

Data Descriptor

Not peer-reviewed version

Modelling the Effect of Investor Sentiments on the Housing Market and Stock Market in Saudi Arabia using Linear ARDL Model

Abdullah Alawajee and Mohd Ismail *

Posted Date: 15 October 2024

doi: 10.20944/preprints202410.0989.v1

Keywords: Investor sentiment; stock returns; real estate returns; ARDL; Saudi Arabia



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data Descriptor

Modelling the Effect of Investor Sentiments on the Housing Market and Stock Market in Saudi Arabia using Linear ARDL Model

Abdullah Alawajee and Mohd Tahir Ismail

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

* Correspondence: perfectexperts@gmail.com

Abstract: The paper presents a comprehensive study of the relationship between investor sentiment and stock market returns in Saudi Arabia from 2009 to 2022. The study uses linear autoregressive distributed lag models (ARDL) to analyze the complex dynamics between variables. ARDL models are particularly valuable for capturing both long- and short-term relationships between variables and allow for dynamic adjustments over time. The study shows that a positive change in investor sentiment leads to higher stock market returns after a lag of two months, but this positive effect reverses in the long-run, suggesting the possibility of overreaction and subsequent adjustments. Money supply has a consistent positive impact on stock returns over time, while real estate market returns have limited short-term impacts on investor sentiment and macroeconomic factors. The ARDL modeling approach is essential to capture the dynamic interaction between investor sentiment and stock returns and reveals the temporary impact of sentiment on market performance. The study extends its analysis to construct a composite investor sentiment index using principal component analysis, tests the stationarity of stock returns and sentiment index through extended Dickey-Fuller and Phillips-Perron unit root tests, and determines the optimal lag length for time series modeling. The ARDL modeling results for stock market returns and real estate returns provide insights into the short- and long-term relationships between investor sentiment, macroeconomic variables, and market returns; in addition, the study provides important insights into the interaction between investor sentiment and stock market returns in Saudi Arabia, with implications for regulatory policy and investment management. The results demonstrate the impact of investor psychology on market performance, which has practical implications for understanding and managing behavioral risk. The study adopts an integrated approach, combining rigorous ARDL models with economic intuition and practical relevance, which is expected to better understand the dynamic interaction between sentiment and returns.

Dataset: <https://www.saudiexchange.sa/>; <https://www.stats.gov.sa/en/843>,
<https://www.sama.gov.sa/en>

Dataset License: license under which the dataset is made available (CC-BY)

Keywords: Investor sentiment; stock returns; real estate returns; ARDL; Saudi Arabia

1. Summary

Investor sentiment plays a vital role in stock market performance. It refers to the overall attitude, mood, and opinion of investors toward a specific stock or the market as a whole. While fundamental analysis and economic indicators provide valuable insights, investor sentiment often has a significant impact on short-term market movements [1]. By staying informed and monitoring market sentiment, investors can make decisions that are more informed and have more confidence in navigating a complex stock market. Investor sentiment is influenced by various factors, including market news, company performance, and general economic conditions. Positive sentiment can lead to increased

buying activity and push up stock prices, while negative sentiment can lead to selling pressure and lower prices [2]. The collective sentiment of investors can add momentum to the market, influencing trading volume and price trends. Investors must be aware of the impact of sentiment on market movements and consider both rational analysis and emotional factors when making investment decisions. While it is impossible to predict market sentiment, understanding investor psychology accurately can provide valuable insights into potential market trends.

The Autoregressive Distributed Lag (ARDL) model is a powerful tool in econometrics for analyzing both long-term and short-term dynamics between variables. By including both lagged values of the dependent variable and lagged values of the independent variables, the ARDL model can capture complex relationships that may be missed by traditional regression models [3].

One of the main advantages of using linear ARDL analysis is its ability to account for both intrinsic and dynamic aspects of the data. By including lagged values of the variables, the model can control for autocorrelation and capture dynamic adjustments that occur over time. This makes the ARDL model particularly useful for analyzing time series data, where the variables may be correlated and evolve over time.

In addition, the ARDL model allows researchers to test whether there is a long-term relationship between the variables. By testing the significance of the coefficients of the lagged variables, researchers can determine whether there is a stable equilibrium relationship between the variables in the long-run. This can provide valuable insights into the underlying dynamics of the data and aid in policy decisions [4].

In addition, the ARDL model is very flexible and can accommodate different types of data and relationships. Whether the data exhibits stationary or non-stationary behavior, the ARDL model can be applied without hesitation. This versatility makes ARDL models a valuable tool for researchers working with different datasets and research questions.

Linear ARDL analysis is a complex but easy-to-use method for analyzing the dynamics between variables in econometric studies. By incorporating lagged values of variables and testing long-term relationships, ARDL models provide valuable insights into the complex dynamics of the data. Researchers from a variety of fields can benefit from using ARDL models to uncover hidden relationships and inform their decision-making processes.

In behavioral economics, investor sentiment reflects general investor attitudes, which are determined by psychological factors, experience, or environmental [5,6]. Through contagion effects, investor sentiment is incorporated into asset valuations and expected returns. This study explores this relationship in the unique context of the Saudi Arabian stock market.

The Saudi market provides an interesting context due to the strong presence of retail investors, its dependence on oil revenues, and its sensitivity to regional stability [7]. With more than 6 million active retail investors, the Saudi market has one of the highest rates of retail investor participation in the world [8]. This makes the market more susceptible to sentiment-related mispricing. Furthermore, oil exports account for more than 70% of Saudi Arabia's revenue, linking market developments to global oil dynamics [9]. Geopolitical events such as the 2017 Qatar crisis and oil price volatility also affect market volatility.

The role of investor sentiment in driving stock market returns has received considerable attention in behavioral finance research. Theoretical models suggest that sentiment can affect returns through two main channels [5,7]. The price pressure hypothesis assumes that sentiment directly affects stock prices, with optimistic investors pushing prices higher and pessimistic investors pushing prices lower. The risk premium perspective assumes that sentiment affects expected returns by changing risk perceptions in pricing.

Empirically, the evidence on the relationship between investor sentiment and returns remains inconclusive. Previous seminal research (see also [10]) provided one of the earliest large-sample studies showing that high investor sentiment predicts lower market returns. By applying an error correction model, they found that sentiment has a long-term effect on returns. Subsequent studies have found that sentiment has a significant effect on various international markets. [11] Showed that consumer confidence, as a sentiment indicator, can predict returns in 18 countries. [5] documented a

negative relationship between sentiment and returns in six major stock markets. [5,12] showed at a cross-sectional level that stocks that are difficult to arbitrate and difficult to value are most affected by investor sentiment. Subsequent work confirmed these results in other markets [13]. also demonstrated that highly subjective stocks are more susceptible to sentiment-related mispricing. [14] showed that the sentiment effect is stronger for stocks with high idiosyncratic volatility.

However, other studies have provided conflicting evidence on the role of sentiment [14,15] found that sentiment has no consistent predictive power for US stock returns [16] document that the causal effect of sentiment on returns in European markets excluding the UK is not significant. [17,18] show that sentiment has a limited impact on US sector returns. While emerging markets appear to be more vulnerable (Schmeling, 2009), Li and Kong (2017) recently found that sentiment does not play a significant role in determining Chinese stock returns.

Empirical evidence from the Middle East remains scarce. [19] conducted an early qualitative study on how investor sentiment affected market activity in Saudi Arabia during the 2008 global financial crisis. Using more advanced techniques, [20] use a Markov switching model to demonstrate that investor sentiment mechanisms help predict returns in the UAE. For the Saudi market, [21] recently studied the impact of sentiment shocks through VAR models, while [22] focused on the predictability of returns using quantile regression.

Methodologically, the autoregressive distributed lag (ARDL) technique is increasingly use for sentiment-return analysis, [23] use the ARDL model to capture the dynamic interaction between investor sentiment and industry stock returns in Saudi Arabia. ARDL and nonlinear ARDL have also used to study the relationship between real estate and stock returns and sentiment in India [24]. In addition to linear methods, Markov switching models have shown promise for modeling emotional states [25].

Recent studies have used sophisticated machine learning techniques. [9] show significant predictability of sentiment at the industry level in China using a long short-term memory (LSTM) approach. [26] applied a random forest algorithm to demonstrate the significant impact of investor attention on Turkish stock returns. By combining sentiment with technical indicators, [13,17] demonstrated improved accuracy of stock return forecasts using deep learning neural networks.

Despite the growing body of research, there remains a gap in the contextual role of investor sentiment in emerging Middle Eastern markets such as Saudi Arabia. Previous studies have limitations in using subjective sentiment measures, simplified empirical models without temporal dynamics, or lack of asymmetric and nonlinear analysis. Advanced computational methods are also not fully utilized. This highlights the need to use sophisticated time series techniques tailored to the Saudi market context to provide strong evidence.

This study uses monthly data from September 2009 to September 2022, covering large fluctuations in the Saudi stock market. The stock return series is from the Tadawul All-Share Index, which reflects the overall market performance.

The composite investor sentiment index is constructed from ten basic financial market variables using principal component analysis. The variables include trading volume, market turnover, number of shares traded, number of trades, stock price volatility, price increase to price decrease ratio, new investor subscription, investor asset size, number of companies with prices above the 20-day moving average, and number of companies with prices above the 50-day average ([5,7,26])

The principal component analysis extracts the relevant information from the ten indicators into orthogonal principal components. The first principal component explains the largest variation in the data and represents the composite sentiment index. Previous studies have shown that this approach can effectively summarize the sentiment information in the indicators compared to simple averaging ([5,7]).

The stationarity of stock returns and sentiment indices is test using extended Dickey-Fuller and Phillips-Perron unit root tests. The identification of the unit root guides the selection of the long-term cointegration framework.

The absence of cointegration would justify modeling the relationship via an unconstrained VAR in first differences. The presence of cointegration requires the use of a vector error correction model

(VECM) to explain the long-run equilibrium or an autoregressive distributed lag (ARDL) approach. ARDL models are often used to study the dynamic interactions between short-run and long-run time series ([23]). The ARDL framework estimates short-run dynamics and long-run equilibrium simultaneously in a databased general-to-specific modeling approach.

The ARDL model has the following form:

$$\Delta SR_t = \alpha + \beta_1 SR_{t-1} + \beta_2 SI_{t-1} + \beta_3 MS_{t-1} + \beta_4 IPI_{t-1} + \beta_5 CCI_{t-1} + \beta_6 GEPU_{t-1} + \sum \gamma_i \Delta SR_{t-i} + \sum \delta_j \Delta SI_{t-j} + \sum \delta_j \Delta MS_{t-j} + \sum \delta_j \Delta IPI_{t-j} + \sum \delta_j \Delta CCI_{t-j} + \sum \delta_j \Delta GEPU_{t-j} + \epsilon_t \quad (1)$$

Where SR and SI are stock returns and investor sentiment; MS, IPI, CCI and GEPU are control variables; Δ represents the first-order difference to capture short-term dynamics; α is a constant; β_1 and β_2 represent long-term multipliers; γ and δ are short-term coefficients; ϵ_t is the error term [15].

The best model is selected by sequential elimination and diagnostic tests. The residuals are verified for normality, serial correlation, heteroscedasticity and stability. CUSUM and CUSUMSQ tests check parameter stability. The significance of the error correction term confirms cointegration.

This rigorous empirical modeling approach will provide targeted insights into the relationship between investor sentiment and Saudi stock returns. It goes beyond simple correlations or qualitative surveys, thus improving on the limitations of previous Saudi research. Modeling both short-term and long-term dynamics is also an advance as it can capture combined effects based on the data.

The findings will have important practical implications for investors, managers, and policymakers in Saudi Arabia. Evidence that sentiment plays an important role will highlight the need to curb excessive risk-taking during market uptrends. This can help inform the design of circuit breakers, margin-lending limits, and other preventive measures. Portfolio managers can improve their market timing skills and alpha through sentiment analysis. Overall, targeted insights specific to the Saudi context will help better understand and manage behavioral risk.

This study attempts to address this gap through rigorous time series analysis. The ARDL framework allows for databased exploration of both long-term and short-term components. Such targeted insights are important for Saudi market participants.

2. Data Description

This section presents the key empirical findings from the time series analysis examining the relationship between investor sentiment and stock market returns as well as real estate market returns in Saudi Arabia. The results are structured around the research objectives and employ a multitude of statistical techniques including unit root tests, cointegration tests, ARDL models and various diagnostic tests.

2.1. Data

The sample of the current study is a daily data from September 2009 to September 2022 to cover enough sample in order to estimate reliable results. The stock market returns are estimated from the Tadawul All Share Index and the real estate market returns from real estate price index. The Tadawul All Share Index is provided by the Saudi Arabia stock exchange (Tadawul: <https://www.saudiexchange.sa/>) and real estate price index data is provided by 'General Authority for Statistics of the Kingdom of Saudi Arabia (<https://www.stats.gov.sa/en/843>).

The data series of international crude oil is taken from the World Bank-Data Bank. The International Crude Oil Volatility series will be created from the crude oil price using the GARCH (p,q) model. As for the control variables, the data on money supply (MS) is gleaned from Saudi Arabia's central bank (<https://www.sama.gov.sa/en>).

2.1.1. Variables Description

Dependent Variables (Market Returns):

- Stock Market Return;
- Real Estate Market Return;

Independent Variables (Sentiment Index Proxies):

1. Housing Market;
2. Stock Market;
3. Energy Market.

Control Variables (Macroeconomy).

A composite investor sentiment index was construct from ten underlying financial market variables using principal component analysis (PCA).

The included variables were: international crude oil price (ICOP), international crude oil volatility (ICOV), real estate primary trading value (REPTVA), real estate secondary trading value (RESTVA), real estate primary trading volume (REPTVO), real estate secondary trading volume (RESTVO), Tadawul All Share Index (TASI), Tadawul Energy Index (TEI), Tadawul real estate management and development trading volume (TREMDITVO), and Tadawul real estate management and development trading value (TREMDITVA).

2.2. Models Estimation

The study used model, the Autoregressive distributive lag model (ARDL):

$$y_t = \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i x_{t-i} + \varepsilon_t \quad (2)$$

In the given equation (2), y is the outcome variable at time t . y_{t-i} is the value of each lag i (from 1 to n) of the dependent variable. The p in equation is the indicator of number of lags followed by the endogenous variable and q expressing the lags number of exogenous variables. Further, x_t is a $k \times 1$ vector of explanatory variable and β is a $k \times 1$ coefficient vector of independent variables. The given a_i is the vector of scalar and ε_t is the equation white noise term with zero mean and finite variance. Given the objective of current study, the basic specification of the ARDL model are as follows:

$$MR_{i,t} = \beta_0 + \sum_{i=1}^p \gamma_1 MR_{t-1} + \sum_{i=0}^q \delta_1 SenT_t + \sum_{i=0}^q \delta_2 ECR_t + \varepsilon_t \quad (3)$$

In the given equation (3.2), MR is the market return i which changes between stock market and housing market returns at time t . On the right-hand side of the equation, β_0 is the equation intercept term and ε_t is the error term with zero mean and finite variance. The p denotes the lag length followed by dependent variable (autoregressive term) and q denotes the independent variables lag length (distributive lags). Further, ECR is capturing control variables at time t . Moreover, γ_1 is obtaining the lag effect ($t-1$) of the dependent variable (market return of stock and real estate) and δ_1 is carrying the coefficient of sentiment index, $SenT$, at time t . Likewise, δ_2 capturing the coefficient estimates of control variables during the sample period at time t .

$$\Delta MR_{i,t} = \rho_0 + \rho_1 MR_{t-1} + \rho_2 SenT_{t-1} + \rho_3 ECR_{t-1} + \sum_{i=1}^p \gamma_i \Delta MR_{t-i} + \sum_{j=0}^q \delta_j \Delta SenT_{t-j} + \sum_{j=0}^q \delta_j \Delta ECR_{t-j} + \mu_t. \quad (4)$$

Similar to Equation (3) in Equation (4), MR is the market return at time t and i changes between stock market return and real estate market return. The given ρ_0 is the intercept term, ρ_1 is capturing the long-run effect of the dependent variable; and $\rho_2 - \rho_3$ are capturing the long-run effects of sentiment index and control variables (exogenous variable), respectively. Further, the variables with Δ are accommodating for the short-run effects. Where the first is obtaining the short-term level and lag effects (j =number of lags followed by $SenT$) of market return at time t . The latter (δ_j) are capturing the short-run level and lagged effects of $SenT$ and ECR at time t . lastly, the equation white noise term is indicate with μ_t . In accordance with [27], the equation (5) specifications can be regroup and summarized as follows:

$$\Delta y_t = \alpha(y_{t-1} + \theta_i X_t) + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + \sum_{i=0}^{q-1} \delta_i \Delta X_{t-i} + \varepsilon_t \quad (5)$$

2.3. Figures, Tables and Schemes

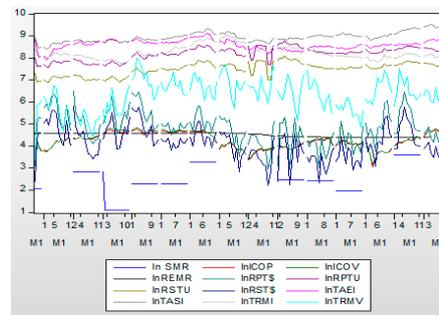


Figure 1. The graph of time series logarithms variables over Sep2009–Sep 2022.

Table 1. This is a table. Variables. A summary of the variables employed in this study is listed in with their respective descriptions.

Table 1. Variables Description.

Variable	Symbol	Description
Dependent Variables (Market Returns)	STR	Stock Market Return
	RER	Real Estate Market Return
Housing Market	Independent Variables (Sentiment Index Proxies)	
	RPTU	Real Estate Primary Trading Volume (Units)
	RPT\$	Real Estate Primary Trading value (\$)
	RSTU	Real Estate Secondary Trading Volume (Units)
	RST\$	Real Estate Secondary Trading value (\$)
Stock Market	TASI	Tadawul All Share Index
	TAEI	Tadawul Energy Index
	TRMI	Tadawul Real Estate Management and Development Index (value)
	TRMV	Tadawul Real Estate Management and Development Index (trading volume)
Energy Market	ICOP	International Crude Oil Price
	ICOV	International Crude Oil Volatility
Control Variables (Macroeconomy)	MS	Money Supply
	IPI	Industrial Production Index
	CCI	Consumer Confidence Index
	GEPU	Global Economic Policy Uncertainty

The PCA results in Table 2 demonstrate that the first principal component explained 31.51% of the total variance and was dominant compared to the other components. An examination of the eigenvectors shows that real estate trading volume variables (lnREPTVO and lnRESTVO) had high positive loadings on PC1, while oil prices and volatility (lnICOP and lnICOV) had negative loadings. The crude oil and real estate factors seem to contrast with each other. Overall, the PCA efficiently summarizes the diverse information from the ten indicators into orthogonal principal components that represent different dimensions of investor sentiment.

Table 2. Principal Component Analysis of Sentiment Index.

Number	Eigenvalue	Difference	Proportion	Cum. Value	Cum. Proportion
1	3.1507	0.1060	0.3151	3.1507	0.3151
2	3.0447	1.1325	0.3045	6.1954	0.6195
3	1.9122	1.0486	0.1912	8.1076	0.8108
4	0.8637	0.3398	0.0864	8.9712	0.8971
5	0.5238	0.1706	0.0524	9.4950	0.9495
6	0.3532	0.2033	0.0353	9.8482	0.9848
7	0.1499	0.1480	0.0150	9.9981	0.9998
8	0.0019	0.0019	0.0002	10.0000	1.0000
9	5E-16	6E-16	0.0000	10.0000	1.0000
10	-3E-17	---	0.0000	10.0000	1.0000

The composite sentiment index (SENT) was constructed by extracting the first principal component. To facilitate valid statistical transformations, an adjustment was made by adding 14.33 to the SENT values before taking natural logarithms.

The optimal lag length for the time series modeling was determined through multiple model selection criteria. The techniques applied were Final Prediction Error (FPE), Akaike Information Criterion (AIC), Schwarz Information Criterion (SC) and Hannan-Quinn Information Criterion (HQIC). The results in Table 3, show that AIC, FPE and LR statistics favored 4 lags, SC indicated 2 lags, while HQIC suggested an intermediate 3 lags. By weighing statistical evidence and economic reasoning, a lag length of 4 was selected for the subsequent ARDL modeling. The choice balances model adequacy with parsimony. The sensitivity analysis considering multiple criteria underscores the complexity of an appropriate lag order selection. Overall, the decision aimed for a dynamically sufficient model that incorporates meaningful economic relationships.

Table 3. Lag Length Selection Criteria.

Lag	LogL	LR	FPE	AIC	SC	HQ
0	1160.293	NA	1.41E-15	-14.32662	-14.19265	-14.27222
1	2558.142	2656.782	7.48E-23	-31.08251	-30.01072	-30.64732
2	2683.122	226.6713	2.92E-23	-32.02636	-30.01675*	-31.21038
3	2769.594	149.3116	1.85E-23	-32.49185	-29.54442	-31.29507*
4	2831.089	100.8369*	1.61e-23*	-32.64707*	-28.76182	-31.0695
5	2868.909	58.7258	1.90E-23	-32.50818	-27.68511	-30.54982
6	2904.885	52.73633	2.32E-23	-32.3464	-26.58552	-30.00725
7	2939.665	47.95729	2.92E-23	-32.16976	-25.47105	-29.44981
8	2979.857	51.9252	3.52E-23	-32.06034	-24.42382	-28.9596

* Indicates lag order selected by the criterion

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

The results in Table 4, indicate a mix of stationary and non-stationary series. Stock market returns (lnSMR) were stationary at level. Real estate returns (lnREMR) and sentiment index (lnSENT) were non-stationary at level but became stationary after first differencing, implying they were integrated of order one, I(1). Other variables like money supply (lnMS) and consumer confidence (lnCCI) demonstrated similar dynamics. The finding of both I(0) and I(1) series provides an appropriate setting for using ARDL-based frameworks.

Table 4. Unit Root Tests.

Variable	Phillips-Perron test		Augmented Dickey-Fuller test		Cum. Proportion
	Adj. t-Stat	Prob.*	t-Statistic	Prob.*	
lnSMR	-3.1461	0.0251	-2.9329	0.0437	Stationary at level
D(lnREMR)	-14.3054	0.0000	-3.9308	0.0023	Stationary at 1st Difference
D(lnSENT)	-18.7670	0.0000	-12.8632	0.0000	Stationary at 1st Difference
D(lnMS)	-14.3095	0.0000	-13.5452	0.0000	Stationary at 1st Difference
lnIPI	-2.9759	0.0393	-3.3243	0.0153	Stationary at level
D(lnCCI)	-3.7496	0.0042	-5.3432	0.0000	Stationary at 1st Difference
lnGEPU	-2.9324	0.0438	-3.3278	0.0152	Stationary at level
lnSMR	-3.1461	0.0251	-2.9329	0.0437	Stationary at level

After the preliminary analysis, ARDL modeling was conducted for stock market returns (LNSMR) as the dependent variable along with sentiment, money supply, industrial production index, consumer confidence index and global uncertainty as predictors. The short-run ARDL model results are presented in Table 5, showing a statistically significant positive coefficient for lagged sentiment (LNSENT(-2)) at the 1% level. This indicates investor sentiment impacts LNSMR with a lag of two periods. Among controls, money supply (LNMS) positively affects LNSMR contemporaneously. Multiple lags of industrial production and consumer confidence also have significant effects.

Table 5. Stock Market ARDL Short- Run.

Dependent Variable: LNSMR				
Selected Model: ARDL(12, 4, 1, 10, 12, 0)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.*
LNSMR(-1)	0.8104	0.0869	9.3221	0.0000
LNSMR(-2)	0.0209	0.1004	0.2078	0.8357
LNSMR(-3)	-0.0597	0.0967	-0.6174	0.5382
LNSMR(-4)	-0.0240	0.0972	-0.2467	0.8056
LNSMR(-5)	0.0686	0.0978	0.7014	0.4845
LNSMR(-6)	0.0592	0.0978	0.6051	0.5463
LNSMR(-7)	-0.0621	0.0971	-0.6398	0.5236
LNSMR(-8)	0.0246	0.0946	0.2604	0.7950
LNSMR(-9)	0.1074	0.0922	1.1646	0.2467
LNSMR(-10)	-0.0794	0.0915	-0.8676	0.3874
LNSMR(-11)	-0.0076	0.0906	-0.0842	0.9331
LNSMR(-12)	-0.1340	0.0705	-1.9015	0.0598
LNSENT	-0.2591	0.3883	-0.6672	0.5060
LNSENT(-1)	-0.2789	0.4834	-0.5770	0.5651
LNSENT(-2)	1.2957	0.4872	2.6594	0.0090
LNSENT(-3)	-0.9780	0.4973	-1.9665	0.0517
LNSENT(-4)	-0.5059	0.4154	-1.2178	0.2259
LNMS	1.5823	0.6732	2.3504	0.0205
LNMS(-1)	-0.8208	0.6460	-1.2706	0.2065
LNIPPI	-0.4808	0.6571	-0.7316	0.4660
LNIPPI(-1)	0.7191	1.0019	0.7177	0.4744
LNIPPI(-2)	-0.1862	1.0972	-0.1697	0.8656
LNIPPI(-3)	-2.1710	1.1584	-1.8741	0.0635
LNIPPI(-4)	1.8360	1.1257	1.6310	0.1057
LNIPPI(-5)	0.0661	1.1456	0.0577	0.9541
LNIPPI(-6)	0.6794	1.1099	0.6122	0.5417

LNIP(-7)	-1.1151	1.1316	-0.9854	0.3265
LNIP(-8)	0.5713	1.1681	0.4891	0.6257
LNIP(-9)	4.0390	1.0942	3.6914	0.0003
LNIP(-10)	-2.0071	1.0914	-1.8390	0.0686
LNCCI	4.5307	20.8547	0.2173	0.8284
LNCCI(-1)	3.3415	56.1661	0.0595	0.9527
LNCCI(-2)	-2.0639	78.7934	-0.0262	0.9791
LNCCI(-3)	-65.8719	85.0868	-0.7742	0.4405
LNCCI(-4)	156.3707	87.9152	1.7787	0.0780
LNCCI(-5)	-181.5774	92.7920	-1.9568	0.0529
LNCCI(-6)	170.8920	90.0547	1.8976	0.0603
LNCCI(-7)	-147.9993	84.4743	-1.7520	0.0825
LNCCI(-8)	58.6632	80.0668	0.7327	0.4653
LNCCI(-9)	110.5427	74.8945	1.4760	0.1428
LNCCI(-10)	-245.9687	67.9468	-3.6200	0.0004
LNCCI(-11)	191.5440	49.0583	3.9044	0.0002
LNCCI(-12)	-64.6304	18.1437	-3.5621	0.0005
LNGEPU	0.0314	0.0734	0.4272	0.6701
C	38.8317	14.1854	2.7374	0.0072
R-squared	0.9000	0.0869	9.3221	0.0000
Adjusted R-squared	0.8607	0.1004	0.2078	0.8357
F-statistic	22.9093	Heteroskedasticity Test: ARCH		
Prob(F-statistic)	0.0000	0.0972	-0.2467	0.8056
Bound Test				
Test Statistic	Value	Sig.	I(0)	I(1)
F-statistic	5.004697	10%	2.08	3.00
k	5	5%	2.39	3.38
		1%	3.06	4.15

Diagnostic tests confirm the absence of serial correlation and heteroscedasticity in the ARDL model residuals. The Bound Test results further verify a long-run cointegrating relationship among the variables at the 1% significance level. This validates proceeding with the long-run ARDL model.

Table 6, show Each coefficient indicates the expected change in the dependent variable (LNSMR) for a one-unit change in the independent variable, holding all other variables constant.

Interpretation of Coefficients:

- For LNSENT: A one-unit increase leads to a decrease of approximately 2.63 units in LNSMR, suggesting that higher sentiment leads to lower market response.

- For LNMS: A one-unit increase results in an increase of approximately 2.76 units in LNSMR, indicating that a higher money supply positively impacts the market.

- For LNIP: A one-unit increase in the industrial production index results in a rise of about 7.08 units in LNSMR, showing strong positive influence.

- For LNCCI: A one-unit increase in consumer confidence leads to a decrease of approximately 44.35 units in LNSMR, suggesting a strong negative relationship here.

- For LNGEPU: This variable has a negligible effect (0.11), indicating it does not significantly influence LNSMR.

- C (Constant): Represents the intercept of the regression model, indicating the value of LNSMR when all independent variables are zero.

- The coefficient CointEq(-1) indicates the error correction term from a cointegration analysis, which is significant ($p < 0.0001$). This suggests that deviations from the long-term equilibrium will correct themselves over time, reflecting a stable long-term relationship among the variables.

The final equation for the error correction model is:

$$EC = \text{LNSMR} - (-2.6334 \cdot \text{LNSENT} + 2.7620 \cdot \text{LNMS} + 7.0757 \cdot \text{LNIPI} - 44.3482 \cdot \text{LNCCI} + 0.1138 \cdot \text{LNGEPU} + 140.8451) \quad (5)$$

Table 6. Stock Market ARDL Long -Run.

Levels Equation				
Case 2: Restricted Constant and No Trend				
Variable	Coefficient	Std. Error	t-Statistic	Prob.*
LNSENT	-2.63343	1.2223	-2.15448	0.0333
LNMS	2.761953	0.830878	3.324137	0.0012
LNIPI	7.07573	2.629152	2.69126	0.0082
LNCCI	-44.3482	14.45311	-3.06842	0.0027
LNGEPU	0.113773	0.275664	0.412723	0.6806
C	140.8451	53.14763	2.650072	0.0092
CointEq(-1)*	-0.2757	0.0454	-6.0753	0.0000

Overall, The long-run coefficients show that sentiment has a significant negative effect on stock returns in the long run, contrasting the short-run dynamics. Money supply and industrial production positively affect stock returns in the long run along with a negative influence from consumer confidence. The long-run causal chain is encapsulated in the level's equation.

The ARDL analysis provides intriguing insights into the temporal effects of investor sentiment on stock returns, with short-term predictive ability but an eventual negative long-run relationship potentially due to market overreaction. Among other variables, money supply demonstrates a consistently positive influence in both the short and long term.

the results in Table 7, Coefficients and Their Significance:

- LNREMR(-1): Coefficient = 0.8850, p-value = 0.0000 (significant)
- LNREMR(-2): Coefficient = 0.0118, p-value = 0.8911 (not significant)
- LNREMR(-3): Coefficient = 0.6177, p-value = 0.0000 (significant)
- LNREMR(-4): Coefficient = -0.5324, p-value = 0.0000 (significant) - Other variables (LNSENT, LNMS, LNIPI, LNCCI, LNGEPU) have high p-values, indicating they are not statistically significant in this model.

1. Intercept (C): - Coefficient = 0.3869, p-value = 0.0823 (marginally significant)

2. Model Fit:

- R-squared = 0.9975 indicates that the model explains approximately 99.75% of the variance in LNREMR.

- Adjusted R-squared = 0.9973, which is a slightly adjusted value that accounts for the number of predictors in the model.

3. Serial Correlation Test:

- The Breusch-Godfrey Serial Correlation LM Test has a p-value of 0.8460, suggesting no serial correlation in the residuals.

4. Heteroskedasticity Test:

- The Heteroskedasticity Test (ARCH) has a p-value of 0.9533, indicating that there is no significant heteroskedasticity.

5. Bound Test for Cointegration: - The F-statistic value of 1.6894 is compared with critical values for different significance levels: - 10% critical value (I(0) = 2.08, I(1) = 3.00) - 5% critical value (I(0) = 2.39, I(1) = 3.38) - 1% critical value (I(0) = 3.06, I(1) = 4.15) - Since 1.6894 is less than the lower bound at all levels, it suggests that there is no long-term relationship among the variables.

Table 7. Real Estate ARDL Short- Run.

Dependent Variable: LNREMR				
Selected model: ARDL(4,0,0,1,0,0)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.*
LNREMR(-1)	0.8850	0.0678	13.0455	0.0000
LNREMR(-2)	0.0118	0.0859	0.1371	0.8911
LNREMR(-3)	0.6177	0.0854	7.2315	0.0000
LNREMR(-4)	-0.5324	0.0680	-7.8355	0.0000
LNSENT	0.0011	0.0062	0.1733	0.8626
LNMS	-0.0042	0.0037	-1.1339	0.2586
LNPI	0.0015	0.0145	0.1050	0.9165
LNPI(-1)	-0.0202	0.0147	-1.3729	0.1718
LNCCI	-0.0318	0.0428	-0.7414	0.4596
LNGEPU	-0.0027	0.0019	-1.4051	0.1620
C	0.3869	0.2212	1.7489	0.0823
R-squared	0.9975	Breusch-Godfrey Serial Correlation LM Test:		
Adjusted R-squared	0.9973	F-statistic	0.3466	Prob. F(4,150)
F-statistic	6151.0599	Heteroskedasticity Test: ARCH		
Prob(F-statistic)	0.0000	F-statistic	0.0478	Prob. F(4,148)
Bound Test				
Test Statistic	Value	Value	Sig.	I(0)
F-statistic	1.6894	1.6894	10%	2.08
k	5	5%	2.39	3.38
		1%	3.06	4.15

2.4. Discussion

The ARDL analysis offers valuable insights into the relationship between investor sentiment and stock market returns in Saudi Arabia. The finding that positive changes in sentiment boost stock returns with a two-month lag indicates that psychology factors do influence investor behavior and market performance. However, the negative association observed in the long-run implies that sentiment-driven mispricing may be subsequently correct, as sentiment effects are arbitrated away over longer horizons. This contrasts with [5,7] who found a persistent negative relationship between lagged sentiment and US stock returns. The Saudi market demonstrates less lasting inefficiencies. The uniformly positive role of money supply aligns with [25], who showed monetary policy shifts significantly predicted Saudi returns. Expansionary policy appears to consistently drive market gains. Moreover, the absence of any marked real estate return response to sentiment contrasts with [27], who found depressed Saudi investor optimism, reduced market activity. The resilience seen here could stem from strict government regulations that curb speculative real estate investing. Overall, while psychology shapes Saudi investor behavior, market forces likely limit sustained mispricing. The nuanced temporal effects highlight the merits of ARDL modeling in capturing the complexity of dynamic relationships. Further research should examine regional differences and the role of oil shocks in triggering sentiment-driven return anomalies. Replication with higher frequency data may also elucidate short-run dynamics.

3. Conclusion

This study offers important insights into the nexus between investor sentiment and stock market performance in Saudi Arabia. The results reveal that positive changes in sentiment positively predict returns over a two-month horizon, indicating an initial psychological influence. However, the long-run negative association points towards subsequent correction of sentiment-induced mispricing. For real estate, investor psychology has limited effects, highlighting the market's resilience. The findings have practical implications for formulating policies regarding speculative activity, implementing

circuit breakers, and designing behavioral investment strategies timed to exploit sentiment-driven anomalies. The study also adds to behavioral finance theories on sentiment contagion and market efficiency within unique emerging market contexts. Extending the analysis to other GCC countries and MENA markets would offer valuable comparative insights into the nuances of investor psychology across the region. Overall, combining rigorous ARDL modeling with economic intuition and practical relevance promises to enhance the understanding of dynamic sentiment-return interactions.

4. Methods

Effective data collection and treatment involve a series of well-defined methods to ensure the integrity and usability of the data. Validation and curation techniques are crucial to maintain quality and minimize noise, resulting in findings that are more reliable. Understanding these processes helps in building strong and credible data-driven conclusions. After the preliminary analysis, Stationarity Analysis the time series properties were examined through Augmented Dickey Fuller (ADF) and Phillip Perrons (PP) unit root tests. ARDL modeling was conducted for stock market returns (LNSMR) as the dependent variable along with sentiment, money supply, industrial production index, consumer confidence index and global uncertainty as predictors.

5. Patents

This section is not mandatory but may be add if there are patents resulting from the work reported in this manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The time series data on stock market returns, housing market re-turns, investor sentiment, and macroeconomic variables used in this study are available from the corresponding author upon reasonable request. The data was obtain from publicly available sources including the Saudi Stock Exchange (Tadawul: <https://www.saudiexchange.sa/>) , the General Authority for Statistics in Saudi Arabia (<https://www.stats.gov.sa/en/843>) , the Saudi Central Bank (<https://www.sama.gov.sa/en>) , World Development Indicators by the World Bank (<https://databank.worldbank.org/>) ,and the global economic policy uncertainty index compiled (<https://doi.org/10.1093/qje/qjw024>) by Baker et al. (2016).

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alabdulqader, K. A., Hannah, G., & Power, D. M. (2013). The appraisal of ordinary shares by Saudi investors. *Research in International Business and Finance*, 29, 233-245.
2. Al-Hajieh, H., Redhead, K., & Rodgers, T. (2011). Investor sentiment and calendar anomaly effects: A case study of the impact of Ramadan on Islamic Middle Eastern markets. *Research in International Business and Finance*, 25(3), 345-356.
3. Alsalman, Z. (2023). Oil price shocks and US unemployment rate: a robustness check. *Applied Economics Letters*, 30(16), 2208-2215.
4. Alsalman, Z., & Herrera, R. (2022). Investor sentiment and sectoral returns in an emerging stock market. *Journal of Behavioral and Experimental Finance*, 30, 100463.
5. Antoniou, C., Doukas, J. A., & Subrahmanyam, A. (2013). Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis*, 48(1), 245-275.
6. Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4), 1645-1680.
7. Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2), 129-52.
8. Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2), 272-287.

9. Chaihetphon, P., & Pavabutr, P. (2010). Price discovery in Indian gold futures market. *Journal of Economics and Finance*, 34(4), 455-467.
10. De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738.
11. Dong, Z., Jiang, L., Liang, Q., & Zhou, W. X. (2021). Investor sentiment and stock returns: A replication study of US markets. *Finance Research Letters*, 38, 101855.
12. Elasrag, H. (2022). Stock market development, economic policy uncertainty, and MENA economic growth. *Economic Analysis and Policy*, 72, 320-339.
13. Guidi, F., Russell, A., & Tarbert, H. (2016). The effect of investor sentiment on the decision to speculate in emerging market assets. *Journal of Behavioral and Experimental Finance*, 9, 100-110.
14. Jansen, W. J., & Nahuis, N. J. (2003). The stock market and consumer confidence: European evidence. *Economics Letters*, 79(1), 89-98.
15. Jebran, K., Chen, S., Ullah, I., & Mirza, N. (2022). Do oil prices and investor fear sentiment predict Islamic equity returns? Fresh insights from a multi-scale perspective. *Resources Policy*, 73, 102302.
16. Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2022). Predicting stock returns using random forest, support vector machines and an ensemble model. *Statistics & Finance*, 2(1), 18-40.
17. Khawaja, A. I., Sun, Q., Mushtaq, R., & Ullah, F. (2022). A Long short-term memory approach to predict sector-level stock returns using investor sentiment. *Empirical Asset Pricing Review*, 2(1), 57-90.
18. Li, T., & Gong, L. (2017). Effect of investor sentiment on stock market returns: An empirical study of Chinese stock market. *Studies in Nonlinear Dynamics & Econometrics*, 21(4), 1-17.
19. Mostafa, M. M., & Elmasry, A. E. A. (2022). The impact of investor sentiment on stock market returns: A deep learning approach. *Managerial Finance*, 48(7), 1106-1132.
20. Nasir, M. A., Wei, Y., Zeng, Y., Shah, S. Z. A., & Ali, M. (2021). Investor sentiment and stock return crash risk: Evidence from international markets. *Economic Modelling*, 106, 105663.
21. Pagan, A., & Soydemir, G. (2001). Response asymmetries in the Latin American equity markets. *Emerging Markets Review*, 2(2), 175-197.
22. Pavabutr, P., & Chaihetphon, W. (2010). Return predictability and discrete bubbles in the Thai stock market. *Emerging Markets Finance and Trade*, 46(6), 83-101.
23. Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance*, 16(3), 394-408.
24. Singhania, P., & Anchalia, J. (2021). Investor sentiment and asset price bubbles: Evidence from housing and stock prices in India. *Decision*, 1-21.
25. Soydemir, G., & Pagan, A. (2001). Response asymmetries in the Latin American equity markets. *Emerging Markets Review*, 2(2), 175-197.
26. Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2), 288-302.
27. Verma, R., & Soydemir, G. (2006). The impact of individual and institutional investor sentiment on the market price of risk. *The Quarterly Review of Economics and Finance*, 46(3), 380-393.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.