GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 20 • Number 1 • 2026

https://doi.org/10.7494/geom.2026.20.1.5

Google Trends as Indicator of Social Preferences: Causality and Intervention in Poland's Housing Market

Mirosław Bełej

University of Warmia and Mazury in Olsztyn, Department of Real Estate and Urban Studies, Olsztyn, Poland

Abstract: The article's primary purpose was to explore the potential of Internet searches for keywords that were related to the polish housing market in order to understand the public's current preferences or reactions to changes in the market environment. The research used data that was downloaded from Google Trend (RSV) from 2010 through 2024. The Granger causality test was then applied to the relationship between RSV and housing prices, and a Bayesian structural time-series model was applied to examine the impact of the external intervention (COVID-19) on the RSV dynamics. The results indicated that significant changes in the market environment could influence fluctuations in interest in housing, as was evidenced by the changes in the online searches. The article respectfully suggests that a more nuanced understanding of market dynamics might be achieved through a thoughtful integration of classical economic data with non-classical Internet data.

Keywords: Google Trend, housing, COVID-19, social behavior

Received: March 31, 2025; accepted: September 23, 2025

E-mail & ORCID iD: miroslaw.belej@uwm.edu.pl, https://orcid.org/0000-0001-8650-6990

1. Introduction

Internet user activity can reflect collective behavior and show the observed population's interests, concerns, and intentions [1]. This idea was the research basis of this paper, in which non-classical data (RSV from Google Trend) helped better understand the essence of the housing market, created precisely by residents with their preferences, moods, or herd behaviors. Classical housing market analyses generally use so-called historical data (sales price, rental price, unemployment). The price of an apartment, for example, becomes a historical fact on the date of the purchase and sale agreement (notarial deed). However, the price of an apartment or house is created earlier in negotiations during the signing of the so-called preliminary agreement (even two months before the actual purchase). As a result, we used the transaction date in the market analysis; i.e., the date of the notarial deed and not the earlier data from the actual settlement of the sale price (the date of the preliminary agreement). If the rest of the data (describing socio-economic conditions) was also historical data, then the current market analysis was time-lagged by dominating the historical data analysis. Overall, the characteristics of the housing market itself generated several different problems in the data analysis of this market:

- limited availability of historical data on housing prices;
- lack of single publicly available database on housing prices;
- paid access to official price register, which is kept in different forms in different administrative units;
- lack of standardized reporting system for real estate price data;
- market reports have different data sources (from deeds, reports from real estate agents, managers or developers, offer prices) and different methods of analysis;
- available housing price reports generally show monthly or quarterly averages;
- heterogeneity of properties, making standardization difficult,
- impact of state intervention on house price dynamics (financial programs);
- impact of national road infrastructure development policy;
- impact of local spatial planning policy;
- lack of holistic spatial thinking (city-wide layout) when it comes to granting decisions on conditions of development investments;
- lack of transparency in housing-rental prices.

Another problem was the generally small number of housing transactions during the time period and region under study as well as the non-uniformity of the data in successive periods of the time-series creation [2], as transactions of similar properties occurred at irregular or random times, thus averaging the prices over successive time-periods of the time series; these did not have the characteristics of similarity as in the case of the other phenomena under study. Rapid changes in the

time-series dynamics in local markets could have resulted from entirely different populations during subsequent periods of the time series. For example, the prices of 20-to-30-year-old residential buildings were mainly used for average prices in the first period, while relatively new 2-to-3-year-old buildings that were located in excellent locations were sold in the next period. Hence, the methodological question arose whether the classical sources of real estate information should not have been expanded with the possibilities that were offered by analyzing society's activity on the Internet.

This paper reinforces the work that was developed in several studies [3–5], the concept of using indicators that describe the current state of social collective behavior, which reflects the current intentions of that community. Nowadays, the Internet is becoming a person's natural environment from where information about the world around us is taken, and Google Trends can be a useful tool for analyzing this phenomenon. According to the principle, the more often a keyword is typed into search engines, the more interested the public is in that topic.

Therefore, the primary goal was to see if searching for keywords on the Internet related to the housing market could provide a basis for understanding the public's current preferences and reactions in this market. In this paper, the relationship between the RSV index and housing prices was determined, and the impact of a global phenomenon (COVID-19) on public interest in the housing market (precisely verified by the number of searches on the web) was examined. In both of the indicated issues, Google Trend was used to measure public interest in the housing market. The study was implemented using the Granger causality test and Bayesian structural time-series model. R software was used for the computations, along with the available packages of this software [6–9].

2. State of Problem

The housing market is the cornerstone of any country's economy, as the housing sector not only boosts investment, construction, architectural, or surveying employment but also provides a sense of security in society [10–12]. A dwelling is also a source of positioning a person in the social hierarchy, since the neighborhood of residence or the type of building can be a condition of membership in a particular social group. A property is part of the social hierarchy. In this sense, societies are increasingly subject to tribalization; that is, the phenomenon of creating closed social groups with shared values or leisure activities. As a result, society is divided into enclaves of wealth or poverty. More on this subject has been presented in a number of publications [13, 14]. The housing market is the basis for the functioning of societies [15, 16] because a basic human need is a roof over one's head and a sense of security. In this context, housing is not a typical market commodity [17] because it connects the user through an emotional bond with a sense of security, protection, identity, or attachment to a place or family

history. Therefore, the housing market should be looked at much more broadly than classical definitions of the transfer of property rights between different entities define it. Since the essence of this market is society and the interrelationships that are taking place within it, attention should be paid to the transformation from an industrial society to an information society. An information society [18–20] is characterized by the extensive use of communications information technology in acquiring, storing, processing, and creating information and data.

Information and communications technology (ITC) is currently fundamental to the development of society and institutions, as it allows information collectors and housing market information providers to carry out the process instantly anywhere in the world using basic tools; i.e., a cell phone or laptop. In this way, a potential home buyer, can analyze the prices of similar properties, the surroundings of the property, the levels of air pollution, the levels of sunshine, the levels of security in the area (assaults, thefts), and even the levels of traffic jams near the selected property without leaving home [21]. This way of data gathering opens up new possibilities for understanding the functioning of this market, precisely by understanding the principles of social reactions during local or global events.

As a rule, the housing market is studied in traditional terms; so, official data sources are used; i.e., national statistics bureaus, price registers, value registers, banking data, cadastral data, official reports, or analyses of markets. This is the basis for reliable global trends or local market behavior assessments. However, this approach has significant limitations, as the data with which we operate is historical and not current. Today's society strongly lives together with information technologies, and their permanent use in everyday life makes it possible to expand the basic sources of real estate data by using current data from the Internet; e.g., Facebook, Flickr, X (Twitter), OpenStreetMap, the web scraping of real estate sales offers, or analyses of Internet search trends (Google Trend).

Google Trends is a freely available online tool that is provided by Google that enables one to input search query terms and provides a series of data depicting the relative interest in that term by other individuals making similar search queries [22]. Therefore, it can be a valuable tool for assessing the popularity or public interest of a given product, topic, or event. Presenting the absolute number of searches for a given topic type would generate difficulties when compared with other keywords [21]. Each data point in the Google Trends database is computed by dividing the total searches from a specific geographic region and the time-period that the total searches cover, which results in a relative popularity measure. This resulting number is then scaled from 0 to 100, with 100 representing the most searched volume within this specific geographic region during a specific time range [22]. Google Trends can therefore show interest in a given keyword not as an absolute value but as a relative relationship to all of the other searches on the Internet at a given time and place.

The Google Trend application, which is used to understand the current preferences of the information society, has been the subject of several studies that have

been carried out in recent years [22–24]. In some of these, a clear correlation was noted between the dynamics of housing market keyword searches and the dynamics of housing prices [23, 25]. The following paper analyzes the predictive power of the RSV index for price dynamics in the search for a good predictor of turning points in the US housing market and a leading indicator of sentiment [5]. Similar studies have confirmed the high correlation between the house price index (the Case–Shiller index) and online public activity, supporting the assumption of a benefit in forecasting prices [26]. The main theoretical contribution of our subsequent work [27] confirms that the freely available information regarding Google user searches can provide an in-depth insight into enriching the generally accepted statistics on supply and demand in the real estate market. This paper postulates that RSV data is an additional demand-side variable for improving the accuracy of real estate price forecasts. A review of the subject revealed that this has emerged as a promising avenue for integrating traditional data with emerging data sources, thereby offering novel insights into the sentiment of prospective buyers within the real estate market [23–25, 28].

3. Methods

In the first stage of the research, data for the complex keyword 'buy a flat' (RTSV index from Google Trend) was downloaded in package 'gtrendR' [29] in R software. In order to remove seasonal variability and hyper-distorted values from the RSV series that was created in this manner, a smoothing procedure was applied that used local regression LOESS (local weighted scatterplot smoothing). LOESS is nonparametric in the sense that the fitting technique does not require an a priori specification of the relationship between the dependent and independent variables. Although it is used most frequently as a scatterplot smoother, LOESS can be generalized very quickly to multivariate data; there are also inferential procedures for confidence intervals and other statistical tests [30]. Two methods were used in the next stage of the research: the Granger causality test in the VAR model, and the Bayesian structural time-series model (BSTS).

The causality test in the Granger sense was an integral part of the VAR model that was used. Vector autoregressive (VAR) models only explain endogenous variables based on their past values; they can also consider external variables such as trends or seasonality, but they do not need to explicitly distinguish between endogenous and exogenous variables [6]. VAR models combine the approaches of traditional econometrics with time-series-analysis methods. These models effectively use the data on the processes that shape the variables under study [31]. A component of the VAR model is a test of causality in the Granger sense, which is based on the principle that one variable causes another variable if it contains valuable information in terms of prediction [32].

According to Granger, there are three basic assumptions in his test [33]:
- the future cannot cause the past – only the past and the present can cause the future;
- if independent variable is deterministically related to another, it does not contribute additional information to analysis;
- all causal relationships have constant direction regardless of passage of time.

The Bayesian structural time-series model (BSTS) combines the Kalman filter, spike-and-slab regression, and Bayesian model averaging. This model is defined by the following equations [34]:

$$y_t = \mu_t + \tau_t + \beta^T x_t + \epsilon_t \tag{1}$$

$$\mu_t = \mu_{t-1} + \delta_{t-1} + u_t \tag{2}$$

$$\delta_t = \delta_{t-1} + v_t \tag{3}$$

$$\tau_t = -\sum_{j=1}^{S-1} \tau_{t-j} + w_t \tag{4}$$

where y_t is the observed value of the target time series at time t, μ_t is the trend level, δ_t is the trend slope, τ_t is the seasonal component, and $\beta^T x_t$ is the regression component incorporating exogenous predictors. The error terms ϵ_t, u_t, v_t, and w_t are independent Gaussian random variables with variances σ_ϵ^2, σ_u^2, σ_v^2, and σ_w^2, respectively.

This technique relies on the BSTS model and the following procedure [9]:
- a linear model of the results that are noted in the experimental group and the control group before and after the intervention is developed;
- the differences in the results that are noted in the experimental group and the control group before and after the intervention are estimated;
- the observed differences (including cumulative differences) are estimated.

Overall, the BSTS model helps to estimate an existing intervention's effects on the time series of a phenomenon under study. Realistically, the model compares the actual time series with a counterfactual forecast; i.e., a prediction from the situation that might have occurred if such interventions had not occurred. Thus, the model can determine the difference between the actual and forecast time series in BSTS to determine the impact of that intervention.

4. Results and Discussion

4.1. Preliminary Analysis

The first stage of the research attempted to analyze the results of Google Trend keyword searches in Poland in the housing market context. The research used the package 'gtrendR' [29] in R software. Using a specialized package in R software allowed

for more-customized and more-selective efforts to retrieve data that was related to keyword searches. The data collection required several restrictions in the application in order to obtain accurate results for people who wanted to buy a flats, were located in Poland, used Polish on the Internet, and searched within the real estate sector (Table 1).

Table 2. Overview of basic Google Trend keyword terms in Poland

Country	Keyword	Geo	Language code	Code	COVID-19 date
Poland	kupię mieszkanie [buy a flat]	PL	pl	Real estate 29	4.03.2020

The important point about this kind of data (Google Trend) is that its value ranges from 0 to 100 and does not specify a nominal count of typing specific words in a particular country or language. The search results are presented in the form of a so-called relative search volume (RSV); thus, the results are relative and not nominal. The RSV is determined in relation to other searches during the same period, resulting in values where the highest RSV value indicates the maximum interest in the selected keyword; in contrast, the lowest RSV value indicates no or little interest in it. A Google trend of 181 monthly RSV values from 2010 through 2024 was downloaded. Due to the fact that in the next stage of the research, however, the RSV will be compared with the dynamics of apartment prices in Poland (quarterly data); the RSV data was also transformed into a quarterly basis. Figure 1 shows the values that were obtained (on a 0–100 scale) for the relevant keywords in the countries that were analyzed, displaying the mean values (boxplot) and outliers.

Figure 1 shows that the median RSV for Poland was approximately 54, with the interquartile range spanning from around 49 to 62. Outliers are visible at both the upper and lower ends of the distribution; several observations reached the maximum value of 100 (reflecting short periods of peak search interest), while others fall below 30 (representing periods of markedly lower relative attention). Overall, the data suggested relatively stable search intensity in Poland, punctuated by occasional peaks of maximum interest and sporadic drops in activity. Figure 2 shows the dynamics of the RSV index in Poland (2010–2024).

Figure 2 shows the RSV dynamics for the keyword 'buy a flat ('kupię mieszkanie' in Polish) between 2010 and 2024 in Poland. Between 2010 and 2015, the initial decline in the RSV stabilized (following the global mortgage crisis in 2008–2009), and moderate increases in the RSV were evident. In the following years (2016–2021), there was a significant increase in interest in the online housing market with the start of the COVID-19 outbreak; this did not stop the increases to date. The decline in interest began in early 2021, reaching its lowest level in early 2023. At that time, there was a significant increase in the main interest rates, a very high level of inflation, a sharp drop in the availability of mortgage loans, and general uncertainty about the economic situation in Poland. The subsequent increase in the RSV index was

most likely due to the introduction of the government's borrower support program (2% Credit), which significantly increased borrower interest, stimulated actual demand for real estate, and increased housing prices.

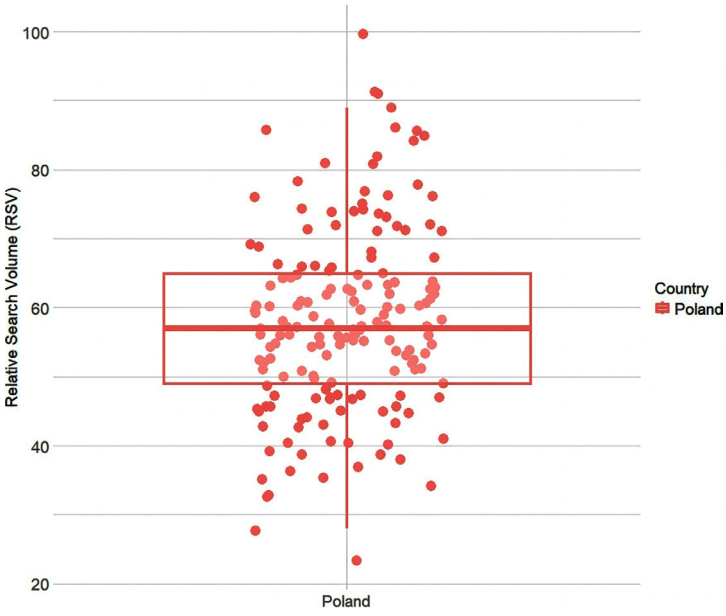

Fig. 1. RSV values for Poland during period of 2010–2024

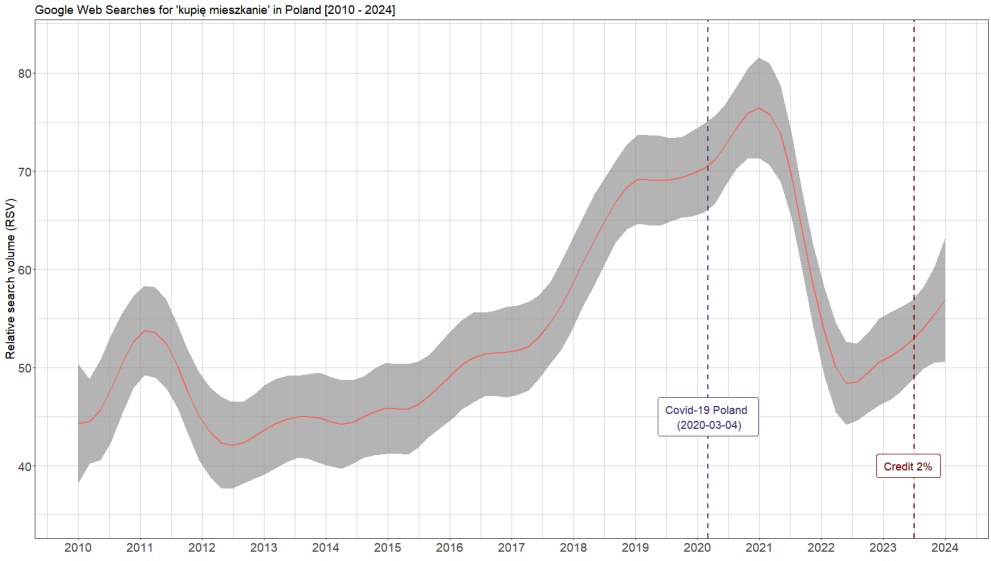

Fig. 2. RSV index dynamics for keyword search 'kupię mieszkanie' ('buy a flat) during period of 2010–2024 in Poland

The dynamics of the RSV indicator (shown in Figure 2) illustrates how important the data that was obtained from Google Trend was as a source of information on the mood and consumer interest in purchasing a dwelling by the Polish public. Interestingly, this indicator's dynamic patterns correlated with important elements (interventions – COVID-19, or the 2% Credit program) of the development of this market in Poland. It seems to be an important tool for tracking current changes in demand in the housing market, as such conclusions can be made with a significant time lag (the date of their publication) in the case of reports on average housing prices. RSV data can help to understand the seasonality of social behavior in the housing market and allow for the forecasting of future trends in transaction prices, as a significant increase in interest generally precedes a significant increase in housing prices.

4.2. Granger Causality Tests

The National Bank of Poland (NBP) database was the primary data source on housing prices; it has maintained a residential property price database (BaRN) since Q3 2006. NBP creates this database from data that is sent by real estate agents and developers who are submitting data to NBP. BaRN includes data from the primary market and the secondary market and offers both prices and transaction prices. The research was based on secondary market transaction prices from 2010–2024 that were averaged for seven cities in Poland: Gdańsk, Gdynia, Łódź, Krakow, Poznań, Warsaw, and Wrocław. At the same time, RSV values for the keyword 'buy a flat' for 2010–2024 were used. Nominal prices (SHP_7), real prices (SHP_7_R), and RSV values are shown in Figure 3. The paper from [3], which forecasted real prices (SHP_7_R) for 2023 (the gray graph in Figure 3), produced a forecast that was virtually in line with the real price trends in 2023. This forecast was made using the RSV indicator from Google Trend (as in this research). It should be noted that the described forecast was based on historical data (2010–2022), and government intervention (2% Credit) was not included in the 2023 projection. Despite this, the 2023 forecast was very similar to the actual behavior of the real prices over this period. This may support the assumption that data from Google Trend (the core of this research) can help predict the housing market's future.

Figure 3 shows the dynamics of three indicators from 2010 through 2024 in Poland. The first indicator (SHP_7) represented the quarterly nominal prices (excluding inflation) of dwellings, while SHP_7_R represented the real prices (taking the inflation levels during the years under study into account). The RSV index represented the current data on public sentiment in the context of the housing market. Due to a different procedure for its creation, the scale from 0 to 100 is on the left side of the graph, while the scale for the housing prices is in Polish currency (PLN) on the right side. The two vertical lines in Figure 3 mark important events for the

Polish housing market: the first line (purple) marks the starting date of the global COVID-19 phenomenon, while the second line (red) denotes local government intervention through the Credit 2% program (i.e., government subsidies for the purchase of housing for the public). The green (SHP_7) and red (SHP_7_R) lines describe the price swaps between 2010 and 2024.

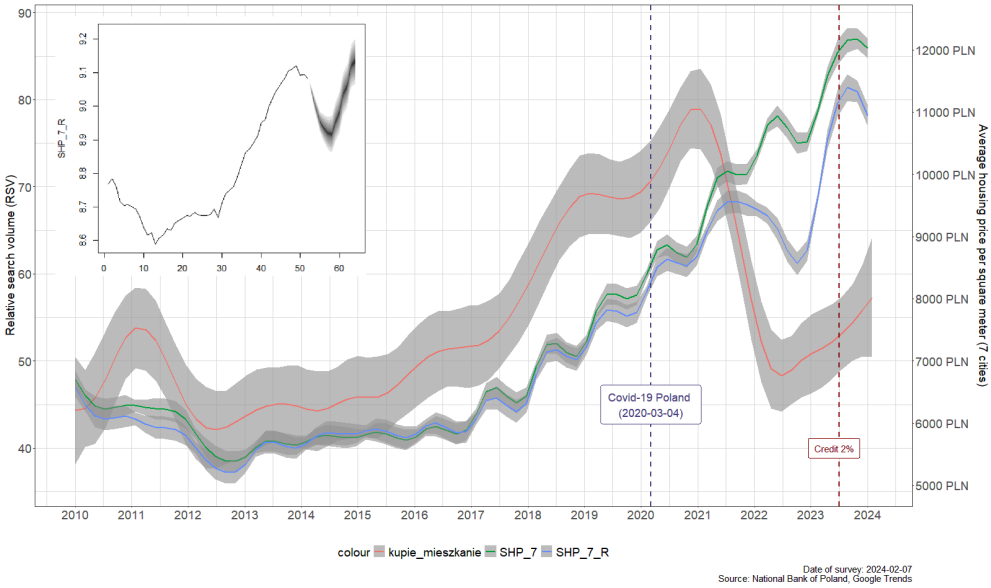

Fig. 3. Average quarterly prices for seven cities in Poland in nominal (SHP_7) and real (SHP_7_R) terms, and RSV index for Google Trend keyword search 'kupię mieszkanie' ('buy a flat')

Between 2010 and 2013, we saw the end of a downward trend in housing prices that started with the global mortgage crisis (2008–2009). Then, nominal and real prices commensurately started an upward trend; so, their upward momentum increased significantly from 2018 onwards. If we look at the RSV indicator at the same time, we can see synergistic behavior with a significant increase in interest in the keyword 'buy a flat' on the web preceding the increases in housing prices by about two years (this can be seen during the period of 2015–2017). The search terms on the web increased knowledge of the phenomenon under investigation and influenced the decision-making process. Many authors [35, 36] have claimed that data resources on the Internet (including social media or various online reviews) unequivocally influence actual decision-making, resulting in financial transactional decisions. One may therefore be interested in the process of housing acquisition, which is preceded by an increase in keyword searches that are related to the acquisition of a flat or house. The genesis of such a research concept is precisely based on the lag of

housing prices concerning RSV dynamics that can be observed in Figure 3. After the onset of the epidemic in Poland (COVID-19), the RSV index decreased significantly, while nominal prices did not fall; however, real prices recorded a sharp decline at the turn of 2022/2023 just to rise sharply again.

The results of Granger causality tests are often misinterpreted as evidence of true causality; in fact, they only indicate a predictive relationship – not necessarily a causal one [37]. Therefore, the rest of this paper will focus more on the predictive effects of this test than on its causal effects. This article attempts to answer the question: Can we measure the predictive power of the time series under study (SHP_7 or RSV)? In the first stage of the study, the two time series (RSV and SHP_7) were prepared for further analysis; they were smoothed using the LOESS function, and the data were normalized so that their values were based on a common scale. The results of these activities are shown in Figure 4.

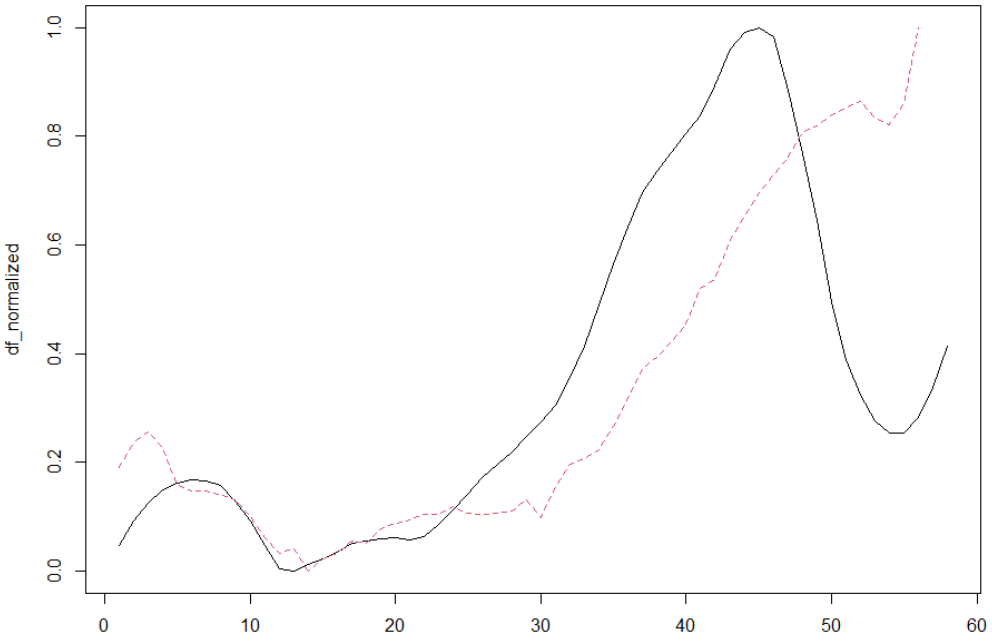

Fig. 4. Dynamics of standardized house prices (red line) and RSV index (black line) from 2010 through 2024 in Poland

The Granger causality test was analyzed using the vector autoregressive modeling (VAR) model, but it only detects predictive relationships – not 'true' causality. The test checks whether information from X improves the forecasting of Y beyond its own past, which Granger described as a precedence or temporal relationship rather than causality in the strict sense.

Table 2 shows the results of the Granger test.

Table 2. Granger test

Time series predictive relationship	*F*-test	*p*-value	Significance
smoothed_SHP_7_R → smoothed_RSV	0.2423	0.9127	–
smoothed_RSV → smoothed_SHP_7_R	2.659	0.04545	*

Note: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

The results that are shown in Table 2 indicate that Granger's *p*-value of causality (smoothed_SHP_7_R → smoothed_RSV) was significantly greater than 0.05. This meant that housing prices in Poland did not improve the predictability of keyword search on the web between 2010 and 2024. In contrast, the reverse situation (smoothed_RSV → smoothed_SHP_7_R) was statistically significant, which meant that the past values of smoothed_RSV provided additional predictive information about apartment prices in Granger's sense. Changes in the RSV index precede changes in the SHP_7 index, because an increase in interest in information about flats on the Internet is observed first, followed by purchasing decisions. The results of this research showed that an important element of market analysis could be an analysis of the society's online behavior (RSV index), which allowed for an examination of the then-current sentiment of the society toward the purchases of flats or investments in the housing market. In the next study stage, the effect of the intervention (an important event) on the dynamics of the RSV index will be analyzed.

4.3. Analysis of the Intervention Effect (COVID-19)

The next stage of the research was undertaken to analyze the RSV index's response to significant changes in the housing market environment. Suppose this index is a good measure of the public's mood or social preferences changes regarding this market. In this case, a significant intervention (environmental collision, geopolitical event) should be noted in the changes in people's behaviors; that is, the RSV index level should change significantly. This article will examine the impact of a global phenomenon (the COVID-19 epidemic) on the number of Internet searches for the keyword 'buy a flat'. At this research stage, the analyses were conducted using a Bayesian structural time-series model (CasualImpact). This model allowed us to answer what would have happened if such an intervention had not occurred. As a result, two time-series trajectories were observed: one representing the historical trajectory of the analyzed phenomenon, while the other predicted the direction of the trajectory had the intervention not occurred. This model generalized the widely used difference-in-differences approach to the time series by explicitly modeling the observed counterfactual time series before and after the intervention. More extensively on this topic can be found in the publication by the authors of the CasualImpact package [9].

The study assumed that the intervention that was analyzed was the COVID-19 epidemic, while the time series in question were the results of the RSV indicator between 2010 and 2024 in Poland. Table 3 shows the results of the study.

Table 3. CausalImpact analysis of intervention for smoothed_RSV – COVID-19 (4.03.2020, Poland)

Effect metric	Effect measure	
	average	cumulative
Actual	0.53	6.84
Prediction (s.d.)	1.2 (0.082)	15.6 (1.069)
Absolute effect (s.d.)	–0.67 (0.082)	–8.74 (1.069)
Relative effect (s.d.)	–56% (3.1%)	–56% (3.1%)

Posterior tail-area probability *p*: 0.001
Posterior probability of casual effect: 99.9%

An analysis of the effect of the intervention (COVID-19) on the RSV index in Poland between 2010 and 2024 (Table 3) showed that, after the intervention, the mean value of the response variable was 0.53. In contrast, the mean predicted value would have been 1.2 if such an intervention had not occurred. This meant that, if there had been no intervention factor in the form of the global COVID-19 epidemic phenomenon, the model would have predicted a further increase in the RSV variable. This represented a further increase in the public's interest in searching for housing-related words on the Internet. The results that are presented here confirmed the previous assumptions that the RSV indicator responded well to a change in the market environment and could show the public's response to the intervention as defined. The analyzed causal effect of the intervention on the examined time series existed, and there was a slight chance ($p = 0.001$) of obtaining this effect by chance; this confirmed the statistical significance of the intervention's causal effect. The results of the causality analysis are best shown by the following chart (Fig. 5).

Figure 5 shows three main panels: original, pointwise, and cumulative. The first panel shows the dynamics of the RSV index (solid line) and its prediction if the intervention (gray vertical line) had not happened (dashed line). The graphical interpretation of the results that are presented in Table 3 confirms that the projected trend of the RSV rate would not have changed its direction of growth had the intervention not occurred (first panel – original). The second panel (pointwise) shows the difference between the actual values (black line in the first panel – original) and the forecast values (blue line in the first panel – original). A clear stabilization of the differences can be observed until the intervention (2% Credit), followed by a sharp

negative reaction; this indicates the actual effect of the intervention. The scale of the deviation is high, which indicates a sustained cumulative effect. The last panel (cumulative) shows the cumulative effect of the intervention. Before the intervention (gray line), this effect was 0, while after the intervention, the line dropped rapidly, and accumulation in the subsequent periods was visible.

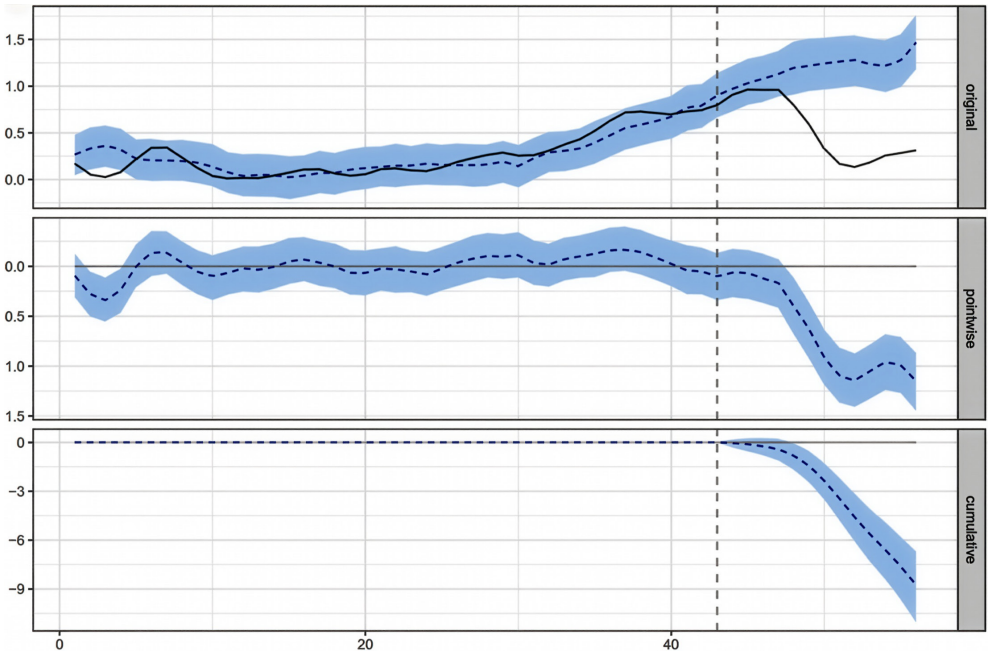

Fig. 5. CasualImpact with Bayesian structural time-series model where impact was date of COVID-19 (4.03.2020, Poland)

These calculations strongly supported the thesis of the value of the RSV indicator in modeling phenomena on the housing market, as it is an indicator that can forecast future investment decisions on the housing market; i.e., the future sales or purchases of a dwelling. At the same time, RSV's great advantage is that it reacts quickly to significant environmental changes. As a result, its predictive value for the future functioning of this market seems quite significant.

5. Conclusion

In the classic approach, housing market analysis is based on official data that is related to the levels of housing (house) prices or rental housing prices, together with a range of data that defines the environment of this market (both in macroeconomic

and local terms). This article considers the use of relative search volume (RSV) in market research to add to classic data elements that diagnose the mood or social preferences regarding the housing market. By creating an RSV index, Google Trends allows one to observe which keywords are relevant to the public during a given time-period. As a result, it is possible to see whether those keywords that are related to housing or those portals that offer housing sales opportunities are the subject of a significant increase or decrease in interest.

The methods that were used in this article were the Granger causality test and the Bayesian structural time-series model. The study showed that RSV dynamics may have been among the causes of housing price movements.

At the same time, it was shown that the global intervention in the form of the COVID-19 pandemic may have influenced previous RSV trends, thus directing public interest toward declines in the real estate market. As a result, it appears that the RSV index may have been a measure of consumer sentiment in the housing market or a predictor of potential changes in the housing market.

It seems that integrating apartment price data and data from modern online media (Google Trend) can contribute to a better understanding of the housing market and allow for the more accurate forecasting of housing prices. This could result in a more accurate data set for investment decisions for individuals and institutions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interests

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The public data in this article include:
- housing prices: www.nbp.pl;
- RSV for keywords ‘kupię mieszkanie’: https://trends.google.com/trends/.

Use of Generative AI and AI-Assisted Technologies

During the preparation of this work, the author used AI-assisted technologies to assist with language editing and improve clarity. After using this tool, the author reviewed and edited the content and take full responsibility for the final version of the paper.

References

[1] Yang S., Santillana M., Kou S.C.: *Accurate estimation of influenza epidemics using Google search data via ARGO*. Proceedings of the National Academy of Sciences of the United States of America, vol. 112(47), 2015, pp. 14473–14478. https://doi.org/10.1073/pnas.1515373112.

[2] Kokot S.: *The analysis of differences in residential property price indices*. Real Estate Management and Valuation, vol. 22(3), 2014, pp. 14–27. https://doi.org/10.2478/remav-2014-0023.

[3] Bełej M.: *Predicting housing price trends in Poland: Online social engagement – Google Trends*. Real Estate Management and Valuation, vol. 31(4), 2023, pp. 73–87. https://doi.org/10.2478/remav-2023-0032.

[4] Bulczak G.M.: *Use of Google Trends to predict the real estate market: Evidence from the United Kingdom*. International Real Estate Review, vol. 24(4), 2021, pp. 613–631. https://doi.org/10.53383/100332.

[5] Dietzel M.: *Sentiment-based predictions of housing market turning points with Google Trends*, [in:] *22nd annual European Real Estate Society Conference*, ERES, Istanbul 2015. https://doi.org/10.15396/eres2015_3.

[6] Pfaff B.: *VAR, SVAR and SVEC models: Implementation within R package vars*. Journal of Statistical Software, vol. 27(4), 2008, pp. 1–32. https://doi.org/10.18637/jss.v027.i04.

[7] Wickham H., Averick M., Bryan J., Chang W., D'Agostino McGowan L., François R., Grolemund G., Hayes A., Henry L., Hester J., Kuhn M., Lin Pedersen T., Miller E., Bache S.M., Müller K., Ooms J., Robinson D., Seidel D.P., ..., Yutani H.: *Welcome to the tidyverse*. Journal of Open Source Software, vol. 4(43), 2019, 1686. https://doi.org/10.21105/joss.01686.

[8] Hackenberger B.K.: *R software: unfriendly but probably the best*. Croatian Medical Journal, vol. 61(1), 2020, pp. 66–68. https://doi.org/10.3325/cmj.2020.61.66.

[9] Brodersen K.H., Gallusser F., Koehler J., Remy N., Scott S.L.: *Inferring causal impact using Bayesian structural time-series models*. Annals of Applied Statistics, vol. 9(1), 2015, pp. 247–274. https://doi.org/10.1214/14-AOAS788.

[10] Adams Z., Füss R.: *Macroeconomic determinants of international housing markets*. Journal of Housing Economics, vol. 19(1), 2010, pp. 38–50. https://doi.org/10.1016/j.jhe.2009.10.005.

[11] Drachal K.: *Causality in the Polish housing market: Evidence from biggest cities*. Financial Assets and Investing, vol. 9(1), 2018, pp. 5–20. https://doi.org/10.5817/FAI2018-1-1.

[12] Lis P.: *Wahania cykliczne rynków mieszkaniowych: Aspekty teoretyczne i praktyczne*. Wydawnictwo Adam Marszałek, Toruń 2012.

[13] Tamari S., Katoshevski R., Karplus Y., Dinero S.C.: *Urban tribalism: Negotiating form, function and social milieu in Bedouin towns, Israel*. City, Territory and Architecture, vol. 3(1), 2016, 2. https://doi.org/10.1186/s40410-016-0031-3.

[14] Kodila-Tedika O., Asongu S.A.: *Tribalism and financial development*. African Governance and Development Institute Working Paper, WP/15/018, May 30, 2015. https://doi.org/10.2139/ssrn.2612429.

[15] Muellbauer J., Murphy A.: *Housing markets and the economy: The assessment*. Oxford Review of Economic Policy, vol. 24(1), 2008, pp. 1–33. https://doi.org/10.1093/oxrep/grn011.

[16] Hendershott P.H., Abraham J.M.: *Patterns and Determinants of Metropolitan House Prices, 1977–91*. NBER Working Paper, no. 4196, National Bureau of Economic Research, Cambridge, MA, 1992. https://doi.org/10.3386/w4196.

[17] Marsh A., Gibb K.: *Uncertainty, expectations and behavioural aspects of housing market choices*. Housing, Theory and Society, vol. 28(3), 2011, pp. 215–235. https://doi.org/10.1080/14036096.2011.599182.

[18] Olszak C.M., Ziemba E.: *The information society development strategy on a regional level*. Issues in Informing Science and Information Technology, vol. 6, 2009, pp. 213–225. https://doi.org/10.28945/1054.

[19] Duff A.S.: *Information Society Studies*. Routledge, London 2013. https://doi.org/10.4324/9781315812250.

[20] Colombo A.W., Karnouskos S., Yu X., Kaynak O., Luo R.C., Shi Y., Leitão P., Ribeiro L., Haase J.: *A 70-year Industrial Electronics Society evolution through industrial revolutions: The rise and flourishing of information and communication technologies*. IEEE Industrial Electronics Magazine, vol. 15(1), 2021, pp. 115–126. https://doi.org/10.1109/MIE.2020.3028058.

[21] Bełej M.: *Does Google Trends show the strength of social interest as a predictor of housing price dynamics?* Sustainability, vol. 14(9), 2022, 5601. https://doi.org/10.3390/SU14095601.

[22] Limnios A.C., You H.: *Can Google Trends improve housing market forecasts?* Curiosity: Interdisciplinary Journal of Research and Innovation, vol. 2, 2021, pp. 1–18. https://doi.org/10.36898/001c.21987.

[23] Castelnuovo E., Tran T.D.: *Google it up! A Google Trends-based uncertainty index for the United States and Australia*. Economics Letters, vol. 161, 2017, pp. 149–153. https://doi.org/10.1016/j.econlet.2017.09.032.

[24] Huarng K.-H., Hui-Kuang Y.T., Rodriguez-Garcia M.: *Qualitative analysis of housing demand using Google Trends data*. Economic Research-Ekonomska Istraživanja, vol. 33(1), 2020, pp. 2007–2017. https://doi.org/10.1080/1331677X.2018.1547205.

[25] Matias Y.: *Nowcasting with Google Trends*, [in:] Kurland O., Lewenstein M., Porat E. (eds.), *String Processing and Information Retrieval*: *20th International Symposium, SPIRE 2013, Jerusalem, Israel, October 7–9, 2013, Proceedings*, Lecture Notes in Computer Science, vol. 8214, Springer, Cham 2013, p. 4. https://doi.org/10.1007/978-3-319-02432-5_4.

[26] Askitas N.: *Trend-spotting in the housing market*. Cityscape, vol. 18(2), 2016, pp. 165–178. https://doi.org/10.2139/ssrn.2675484.

[27] Rizun N., Baj-Rogowska A.: *Can web search queries predict prices change on the real estate market?* IEEE Access, vol. 9, 2021, pp. 70095–70117. https://doi.org/10.1109/ACCESS.2021.3077860.

[28] Bełej M.: *Analysis of the relationship between COVID-19 infections and web-based housing searches*. Real Estate Management and Valuation, vol. 30(4), 2022, pp. 89–102. https://doi.org/10.2478/remav-2022-0031.

[29] Massicotte P., Eddelbuettel D.: *gtrendsR: Perform and display Google Trends queries*. R package version, https://cran.r-project.org/web/packages/gtrendsR/index.html [access: March 31, 2025].

[30] Jacoby W.G.: Loess: *A nonparametric, graphical tool for depicting relationships between variables*. Electoral Studies, vol. 19(4), 2000, pp. 577–613. https://doi.org/10.1016/S0261-3794(99)00028-1.

[31] Kusideł E.: *Modele wektorowo-autoregresyjne VAR: Metodologia i zastosowania*. Absolwent, Łódź 2000.

[32] Hmamouche H.: *NlinTS: An R package for causality detection in time series*. The R Journal, vol. 12(1), 2020, pp. 21–31. https://doi.org/10.32614/RJ-2020-016.

[33] Granger C.W.J.: *Investigating causal relations by econometric models and cross-spectral methods*. Econometrica, vol. 37(3), 1969, 424. https://doi.org/10.2307/1912791.

[34] Scott S.L., Varian H.R.: *Predicting the present with Bayesian structural time series*. International Journal of Mathematical Modelling and Numerical Optimisation, vol. 5(1–2), 2014, pp. 4–23.

[35] Cajita M.I., Whitehouse E., Budhathoki C., Hodgson N.: *Association between Internet use and decision-making preference in older adults*. Gerontechnology, vol. 14(2), 2016, pp. 97–104. https://doi.org/10.4017/gt.2016.14.2.008.00.

[36] Lagan B.M., Sinclair M., Kernohan W.G.: *What is the impact of the Internet on decision making in pregnancy? A global study*. Obstetric Anesthesia Digest, vol. 32(4), 2012, pp. 211–212. https://doi.org/10.1097/01.aoa.0000422689.35910.77.

[37] Maziarz M.: *A review of the Granger-causality fallacy*. The Journal of Philosophical Economics: Reflections on Economic and Social Issues, vol. 8(2), 2015, pp. 86–105. https://doi.org/10.46298/jpe.10676.