



## Research article

## Housing market forecasts via stock market indicators

Varun Mittal<sup>a</sup>, Laura P. Schaposnik<sup>b,c,\*</sup><sup>a</sup> James B. Conant High School, Schaumburg, IL, USA<sup>b</sup> University of Illinois, Chicago, IL, USA<sup>c</sup> University of Oxford UK

## ARTICLE INFO

## Keywords:

Stock market indicators

Housing market

Trend study

## ABSTRACT

Through the reinterpretation of housing data as candlesticks, we extend *Nature Scientific Reports* article by Liang and Unwin [LU22] on stock market indicators for COVID-19 data, and utilize some of the most prominent technical indicators from the stock market to estimate future changes in the housing market, comparing the findings to those one would obtain from studying real estate ETF's. By providing an analysis of MACD, RSI, and Candlestick indicators (Bullish Engulfing, Bearish Engulfing, Hanging Man, and Hammer), we exhibit their statistical significance in making predictions for USA data sets (using Zillow Housing data) and also consider their applications within three different scenarios: a stable housing market, a volatile housing market, and a saturated market. In particular, we show that bearish indicators have a much higher statistical significance than bullish indicators, and we further illustrate how in less stable or more populated countries, bearish trends are only slightly more statistically present compared to bullish trends.

## 1. Introduction

The mathematical study of housing markets dates back to the 1970s [1], and its prominence has seen a significant increase in the last decades, as consumers are increasingly looking at the right time to make a significant investment [2]. Whilst classical studies have included complex models and economic examinations to varying levels of success [3], after the housing market crashed in 2008 it became clear that new models are necessary to understand both the local and global behavior of housing markets. Since then, more novel approaches begun to be implemented [4], but these are yet to become as robust or successful as they could potentially be [5]. Purchasing a house is never easy, and buying at the optimal time is a significant decision. As a result of COVID, the housing market has become more volatile and has fluctuated significantly [6]. Therefore, we propose the utilization of technical indicators that are commonly applied in the stock market to the housing markets to identify both bearish and bullish trends in the housing market, where

- (a) a trend is considered **bullish** if it corresponds to an increase in value for a time series.
- (b) a trend is considered **bearish** if it corresponds to a decrease in value for a time series.

In a recent *Nature Scientific Reports* article, Liang and Unwin proposed to use stock market indicators to understand trends in COVID-19 cases in the US [7]. Inspired by this work, we shall consider Zillow's data set of the USA's housing market and study it

\* Corresponding author at: University of Illinois, Chicago, IL, USA.  
E-mail address: [schapos@uic.edu](mailto:schapos@uic.edu) (L.P. Schaposnik).

<https://doi.org/10.1016/j.heliyon.2023.e16286>

Received 20 October 2022; Received in revised form 10 May 2023; Accepted 11 May 2023

Available online 18 May 2023

2405-8440/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Market type	Saturated	Volatile	Stable
Country	Japan	Canada	Germany

Fig. 1. Market types and representative countries studied.

through examining the usability of statistical indicators usually seen in the stock market such as MACD Bearish, MACD Bullish, RSI Bearish, RSI Bullish, Hanging Man, and Hammer. These indicators can be used to identify trends (both bearish and bullish) that can be vital when it comes to house purchasing decisions. In what follows we shall show via these indicators, trends (both bearish and bullish) can be detected and used to understand housing prices.

For each indicator we calculate the statistical significance using the Wilcoxon test, and show in particular that while both RSI indicators were statistically significant, only the bearish indicator is significant for MACD. Finally, we extend our analysis to other countries, with the intent of being able to make insights in various types of markets. To be encompassing and robust, we considered different types of markets (Saturated, Volatile and Stable) and studied the relevance of the indicators within each setting, by considering representative countries in each category as shown in Fig. 1 (TABLE) below.

In particular, we show that

- across all countries, bearish indicators were much more statistically significant compared to bullish indicators, for the vast majority of indicators.
- the difference in the statistical significance of bullish and bearish trends isn't nearly as large in saturated and volatile housing markets.
- when using stock market indicators directly on housing prices, the trends of the housing price can be predicted more accurately than when considering the indicators on housing ETFs.

To carry out our study, we considered RSI, MACD, and Candlestick Analysis conducted through R. After grouping Zillow's data [8] (about average median housing sales per week in a region) into Heikin Ashi candlesticks, we scanned the data for the four candlestick indicators mentioned above and used the Wilcox, and later performed a similar procedure for the other indicators. After introducing some background in Section 2, we dedicate Section 5 and Section 6 to the main findings of our work, which can be seen in two different directions in terms of the indicators themselves and their statistical significance in various markets. To highlight the utility of our approach vs any standard stock market study of Housing ETFs, we compared our findings to what one can deduce through the same indicators on the most used Housing ETFs, namely VNQ (Vanguard Real Estate ETF), SCHH (Schwab US REIT ETF), and XLRE (Real Estate Select Sector SPDR Fund). Finally, we expand on the analysis and applications of the above findings in Section 7. Finally, we also attached our code for the project on Github [9].

## 2. Background: stock market indicators

Researchers have considered a large number of ways that stock market prices can be forecasted, for example by speculating a stock's intrinsic value, employing complex algorithms, and speculating human behavior [10], or by using some complex algorithms such as rolling window analysis and deep learning [11]. Moreover, recently there has been some success around more simple but nuanced indicators, popularly known as candlestick analysis, MACD, and RSI [12]. Stock indicators are being used increasingly to identify trends in the stock market, as traders are looking to maximize their profits by trying to gain insight on the future of the market.

The current research is inspired by the work in [7] where stock market indicators were used to forecast COVID-19 cases, and thus we shall dedicate this section to review some of the main results which shall prove useful for our research and for comparison with our own findings. The indicators considered in [7] were candlestick analysis, RSI and MACD indicators. After giving a mathematical definition of those indicators for a time series of datapoints, the authors measured their statistical significance by first considering the null hypothesis, the probability that an external circumstance was the cause of a bearish or bullish trend, not an indicator (recall that an indicator is described as **bullish** if that indicator predicts a future decrease while an indicator is described as **bearish** if it describes a future increase in prices). This probability was denoted as a  $p$ -value, and it was then ascertained that lower  $p$ -values translated to higher statistical significances as that meant the probability of a trend being forecasted by an indicator was higher in comparison to a random external event. After considering this hypothesis, they used the Wilcoxon test for R to determine these  $p$ -values across all of the indicators. This allowed them to make conclusions about the applicability of indicators.

In what follows we will give a brief overview of the work in [7] on candlestick analysis, RSI and MACD indicators introduced for the timed datasets of COVID-19 cases, and we will then extend to the timed housing data of [8] on average median housing sales per week in a region.

### 2.1. Candlestick analysis

The first type of indicators considered were candlestick indicators, which included Hammer, Bullish Engulfing, Dark Cloud Over, Bearish Engulfing, and Hanging Man.

Weekly candlesticks were constructed with the real body constituting the region marked by the endpoints of the opening and closing values (covid cases), and the upper and lower limits being the maximum and minimum cases over the time interval respec-

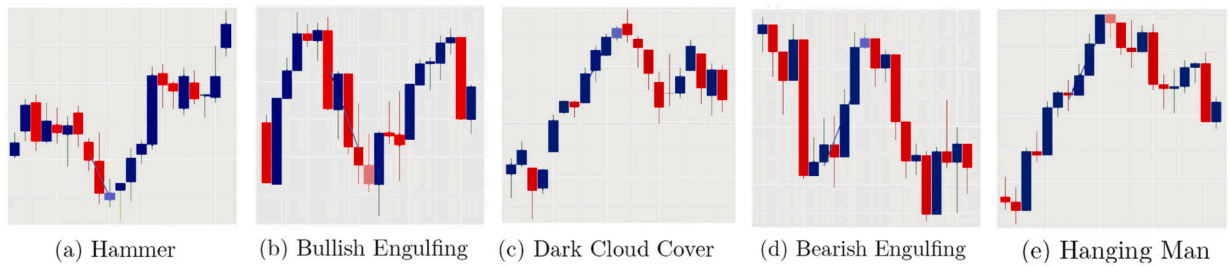


Fig. 2. Candlestick patterns studied in [7]. (a) Hammer pattern; (b) Bullish Engulfing pattern; (c) Dark Cloud Cover pattern; (d) Bearish Engulfing pattern; (e) Hanging man pattern.

tively. Then, the authors developed R code was developed to identify the candlestick patterns shown in Fig. 2 within the dataset to identify the frequencies and predictive capabilities of each of these patterns occurred.

### 2.2. RSI

The RSI (Relative Strength Index), is a popular indicator that is similar to considering the recent “momentum” of a time series through average rate of increases and decreases in a market, in this case the market in consideration (housing, COVID, or stock). The indicator is constructed by dividing the closing values  $C_t$  over some period into two sets:

- The set  $G_t$  in which the series increased as shown in Eq. (1):

$$G_t = \frac{C_t - C_{t-1}}{C_t}, C_t > C_{t-1}. \tag{1}$$

- The set  $D_t$  in which the series decreased as shown in Eq. (2):

$$D_t = \frac{C_{t-1} - C_t}{C_t}, C_t < C_{t-1}. \tag{2}$$

From Eq. (1)- Eq. (2) one can compute in Eq. (3) the averages  $A_t$  and  $B_t$  using the EMA  $V_n$  over  $n$  periods with a smoothing factor of  $\frac{1}{n}$  leading to

$$A_t = V_n[G_t] \text{ and } D_t = V_n[D_t]. \tag{3}$$

Then, the RSI<sub>*t*</sub> indicator at time  $t$  is defined in Eq. (4) as follows

$$RSI_t = 100 - \frac{100}{1 + \frac{A_t}{D_t}} \tag{4}$$

Within the stock market, the indicator becomes useful when the RSI value exceeds a lower and upper threshold, as that is when the stock in consideration becomes overbought or oversold, as seen in Fig. 3. If the RSI value falls below the lower threshold, that signifies a stock being oversold, which signals a bullish trend as stock prices will soon increase in price. Conversely, if the RSI value exceeds an upper threshold, that signifies a stock being overbought, signaling a bearish trend as stock prices will fall.

### 2.3. MACD

The MACD indicator (Moving Average Convergence Divergence) is based on a moving average and a signal line. In MACD and signal line values for the weekly candlesticks. By considering weekly COVID-19 cases, [7] identified when the MACD values crossed the signal values from above (bearish trend) and when the MACD values crossed the signal lines from below (bullish trend), and calculated the statistical significance of both of these indicators. The MACD is obtained using two exponential moving averages (EMAs), calculated over two periods of differing length  $n$ . Specifically, for a given dataset of length  $n$ , labeling the closing values by  $C_i$ , the EMA  $V_n$  is calculated in Eq. (5) recursively as

$$V_i[C_i] = \begin{cases} C_1 & \text{if } i = 1; \\ sC_i + (1 - s)V_{i-1} & \text{if } i > 1, \end{cases} \tag{5}$$

where  $S = \frac{2}{n+1}$  is a smoothing factor. Thus,  $V_n$  can be seen as the exponential average over  $n$  intervals, which by substitutions can be expressed as in Eq. (6) as

$$V_n = R[C_n + (1 - R)C_{n-1} \dots (1 - R)^{n-1}C_1] \tag{6}$$

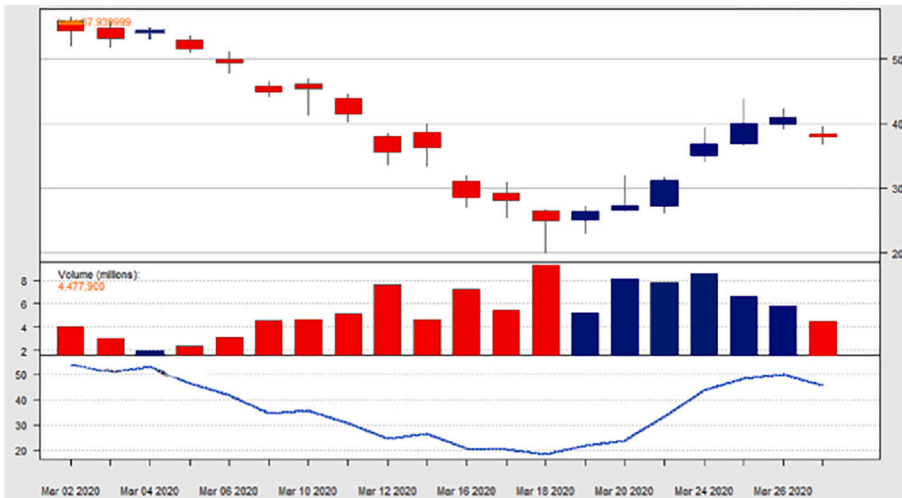


Fig. 3. RSI indicator on Zillow Stock.

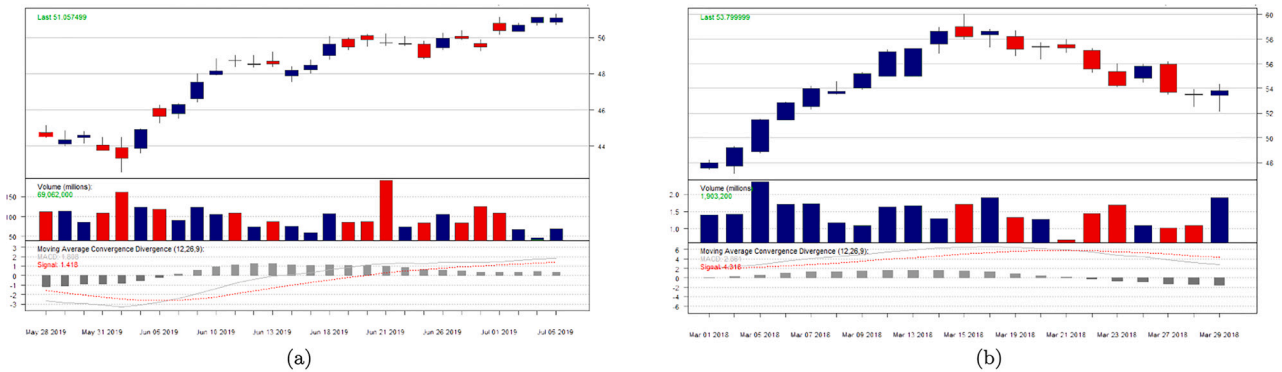


Fig. 4. Two visualizations demonstrating the use of the MACD indicator on (a) Zillow; and (b) Apple Stock respectively.

Note that the coefficient of each term decreases exponentially for earlier values in the time series, thus giving greater weighting to more recent data, hence the name. Given the EMA, the MACD is defined in Eq. (7) as the difference between a longer period average  $n_2$  and a shorter period average  $n_1$  (where by convention  $n_1 < n_2$ ) as follows

$$MACD_{(n_1, n_2)} = V_{n_1} - V_{n_2} \tag{7}$$

In Fig. 4 one can see the MACD indicators in action. As depicted in the graphs a bullish trend with MACD occurs when the MACD line crosses the signal line from below, and a bearish trend within MACD occurs when the MACD line crosses the signal line from above.

### 2.4. Statistical significance

As mentioned above statistical significance values were calculated in [7] using the Wilcoxon test for R, after consideration of the null hypothesis. Through this robust test to measure the significance (due its non-reliability on normalization and it is a non-parametric method) the authors showed that both bearish and bullish indicators were significant in consideration of COVID cases. Moreover, whilst their study highlighted that the 3 bullish trends given by Hammer, Bullish Engulfing, and Bullish MACD, and bearish trend given by Bearish MACD were statistically significant, they shoed that with high  $p$ -values, were determined to be statistically insignificant in forecasting COVID cases.

Along the same lines of research in what follows we shall consider housing market trends through the above indicators, as well as through the Heikin Ashi Candlestick indicator, which we shall introduce and do a statistical study of in the upcoming sections.

### 3. Heikin Ashi candlesticks

In order to take into account individual data points that alter the statistical significance levels of an indicator, and hence to be able to notice trends more clearly, it is sometimes useful to consider **Heikin Ashi candlesticks**. These candlesticks are similar to the

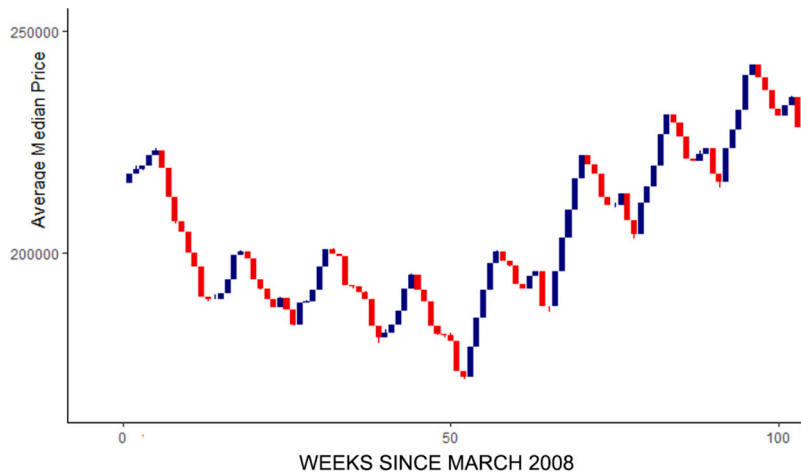


Fig. 5. Heikin Ashi candlesticks for Housing Prices.

normal candlestick described before, except that the individual bars are calculated differently in Eq. (8) - Eq. (11) by defining the following:

$$\text{Close}_n := 1/4[\text{Open}_n + \text{High}_n + \text{Low}_n + \text{Close}_n]; \quad (8)$$

$$\text{Open}_n := 1/2[\text{Close}_{n-1} + \text{Open}_{n-1}]; \quad (9)$$

$$\text{High}_n := \text{Max}[\text{High}_n, \text{Open}_n, \text{Close}_n]; \quad (10)$$

$$\text{Close}_n := \text{Min}[\text{High}_n, \text{Low}_n, \text{Close}_n]. \quad (11)$$

It is important to notice how the terms for calculating the open values are recursively taking the mean of the previous open and close values. This lends itself to better forecasting because it signifies a continuity within a trend. For example, with a normal candlestick analysis, there can be certain days that contradict the trend, which makes analyzing and identification of trends difficult (e.g. see third blue candle surrounded by red candles in Fig. 3). When using Heikin Ashi candlesticks, these values are being recursively calculated, and thus it has the effect of smoothing values over time, such that trends will persist and be more apparent for analysis. This solved our problem of the contradictory points, and can also be compared to the smoothing factor in [7] in the mathematical expression of the RSI indicator.

In summary, normal candlesticks can fluctuate more within a certain trend, however, since Heiken Ashi relies on averages, trends are more continuous and observable for analysis as shown in Fig. 5. This lends itself better to analysis and heightens the statistical significance of indicators, since it makes it more clear that indicators were more likely the sole thing connected to a change, rather than a random external factor.

The statistical significance is assisted because it centralizes around a  $p$ -value, or the probability that a random external event is the cause of a trend occurring. Therefore, by eliminating small contradictory points, the potential of the  $p$ -value being large is made smaller as it isolates the trend as causing the change.

#### 4. Housing data vs ETF data

In order to demonstrate the utility of stock market indicators for analyzing average median housing sales per week in a region from [8], we shall consider here another type of data driven from housing prices which one could use to forecast housing market trends, which is given by housing Exchange-Trade Funds (ETF's). These are a popular type of pooled investment security that can be purchased or sold on a stock exchange similar to a normal stock, but which is formed from a weighted set of many stocks. To compare our previous analysis of average median housing sales per week in a region from [8] and ETF's, we shall consider here *Vanguard Real Estate Index Fund ETF* (VNQ) because of its high AUM (Assets Under Management) and relative prominence in the market, for which a candlestick chart appears in Fig. 6 below.

Since ETFs are easier to analyze and study than larger datasets, it is natural to wonder whether studying ETFs alone (without considering housing prices as our study did) would be sufficient for accurate trend prediction. To study this, we have considered the stock market indicators described in Section 2 to VNQ's candlestick chart.

To understand the relevance of Zillow's dataset [8] (and in particular of the average median housing sales per week in a region) we studied the statistical significance of the different stock market indicators considered in our paper when applied to VNQ's candlestick chart, for which a table of  $p$ -values appears in Fig. 7 (TABLE) (the probability that a external factors caused the reversal being studied, and not the indicator – more on this in Section 5).

These probabilities were calculated with the Wilcoxon test in R, after scanning the code for each of these indicators (MACD, RSI, and Candlestick Indicators which included Bearish and Bullish Engulfing, Hanging Man, and Dark Cloud Over) – the definition of

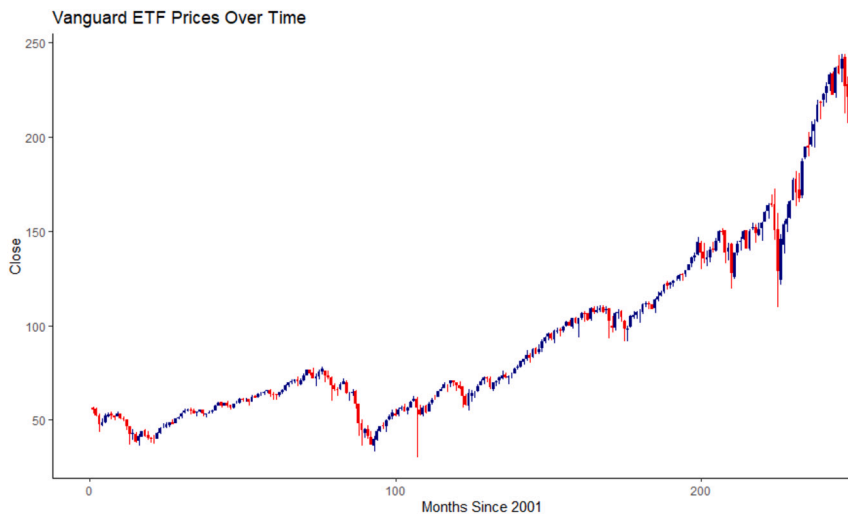


Fig. 6. A visualization showing a candlestick chart of Vanguard Real Estate Index Fund ETF since 2001. Similar to housing prices, this real estate ETF has increased over time.

Indicator	Bullish	Bearish
MACD	0.9999	$2.5564 \cdot 10^{-88}$
RSI	0.9080	$3.2399 \cdot 10^{-6}$
Engulfing	0.9997	$1.8713 \cdot 10^{-13}$
Hanging Man	-	0.3937
Dark cloud over	-	0.0022

Fig. 7. (TABLE)  $p$ -values for the different indicators when applied to VNQ's candlestick chart.

these indicators is provided in Section 2. In particular, recall that since we want to observe the efficacy of indicators, high statistical significance is associated with a low  $p$ -value for any indicator. As shown in Fig. 7 (TABLE), all of the bearish indicators have low  $p$ -values (signifying high statistical significance) while all of the bullish indicators have high  $p$ -values (signifying low statistical significance). As a result, one can see that studying ETFs alone would not lead to enough insight into predicting housing market trends since bullish trends could not be foreseen with the usage of technical indicators.

From the above analysis, one can see that it becomes of much interest to perform the study of stock market indicators directly on housing data, and thus we shall dedicate the following sections to do so by considering actual housing market data for countries across the globe within varying housing markets, including saturated, volatile, and stable housing markets. Our full results, summarized in Section 5, can be seen with Fig. 12. In particular, we shall demonstrate how some bullish indicators have low  $p$ -values (correlating to a high statistical significance), for example MACD Bullish in Canada has a  $p$ -value of 0.11, implying that more insights could be made with bullish indicators through the consideration of housing data rather than ETF prices.

### 5. Housing market trends

In order to understand housing market trends via stock market indicators, we shall consider Zillow Data set [8] (about average median housing sales per week in a region) and the information presented there on average median housing sales per week in a region. The data used together with an approximation through a lineal regression are shown in Fig. 8.

In order to compare our study of the US housing market with global markets, we shall consider in this paper the following three different types of housing markets:

- (a) A **saturated** housing market has a lot of activity and is usually correlational with a large population. These markets, with their large-scale activity, have high volume, and generate large numbers of transactions from both purchasers and sellers.
- (b) A **volatile** housing market is one that changes quite frequently. For one period of time, it may be undergoing a bearish trend, but then it undergoes a sudden reversal in a short span of time and it is now growing in value. One purpose of our study is being able to forecast these reversals with technical indicators typically found in the stock market.
- (c) A **stable** housing market is one that has sustained trends and doesn't plunge or proliferate suddenly in value across a time period.

We shall study the above markets by considering representatives from each set, and look at the housing data from Germany (a stable housing market), Canada (a volatile housing market), and Japan (a saturated housing market), thus allowing for our analysis to yield powerful insights into not one but multiple different markets. In order to understand which indicators can be used within

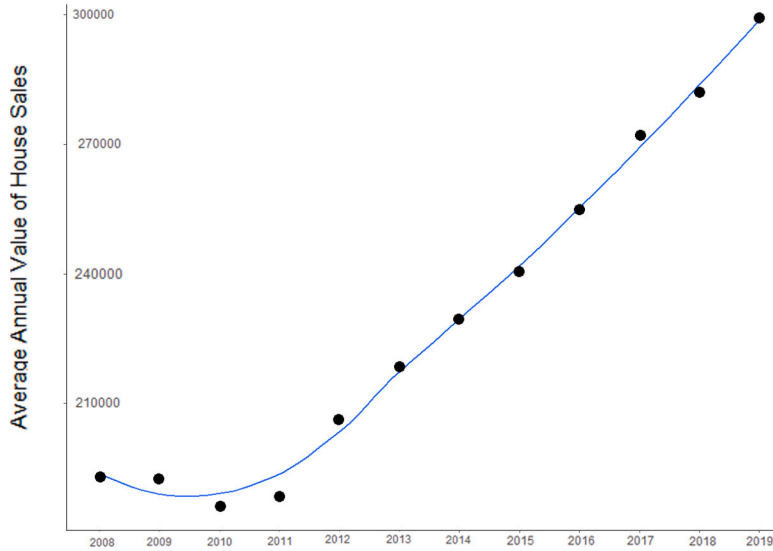


Fig. 8. Housing Prices Over Time (original data graph) plotted through Zillow's dataset [8].



Fig. 9. Housing Prices Over Time Heikin Ashi Candlesticks, obtained via geom\_candlestick() function in R with appropriate candlestick data.

each setting, we considered the significance of all of our indicators by using the Wilcoxon test to measure the statistical significance of our indicators.

Following standard conventions, we considered the null hypothesis being that a trend occurred randomly rather than an indicator inducing it. We called this probability the  $p$ -value, so the larger the  $p$ -value, the less statistically significant our indicator is – and the closer the  $p$ -value was to 0, the more statistically significant our indicator is, since there is less of an inherent random event changing the housing prices compared to the indicator taking effect.

We also utilized historical housing data from Germany, Canada, and Japan [13]. We grouped our data in the form of monthly Heikin Ashi candlesticks for the Zillow (USA) data as shown in Fig. 9, and then yearly candlesticks for the various countries and utilized these datasets to substantiate and support our claims.

5.1. Predictions via RSI

In what follows we shall consider the RSI indicator on the Zillow Data Set [8] as a way to forecast trends, both bearish and bullish. We developed R code to produce RSI values of monthly candlesticks, and observed viable lower and upper thresholds respectively. We then produced a model that would identify the frequency of both RSI Bearish and Bearish indicators, noticing two interesting trends:

- Except for the US housing market with Zillow Data, all RSI Bearish indicators were not statistically insignificant, with  $p$ -values larger than 0.9.

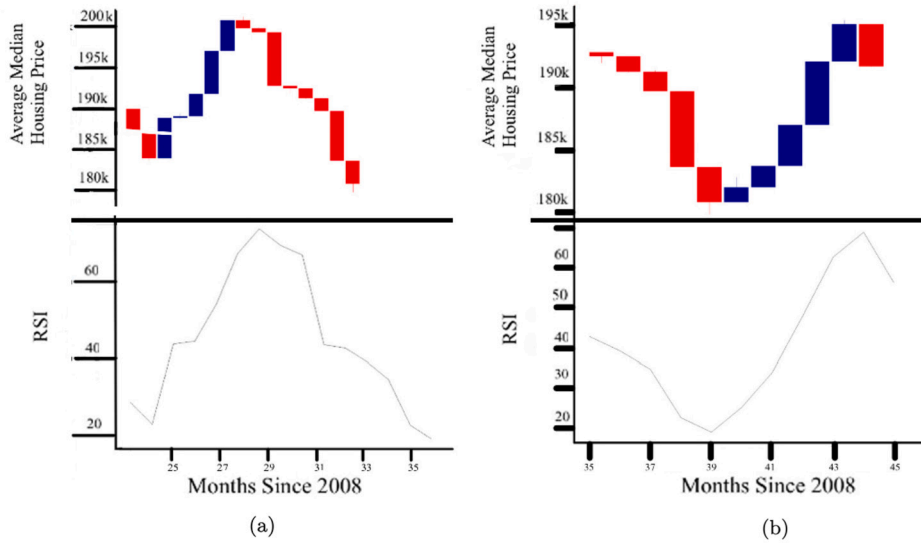


Fig. 10. RSI Trends. (a) the RSI value crosses an upper threshold, corresponding to a bearish trend, or a decrease in housing prices; (b) the RSI value crosses a lower threshold, corresponding to a bullish trend, or an increase in housing prices.

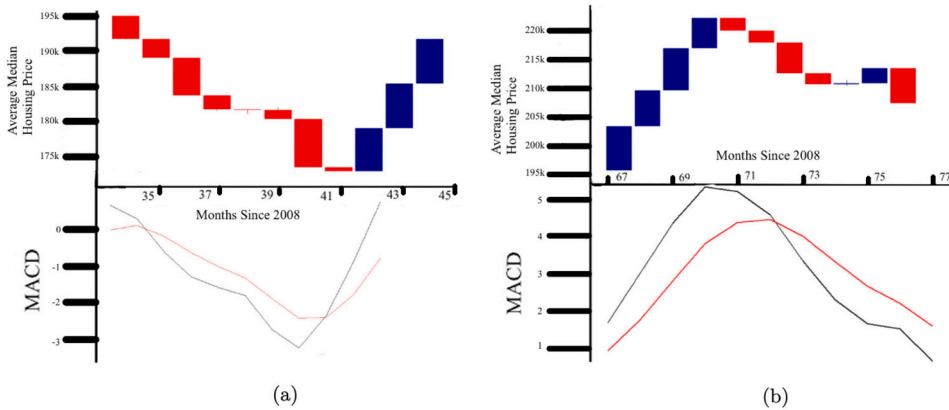


Fig. 11. MACD indicator on Zillow (USA) housing data. (a) First 45 months since 2008, downward trend; (b) Last 30 months since 2008, upward trend.

- When considering the indicator on all types of markets, RSI is the only non-candlestick indicator across all 4 countries in examination that had a statistically significant bearish indicator. Moreover, bullish indicators were much more statistically significant than bearish trends.

The RSI indicator for the Zillow dataset of [8] is shown in Fig. 10: an example of a bearish trend is indicated by the RSI value crossing an upper threshold (as in the left visualization), and a bullish trend is indicated by the RSI value crossing a lower threshold (as in the right image).

### 5.2. Predictions via MACD

The MACD indicator is obtained via a correlation of the moving average and a signal line obtained through the timed data series. By developing R code that produces MACD and signal lines for Zillow’s dataset [8], we could combine it with the visualization via monthly candlesticks. When the MACD crossed the signal line from above it signified a bearish reversal, and when it crossed from below it signified a bullish reversal, and this is shown in Fig. 11, where a bearish trend can be seen in the left visualization, and a bullish trend in the right visualization.

As before, *p*-values were calculated for each indicator using the Wilcoxon test over monthly intervals as described above. From our study we see different levels of accuracy obtained in our predictions:

- MACD signals were the most frequently occurring signals that our R code detected. This is salient since it might be the most ubiquitous for consumers to utilize for their benefit.



Indicator	Country	Type of Market	P-Value
Bearish Engulfing	Germany	Stable	0.17
Bullish Engulfing	Germany	Stable	0.99
Hammer	Germany	Stable	0.94
MACD Bearish	Germany	Stable	0.97
MACD Bullish	Germany	Stable	0.24
RSI Bullish	Germany	Stable	0.96
RSI Bearish	Germany	Stable	$2.2 * 10^{-16}$
Hanging Man	Germany	Stable	0.000545
Bullish Engulfing	Canada	Volatile	0.82
Hammer	Canada	Volatile	0.78
MACD Bearish	Canada	Volatile	0.93
MACD Bullish	Canada	Volatile	0.11
RSI Bullish	Canada	Volatile	0.99
RSI Bearish	Canada	Volatile	$2.1 * 10^{-15}$
Bearish Engulfing	Canada	Volatile	0.14
Hanging Man	Canada	Volatile	0.22
Bullish Engulfing	Japan	Saturated	0.95
Hammer	Japan	Saturated	0.70
MACD Bearish	Japan	Saturated	0.97
MACD Bullish	Japan	Saturated	0.35
RSI Bullish	Japan	Saturated	0.99
RSI Bearish	Japan	Saturated	0.000015
Bearish Engulfing	Japan	Saturated	0.95
Hanging Man	Japan	Saturated	0.70

Fig. 12. (TABLE) A table of all of the determined *p*-values, broken down by country and market that underwent analysis.

- The statistical significance of the MACD Bearish and MACD Bullish indicators is very large: whilst MACD Bullish has a low *p*-value with high statistical significance ( $7.384 * 10^{-15}$ ), MACD Bearish has a high *p*-value with low statistical significance (0.99). This was also observed for the RSI values, with the exception being the US housing market.

### 5.3. Predictions via Heikin Ashi

We examined the dataset with respect to 4 different candlestick patterns via Heikin Ashi candlesticks, Hammer, Hanging Man, Bearish Engulfing, and Bullish Engulfing. We then scanned the data with R code for these indicators, and then used the Wilcoxon test reiterated above to determine the statistical significance. From our study we saw different levels of accuracy obtained in our predictions, allowing us to infer certain candlestick pattern indicators were more viable with respect to others:

- Both within the US and as well as for the other housing markets, bearish trends were less statistically significant compared to bullish trends.
- Within the US, as with the previous indicators, the Bearish signal (Hanging Man) is determined to be statistically insignificant with a *p*-value of 0.97, but the Bullish signal is determined to be a viable indicator and statistically significant indicator with a *p*-value of 0.05.
- Across the other markets studied, the converse was observed: *p*-values for bearish indicators were lower (high statistical significance) and the *p*-values for bullish indicators were higher (low statistical significance). A table of statistical significance for all of the indicators is shown in Fig. 12 (TABLE), and will be the focus of our analysis in later sections.

## 6. Discussion on housing market via stock market indicators

Having determined all of the *p*-values and the statistical significance values for the indicators considered in the present paper (see Fig. 12), we shall describe here several interesting trends that provide insight into the behavior of the US housing market as well as the three different global housing markets we considered. For context, recall that we selected and analyzed the US housing market, the German housing market (representative of a stable economy), the Canadian housing market (representative of a volatile economy), and the Japanese housing market (representative of a saturated market).

Through our analysis of the MACD indicator on housing data, and in accordance with Fig. 12, one could observe the following:

- This indicator followed the pattern that is emerging within our analysis that exhibited high statistical significance for bullish trends and low statistical significance for bearish trend, but only for the United States.
- For other countries, the above scenario was reversed: bearish MACD indicators were more statistically significant compared to bullish trends.

From the above, one can see that in particular while MACD Bullish can be used by both consumers and sellers in the housing market, people who are looking to forecast bearish reversals in the housing market should not look to utilize this indicator. But the most important thing to realize is the difference in the statistical significance between the MACD Bearish and Bullish.

In more populated, saturated, and volatile markets, the difference in statistical significance of the MACD Bearish and Bullish is smaller, as exemplified by Japan. In Japan, the difference between MACD Bearish and Bullish is 0.62, but in Germany is significantly larger, hovering around 0.73. This variance in difference is a salient insight because it exhibits that indicators, both statistically significant and insignificant, can be evaluated differently. In a market with a large difference, the failures and effectiveness of indicators can be more established (due to the higher polarization of  $p$ -values), but this is not the case with more inherent randomness with a smaller difference and decreased polarization. Moreover, this trend is persisting when the significance of the MACD is at a constantly low value ( $p$ -value > 0.9). This signifies that MACD Bearish trends are less statistically significant in more crowded and volatile housing markets, something that could provide important information for consumers in that region looking to save money.

From a different perspective, the RSI indicator proved to uphold the trends we observed throughout our examination. This trend, as reiterated above, is that bullish indicators were less statistically significant in comparison to bearish indicators, across all three housing markets (studied through their representative countries), except for the United States.

In the United States, as exhibited by the table in Fig. 12, RSI is very significant statistically in terms of both bullish and bearish. The United States case is interesting because the US can be justified both as a populated yet stable market, so further research into the US housing market can lend itself to even more intriguing analysis. From our above analysis, the following main takeaways are important:

- For the United States, RSI is the only statistically significant and viable indicator for bearish trends. This is important for buyers who are looking to purchase at the optimal price.
- The  $p$ -values for RSI were mostly similar across the three markets under consideration, which signifies that the circumstances of a market may not dictate RSI's efficacy. This makes sense considering RSI's emphasis on recency over long term processes.

Given the above analysis, it seems almost contradictory that the RSI indicators can have a low statistical significance. After all, RSI's focus on recency and its practicality when identifying trends should lend itself to higher statistical significance. However, this is where one needs to remember that external circumstances still must be considered, as not all markets are conducive to high statistical significance (low  $p$ -values).

Finally, the most appealing analysis is the one that can be done Heikin Ashi candlesticks:

- Similar to RSI and MACD and as exemplified in the table of Fig. 12, the bearish indicators (Hanging Man, Bearish Engulfing) were more statistically significant than the bullish indicators (Hammer, Bullish Engulfing).
- The above trend did not continue in the housing markets for Japan and the US, which constituted our saturated/popular housing markets in our study.
- In the US and Japan, both bearish and bullish indicators had low statistical significance, compared to the bearish being of high statistical significance and the bullish being of low statistical significance.

This is interesting because it provides insight into the applicability of these indicators in more populated markets. This could help consumers understand which indicators not to use, as that is just as important as knowing which ones to use. The main takeaways can be summarized as follows for Heikin Ashi Candlesticks indicators:

- The bearish indicators were more significant compared to the bullish ones, which is consistent with our other results.
- For more populated housing markets such as the US and Japan, neither bullish nor bearish candlestick indicators were statistically significant. Hence, these indicators are less likely to be significant with more volume in the market, as that proliferates any inherent randomness.

## 7. Concluding remarks

In the present paper we have analyzed datasets provided by Zillow [8] and FRED [13] for Japan, Germany, and Canada's historical housing prices. We then used technical indicators typically utilized in the Stock Market to identify trends with respect to our data. To determine the usability and frequency of technical indicators, we analyzed 4 candlestick patterns, we used the Wilcoxon test, developed in R code to calculate a  $p$ -value that signified statistical significance.

An application of this research study is being able to reliably forecast future trends using these stock market indicators. Fig. 13 and Fig. 14 demonstrate the capacity of forecasting and its usability to understand housing market behaviors both within the US as well as for other global housing markets.

Through the study of candlestick indicators via Heiken Ashi (Hammer, Hanging Man, Bearish Engulfing, Bullish Engulfing) we found the following behaviors:

- (I) Populated/Saturated housing markets have low statistical significance values due to more randomness present in a more active market.

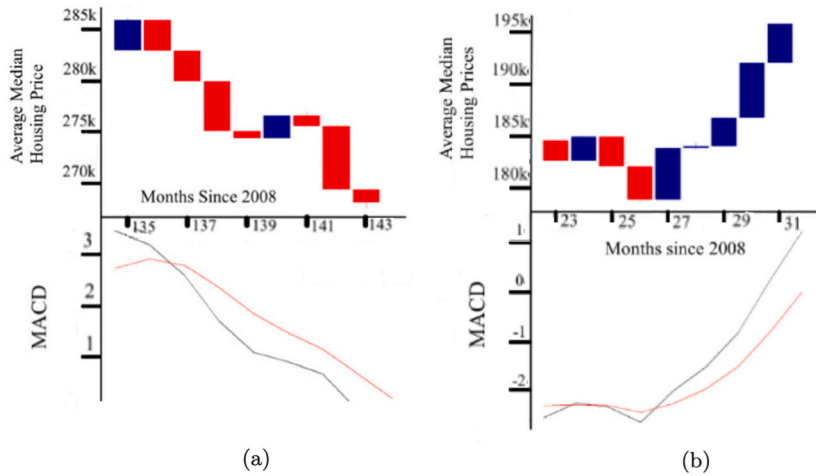


Fig. 13. A visual demonstration of the predictive capacity of the MACD indicator for (a) a bearish trend; and (b) a bullish trend.

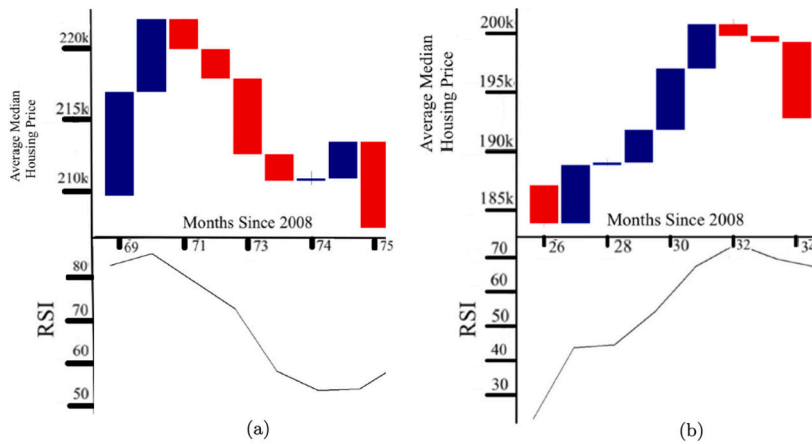


Fig. 14. A visual demonstration of the predictive capacity of the RSI indicator for (a) a bearish trend; and (b) a bullish trend.

- (II) For the US housing market in particular, we found that bearish trends were less statistically significant compared to bullish trends.
- (III) The converse of (I) and (II) is true when we examined the data for other types of markets through their representative countries, with bearish trends being less statistically significant compared to other countries.
- (IV) The difference in the significance of bearish and bullish trends varies depending on the market under consideration. In more volatile and saturated housing markets such as the US, Canada, and Japan, this difference is smaller as there is more randomness to decrease polarization, but in a stable market such as Germany this is not the case.

For nearly all of our analysis, RSI remained fairly strong in terms of its applicability. We conjecture this is the case because whilst most indicators analyze data over a large period of time, RSI places its emphasis on recent trends. Hence, a strong RSI lends itself to a bullish trend, and a weak RSI signifies a bearish trend. In contrast, all of the other indicators such as MACD, Hanging Man, and Bullish Engulfing, an emphasis on later trends accounts for the decrease in statistical significance since there is a larger probability of external factors playing an important role.

Using the Wilcoxon test in R, we were able to determine the  $p$ -value of all of our indicators as shown in the table of Fig. 12. Through this study, one can deduce which indicators should be used in each type of market, and this is summarized in Fig. 15 below, where we present which indicators would be viable for various markets. This comes to embody the final results of our study, where we present the most effective indicators.

It should be noted that the RSI Bearish was statistically significant across all markets, signaling that is an indicator that can be used in practically any case study. Moreover, from our study one can see that the difference in statistical significance between bearish and bullish indicators is smaller, as exemplified by Japan.

In Japan, the difference between MACD Bearish and Bullish is 0.62, but in Germany is significantly larger, is 0.73. This means that in more crowded and volatile markets it can be more confusing and inherently variable about not only the success of statistically significant indicators but the failures of statistically insignificant indicators, whereas this notion is minimized as the difference gets

Type of Market	Viable indicator
Stable (Germany)	RSI Bearish Hanging Man
Volatile (Canada)	MACD Bullish RSI Bearish Bearish Engulfing
Saturated (Japan)	RSI Bearish

Fig. 15. Indicators that can be used for the markets and countries analyzed in our study, determined this by parsing through our  $p$ -values in Fig. 12 and determining the lowest  $p$ -values per market and country, as these would yield the most statistically significant indicators for each case.

larger in stable markets. This is because lower and higher  $p$ -values dictate a more concrete sense of failure and success for statistically insignificant and significant indicators respectively.

The analysis of 6-month forecasts through the usage of MACD and RSI as shown in Fig. 13 and Fig. 14, depict how consumers can forecast prices with these indicators to their benefit. Within the MACD analysis, observe that for the left images, consumers can deduce that for this period of 6 months (in this case Months 135-143 from 2008, so February to July 2020), prices will decrease, due to the fact the black MACD line crossed the red signal line from above. Similarly, for the right images consumers can deduce that for this period of 6 months (in this case Months 23-31 from 2008, so November 2009 to April 2010), prices will increase, due to the fact the black MACD line crossed the red signal line from below. This analysis can be extended to the RSI indicator, where one can observe that from the left graph consumers can deduce that for this period of 6 months (in this case Months 69-75 from 2008, so August 2013 to March 2014), prices will decrease, due to the fact the RSI values crosses an upper threshold. Similarly, for the bottom visualization consumers can deduce that for this period of 6 months (in this case Months 26-34 from 2008, so April 2011 to September 2011), prices will increase, due to the fact the black MACD line crossed the red signal line from below.

In our study, we did an extensive analysis on utilizing stock market indicators upon housing data from a variety of housing markets, ranging from volatile to stable housing markets in Canada, Germany, Japan, and the US. An interesting future application is rentals/apartments, which could be of immense interest with the application of indicators because the renting market is even more volatile than the housing market [14], since there is a plethora of other external factors being inherently present due to the rental being a temporary residence (for example, holiday season variables, or reasons for sublet, e.g., a number of college students take rentals when doing internships at a company headquartered at a city).

In the last decade, there has been significant innovation in stock market analysis and the evolution of the housing market. In the past, the stock market has been analyzed with less nuanced techniques such as intrinsic value and speculating human behavior [10]. However, this changed drastically with the rise of candlestick analysis, as well as the emergence of the MACD and RSI indicators [15]. Moreover, the onset of the COVID-19 pandemic has meant that the housing market has transformed significantly [16]. This inspired us to present the application of stock market indicators on the housing market with data from a variety of housing markets, so people, no matter what type of market they are in, can know which indicators are significant and how trends can be forecasted, which is salient in saving millions of dollars within the housing market [17].

Finally, it is important to note that bearish trends were less statistically significant in the US but more significant in other countries. It would be thus very interesting to analyze whether this would be true across other time series, such as the rental market described earlier. But in the meantime, it is significant to realize that there are certain stock market indicators that can forecast future trends in housing prices.

#### CRediT authorship contribution statement

**Varun Mittal:** Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

**Laura P. Schaposnik:** Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

#### Declaration of competing interest

I have no relevant interest(s) to disclose.

#### Data availability

Data associated with this study has been deposited at Mittal and Schaposnik's repository, Varun121322/housing-market-forecasts-via-stock-market-indicators [9]: An analysis of housing market forecasts with popular stock market indicators, including macd, rsi, and candlestick indicators," 2022.

## Acknowledgements

The work of Laura P. Schaposnik is partially supported through the NSF grants CAREER DMS 1749013, NSF FRG Award DMS-2152107 and a Simons Fellowship. The work appearing here was finished whilst LPS was a RP at SLMath, and their support is greatly acknowledged. The content on this manuscript is not intended to provide legal, financial or real estate advice. It is for information purposes only.

## References

- [1] Andrew M. Kirby, Housing market studies: a critical review, *Transactions of the Institute of British Geographers* 1 (1) (1976) 2–9.
- [2] Cynthia M. Gong, Colin Lizieri, Helen X.H. Bao, “Smarter information, smarter consumers”? Insights into the housing market, *J. Bus. Res.* 97 (2019) 51–64.
- [3] David T. Carruthers, Housing market models and the regional housing system, *Urban Stud.* 26 (2) (1989) 214–222.
- [4] Ahmed Khalafallah, Neural network based model for predicting housing market performance, *Tsinghua Sci. Technol.* 13 (S1) (2008) 325–328.
- [5] R.M. Kirwan, D.B. Martin, Some notes on housing market models for urban planning, *Environ. Plann. A, Econ. Space* 3 (3) (1971) 243–252.
- [6] Sitian Liu, Yichen Su, The impact of the Covid-19 pandemic on the demand for density: evidence from the u.s. housing market, *Econ. Lett.* 207 (2021) 110010.
- [7] Yi Liang, James Unwin, Covid-19 forecasts via stock market indicators, *Sci. Rep.* 12 (1) (2022) 13197.
- [8] Research Zillow, Zillow data. Zillow Research Data set, Mar 2021.
- [9] Varun121322, Varun121322/housing-market-forecasts-via-stock-market-indicators: An Analysis of Housing Market Forecasts with Popular Stock Market Indicators, Including Macd, Rsi, and Candlestick Indicators, GitHub, 2022.
- [10] Marcelle Chauvet, Simon Potter, Coincident and leading indicators of the stock market, *J. Empir. Finance* 7 (1) (2000) 87–111.
- [11] Daiki Matsunaga, Toyotaro Suzumura, Toshihiro Takahashi, Exploring graph neural networks for stock market predictions with rolling window analysis, 2019.
- [12] Terence Tai-Leung Chong, Wing-Kam Ng, Venus Khim-Sen Liew, Revisiting the performance of macd and rsi oscillators, *J. Financ. Risk Manag.* 7 (1) (2014) 1–12.
- [13] St. Louis Federal Reserve Bank, Federal Reserve Bank of St. Louis Housing Data, FRED, 2022.
- [14] Kyung-Hwan Kim, Chang-Moo Lee, Young-Man Lee, Chapter 12: Rental housing system and housing market volatility: Monthly rent-based vs. asset-based systems, Feb 2014.
- [15] Rinni Bhansali, Laura P. Schaposnik, A trust model for spreading gossip in social networks: a multi-type bootstrap percolation model, *Proc. R. Soc. A* 476 (2235) (2020) 20190826.
- [16] Ka Shing Cheung, Chung Yim Yiu, Chuyi Xiong, Housing market in the time of pandemic: a price gradient analysis from the Covid-19 epicentre in China, *J. Financ. Risk Manag.* 14 (3) (2021).
- [17] Gary V. Engelhardt, Christopher J. Mayer, Intergenerational transfers, borrowing constraints, and saving behavior: evidence from the housing market, *J. Urban Econ.* 44 (1) (1998) 135–157.