular, we employ the *COINr* (W. Becker, 2021a) package as it has sophisticated tools for composite indicator construction. It was developed by the European Commission's Joint Research Centre to ease the development of high-quality indices.

4.1 Data

This subsection is three-fold. First, we describe the dataset; second, we proceed with exploratory data analysis; third, we explain necessary data treatment procedures. These steps aim to characterize the underlying data of the composite and to treat data errors.

4.1.1 Data Description

Our data consists of roughly 46 000 second-hand residential property transactions in Oslo Municipality from January 2017 to May 2021. The sample constitutes sales brokered through FINN's classified advertisement website. In total, FINN is the intermediary of approximately 70 percent of all second-hand property transactions in Norway (Eie, 2021).

Information about dwellings listed on FINN is collected through web crawling. The data contains details on list prices, list dates, sale dates, and snapshot data from 2019, which captures changes in property listings. A snapshot exists for each revision made to a listing. We also aggregate data from The Norwegian Mapping Authority (NMA). The NMA provides additional details on all individual property transactions, such as the date of ownership change, usable area, dwelling type, and payments. Table 4.1 contains a description of data features and their respective sources.

Table 4.1: Description of Data Features and Sources

Variable	**Description**	**Source**
address_id	Unique dwelling ID	NMA
unit_type	Dwelling type	NMA
official_date	Date the dwelling changed ownership	NMA
official_price	Price the dwelling was sold for (NOK)	NMA
useable_area	Useable area (m^2)	FINN
register_date	Date the dwelling was listed for sale on FINN	FINN
asking_price	Seller's asking price (NOK)	FINN
sold_date	Date the dwelling was sold	FINN
borough	Residential borough in Oslo	FINN

NMA: Norwegian Mapping Authority

4.1.2 Exploratory Data Analysis

Table 4.2 shows summary statistics of numeric variables. The transaction value of dwellings sold (official_price) ranges from roughly 0.5 to 71 million, with a mean of 6 million and a median of 5 million. The usable area varies from 0 to 677 square meters, with a mean of 86 and a median of 70 square meters. Both distributions have significant differences between the mean and median, implying positive skew. Moreover, looking at the square meter price statistics, we observe the presence of erroneous data. Dwellings with a usable area of 0 square meters should not exist, and the maximum square meter price also looks questionable. Hence we are required to go through a process of data treatment before constructing the composite indicator.

Table 4.2: Summary Statistics for Numerical Variables

	Min	**Max**	**Mean**	**Median**	**Std.dev**
asking_price	1,100,000	80,000,000	5,880,000	4,850,000	3,420,000
official_price	480000	71,100,000	5,990,000	4,950,000	3,410,000
usable_area	0	677	86	70	55
sqm_price	0	760,300	74,600	74,700	21,200

4.1.3 Data Treatment

We adopt suppression rules implemented in SSB's housing models to ensure satisfactory data quality. This step equals SSB's procedures in modeling housing statistics and is necessary to handle the evident data errors, such as transactions with a square meter price of 0. Table 4.3 shows criteria on square meter prices and usable areas for apartments and houses. We remove all transactions where apartments have a square meter price below 10,000 and above 200,000 NOK. The limit for a house is 5,000 to 150,000 NOK. The usable area must range from 50 to 550 square meters for houses and 12 to 350 for apartments. The suppression rules reduce the sample size by about 200 transactions (0.45 percent).

Table 4.3: Suppression Rules for Square Meter Prices and Usable Areas

	Usable Area (m^2)	**Price/m^2 (NOK)**
House	50 - 550	5,000 - 150,000
Apartment	12 - 350	10,000 - 200,000

Note. Adapted from *Modell for beregning av boligformue* by Medby and Takle (2021)

4.2 Indicators

This subsection introduces the indicators that constitute the index and how we compute them. Potential indicators are those identified in Subsection 2.2 through the literature review and contemporary indices. While each indicator represents a particular phenomenon, we grouped them under the hypothesis that they, in broad terms, describe different conceptual areas. That is demand, supply, and price negotiation. Accordingly, the final composite indicator should explain these domains. However, the question of how to select indicators also depends on their quality and whether it is reasonable to compile them together. Thus the selection process is guided by OECD's quality framework and multivariate analysis. Finally, we use descriptive statistics to summarize the indicators created.

4.2.1 Quality Assessment

To evaluate if we can use the ten indicators in constructing our index, we assess how well each indicator conforms to six quality dimensions. We summarize the analysis in Table 4.4, while Appendix A1 describes the complete evaluation. The result is that we have to omit Waived Contingencies, Listing Views, Expired Listings, and Competing Bids. The main reason is accessibility, which relates to how easily we can locate specific data required to construct the indicators. For example, contingencies and bids are only accessible to the participants in particular bidding rounds, and thus not available to the public.

Table 4.4: Quality Assessment

	Accessibility	**Relevance**	**Accuracy**	**Timeliness**	**Availability**	**Coherence**
Sale-to-List Price Ratio	✓	✓	✓	✓	✓	✓
Listings With a Price Cut	✓	✓	✓	✓	✓	✓
Waived Contingencies	–	✓	✓	✓	✓	✓
Days-on-Market	✓	✓	✓	✓	✓	✓
Listing Views	–	✓	✓	✓	✓	✓
Sales Volume	✓	✓	✓	✓	✓	✓
Expired Listings	✓	✓	–	✓	✓	✓
Competing Bids	–	✓	✓	✓	✓	✓
Absorption Rate	✓	✓	✓	✓	✓	✓
Home Inventory	✓	✓	✓	✓	✓	✓

✓: Quality criteria *satisfied*, – : quality criteria *unsatisfied*

Quality also implies that each indicator has a "fitness of use" (OECD, 2008, p.42). That is, the index derived from the indicators meets the needs of the buyers, sellers, and policymakers who might use it. Accordingly, we argue that we should select one indicator to explain each of the three domains of market temperature: supply, demand, and price negotiation. The rationale is two-fold. First, it will enhance interpretability. When communicating the index, it will be more manageable to decompose an index consisting of three rather than the remaining six indicators. Second, when two or more indicators partially measure similar domains, we might introduce an element of double counting into the composite index. That is, it can cause an imbalance in what the index actually measures.

In terms of price negotiation, the choice is between the Sales-to-List Price Ratio and the Number of Listings With a Price Cut. Several reasons substantiate using the former. All but one index previously reviewed uses the Sales-to-List Price Ratio, while Zillow is the only one that uses Listings With a Price Cut. The reason is perhaps that the interpretation of the former indicator is more valuable. Since it describes how close selling prices are to asking prices, the index can, for example, help buyers estimate how much they have to bid to win. Besides, the Sales-to-List Price Ratio is a more explicit quantification of the leverage between market participants. That is, a positive discrepancy between the prices indicates a hotter market, while a negative implies a colder. Thus, we select the Sales-to-List Price Ratio to indicate price negotiation in the real estate market.

Although it is evident that supply influences them, Days on Market and Sales Volume indicate demand-side pressure. Days on Market is a frequently used statistic to describe how long it takes to sell a home. All indices reviewed include the indicator, making it the most popular among the two. While the Sales Volume provides information about market demand, it is less valuable without the context of available supply. Moreover, the interpretation of Days on Market is both straightforward and especially applicable for end-users, since it helps gauge the pace of the market. When Days on Market is low, homes sell quickly, which suggests a hot market. Thus, Days on Market is the most appropriate to indicate demand pressure.

In terms of the supply side, the choice is between the Absorption Rate and Home Inventory. While Home Inventory describes the total number of dwellings available for sale, the

Absorption Rate indicates a relative figure between the sales rate and the home inventory. Both indicators inform market participants about available prospects; however, we argue that it is more beneficial to know whether the opportunities for buyers are increasing or decreasing. For instance, a rate of 0.5 means that 50 percent of the available housing supply sells in a single month. Thus, a high rate indicates growing competition for the available supply. Subsequently, we select the Absorption Rate to indicate relative supply.

4.2.2 Indicator Construction

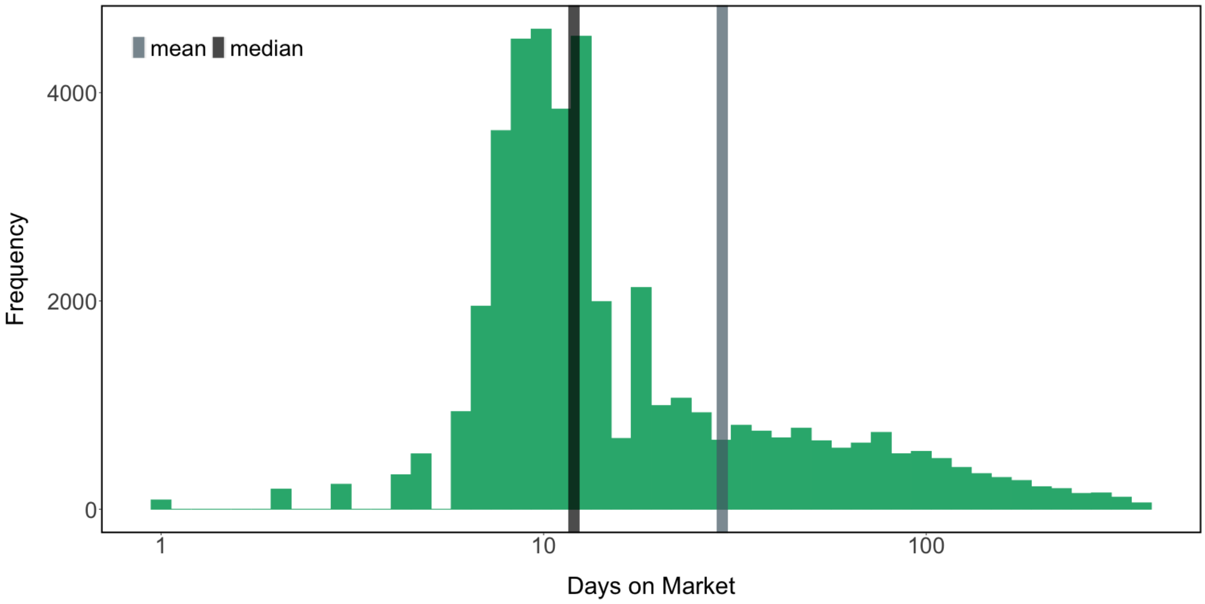
In the following, we describe the process of computing the three indicators: Days on Market, Sale-to-List Price Ratio, and the Absorption Rate. We compute the indicators on a monthly time granularity to ensure sufficient data underlying each estimate over time. We apply separate suppression rules for the Sale-to-List Price Ratio and Days on Market to promote accuracy and reduce the impact of outliers. These suppression rules are guided by Zillow's methodology in constructing their Buyer-Seller Index. Suspicious observations are removed individually in the computation of each indicator without excluding them from the overall sample. For example, a transaction's listing days might be correct, even though the sale price is questionable. Finally, as we summarize each indicator monthly, we discuss appropriate measures of central tendency.

4.2.2.1 Days on Market

We define Days on Market as the median number of days a listing is active on FINN before an offer is accepted or the seller delists the dwelling. We begin by computing the daily time difference between dwellings' list dates and sold dates. If a listing is active for more than 365 days or sells before or within the same day it is listed, we omit it from the estimation. The suppression rule ensures the accuracy of the variable. For example, 48 observations have negative active days on the market, which suggests data errors. In total, the filter excludes 344 observations (0.75 percent of the data). However, this is variable-specific, implying that we do not omit data from the overall sample.

After the suppression, we want to summarize the measure of central tendency to create the indicator. Two feasible options are to use the mean or median. The decision depends on the distribution of the data. Figure 4.1 shows a histogram of Days on Market for the full sample, with the vertical lines showing the mean and median. Note that the x-axis is log-scaled. We observe a positive skew and that the median is better at providing the central location of the data. Appendix A2 confirms a large skew in the distribution for the individual months. Thus, we compute the monthly median to yield the final indicator of Days on Market.

Figure 4.1: Histogram of Dwellings' Active Days on Market



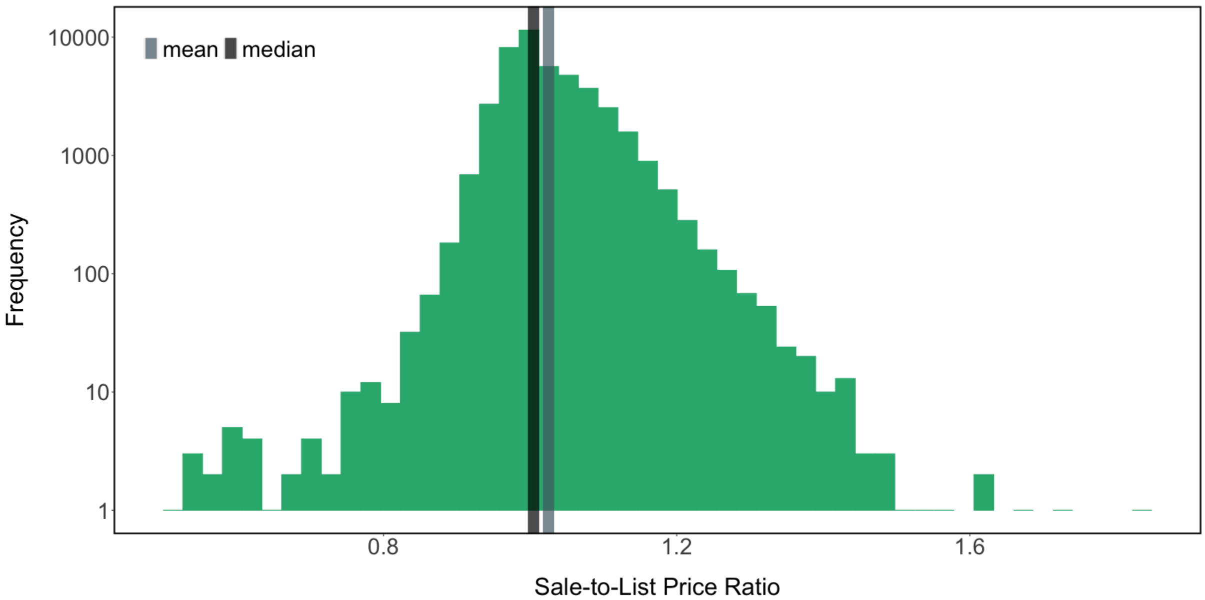
Note. The x-axis is log-scaled.

4.2.2.2 Sale-to-List Price Ratio

We define the Sale-to-List Price Ratio as the mean ratio between the official sale price registered by the NMA and the listing price on FINN for dwellings sold within a specific month. We compute the Sale-to-List Price Ratio by dividing sale prices by listing prices for each transaction. If a dwelling's sale-to-list price ratio exceeds two or is less than 0.5, we omit it. For example, a house selling for more than 4 million, when the asking price was 2 million, would be removed. This variable-specific suppression rule minimizes arm-in-arm transactions where buyers and sellers might have an identity of interest. In total, we exclude ten observations.

Again, we want to summarize the measure of central tendency. Figure 4.2 shows the distribution of the Sale-to-List Price Ratio for the sample. Note that the y-axis is log scaled. In contrast to Days on Market, the distribution is a lot closer to normal, where both the mean and the median are close to the data's central location. Appendix A2 shows the skew for the monthly distributions, which is generally uniform with the overall distribution. Selecting the mean in this context will be beneficial because it makes it easier to make distinctions between months. That is, using the median would yield numerous months a Sale-to-List Price Ratio of 1.00. Conversely, the mean would produce more variation. Hence, we compute the final indicator as the monthly mean Sale-to-List Price Ratio.

Figure 4.2: Histogram of Dwelling's Sale-to-List Price Ratio



Note. The y-axis is log-scaled.

4.2.2.3 Absorption Rate

The Absorption Rate is the third and final indicator. We compute the indicator by dividing the total number of dwellings sold within a month by the sum of all active listings during that month. For example, if 2000 dwellings was listed during a month and the monthly transaction volume was 500 homes sold, the rate is 0,25. That means, 25 percent of the homes listed were sold.

4.2.3 Multivariate Analysis

This subsection delves into the statistical properties of the indicators to assess if it makes sense to compile them into an index. We begin by testing the indicators for correlation using the Pearson Coefficient. Then, we use Cronbach's Alpha to quantify how well they measure the construct of market temperature. Finally, we apply principal component analysis to test whether the indicators are associated with a single latent phenomenon and to which degree they contribute to the overall variation.

4.2.3.1 Correlation

Table 4.5 shows the correlation matrix for the three indicators. The analysis verifies that Days-on-Market has an inverse relationship with market temperature. Moreover, we observe absolute coefficients ranging from 0.53-0.71. A moderate correlation is expected as the indicators intend to measure the same phenomenon. However, there is an imbalance, where Days on Market and the Absorption Rate are more collinear than their relationship with the Sale-to-List Price ratio. The reason is conceivably due to the indicators explaining supply and demand characteristics, while the Sale-to-List Price ratio relates more to price negotiation. Given the higher correlation, we must consider whether we should, at a later stage, give the indicators equal weight or reduce the weight of the pair when we aggregate to a composite.

Table 4.5: Correlation Matrix of the Three Indicators

	Days on Market	**Sale-to-List Price Ratio**
Sale-to-List Price Ratio	-0.53	
Absorption Rate	-0.71	0.53

4.2.3.2 Cronbach Coefficient Alpha

From the correlation analysis, we know there is internal coherence due to the covariances. However, the coefficient alpha, which is the function of the average pairwise correlations across the three indicators, helps quantify the consistency. We compute the Cronbach Coefficient Alpha to be 0.81, which is ideal because it exceeds the threshold of 0.7-0.8 discussed in Subsection 3.1.2. Our alpha implies that the three indicators are internally consistent in explaining the phenomenon of market temperature. Besides, it is further evidence that the indicators measure a unidimensional construct. Nevertheless, to confirm the fact, we have to apply PCA.

4.2.3.3 Principal Component Analysis

Table 4.6 shows the eigenvalues of the three indicators. To ensure an accurate result we normalize the indicators to have a mean of 0 and a standard deviation of 1 prior to computation. The eigenvalues indicate the amount of variation maintained by each component. The result shows that our data is rather homogenous, whereby the first component identifies a single latent dimension that captures almost three-quarters of the variance in all indicators. The second and third component explains 18, and 10 percent, respectively.

Table 4.6: Eigenvalues of the Three Principal Components

Component	Eigenvalue	Variance (%)	Cumulative variance (%)
1	2.18	73%	73%
2	0.53	18%	90%
3	0.29	10%	100%

Table 4.7 shows loadings extracted from the varimax rotation of the two first principal components. Note that we reverse the direction of Days on Market to positively correlate with the other indicators. The first loading vector gives roughly equal weight to all indicators, with the Absorption Rate and Days on Market loaded slightly more than the Sale-to-List Price Ratio. The difference implies that the latter indicator has less correlation with the latent factor. The second loading vector gives the greatest weight to the Sale-to-List Price Ratio. Since the two other indicators are negative, they have an inverse relationship with Sale-to-List Price along the second dimension.

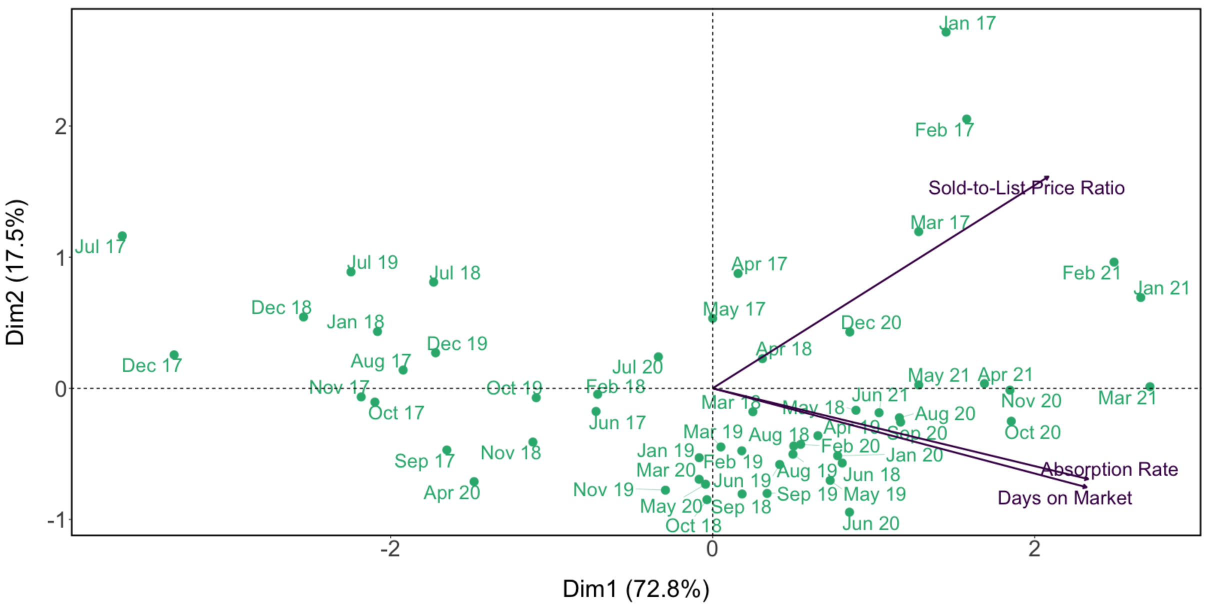
Table 4.7: Loadings on PC1 and PC2

	PC1	PC2
Days on Market	0.60	-0.39
Sale-to-List Price Ratio	0.54	0.84
Absorption Rate	0.60	-0.36

Note. Days on Market is reversed to have a positive correlation with the other indicators

Figure 4.3 plots the first two principal components. Since the first component explains most of the variation in the data, it should generally correspond to a proxy of market temperature. The interpretation of the loadings suggests that months to the left are increasingly cold. Frequent occurrences in this area are winter months and July. The latter half of 2017 and the beginning of 2018 appear to be colder than other years. Conversely, periods to the right side of the plot are hotter, with a relatively extensive presence of months in 2021. Anecdotal evidence seems to coincide with these trends (see Iversen, 2017; Eiendomsverdi, 2019; Eiendom Norge, 2021). Occurrences of high Sale-to-List Price Ratios are present in the upper quadrants, while months in the lower quadrants have relatively greater Absorption Rates and fewer Days on Market.

Figure 4.3: Principal Component Analysis Biplot



To conclude the multivariate analysis, the Pearson correlation and Cronbach's Alpha indicate that the statistical structure is coherent to explain market temperature. The principal component analysis further shows that compiling the indicators will retain considerable variation in the data. Moreover, PCA statistically confirms a single latent dimension in the selected indicators, which justifies summarizing the indicator to an index.

4.2.4 Descriptive Statistics

Table 4.8 shows descriptive statistics for the three indicators. We observe that the monthly median selling time is 12.5 days. The median listing period reaches about 27 days in the coldest month and 9 days in the hottest month. Moreover, a dwelling's sale price is two percent higher than the listing price on average, but in hotter months, selling prices exceed the asking price by up to 9 percent. In the coldest month, the average home sells for just 1 percent below the suggested price. Finally, the Absorption Rate shows that 42 percent of the market clears each month on average. In extreme months, clearance ranges roughly 20 percent from the mean.

Table 4.8: Descriptive Statistics of the Three Indicators

Indicator	Min	Max	Mean	Median	Std.dev	Skew
Days on Market	9	27	13.24	12.5	3.99	1.39
Mean Sale-to-List Price Ratio	0.99	1.09	1.02	1.02	0.02	1.11
Absorption Rate	0.22	0.60	0.42	0.43	0.08	-0.49

Days on Market and the Sale-to-List Price Ratio are positively skewed. The skew tells us that sellers can expect frequent months where homes sell quickly for the asking price. Nevertheless, some periods will require a considerable number of days on the market before dwellings sell. Although sellers can also expect to receive bids way above the asking price in a few instances. Finally, we observe that the Absorption Rate has a relatively more normal distribution.

4.3 Normalization

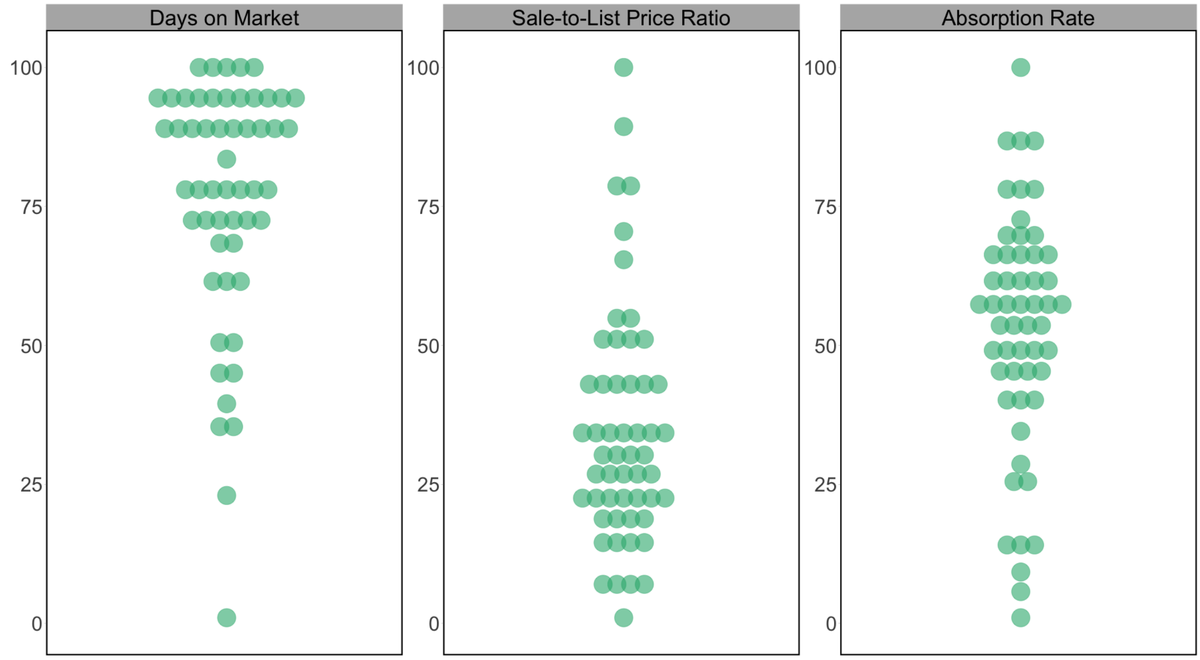
While the multivariate analysis confirms that aggregation is sensible, we first have to make the indicators comparable. Thus, we linearly transform each indicator to the inclusive interval between 1 and 100 using min-max normalization. The minimum is set to 1 rather than 0 because several aggregation methods conflict with 0 or negative values. The following discussion justifies the choice of min-max in favor of z-scores and rank normalization.

The first reason that substantiates min-max aggregation is that we want an index that distinguishes the degree of market temperature. In order to capture how much "hotter" a market is from month to month, it is essential to keep information about absolute differences. For example, if the highest average Sale-to-List Price Ratio is 1.09 and the second highest is 1.06, the ranking method would yield ranks 1 and 2, respectively. Now, if the third and fourth highest ratios are 1.05 and 1.04, with ranks 3 and 4, the comparatively larger difference between 1.09 and 1.06 would be lost. Retaining this information is only possible with linear methods, such as min-max and z-score.

Second, we want an index that is easy to communicate to users. Min-max allows us to transform indicators to a stable range that is easy to disseminate and stable over time. In contrast, using z-scores can result in an indefinite scale that can change over time, and interpretation can be challenging. That is, the scale depends on the indicator's distributions, which will vary for each new observation to the index. Min-max mitigates this problem since the scale will be constant between 1 and 100. Hence, we consider min-max as more applicable than z-scores and rank.

Yet, there is a disadvantage to using min-max normalization. A notable shortcoming is that extreme values can dominate the indicator scales and reduce each indicator's discriminatory power. For example, we can see the effect in Figure 4.4, which shows the normalized indicator distributions. The presence of outliers and skewness in Days on Market and Sale-to-List Price Ratio has a distortion effect on the normalized scores. The indicator scores for Days on Market are considerably higher than the other indicators because the minimum value is an outlier. Conversely, most scores for the Sale-to-List Price Ratio are below 50. We do not consider these outliers as unreliable. However, being aware of the effect is important as it can distort the aggregation process.

Figure 4.4: The Distribution of Normalized Indicators Using Min-Max



The discussion reveals that rank normalization has a significant impact on scores. For example, the two lowest values of Days on Market are roughly 1 and 25, with an absolute distance of 24. In contrast using rank normalization, this distance would be 1. The large discrepancy in distances consequently affects the aggregated composite index.

4.4 Weighting

We apply a weighting scheme based on equality. That is, all indicators weigh one-third in forming the composite index. Statistical approaches such as principal component analysis and correlation neutralization are plausible alternatives. PCA is a viable option because of its statistical properties and objectivity. Due to the strong correlation between Days on Market and Absorption Rate, a neutralization method is also viable. However, these methods produce contrary results, where the main difference is the relative weighting of the Sale-to-List Price Ratio. Table 4.9 shows the difference between the two methods.

Table 4.9: Weights from PCA and Correlation Neutralization

Indicator	**Cor. Neutralization**	**PCA**
Days on Market	29.7%	35.6%
Sale-to-List Price Ratio	41.0%	28.8%
Absorption Rate	29.3%	35.6%

Since the correlation neutralization method seeks to correct for the higher bivariate correlation between Days on Market and the Absorption Rate, the process yields a higher weight for Sale-to-List Price Ratio (41 percent). Conversely, because the Sale-to-List Price Ratio correlates less with the first component in the PCA, it's weight is slightly undermined (28.8 percent). Appendix 3 describes the complete procedure to neutralize the correlation effect, while the PCA weights are extracted from the loadings of the varimax rotation in Subsection 4.2.3. We compute weights according to the proportion of the variance explained by the indicator in the component. For example, the sum of squared loadings equals the total variance explained by the component, which is $0.6^2 + 0.54^2 + 0.6^2 \approx 1,01$. Since Days on Market's loading is 0.6, the weight is subsequently $0.6^2/1.01 \approx 35,6$.

From a statistical point of view, we can justify using the weights elicited from PCA. The reason is that the first component fulfills the requirements brought forward in Section 3.3; the eigenvalue is greater than one and explains more than 60 percent of the total variance. Yet, it might impose double-counting because two indicators express relatively more of a distinct aspect of market temperature. Subsequently, neutralizing the effect is appealing, but the method does not necessarily guarantee each indicator's equal contribution. These drawbacks favor the decision to use equal weights. First, the rather similar loadings suggest that equality is appropriate. Second, inequality expresses specific priorities associated with the concepts the indicators measure, which we have scarce evidence for. Third, equality makes it easier to communicate and interpret the index. Thus, we can justify using equal weights.

4.5 Aggregation

We aggregate the normalized indicators to a composite index using the weighted geometric mean. Arithmetic and harmonic mean were viable options. The following discussion substantiate the decision.

The selection of a aggregation method depends on how much compensability we want in the index. While arithmetic aggregation assumes constant trade-offs between indicators, harmonic aggregation offers minor compensability. The geometric aggregation method is in between the extremes with partial compensability. Arithmetic aggregation is problematic because we have no theoretical basis to justify that a unit can compensate for the loss in one indicator with a gain in another. The developers of *The Human Development Index* (HDI) used the same argument when they substituted arithmetic with geometric aggregation in 2010 (Kovacevic, 2010). On the other side, harmonic aggregation can be considered too extreme in penalizing low values. The in-between solution we elect is thus geometric aggregation.

Table 4.10 illustrates the phenomenon of compensability. The three columns to the right show normalized indicator values (using min-max), while the three to the left show aggregated index scores for each method. Each row shows market conditions in July 2018 and 2020. We observe that July 2018 has a large discrepancy between indicator values. For example, the Sale-to-List Price Ratio is 27, while the Absorption Rate is 1. For July 2018, we see that arithmetic aggregation fully compensates for the low Absorption Rate with a high Sale-to-List Price Ratio. Accordingly, the method yields the highest index score among the three methods. The geometric and harmonic techniques offer diminishing returns for lower indicator values, but the latter is more extreme in decreasing the score. In other words, we observe that reducing compensability penalizes months with large differences between indicator values. Conversely, index values are roughly equal when indicator values are coherent. July 2020 reflects this fact.

Table 4.10: The Impact of Different Aggregation Methods on Index Scores

	Arithmetic Index Score	Geometric Index Score	Harmonic Index Score	Days on Market	Sale-to-List Price Ratio	Absorption Rate
Jul 2018	14	7.23	2.71	14	27	1
Jul 2020	22.17	20.47	19.09	17.5	35	14

4.6 Robustness

After computing a sensible index where each methodological step is carefully considered, a validation step is necessary to assess the quality of the index in terms of its robustness. The following section describes the the methods we use to validate the index. We apply uncertainty and sensitivity analysis to assess bias in our model and ensure transparency. Further, we evaluate the influence of the underlying indicators by analyzing the effect of sequentially removing each indicator from the index. We present results in Section 5.

4.6.0.1 Uncertainty and Sensitivity Analysis

During the composite index construction, we make numerous decisions that introduce uncertainty to the index. In this paper, we limit the types of uncertainty to 3 input factors: (1) Normalization, (2) Aggregation, and (3) Weighting. Table 4.11 shows the alternative methods within each input factor which we consider as plausible alternatives. These alternatives are used as input uncertainties in the Monte Carlo simulation. Appendix A4 summarises alternative weighting schemes and the formulas used for the alternative methods for normalization and aggregation.

Table 4.11: Input Factors in the UA and SA

	Input Parameters		
Normalization	Rank	Min-Max	Z-score
Aggregation	Arithmetic mean	Geometric mean	Harmonic mean
Weighting	Equal Weight	Factor Analysis	Neutralization

We conduct UA and variance-based SA using the Monte Carlo framework presented in Section 3.6. The simulation involves regenerating the index several times, each time randomly sampling one of the three alternatives for each input uncertainty d described in table 4.11. Since the simulation is generally inexpensive in terms of computational power, we generate output for UA and SA simultaneously using a single simulation. We use a base sample size of $N = 2000$ to ensure accurate estimates. Since capturing interaction effects require a high number of replications, we apply the total-order estimator from Jansen (1999) (see section 3.6), which requires a total of $N_T = N(d+2) = 2000(3+2) = 10000$ replications. For each replication, we compute index ranks, $Rank(CI_m)$ and the mean absolute rank change, R_s, relative to our nominal model.

We apply index ranks to conduct UA. Ranks are used because indicator scores can be very inconsistent depending on the methodology, while ranks are much more stable. After completing the simulation, we compute the mean, the median and confidence intervals of each month's rank across all simulations. These statistics are then compared to the ranked index scores from our nominal model to quantify the uncertainty of the model.

We proceed with SA to identify how much output uncertainty is caused by each input uncertainty and to detect interaction effects. Using the mean absolute rank change for all replications, we decompose the total variance into two measures of uncertainty: *first-order sensitivity index* (S_i) and *total-order sensitivity index* (S_{Ti}). The first-order sensitivity index measures the output uncertainty caused by input uncertainty d alone, while the total-order sensitivity index (S_{Ti}) measures the total contribution to the variance caused by input uncertainty d alone and its interaction effects. We use these sensitivity indices to quantify main effects and interaction effects for all input uncertainties.

4.6.0.2 Indicator influence

Finally, we explore the effect of removing an indicator from the composite. This step is necessary to assess whether the indicators are influential in measuring the concept of market temperature. To estimate the influence of each indicator, we sequentially remove one indicator from the model and regenerate the index results. In total, we generate 3 models, each using only two of the indicators at a time. Then, we compare the regenerated models with the nominal model by computing the mean absolute rank change between the regenerated results and the nominal results. To evaluate the influence of each indicator we compare the mean absolute rank change for each regenerated model.

5 Results

This section presents and evaluates the composite indicator. We begin by visualizing the index and indicator scores over time. Then, we describe and quantify common market conditions of hot and cold periods by assessing the underlying indicators. Finally, we evaluate the robustness of the index through uncertainty and sensitivity analysis. Before describing the results in detail, we highlight significant findings as follows.

In hot months, the Absorption Rate is 50 percent, The Sale-to-List Price Ratio is 1.05, and Days on Market is roughly ten days. Conversely, the Absorption Rate is 30 percent in cold months, sale and list prices are equal, and Days on Market reach about three weeks. We find 2021 to be the hottest period in the sample, closely followed by the beginning of 2017. Yet, the latter half of 2017 was the coldest. Decomposing the index to its indicators, we find that extreme Sale-to-List Price Ratios caused the initial heat of 2017, while slow sales and low sale-to-list prices induced the cooling. Moreover, we uncover a strong seasonal pattern in which July and December are consistently colder than average while the spring is warmer. The Absorption Rate appears to be the main driver for the seasonality.

The uncertainty analysis shows that most months have a rank close to the median value of a distribution that concedes uncertainty in aggregation, weighting, and normalization methods. Thus, the nominal model, provides a picture of market temperature that is generally not biased. However, certain months have considerable variance, such as in January 2017, where the rank can shift from the 10th position to the 32nd. The sensitivity analysis reveals that such variation is mainly due to normalization, which is the only significant uncertainty factor. Nevertheless, when accounting for various normalization and aggregation methods, the mean absolute rank change is just 2. Finally, we show that all indicators are influential, but that the Sale-To-List-Price Ratio is most significant for the index.

5.1 The Index

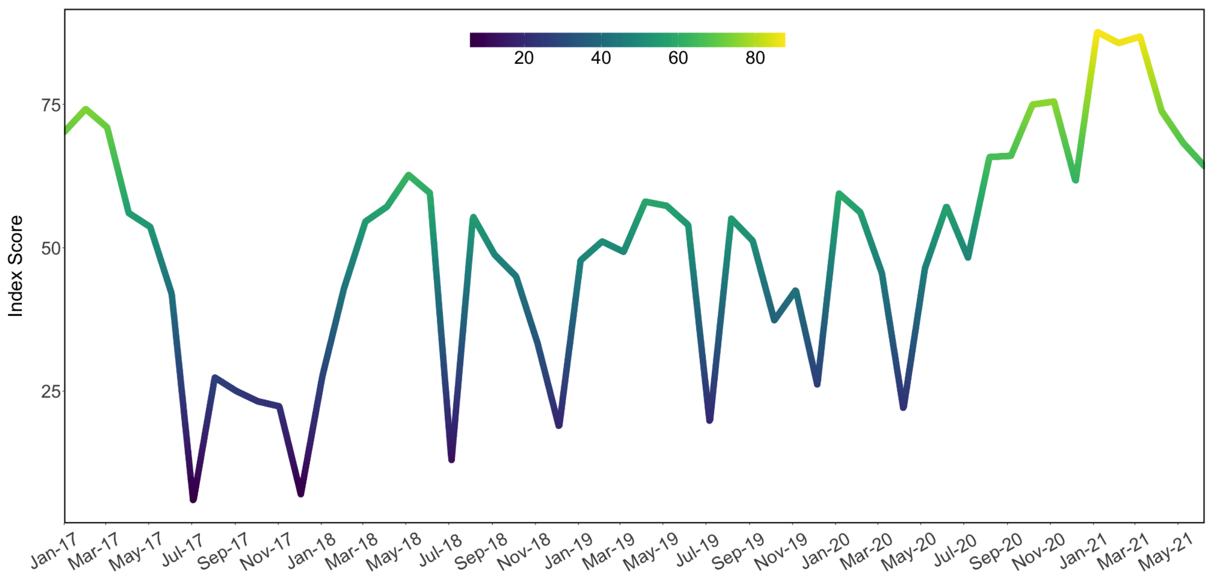
The composite indicator has a hypothetical range from 1 to 100, of which larger values indicate a hotter market. In the period from January 2017 to May 2021, index values range from 5.7 to 87.6, as shown by the summary statistics in Table 5.1. We define months within the interquartile range as neutral markets, while months with scores exceeding the upper quartile as hot and the lower as cold. Appendix A5 contains a complete table of index and indicator scores.

Table 5.1: Summary Statistics of the Composite Indicator

Min	Max	Mean	Median	Std.dev	Q.25	Q.75
5.7	87.6	49.8	53.8	20.1	38.5	62.4

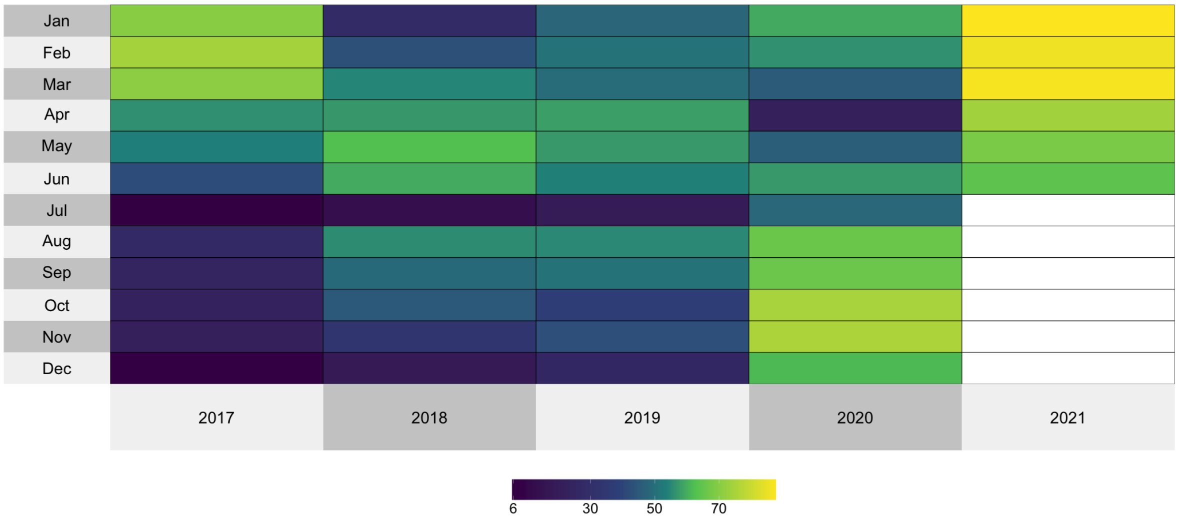
Figure 5.1 shows a line chart of the composite on a monthly basis. We observe an evident seasonal trend and volatile markets at the beginning and end of the sample period. January through March 2017 was particularly hot, but temperatures quickly declined to an all-time low in July. The market remained cold until 2018. In the following two years, the general tendency was a neutral market, where buyers and sellers had roughly equal, albeit interchanging, bargaining power. Then, the trend turned from the latter half of 2020, and the market experienced a significant increase in market heat. October 2020 surpassed February 2017 as the hottest month, and prolonged heat reached an all-time high in January 2021. The following two months were similarly hot before the market began to cool down in April and May.

Figure 5.1: Line Chart of the Composite Indicator Over Time



We proceed by assessing the apparent seasonality. Figure 5.2 illustrates Oslo's market temperature in a heatmap, where brighter colors indicate a higher score. Foremost, we observe the well-documented (Ngai and Tenreyro, 2014; Krogsveen, 2020) seasonal pattern of real estate activity , where July and December are relatively cold. Conversely, the spring is usually hot. However, there are some apparent exceptions; we see that 2020's spring was abnormally cold. Especially evident is the off-pattern temperature in April.

Figure 5.2: Heatmap of the Composite Indicator Over Time



5.2 The Indicators

Plotting the index gives a comprehensive overview of the historical market state. Yet, it is necessary to assess the underlying indicators to describe common market conditions of hot and cold periods. In the following, we explore normalized and raw indicator values.

Figure 5.3 shows a heatmap of the normalized indicator scores. The notable difference in the plot's gradient across indicators is due to outliers. For example, the Sale-to-List-Price dimension is generally dark due to January 2017 being exceptionally high. Similarly, Days on Market is abnormally bright because of July 2017. The distortion effect is caused by the min-max aggregation, as discussed in Section 4.3.

The indicators are rather consistent in highlighting the temperature variations. That is, they generally show the same trends. For example, each indicator shows a cold period between July and December 2017 and heat in the last year of the sample. Nevertheless, there are some indicator-specific differences. For example, extreme fluctuations in the Sale-to-List Price Ratio is the main driver behind the significant market heat and subsequent temperature drop in 2017. Moreover, we notice that the Absorption Rate, accompanied by Days on Market, particularly captures the seasonal trend.

Figure 5.3: Heatmap of the Indicators Composing the Index

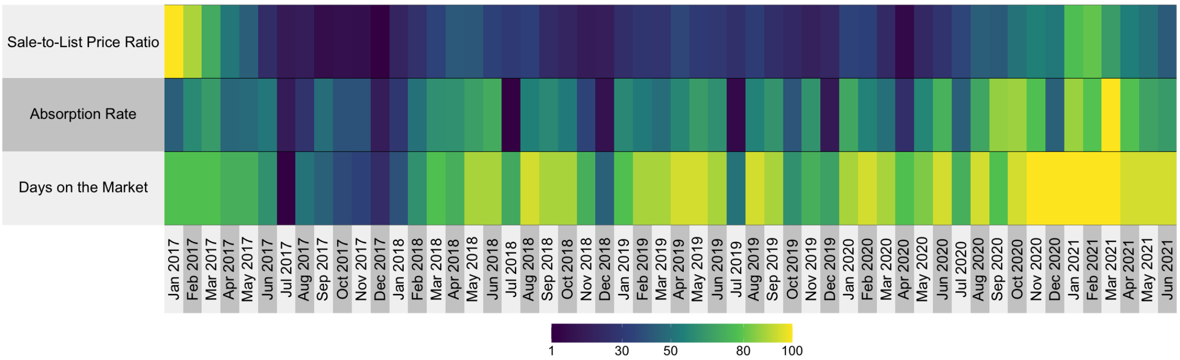
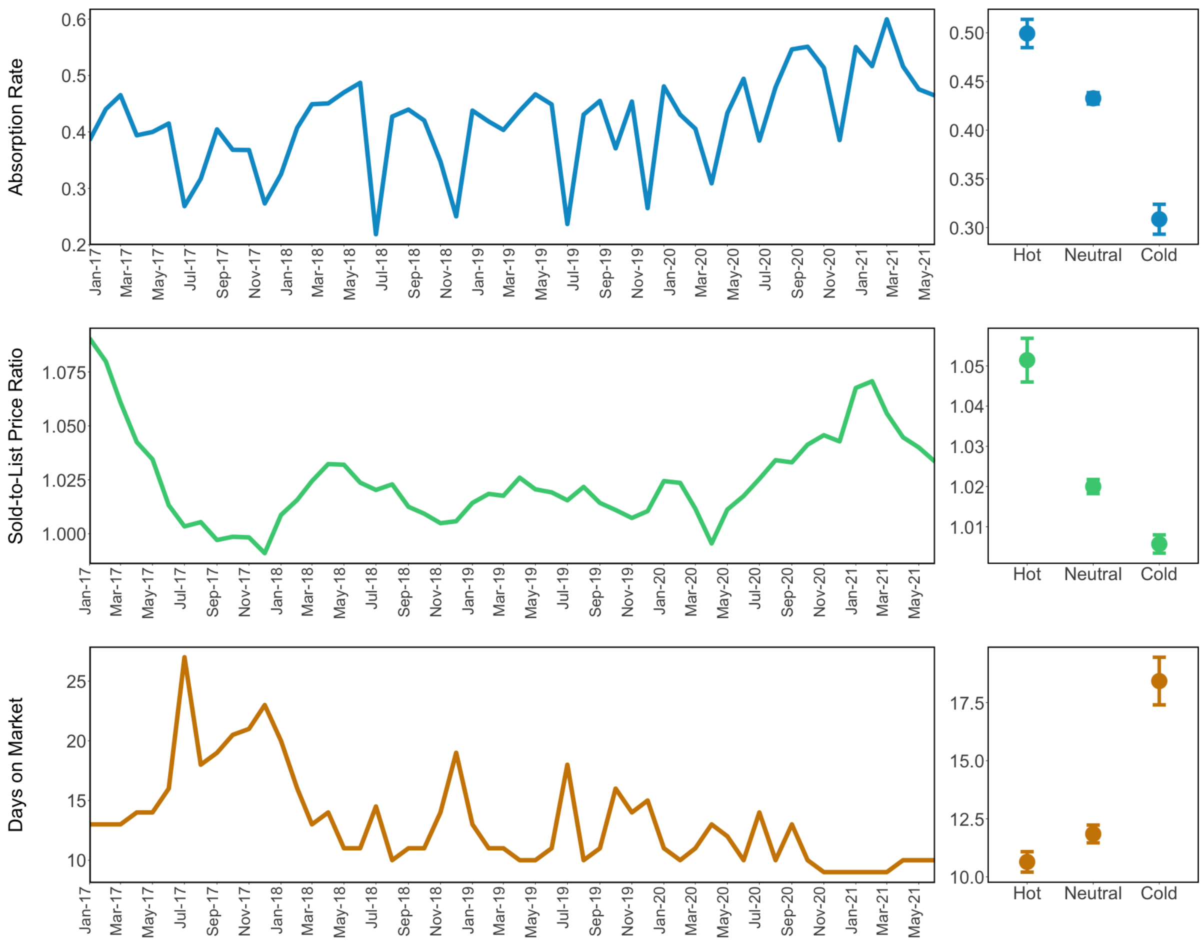


Figure 5.4 describes the magnitude between hot and cold periods through raw indicator values. The blue line shows actual Absorption Rates over time, the green line shows Sale-to-List-Price Ratios, and the red line shows Days on Market. The sub-plots on the right side summarizes the average indicator values by hot, neutral, and cold markets and their respective standard errors.

From the right side plot, we discern that the Absorption Rate is on average roughly 30 percent in cold months and 50 percent in hot months. The blue line plot indicates a surprisingly neutral market at the beginning of the sample. The finding contrasts both the Sale-to-List Price Ratio and Days on Market, which perform exceptionally well in the first months of 2017. However, the Absorption Rate's trend coincides with the other indicators from 2020, when activity grew significantly, reaching an all-time high of 60 percent in March 2021. That is, almost two-thirds of the housing supply was purchased in a single month.

Figure 5.4: Line Plots of Raw Indicator Values



Continuing with the Sale-to-List Price Ratio, we see that sale prices exceed listing prices by roughly 5 percent in hot months, while they are equal in cold months. The green line plot indicates significant volatility in prices at the onset of the sample. Homes sold from January to March 2017 exceeded asking prices by roughly 9, 8, and 6 percent, respectively. The following three years were colder, with prices diverging between -1 to 3 percent from the asking price. However, from May 2020, offers increased significantly, reaching a high of 7 percent above asking in February 2021.

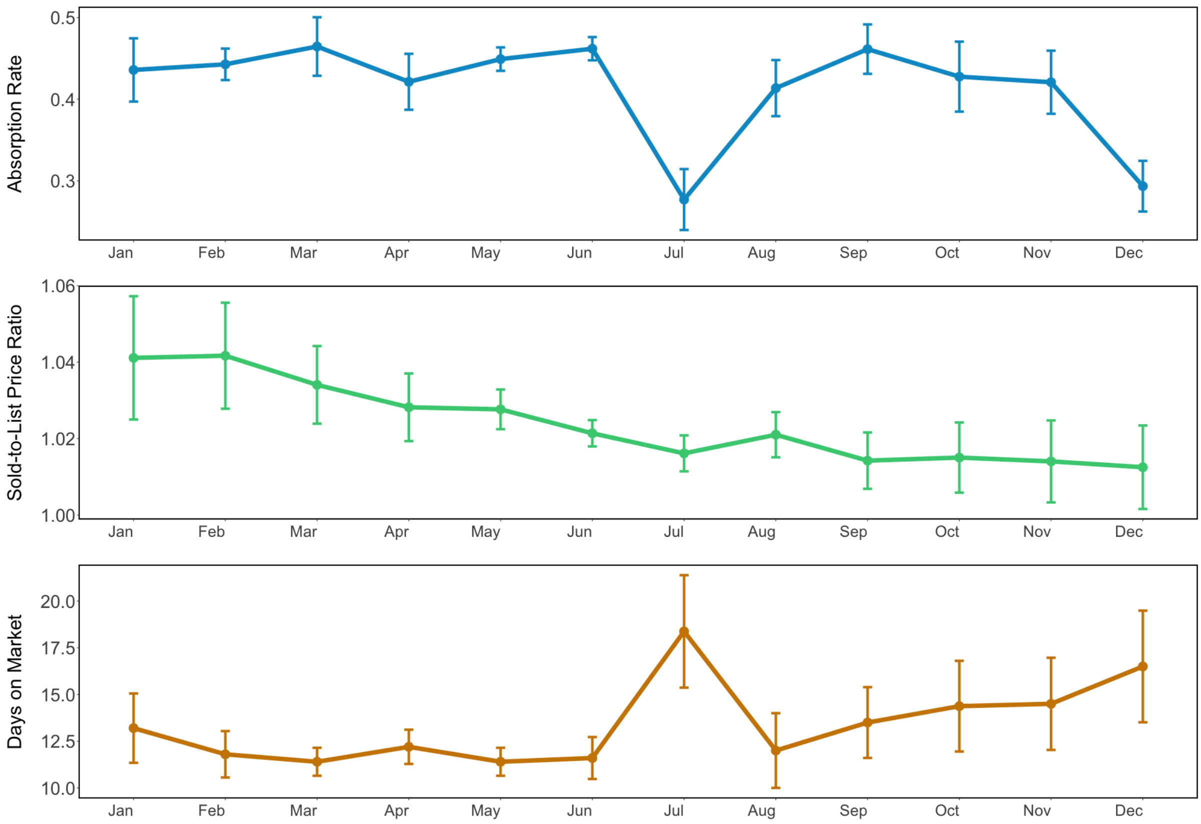
Lastly, Days on Market is, on average, roughly ten days in hot periods and three weeks in cold periods. However, neutral markets are close to hot, requiring just 12 listing days before an offer is accepted. From the brown line plot, we better understand why the latter half of 2017 was especially cold; it took two to three times as many days to sell a home compared to hotter months. In July 2017, the median listing days peaked at 27 days, while January through March 2021 required only nine days to sell.

Finally, we assess the seasonal pattern. Figure 5.5 shows each indicator's monthly mean and standard errors. Starting with the Absorption Rate, it is clear that demand relative to supply is roughly equal throughout the year, with two exceptions. In July and December, the average rate drops approximately 20 points.

While July is seemingly worst in terms of Absorption Rates, the Sale-to-List Price Ratio has a distinctive pattern. We observe that price ratios roughly decline month over month throughout the year, dropping from a mean of 1.04 in January to 1.01 in December. The pattern is influenced by the abnormal heat at the beginning of 2017. Yet, we can be reasonably confident that sale prices exceed listings prices by a greater extent during the spring than the last quarter of the year.

We observe a similar, albeit inverse, pattern between Days on Market and the Absorption Rate. Homes sell quickly during the spring, averaging around 11 to 15 days, while they linger between 15 and 22 days in July. As the weather cools, Days on Market increase month over month from roughly 12 days in August to 16 days in December. However, we note substantial standard errors in the third and fourth quarters, implying higher volatility in the period.

Figure 5.5: Monthly Means and Standard Errors for Each Indicator



5.3 Robustness

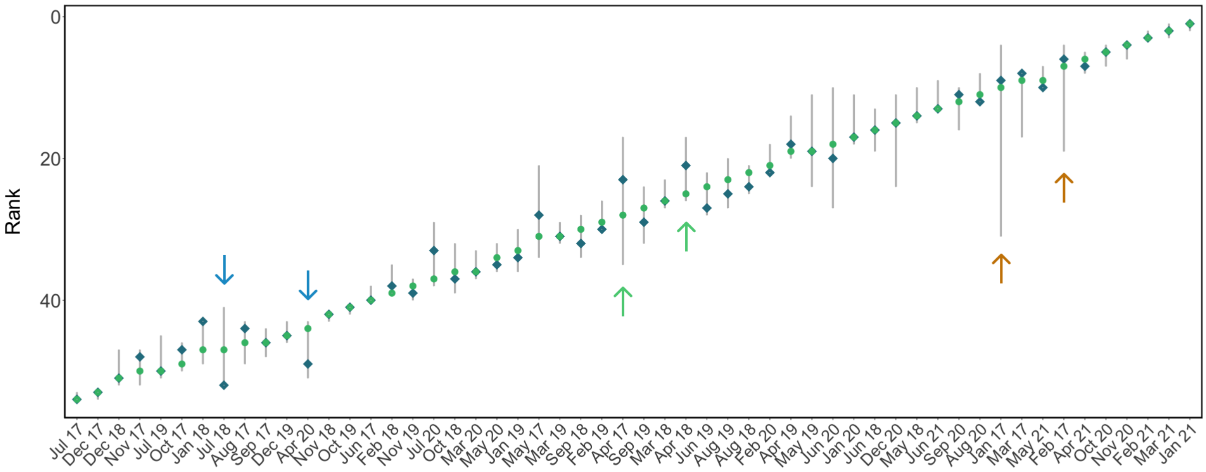
5.3.1 Uncertainty Analysis

Figure 5.6 displays the median rank (circle) and fifth and 95th percentiles (bounds) of the rank across all simulations. Diamonds represent the month's actual ranks. We order the months according to their median rank from all simulations. Consequently, the hottest months across all simulations are to the far right, while the coldest are to the far left.

The analysis shows that most months have a rank close to the median, implying that the nominal model provides a generally unbiased gauge of market temperature. However, we observe instances where actual ranks fluctuate up to 5 ranks from the median. For example, we see that the nominal model consider April 2017 and April 2018 (green arrows) hotter than the median. Accounting for uncertainties in the weighting scheme, normalization, and aggregation method, April 2017 and 2018's performance falls five ranks. On the other hand, we see that the nominal model consider July 2018 and April 2020 (blue arrows) colder than their respective median rank.

The uncertainty analysis provides confidence to the assessment of 2021 having the hottest months in the sample. Similarly, July and December 2017 are indeed the coldest. However, January and February 2017 (brown arrows) are among the hottest months in the nominal model, but they have a considerable variance when accounting for changes in input parameters. The extent of their percentile bounds shows that certain combinations of input parameters severely penalize their performance.

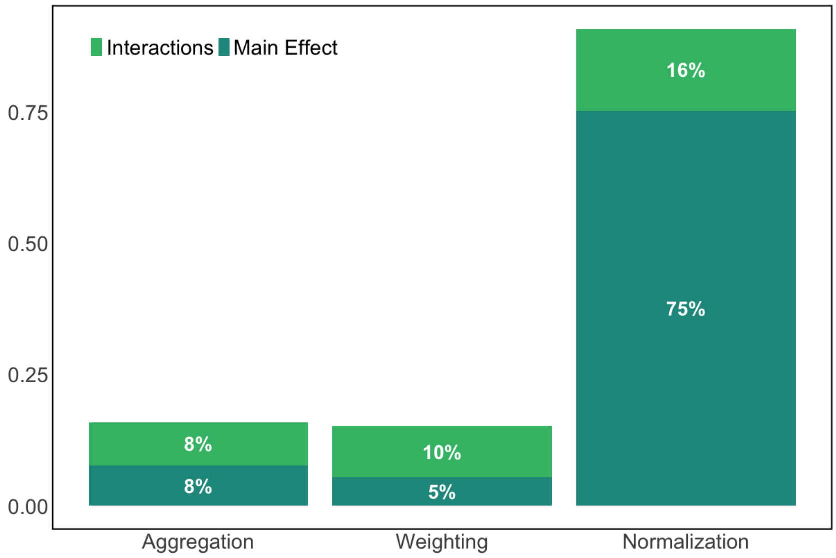
Figure 5.6: Uncertainty Analysis of Index Rankings



5.3.2 Sensitivity Analysis

Figure 5.7 shows the sensitivity indices of the first order, S_i (main effect), and the interaction effect ($S_{Ti} - S_i$), for the average shift in month's ranks (R_s) with respect to the nominal model. The total height of each bar represents the total uncertainty S_{Ti}, caused by each input uncertainty on its own and its interactions. Dark and light green bars shows the main effects and interaction effects of each input uncertainty, respectively.

Figure 5.7: Sensitivity Analysis of Index Rankings



We find that the input uncertainties individually explain $8 + 5 + 75 = 88$ percent of the total variance. The trigger for the normalization method is by far the most critical input uncertainty, explaining 75 percent of the output variance. Aggregation methods and weighting schemes are insignificant, individually explaining 8 and 5 percent, respectively. The remaining 12 percent of the output variance that we can not explain by the main effect is the interaction between the input uncertainties themselves. Looking at the interactions, we see that normalization has the most potent interaction effect with the other factors. For weighting, the extent of interaction is even greater than the main effect, and for aggregation, the interaction effect and main effect is equal. Note that the sum of the total effect equals a number greater than one due to the existing interactions among the factors (OECD, 2008).

The uncertainty analysis revealed that January and February 2017 have wide percentile bounds, which warrant further analysis. Table 5.2 shows the shift in January 2017's rank relative to the nominal model by different normalization and aggregation methods with weights fixed equal to the nominal model. We observe that arithmetic aggregation, in combination with min-max or z-score normalization, favors the month. Z-scores offer a slight disservice when combined with the less compensatable aggregation methods. In general, rank normalization is especially penalizing. Rank normalization can shift January 2017's position down by 15-22 ranks. A further analysis of the output data proves that ranking the Sale-to-List Price Ratio leads to a deterioration in the overall rank. The reason is that the mean Sale-to-List Price Ratio for January is 1.09, an evident outlier in the dataset, which z-score and min-max normalization generally reward.

Table 5.2: The Shift in January 2017's Ranks by Normalization and Aggregation Methods

	Rank	Min-Max	Z-score
Arithmetic mean	-15	1	1
Geometric mean	-19	*	-1
Harmonic mean	-22	0	-4

* = The Nominal Model

We generalize the preceding analysis to show the mean absolute shift in all month's ranks by normalization and aggregation with weights fixed equal to the nominal model. The mean shift is roughly two ranks with a standard deviation of 0.8. From Table 5.3, we observe that rank normalization has the most significant effect. The method shifts the month's rank by roughly three places across all aggregation methods.

Table 5.3: The Shift in all Month's Ranks by Normalization and Aggregation methods

	Rank	Min-Max	Z-score
Arithmetic mean	2.9	1.6	1.6
Geometric mean	2.9	*	1.5
Harmonic mean	3.2	1.4	1.4

* = The Nominal Model

Finally, we explore the effect of removing a single indicator from the composite. Table 5.4 shows that removing the Sale-to-List Price Ratio would yield the most significant mean absolute rank change of over five. Subsequently, it has the most impact on the composite indicator. The figure for Days on Market and Absorption Rate is notably less, at roughly three and two ranks, respectively.

Table 5.4: Mean Absolute Rank Change by Removing a Single Indicator

Indicator Removed	**Impact**
Days on Market	2.15
Sale-to-List Price Ratio	5.3
Absorption Rate	3.19

6 Discussion

As we stated initially, the primary purpose of this paper is to develop a robust measure of market temperature to improve the informational efficiency in the Norwegian real estate market. The results reveal that our methodology yields a robust index. However, assessing the implication of the index in view of supporting decision-makers and market participants remains. The first logical step is to examine the composite indicator's explanatory power to consider whether it measures the phenomenon it intends to. Thus, we begin this section by correlating the index with two related concepts; home appreciation rates and news-based sentiments of market performance. Then, we offer an in-depth discussion on how the index can improve informational efficiency. Finally, we present limitations and topics of further research.

6.1 Explanatory power

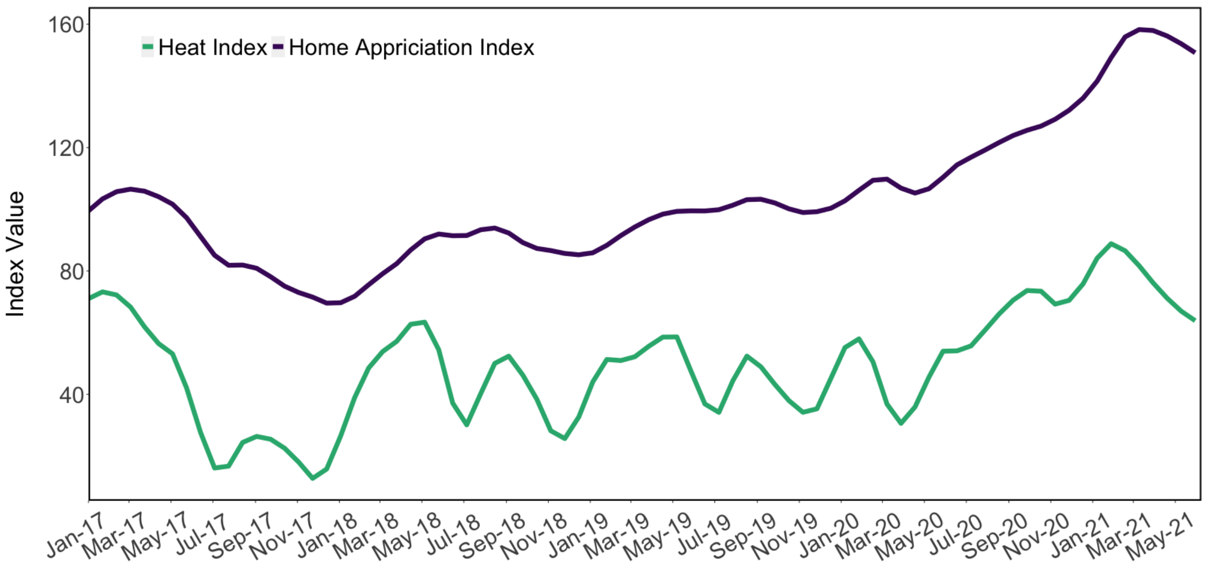
To assess the explanatory power of the composite indicator, we can link our index with related concepts or indices. Comparing contemporary indices from Zillow, Redfin, and Realtor would be interesting. However, their exact methodological procedures and data are unavailable. Another option is to correlate our index with measures that should have relevancy. For example, Carrillo (2013) evaluates his index on sellers' bargaining power by assessing how well it conforms with home appreciation rates and popular perceptions about market "heat".

Although home appreciation is not suitable to depict the relative bargaining power between market participants alone, we assume bargaining power to have a relationship with prices. That is, we expect hot months in our composite to conform with higher prices. To test a facet of the explanatory power of the index, we can thus correlate our index with the *Oslo House Price Index* disseminated by Real Estate Norway (2021). Another approach is to explore how well our index conforms with how the media describes the market. We develop a proxy of news sentiment by measuring the number of articles written about a "hot housing market" over time in Oslo. Even though we expect the correlation analysis to reveal positive relationships with both measures, it is critical to note that it does not imply causality. That is, an increase in our index does not lead to a rise in the number of articles written about the housing market or prices, and vice versa.

6.1.1 Comparison with a Home Appreciation Index

Figure 6.1.1 compares our index with the Oslo House Price Index. The green line represents our index, while the purple line illustrates the price development. We observe that our index generally coincides with the ups and downturns of the home appreciation rates. Both home prices and market heat fell throughout 2017 before the increase in 2018. In addition, both indices experienced an abnormal drop in April and May 2020 and high temperatures in 2021. Correlation analysis reveals that the indices have a bivariate correlation of 0.71. However, our index especially indicates varying market conditions from 2018 to 2020. We assume that some of the differences arise from the seasonal adjustment in the price index. Nevertheless, there is an inverse trend between the indices in the period. The discrepancy highlights that our index explains a broader range of market conditions than just price developments.

Figure 6.1: Comparison with a Home Appreciation Index

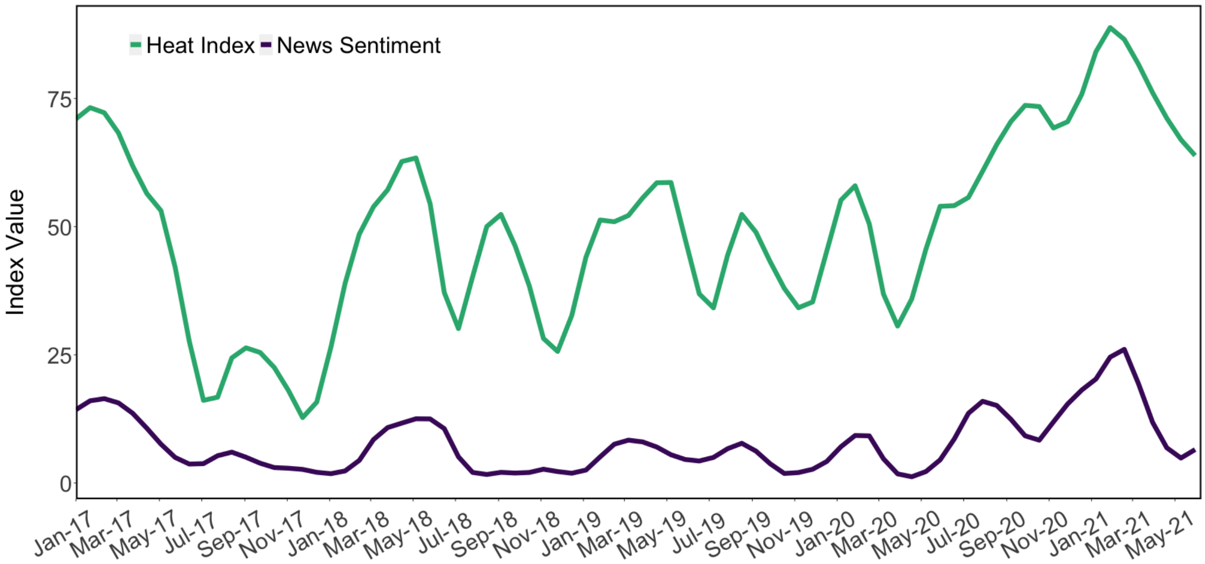


6.1.2 Comparison with a News-Based Sentiment Analysis

Figure 6.1.2 illustrates how well our index conforms with a news-based sentiment analysis. The green line illustrates the index scores, while the purple line shows the news sentiment.

The sentiment's index value represents the number of articles written about a "hot housing market in Oslo". Appendix A6 shows how we derive the result. We observe that our index generally coincides with aligning peaks and lows, which is confirmed by the bivariate correlation of 0.65. The sentiment indicates that the beginning of 2017 was hot; then, the market was balanced until the third quarter of 2020. Further, the sentiment shows that the market was heating up towards an all-time high at the beginning of 2021, which fits with the hottest months in our index. Besides, both indices exhibit a subsequent drop from March to May 2021.

Figure 6.2: Comparison with a News-Based Sentiment Analysis



To conclude, the moderate to strong relationships between our index and the related concepts of price and sentiments are further evidence that our index measures the phenomenon it intends to.

6.2 Implications

With evidence of a robust index that indeed measures market temperature, we can proceed with the discussion of how our index can improve informational efficiency. The general value of the index is in its ability to summarize diverse market conditions in a more efficient manner than is possible with a collection of separate indicators. The index has numerous expected applications, depending on the party of interest. However, the prevailing benefit is that of decision support. To exemplify how the index can guide decision-makers, we will in the following discuss the implication of market participants' bargaining power over time, across regulatory and economic shocks, between geographical areas, and among dwelling sizes. The intention, however, is not to assess causal associations but rather to comment on observed circumstances.

6.2.1 Seasonal Temperature

In this subsection, we explore how buyers and sellers can use the apparent seasonality of the index as decision support. The variation in bargaining power across months provides insight for market participants, such as when they are more likely to negotiate a better deal.

Our index illustrates that July and December are colder than average. The main reason is less activity, as described by the Absorption Rate. For example, the rate drops from an average of roughly 45 to 30 percent in July, implying fewer transactions of the available supply. Fewer transactions indicate less demand which reduces the bargaining power of sellers. It is conceivable that the seasonality manifests itself in fewer open house attendees, fewer bids, and thus less competition. Accordingly, buyers are better positioned for negotiation, which can constitute lower prices, practical reservations, or other non-monetary benefits.

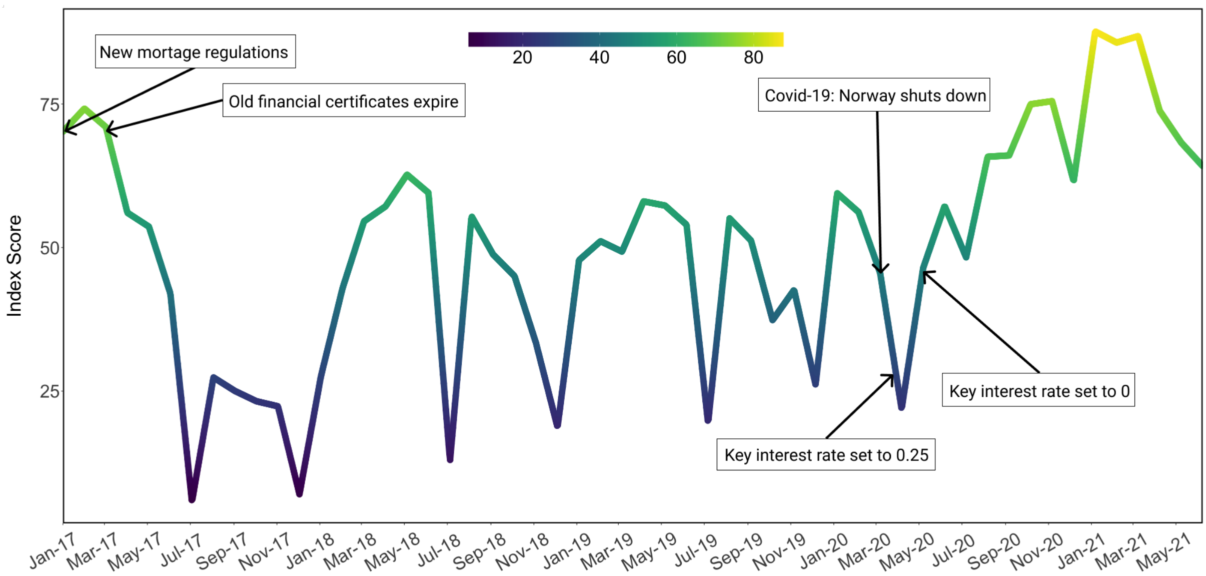
The Sale-to-List Price Ratio supports the notion. For example, prices range from roughly 2-4 percent above asking in January, while in December, the range is between 0-2 percent. Moreover, December's decrease in demand correlates with more days listed. Sellers can expect to wait two to three weeks before accepting an offer between July and December, but only one to two weeks in the first and second quarters. Hence, if buyers' utility depends on efficient transactions, they should avoid listing their property in the latter part of the year.

Knowing the temperature of the market can thus improve informational efficiency. For example, if buyers know their current bargaining power and its impact, they might better predict which homes they can afford. Subsequently, they can evade spending time prospecting homes that are likely to be out of reach because bidding wars render them too expensive. Moreover, it can aid buyers with deciding a bid strategy, including whether they should consider putting in or waiving reservations. Similarly, sellers can be more efficient in assessing bids and the likelihood of receiving more appealing offers. Besides, having an indication of how long it will take to sell the property can support planning and financing activities.

6.2.2 Temperature Through Shocks

Our index can provide valuable insight into the association between shocks and variations in bargaining power. During the sample period, three significant events impacted the housing market. First, In January 2017, the Norwegian authorities imposed new regulations for how banks should assess mortgage applications, effectively reducing financing for most people (Regjeringen, 2016). Second, COVID-19 provoked strong infection control measures from March 2021, resulting in a country-wide shutdown (Lund, 2021). Third, Norges Bank cut the key interest rate from 1.5 percent to 0 in just a few months after the pandemic began (Lindquist, 2021). Figure 6.3 illustrates the major events related to mortgage regulations and the pandemic. In the following, we interpret how these events influenced the market.

Figure 6.3: Temperature Through Shocks



6.2.2.1 Covid-19

On March 12th, the Corona pandemic struck Norway, and the government announced a national lockdown (Lund, 2021). Restrictions had an apparent impact on the housing market. Although the spring is usually hotter than average (Krogsveen, 2020), 2020 was quite the contrary; March, April, and May were the coldest months of the year. The main drivers were less activity, as depicted by the Absorption Rate and a sharp drop in the Sale-to-List Price Ratios. Corona had the most significant impact in April, where the Sale-to-List Price Ratio and Absorption Rates fell to the month's all-time-lows of 0.99 and 30 percent, respectively.

In an effort to minimize the economic impact of the coronavirus, Norges Bank cut the key interest rate from 1.5 to 0 percent in May (Norges Bank, 2020). As a result, the average mortgage rate fell from 3.0 to 1.9 percent in the fourth quarter of 2020 (Norges Bank, 2020). The lower rate made it cheaper to service mortgage debt while reducing the return on alternative investments. Thus, it became more attractive to invest in real estate, which our index indicates through the rapid increase in market heat.

6.2.2.2 Mortgage Regulations

We can mainly attribute the significant market heat and subsequent temperature drop in 2017 to a specific incident. After years of rapid growth in Oslo and increasing debt, the Norwegian authorities imposed new regulations to create a more sustainable housing market and reduce debt ratios. From January 2017, the new regulations required that:

- banks could not grant mortgages if customers' total debt exceeded five times gross annual income,
- the debt-to-asset ratio could not exceed 85 percent of the home value,
- the borrower's debt service capacity must endure an interest rate increase of five percentage points, and
- forty percent equity requirements for secondary homes in Oslo (Regjeringen, 2016).

Although the new regulations applied from January 1st, financing certificates last for three months, implying that many buyers had until March 2017 to enter the market before the new rules would apply. Subsequently, the beginning of 2017 experienced abnormal buying pressure as actors rushed into the housing market before expiry (Iversen, 2017). After expiration, the index value dropped sharply. Marked conditions remained in favor of buyers until 2018, when the market stabilized.

Our findings indicate that the regulations calmed the market for a certain period. However, we see significant volatility in the Sale-to-List Price Ratio. Our results describe a year-over-year difference of roughly eight percentage points between January 2017 and 2018. Surprisingly, buying activity remained strong, with comparable Absorption Rates in both years. These findings suggest that regulations helped restrain home price appreciation for a period without significantly influencing market activity. Nevertheless, the recent market heat might indicate that the financing constraints are inconsequential in ensuring a sustainable housing market in the long term.

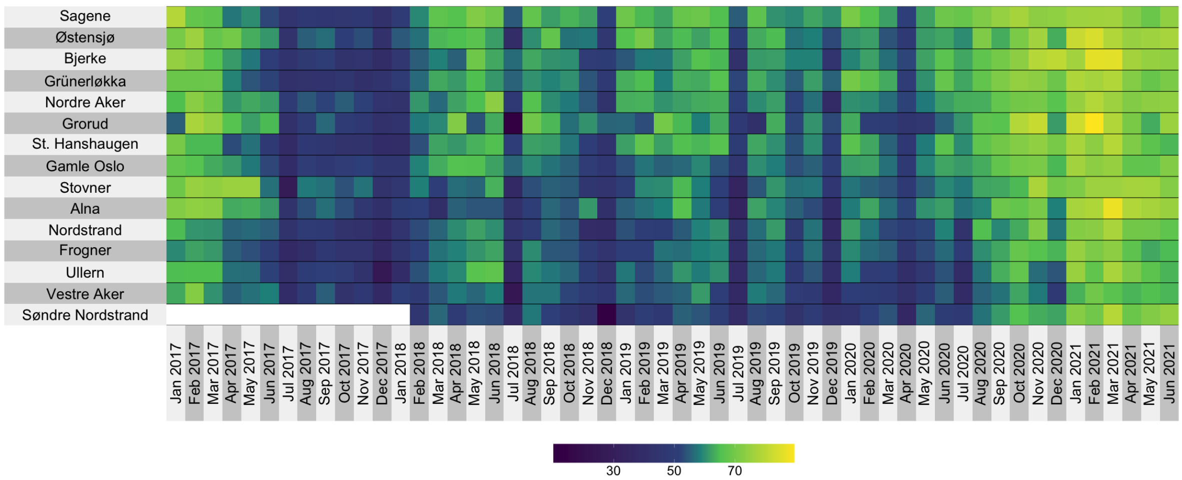
Although the index itself is not suited to support causal claims about policy significance, it allows for assessing whether trends are in accordance with intended impacts. The index provides comprehensive snapshots of variations in bargaining power which can be helpful in communication with the public or administration. For example, the snapshots can describe the extent of unfair conditions between the supply and demand side, which can inform strategic policy objectives and initiatives. Besides, the methodology can be applied to estimate bargaining power in cities or boroughs to support benchmarking and assessing regional differences. The following section presents an example of this use case.

6.2.3 Performance Across Geographical Areas

The variation in market heat across geographical areas can be valuable for decision-makers and market participants. By comparing the difference in competitiveness, buyers can, for example, better determine which city or neighborhood they are more likely to place a successful bid. This subsection exemplifies how we extend our methodology to benchmark Oslo's boroughs. Figure 6.4 illustrates a heatmap of a borough-level index, where we arrange boroughs in descending order according to their average index score. In general, the boroughs follow the same pattern as Oslo. The first half of 2017 is cold, and 2021 is hot. In addition, the same seasonal patterns exist, where July and December are cold. However, the relative market heat between boroughs differs significantly, with an average index score ranging from 45 to 62.

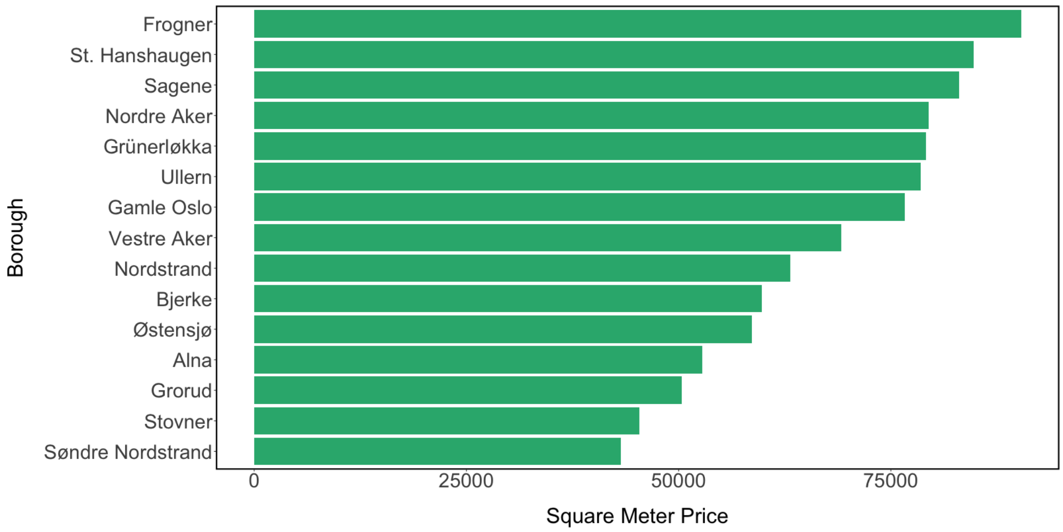
We find Sagene, Østensjø and Bjerke to be the three hottest boroughs, while Ullern, Vestre Aker and Søndre Nordstrand the coldest. A notable drop in temperature in 2017, conceivably driven by new regulations, appears to have a more considerable impact on the latter boroughs. The circumstance is apparent from the lengthy period of cold months from April 2017 to August 2020. Conversely, The former boroughs quickly recovered from 2017's events and are consistently hot, albeit with seasonal downswings, the remaining period.

Figure 6.4: Heatmap of Oslo's Boroughs



Interestingly, we find notable differences by comparing the borough's square meter prices with their temperature, as shown in Figure 6.5. We find that expensive areas such as Frogner do not necessarily imply high competitiveness. Despite being the most desirable area, our index reveals that it is one of the colder areas in the city. The interpretation is that although prices are high, sellers should expect slower sales and sale prices closer to asking prices compared to other areas. Conversely, Østensjø is considered hot despite being cheap. The index thus suggests that buyers can expect less bargaining power when negotiating a deal in that area.

Figure 6.5: Square Meter Price of Oslo's Boroughs



However, it is important to note that the temperature for a borough is not necessarily representative of all dwellings in that area. Housing is a broad term because it implies a unit of accommodation that varies significantly in size, standard, and form. Yet, housing statistics rarely differentiate between such attributes. Thus, we examine performance patterns across small, medium, and large dwellings in the next section.

6.2.4 Performance Among Dwelling Sizes

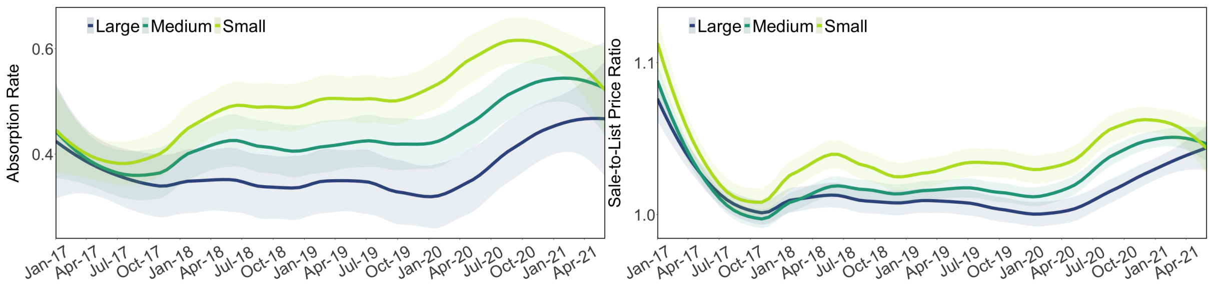
Eiendomsverdi (2018) proposes that small dwellings experience the most immediate impact of changing housing trends. They argue that market participants who sell small apartments usually aim to "size up" to larger homes. Hence, when the market experiences a shock, such as significant interest rate cuts, activity initially concerns smaller dwellings. When the activity of smaller dwellings settles down, the demand for medium and larger-sized homes will accordingly increase. In this way, the behavior manifests a lag where small homes experience changing conditions first and large homes last.

According to the premise, smaller homes should break trends before larger homes. Figure 6.6 seems to support the notion. The left-side plot illustrates the Absorption Rate of small, medium, and large dwellings over time. We consider homes small if less than 45 square meters and large greater than 90. Anything in between is medium sized. The right-side plot shows the Sale-to-List Price Ratio for equal sizes. To aid with seeing patterns, we use the smoothed conditional mean. From the left-side plot, we discern that when the market began to heat up in 2018, the Absorption Rate of small and medium homes quickly increased, while larger homes' ratios remained low. Equivalently, the interest rate adjustments in 2020 had a more immediate impact on smaller homes.

As the market began to cool down at the beginning of 2021, the Absorption Rate of small dwellings has declined first. The phenomenon is especially evident as we see the trendline of small homes crossing that of medium dwellings at the end of the sample. The right-side plot illustrates the same phenomenon. While the Sale-to-List Price Ratio decreases for small and medium-sized homes, the growth rate of large houses is still positive.

Thus, in the event of impactful policy changes or economic shocks, a sudden change in temperature does not manifest itself equally for all dwelling sizes. Therefore, different strategies could be applied depending on if you own a small or big home. Owners of small apartments should perhaps be quick to capitalize on a positive trend-change, while owners of larger homes could benefit from waiting a couple of months before selling.

Figure 6.6: Absorption Rates and Sale-to-List Price Ratios by Dwelling Size



6.3 Limitations and Further Research

This section will discuss three central issues in constructing an index that measures market temperature, including the selection of underlying indicators, normalization methods, and the data sample.

First, due to limited accessibility, we could not include Waived Contingencies, Listing views, and Competing Bids in the composite. The limitation might be problematic in the sense that our index can be susceptible to an "omitted variable bias" (Hardeman et al., 2013). The bias implies that variables not included in the index might drive the results.

That is, their inclusion could significantly alter the findings. The lack of data makes it hard to reject or affirm the bias. However, we argue that our composite is unlikely to suffer significantly. The reason is that the omitted indicators capture the same information as those included in the composite. For example, competing bids likely correlate strongly with Days on Market. Many bids yield quick sales, and few bids extend days on the market.

The second limitation relates to the methodological step of normalizing indicators. Sensitivity analysis revealed normalization to be the only significant uncertainty factor in the model. The rationale of min-max scaling was two-fold. First, it retains information on absolute differences; second, it transforms indicators to a static range that is easy to disseminate. However, the trade-off is a distortion effect. For instance, an extreme Sale-to-List Price Ratio rendered the majority of months to have an indicator score below 50. Although we do not consider the outliers unreliable, we know that outliers limit the overall index scores. We could mitigate the distortion effect by using z-scores or rank normalization. However, applying these normalization techniques impose other implications that we deem more limiting for the overall purpose of the index. For instance, z-score normalization inflicts uneven indicator scales.

The third limitation concerns the homogeneity of the sample period. Our sample consists of transactions between January 2017 and May 2021, which is a period that we generally consider hot. Although there is notable volatility within the underlying indicators, the overall market has been relatively stable since the financial crisis in 2008. Thus, we want to emphasize that although July 2017 is considered the coldest month, it only implies the coldest month in the sample. If we compare July 2017 with market conditions in October 2008, the connotation of cold would no longer apply. For that month, Norwegian housing prices fell by more than 4 percent. Moreover, 40 percent fewer homes were sold than the previous year (Dreyer, 2018). Including the observation in the model would distort all months in the original sample to be considered hot relative to October 2008.

These limitations and the general complexity of constructing a composite indicator suggest that the findings in this paper are not conclusive. Contrarily they open the door for further research. Furthermore, it would be interesting to explore variations in bargaining power between Norway's major cities. Particularly interesting is Stavanger, which has a market especially influenced by the oil industry. Hence a topic of research could be the comparative impact of 2014's oil crisis on bargaining power across Stavanger and Oslo. On that account, we have yet to examine how the methodology holds up when applying the index to cities with considerably less transaction data. Thus such a topic also introduces the opportunity to study measures to maintain robustness for small sample sizes. Finally, it would be interesting to explore if the index can be an influential parameter in a price prediction context.

7 Conclusion

In this paper, we develop a composite indicator that estimates the relative bargaining power between buyers and sellers in Oslo's residential real estate market. We compute the index monthly from January 2017 to May 2021. It constitutes a geometric mean of three min-max scaled indicators: the Absorption Rate, Days on Market, and the Sale-to-List Price Ratio. These metrics denote domains of supply, demand, and price negotiation, which collectively delimit and quantify the phenomenon of market temperature. The composite lies on the inclusive interval of 1 to 100, of which larger values indicate a hotter market.

The prevailing purpose of the index is to improve the informational efficiency in the Norwegian real estate market. The goal necessitates a simple but robust measure. We achieve simplicity through comprehensible indicators and methodological decisions that are straightforward to interpret. Yet, it does not prevail over robustness. By means of a Monte Carlo experiment, we conduct uncertainty analysis and variance-based sensitivity analysis. The uncertainty analysis shows that most months have an index score close to the median value of a distribution that concedes uncertainty in aggregation, weighting, and normalization methods. Thus, the nominal model provides an estimate that is generally not biased. The sensitivity analysis reveals that 75 percent of the total output variance is due to normalization. Subsequently, it is the only significant uncertainty factor. In addition, all indicators are influential for the index score. These evaluations form evidence that our index is a robust estimate.

We quantify average market conditions of hot and cold periods as follows. In hot months, the Absorption Rate is 50 percent, The Sale-to-List Price Ratio is 1.05, and Days on Market is roughly ten days. Conversely, the Absorption Rate is 30 percent in cold months, sale and list prices are equal, and Days on Market reach about three weeks. Moreover, we find 2021 to have the hottest period, reaching an all-time high in January 2021. The latter half of 2017 was the coldest. We uncover a strong seasonal pattern in which July and December are consistently colder than average, while the spring is warmer. The Absorption Rate is the main driver for this seasonality.

To assess the explanatory power of the composite indicator, we correlate the index with a news-based sentiment of market performance and home appreciation rates in Oslo. We find moderate and strong relationships with both indices, implying that our composite measures the phenomenon it intends to. Subsequently, the index has numerous expected applications, depending on the party of interest. To exemplify how the index can assist market participants, we show that understanding contemporary bargaining power provides insights that can guide prospecting, bid, and sales strategies. Then, we illustrate that the composite can help decision-makers monitor the impacts of particular policies, with examples from new mortgage regulations and Covid-19 restrictions. Next, we extend the application of the index to measure market heat in Oslo's boroughs to demonstrate its benchmarking capabilities. Finally, we show that changes in market temperature do not manifest themselves equally for all dwelling sizes.

References

Arrow, K. J., Raynaud, H., et al. (1986). Social choice and multicriterion decision-making. *MIT Press Books*, *1*.

Bandura, R. (2008). A survey of composite indices measuring country performance: 2008 update. *New York: United Nations Development Programme, Office of Development Studies (UNDP/ODS Working Paper)*.

Bandura, R. (2011). Composite indicators and rankings: Inventory 2011. *New York: United Nations Development Programme, Office of Development Studies (UNDP/ODS Working Paper)*.

Becker, L. S. (2019). *Zillow: What is zillow's buyer-seller index, and how is it computed?* Retrieved from https://www.zillow.com/research/understanding-the-zillow-buyer-seller-index-2883/

Becker, W. (2021a). Coinr: Composite indicator construction and analysis [Computer software manual]. Retrieved from https://cran.r-project.org/web/packages/COINr/vignettes/Overview.html (R package version 0.6.1)

Becker, W. (2021b). *Composite indicator development and analysis in r with coinr: Chapter 10 aggregation.* Retrieved from https://bluefoxr.github.io/COINrDoc/aggregation.html

Becker, W. (2021c). *Composite indicator development and analysis in r with coinr: Chapter 14 sensitvity analysis.* Retrieved from https://bluefoxr.github.io/COINrDoc/sensitivity-analysis.html

Becker, W. (2021d). *Composite indicator development and analysis in r with coinr: Chapter 5 multivariate analysis.* Retrieved from https://bluefoxr.github.io/COINrDoc/multivariate-analysis.html

Becker, W. (2021e). *Composite indicator development and analysis in r with coinr: Chapter 9 aggregation.* Retrieved from https://bluefoxr.github.io/COINrDoc/normalisation.html

Becker, W. (2021f). (Personal communication, 14.12.2021).

Boylaud, O., & Nicoletti, G. (2000). Regulation, market structure and performance in telecommunications. , *237*.

Carrillo, P. E. (2013). To sell or not to sell: Measuring the heat of the housing market. *Real estate economics*, *41*(2), 310–346.

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and psychological measurement*, *64*(3), 391–418.

Decancq, K., & Schokkaert, E. (2016). Beyond gdp: Using equivalent incomes to measure well-being in europe. *Social indicators research*, *126*(1), 21–55.

De Muro, P., Mazziotta, M., & Pareto, A. (2011). Composite indices of development and poverty: An application to mdgs. *Social indicators research*, *104*(1), 1–18.

Dreyer, C. (2018). *Finanskrisen på norsk.* Retrieved from https://eiendomnorge.no/blogg/finanskrisen-pa-norsk-article599-923.html

Eie. (2021). *Verdiovervåking.* Retrieved from https://eie.no/eiendom/verdiovervakning

Eiendom Norge. (2021). *Nedgang i boligprisene i november.* Retrieved from https://eiendomnorge.no/nyheter/nedgang-i-boligprisene-i-november-article2178-919.html

Eiendomsverdi. (2018). *The policy rate.* Retrieved from https://eiendomnorge.no/blogg/lardommer-fra-2017-og-perspektiver-for-2018

-article714-923.html?fbclid=IwAR2d8E6dJT6kVLJiMp3UHfPuQgifW4zBm_WFM9Td6n9FVSCmHEaBj7EG7N0

Eiendomsverdi. (2019). *Lærdom fra boligmarkedet i 2018.* Retrieved from https://eiendomnorge.no/blogg/lardom-fra-boligmarkedet-i-2018-article462-923.html

European Commission. (n.d.-a). *Composite indicators. step 1: Theoretical framework.* Retrieved from https://knowledge4policy.ec.europa.eu/composite-indicators/10-step-guide/step-1-theoretical-framework_en

European Commission. (n.d.-b). *Composite indicators. step 5: Normalisation.* Retrieved from https://knowledge4policy.ec.europa.eu/composite-indicators/10-step-guide/step-5-normalisation_en

European Commission. (n.d.-c). *Composite indicators. step 6: Weighting.* Retrieved from https://knowledge4policy.ec.europa.eu/composite-indicators/10-step-guide/step-6-weighting_en

Freudenberg, M. (2003). Composite indicators of country performance: a critical assessment.

Glen, S. (2021). *Correlation coefficient: Simple definition, formula, easy steps.*

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social indicators research*, *141* (1), 61–94.

Hardeman, S., Van Roy, V., Vertesy, D., & Saisana, M. (2013). An analysis of national research systems (i): A composite indicator for scientific and technological research excellence. *JRC scientific and policy reports-EUR 26093 EN*.

Herath, S., & Maier, G. (2015). Informational efficiency of the real estate market: A meta-analysis.

Iversen, K. (2017). *Hvorfor falt boligprisene i 2017?* Retrieved from https://dnbeiendom.no/altombolig/kjop-og-salg/boligpriser1/2017/hvorfor-falt-boligprisene-i-2017?fbclid=IwAR0S8NpjygEHMHp0RkXKZKsDIsQVjJ6oFXhi5LTo5GuFuDC8ZRTmBZw7pzk

Jansen, M. J. (1999). Analysis of variance designs for model output. *Computer Physics Communications*, *117* (1-2), 35–43.

Kovacevic, M. (2010). Measurement of inequality in human development–a review. *Measurement*, *35*, 1–65.

Krogsveen. (2020). *Når på året er det best å kjøpe bolig?* Retrieved from https://www.krogsveen.no/magasin/nar-pa-aret-er-det-best-a-kjope-bolig

Lindquist, M. S. V. B. W. E., K. (2021). Boligmarkedet i pandemiåret. *Staff memo*, *6*. Retrieved from https://www.norges-bank.no/contentassets/2d89852082084e2a8bd27f5d880b5fcb/staff_memo_6_21---boligmarkedet-i-pandemiaret-2020.pdf?v=06%2F17%2F2021082509&ft=.pdf&fbclid=IwAR0FVYxTfP4Pa1b61qdV6Bd5owGeObaa7JbTEnygmKvYCpA6LesFKRI-ZLI

Lund, O. (2021). Snart ett år siden norge stengte ned: – vi visste det ville ta tid, men snakket ikke så høyt om det. Retrieved from https://www.tv2.no/a/13140801/

Maggino, F., & Ruviglioni, E. (2009). Obtaining weights: from objective to subjective approaches in view of more participative methods in the construction of composite indicators. *Proceedings NTTS: new techniques and technologies for statistics*, 37–46.

Medby, P., & Takle, M. (2021). *Modell for beregning av boligformue. oppdatert med tall for 2020.* Retrieved from https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/modell-for-beregning-av-boligformue--448122?fbclid=

IwAR3uOocAy2tURSZskZrtOn1ujZbrj5wikw3xWflltQLYzpizOmXcW2HAvpo
Miller, N., & Sklarz, M. (2012). Integrating real estate market conditions into home price forecasting systems. *Journal of Housing Research*, *21*(2), 183–213.
Munda, G. (2016). Multiple criteria decision analysis and sustainable development. In *Multiple criteria decision analysis* (pp. 1235–1267). Springer.
Munda, G., & Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, *41*(12), 1513–1523.
Ngai, L. R., & Tenreyro, S. (2014, December). Hot and cold seasons in the housing market. *American Economic Review*, *104*(12), 3991-4026. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.104.12.3991 doi: 10.1257/aer.104.12.3991
Nickolas, S. (2021). *What do correlation coefficients positive, negative, and zero mean?* Retrieved from https://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-coefficient-positive-negative-or-zero.asp
Norges Bank. (2020). *Norges bank og utbruddet av koronaviruset (covid-19).* Retrieved from https://www.norges-bank.no/tema/aktuelt/norges-bank-og-koronaviruset/
Novy-Marx, R. (2003). *The microfoundations of hot and cold markets.* Center for Research in Security Prices, Graduate School of Business,University of Chicago.
Nunnally, J. C. (1978). *Psychometric theory 2nd ed.* Mcgraw hill book company.
OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide.* OECD publishing.
Oslo Economics. (2020). *Komparativ analyse av boligomsetningen i norge og andre land/.* Retrieved from https://www.regjeringen.no/contentassets/79704749185444af9cad00cb5f95c17c/oslo-economics.pdf
Paruolo, P., Saisana, M., & Saltelli, A. (2013). Ratings and rankings: voodoo or science? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(3), 609–634.
Puy, A., Becker, W., Piano, S. L., & Saltelli, A. (2020). A comprehensive comparison of total-order estimators for global sensitivity analysis. *International Journal for Uncertainty Quantification*.
Real Estate Norway. (2021). *Oslo house price index.* Retrieved from https://eiendomnorge.no/housing-price-statistics/statbank/
Realtor. (2021). *Market hotness index.* Retrieved from https://www.realtor.com/research/reports/hottest-markets/
Redfin. (2021). *How redfin compete score works.* Retrieved from https://www.redfin.com/how-compete-score-works
Regjeringen. (2016). *Fastsetter ny boliglånsforskring.* Retrieved from https://www.regjeringen.no/no/dokumentarkiv/regjeringen-solberg/aktuelt-regjeringen-solberg/fin/pressemeldinger/2016/fastsetter-ny-boliglansforskrift/id2523967/
Regjeringen. (2021). *Nou 2021: 7.* Retrieved from https://www.regjeringen.no/no/dokumenter/nou-2021-7/id2856737/?ch=3&fbclid=IwAR2xNI4ckK3SXdridmMDYpZDTQLhrvaZD6kOgVsIWmogk7Ay2NXgzvWTdWo
Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *168*(2), 307–323.

Saisana, M., & Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development* (Vol. 214). Citeseer.

Saisana, M., & Tarantola, S. (2016). *State-of-the-art report on current methodologies and practices for composite indicator development.* Retrieved from `https://www.researchgate.net/publication/305392511_State-of-the-art_report_on_current_methodologies_and_practices_for_composite_indicator_development` doi: 10.13140/RG.2.1.1505.1762

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social indicators research*, *81*(1), 65–77.

SSB. (2021a). *Strategi for ssb.* Retrieved from `https://www.ssb.no/omssb/om-oss/strategier`

SSB. (2021b). *Tabell 3: Makroøkonomiske hovedstørrelser. kvartal og år. løpende priser. millioner kroner.* Retrieved from `https://www.ssb.no/nasjonalregnskap-og-konjunkturer/nasjonalregnskap/statistikk/nasjonalregnskap`

Ten Berge, J. M., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*(4), 613–625.

UNDP. (n.d.). *Human Development Reports: Why is the geometric mean used for the HDI rather than the arithmetic mean?* Retrieved from `http://hdr.undp.org/en/content/why-geometric-mean-used-hdi-rather-arithmetic-mean`

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in cronbach's alpha. *Leisure Sciences*, *39*(2), 163–173.

Wilhelmsson, M. (2008, 10). The evidence of buyer bargaining power in the stockholm residential real estate market. *Journal of Real Estate Research*, *30*, 475-500.

Appendix

A1 Quality Assessment

We apply OECD's quality framework, as outlined in Section 3.1, to assess which indicators constitute potential candidates for the composite. The following analysis will discuss each instance where an indicator does not conform with a quality dimension. Thus, a lack of discussion implies that the indicator fulfills the requirement.

A1.1 Accessibility

Waived Contingencies, Listings View, and Competing bids are three indicators of which data is not readily available. Waiving contingencies is a strategy for buyers to make their offer more appealing by giving up certain rights, such as the right to exit the transaction if they cannot secure financing. Contingencies are not a common phenomenon in Norwegian real estate transactions. A proxy could be reservations included with bids. Nevertheless, information concerning bids is not available to the public, thus limiting our ability to construct both Waived Contingencies and Competing bids. Listings views, however, is a statistic collected by FINN but not shared on real estate listings. Since data from FINN originates from scraping listings, the indicator is not accessible.

A1.2 Accuracy

Expired Listings is an indicator of which we cannot ensure the credibility of the data source. The indicator necessitates snapshot data for all listings in the sample period. Then, if a property is listed but not sequentially sold, it counts as expired. However, realtors frequently delist and relist dwellings for short periods due to numerous reasons, such as editing or making the listing appear newer. Thus the indicator can be rather inaccurate.

A2 Indicator Skewness

Table A2.1: Indicator Skewness

Month	Days on Market Skew	Sale-to-List Price Ratio Skew
Jan-17	2.39	0.59
Feb-17	3.86	0.62
Mar-17	4.83	0.8
Apr-17	4.03	0.08
May-17	4.06	1.18
Jun-17	3.48	1.37
Jul-17	2.53	1.87
Aug-17	2.94	1.08
Sep-17	2.87	0.78
Oct-17	2.63	1.38
Nov-17	2.25	1.36
Dec-17	2.23	1.26
Jan-18	2.13	2.14
Feb-18	2.1	0.37
Mar-18	2.39	1.13
Apr-18	2.61	1.11
May-18	3.76	1.57
Jun-18	5.19	1.67
Jul-18	4.02	1.05
Aug-18	4.13	0.9
Sep-18	4.12	1.09
Oct-18	4.25	0.67
Nov-18	3.27	1.01
Dec-18	2.35	2.32
Jan-19	2.42	1.58
Feb-19	2.55	1.66

Table A2.1 continued from previous page

Mar-19	3.4	1.64
Apr-19	4.37	1.21
May-19	4.16	0.85
Jun-19	4.58	0.99
Jul-19	2.77	2.03
Aug-19	3.37	3.21
Sep-19	2.99	0.76
Oct-19	3.14	1.41
Nov-19	3.22	1.04
Dec-19	2.88	2.1
Jan-20	2.62	0.52
Feb-20	2.77	0.71
Mar-20	3.66	1.13
Apr-20	3.97	0.88
May-20	4.26	2.14
Jun-20	3.85	0.98
Jul-20	2.7	1.01
Aug-20	3.12	0.95
Sep-20	3.81	0.68
Oct-20	4.89	1.3
Nov-20	3.96	1.49
Dec-20	4.1	0.01
Jan-21	3.99	0.57
Feb-21	5.35	0.6
Mar-21	7.78	0.87
Apr-21	8.72	1.1
May-21	6.67	0.74
Jun-21	4.16	1.18

A3 Electing Weights With Correlation Neutralization

The indicator's coefficients of determination is shown A3.1 below.

Table A3.1: Coefficients of Determination

	Days on Market	**Sale-to-List Price Ratio**
Sale-to-List Price Ratio	0.276	
Absorption Rate	0.502	0.286

If we want to correct for the higher correlation between Days on Market and the Absorption Rate, we begin by computing the mean of the coefficients for the given set of indicators: $(0.502 + 0.276)/2 = 0.389$, $(0.502 + 0.286)/2 = 0.394$ and $(0.286 + 0.276)/2 = 0.281$. Then we proceed to make the weight for a given indicator inversely proportional to the given mean:

$$w_{days} = (0.389 + 0.394 + 0.281)/0.389 \approx 2.74 \tag{.1}$$

$$w_{abs} = (0.389 + 0.394 + 0.281)/0.394 \approx 2.7 \tag{.2}$$

$$w_{sale} = (0.389 + 0.394 + 0.281)/0.281 \approx 3.79 \tag{.3}$$

Table A3.2 below shows the weights after rescaling so that the sum of the weights equal 1.

Table A3.2: Weights After Neutralizing Uneven Bivariate Correlations

Indicator	Weight
Days on Market	29.7%
Sale-to-List Price Ratio	41.0%
Absorption Rate	29.3%

A4 Sensitivty and Uncertainty Analysis

Let CI_m be the composite index value for month $m, m = 1..., M$,

$$CI_m = f_{rs}(I_{1m}, I_{2m}, ..., I_{Qm}, w_{s1}, w_{s2}, ..., w_{sQ}) \tag{.4}$$

estimated by the weighting model $f_{rs}, r = 1, 2, 3, s = 1, 2, 3$, where r represents the aggregation type and s represents the weighting method as shown in Table A4.3 and A4.2, respectively. The composite indicator is computed using Q normalized indicators $I_{1m}, I_{2m}, ..., I_{Qm}$ for a month with weight $w_{s1}, w_{s2}, ..., w_{sQ}$ depending on the weighting method. I_{Qm} is normalized by the methods in Table A4.1.

Table A4.1: Normalization Method

n	Normalization	Estimation
1	Rank	Equation 3.5
2	Min-Max	Equation 3.6
3	Z-score	Equation 3.7

Table A4.2: Weighting Method

s	Weighting	Estimation
1	Equal Weights	Section 3.3.1
2	PCA	Section 4.4
3	Correlation Neutralization	Section 4.4

Table A4.3: Aggregation Method

r	Aggregation	Estimation
1	Arithmetic	Equation 3.10
2	Geometric	Equation 3.11
3	Harmonic	Equation 3.12

A5 Indicator Scores

Table A5.1: Index and Indicator Scores

Date	Index	Days on the Market	Sale-to-List Price Ratio	Absorption Rate
January 21	87.63	100.0	77.1	87.2
March 21	86.81	100.0	65.4	100.0
February 21	85.69	100.0	80.2	78.4
November 20	75.49	100.0	55.4	77.7
October 20	74.96	94.5	51.0	87.4
February 17	74.20	78.0	89.4	58.6
April 21	73.82	94.5	54.4	78.3
March 17	70.98	78.0	70.5	65.0
January 17	70.17	78.0	100.0	44.3
May 21	68.28	94.5	49.7	67.8
September 20	66.04	78.0	42.9	86.2
August 20	65.81	94.5	43.8	68.8
June 21	64.18	94.5	43.1	64.8
May 18	62.70	89.0	41.8	66.3
December 20	61.52	100.0	52.5	44.4
June 18	59.53	89.0	33.5	70.8
January 20	59.47	89.0	34.2	69.1
April 19	58.04	94.5	35.8	57.8
May 19	57.31	94.5	30.4	65.5
June 20	57.25	94.5	27.3	72.6
April 18	57.16	72.5	42.0	61.3
February 20	56.16	94.5	33.4	56.1
April 17	56.06	72.5	52.2	46.6
August 18	55.47	94.5	32.7	55.2
August 19	55.08	94.5	31.5	56.1
March 18	54.56	78.0	34.2	60.9

June 19	53.96	89.0	29.0	60.8
May 17	53.63	72.5	44.2	48.1
September 19	51.20	89.0	24.2	62.4
February 19	51.10	89.0	28.3	52.9
March 19	49.29	89.0	27.4	49.0
September 18	48.81	89.0	22.4	58.4
July 20	48.29	72.5	35.2	44.1
January 19	47.80	78.0	24.2	58.0
May 20	46.41	83.5	21.1	56.8
March 20	45.59	89.0	21.5	49.6
October 18	45.02	89.0	19.2	53.4
February 18	42.86	61.5	25.6	50.1
November 19	42.63	72.5	17.2	62.1
June 17	41.94	61.5	23.1	52.0
October 19	37.35	61.5	20.9	40.6
November 18	33.35	72.5	14.8	34.5
January 18	27.60	39.5	18.6	28.6
August 17	27.38	50.5	15.3	26.6
December 19	26.10	67.0	20.4	13.0
September 17	25.02	45.0	7.1	49.3
October 17	23.23	36.8	8.6	39.9
November 17	22.35	34.0	8.3	39.8
April 20	21.85	78.0	5.5	24.4
July 19	19.40	50.5	25.4	5.7
December 18	18.69	45.0	15.7	9.2
July 18	12.81	69.8	30.1	1.0
December 17	7.04	23.0	1.0	15.2
July 17	5.70	1.0	13.3	13.9

A6 News-Based Sentiment Analysis

We use Retriever's ATEKST database to access media archives and media analysis. It provides information from 15 Norwegian media sources listed in Table A6.1. The query ensures that we include articles describing hot real estate conditions in Oslo and excludes mentions of cold conditions. We manually omit irrelevant and duplicate articles. The query below returns results from from January 2017 to June 2021. Note that the asterisk sign ensures that all inflections are included in the search.

```
(Boligmarked*) AND (Oslo*) AND (het* OR brennhet* OR hot OR sterkt OR opphet*
OR "selgers marked" OR "hoy temperatur" OR stig* OR sterk* OR steg OR prisvekst*
OR oppgang OR pristopp OR prisokning OR vekst*) AND (boligpris*) AND NOT
(fall* OR boligprisfall OR kald* OR kulde OR falt OR svak* OR prisnedgang
OR nedgang OR prisfall)
```

Table A6.1: News Sources included in the query

Source	Medium
Dagens Næringsliv	Print
Finansavisen	Print
Dagsavisen	Print
Aftenposten	Print
Klassekampen	Print
VG	Print
Dagbladet	Print
E24	Web
ABC nyheter	Web
Nettavisen	Web
Adresseavisen	Web
NRK	Web
Vårt Oslo	Web
TV2	Web