UCD DUBLIN

UCD Michael Smurfit Graduate Business School

FIN41660 Financial Econometrics

Individual Practical Assignment in Python 2024

Professor: Alessia Paccagnini

Name: Tejas Deepak Landge

Student No.: 24218694

Submission Date: 19 December 2024.

UCD Michael Smurfit Graduate Business School

Declaration

I hereby declare that the project titled **"Analyzing the Relationship Between Housing Prices and GDP Using Time Series Econometric Methods"** is my original work and has not been submitted, in part or full, for any other degree or diploma.

All work presented in this report has been carried out by me, except where explicitly stated otherwise. The data sources, tools, and methodologies have been appropriately acknowledged, and any assistance received during the project has been duly recognized.

This report complies with the academic and ethical standards required by **UCD Michael Smurfit Graduate Business School**.

Name: Tejas Deepak Landge
Student No.: 24218694
Date: 19 December 2024.

Table of Contents

Analyzing the Relationship Between Housing Prices and GDP in the U.S Using Time Series Methods

A Study Using Time Series Methods for The Period (2012-2024)

Abstract:

The research study uses the **Federal Housing Finance Agency's (FHFA) House Price Index** as a proxy for housing prices to examine the relationship between **real GDP** and housing prices in the United States from 2012 to 2024. Housing prices were **standardized** before analysis to improve interpretability and remove scale discrepancies. The study makes use of sophisticated time series econometric methods, such as **Ordinary Least Squares (OLS)** Regression, **ARIMA** model for autocorrelation correction, and **GARCH** model for volatility analysis. The findings indicate a robust positive correlation between GDP and standardized house prices, as well as indications of **volatility clustering** in the residuals. For **economic policymakers**, GDP and conditional volatility forecasts offer important insights by showing steady economic growth and ongoing but moderate volatility in the future.

Introduction:

Numerous studies have examined the relationship between **real estate prices** and **macroeconomic indicators** such as **GDP** because of its significant impact on **economic stability**. A major influence on **aggregate demand**, **investment**, and overall **economic growth** is exerted by **housing markets**. **Consumer confidence** and the state of financial institutions are both reflected in changes in **house prices**, which frequently function as markers of large **economic cycles**.

During the **2008 Global Financial Crisis**, when a collapse in **real estate prices** precipitated a **global economic slowdown**, the importance of **housing markets** was prominently illustrated. This incident demonstrated the close relationship between **housing markets**, **household wealth**, **credit flows**, and **financial stability**. On the other hand, increased **housing costs** can boost **economic activity** by raising **household net worth**, promoting **consumer spending**, and attracting **investment** in the **real estate** and **construction industries**.

Empirical research highlights that **economic cycles** and **housing price dynamics** are frequently correlated. For example:

- In periods of **economic expansion**, increasing **income levels** and **favourable credit conditions** drive up **housing prices**.

- Conversely, during **recessions**, declining **demand** and **credit constraints** often result in falling **property prices**, which further depress **economic growth** through **negative wealth effects**.

Housing prices are increasingly seen as a **leading indicator** of **economic performance** because of its significant correlation. In addition to reflecting **imbalances between supply and demand**, changes in **consumer behaviour**, **confidence** in the **financial system**, and **policy interventions** like **interest rate changes** or **government housing initiatives** are also reflected in changes in **housing prices**.

A relevant study by Klarl (2016) [1] used wavelet coherence analysis to examine the dynamic relationship between GDP and house prices in the U.S. The study discovered that home prices frequently drive changes in GDP, especially during times of economic crisis, highlighting the significance of property markets as a key indicator of macroeconomic trends.

Building on this previous relationship, the goal of this study is to examine the **dynamic relationship** between **real GDP** and **housing prices** in the **United States between 2012 to 2024**. By applying advanced econometric methods such as **OLS Regression**, **ARIMA modelling**, and **GARCH analysis**, the research seeks to:

1. **Quantify** the **impact of housing prices** on **GDP**.
2. **Address issues** such as **residual autocorrelation** and **volatility clustering**.
3. **Forecast future GDP trends** and associated **volatility**.

Understanding this relationship is critical for **policymakers**, **investors**, and **economists**, as it offers valuable insights into **economic stability** and future **risks**.

Literature Review:

The correlation between housing prices and macroeconomic metrics like GDP has been thoroughly examined due to its significance in comprehending economic growth and stability. The importance of housing markets as indicators and drivers of economic cycles has been highlighted by earlier research.

1. **Relationship Between Housing Prices and GDP:**
   Klarl (2016) [1] provided a comprehensive analysis of the dynamic relationship between housing prices and GDP in the United States using the wavelet coherence analysis. According to the report, housing prices frequently outpace GDP, especially in times of economic crisis like Global Financial Crisis of 2008. The significance of housing markets in affecting loan flows, household wealth, and aggregate demand is highlighted by this link.

Other empirical studies emphasize the cyclical nature of the housing prices and GDP. For example:

- Housing prices rise in times of economic expansion due to favourable lending conditions and growing incomes, which feeds back positively into GDP.
- In contrast, tighter loan requirements and dwindling demand during recessions lead to a decline in real estate values, which exacerbates economic slowdowns through adverse wealth impacts.

These findings highlight the dual role of housing prices as both an economic driver and a reflection of macroeconomic trends.

2. **Research Gaps:**

   Despite extensive research, several gaps remain in the literature:

   - **Time Period:** Most of the research focuses on data collected before 2010, which leaves a knowledge gap about the post-crisis era and the impact of COVID-19 pandemic on GDP and housing prices.
   - **Standardization:** Only a small number of research have used standardized housing prices data, which improves the results' comparability and interpretability.
   - **Comprehensive Analysis:** Some of research examines the linear relationship and volatility dynamics between housing prices and GDP using variety of econometric methodologies, including OLS, ARIMA, and GARCH.

3. **Contribution of This Study:**

   This study builds on existing literature by addressing these gaps:

   - Using standardized housing prices to improve interpretability.
   - Analyzing the data from 2012 to 2024, covering the post-financial-crisis period and COVID-19 era.
   - Combining OLS Regression, ARIMA models, and GARCH analysis to provide a comprehensive understanding of the relationship between housing prices and GDP.
   - Offering insights into future GDP trends and volatility, with implications for a policymakers and investors.

Methodology:

Data Collection:

This study uses quarterly data spanning from 2012 to 2024. The data sources are:

1. **Real GDP:** Based on precise and reliable macroeconomic data statistics from the **Federal Reserve Economic Data (FRED)**.
2. **Housing Prices:** The **Federal Housing Finance Agency (FHFA) House Price Index** is used as a proxy for housing prices in the U.S. This index tracks changes in home prices across the country and serves as a reliable measure of housing market trends.

The data is aggregated at a **quarterly frequency** to ensure consistency and reduce noise that may exist in higher-frequency datasets. To ensure consistency in the analysis and to eliminate the effects of scale differences, the FHFA House Price Index is **standardized.** Standardization transforms the variable to have a mean of zero and a **standard deviation of one**, improving the interpretability and stability of the regression model.

The standardization formula is as follows:

$$X_t^* = \frac{X_t - \bar{X}}{\sigma_X}$$

Where:

- X_t: Original housing prices at time *t*.
- $\bar{X}$: Mean of the housing prices.
- σ_X: Standard deviation of the housing prices.
- X_t^*: Standardized housing prices.

Ordinary Least Squares (OLS) Regression:

The OLS Regression is used to estimate the linear relationship between Real GDP (*Y*) and Housing Prices (*X*).

The model is specified as follows:

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

Where:

- Y_t: Real GDP at time *t*.
- X_t: Standardized FHFA House Price Index at time *t*.
- β_0: Intercept term.
- β_1: Coefficient measuring the impact of housing prices on GDP.
- ϵ_t: Error term.

The OLS regression assumes:

1. **Linearity:** The relationship between GDP and Housing Prices is linear.
2. **No Multicollinearity:** The independent variable is not correlated with itself.
3. **Homoscedasticity:** The variance of the residuals remains constant over time.
4. **No Autocorrelation:** Residuals are not correlated with their past values.

Residual Diagnostics:

To validate the OLS Regression results, diagnostics tests are conducted:

1. **Jarque-Bera Test:** Assesses whether residuals are normally distributed.
2. **Durbin-Watson Test:** Checks for autocorrelation in the residuals.
3. **Residual Plots:** A visual inspection of residual histograms and ACF/PACF plots is performed to confirm the assumptions.

If residuals exhibit autocorrelation, further time series models (e.g., ARIMA) are applied for correction.

Autoregressive Integrated Moving Average (ARIMA) Model:

The ARIMA model is employed to account for autocorrelation in the residuals. The ARIMA(1,0,0) model is specified as:

$$Y_t = \emptyset_t \, Y_{t-1} + \epsilon_t$$

Where:

- Y_t: Residual value at time *t*.
- $\emptyset_t$: Coefficient of the first lag of residuals.
- ϵ_t: White noise error term.

Steps for ARIMA implementation:

1. **Identify Model Order:** The ACF and PACF plots are analyzed to determine the lag order (p, d, q).
2. **Model Fitting:** ARIMA(1,0,0) is fitted to the residuals.
3. **Validation:** The ACF and PACF plots of the residuals are inspected to confirm the exhibit white noise.

The ARIMA(1,0,0) model was chosen to address residual autocorrelation. The model assumes that the series is stationary, which was verified prior to implementation. Residual diagnostics were conducted to ensure that the ARIMA model produced white noise residuals.

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model:

The GARCH(1,1) model is applied to analyze the volatility clustering present in the residuals. The GARCH model estimates the conditional variance as follows:

$$\sigma_t^2 = \omega + \alpha_1 \, \epsilon_{t-1}^2 + \beta_1 \, \sigma_{t-1}^2$$

Where:

- σ_t^2: Conditional variance at time *t*.
- ϵ_{t-1}^2: Lagged squared residuals.
- σ_{t-1}^2: Lagged conditional variance.
- ω: Constant term.
- α_1 , β_1: Coefficients of GARCH model.

Steps for GARCH Implementation:

1. **Fit GARCH Model:** The GARCH(1,1) model is applied to the ARIMA residuals.
2. **Interpret Coefficients:** Significant α_1 *and* β_1 values confirm the presence of volatility clustering.
3. **Forecast Volatility:** Future conditional volatility values are forecasted for the next eight quarters to assess the economic uncertainty.

The GARCH(1,1) model was implemented to capture volatility clustering in the residuals, which is common in macroeconomic time series data. The model assumes that the volatility depends on both the past shocks and the past volatility, making it suitable for analyzing risk persistence.

Forecasting:

Forecasting is performed for both GDP and Volatility:

1. **GDP Forecast:** The AIRMA model predicts GDP for the next eight quarters based on historical patterns.
2. **Volatility Forecasts:** The GARCH model forecasts conditional volatility, highlighting potential periods of economic instability.

These forecasts offer insights into future economic trends and risks.

Results and Discussion:

OLS Regression Results:

To investigate the relationship between real GDP and house prices, Ordinary Least Squares (OLS) Regression was used. Standardization of the FHFA House Price Index was done to make it easier to understand and to manage scalability problems. Because of standardization, the coefficient is guaranteed to accurately reflect how a one unit change in the standardized housing prices affects the real GDP.

The OLS Regression model is specified as follows:

$$Y_t = \beta_0 + \beta_1 X_t^* + \epsilon_t$$

Where:

- Y_t: Real GDP at time *t*.
- X_t^*: Standardized FHFA House Price Index at time *t*.
- β_0: Intercept term.
- β_1: Coefficient for standardized housing prices.
- ϵ_t: Error term

Key Results:

The regression results are summarized as follows:

1. **R-Squared:** The model achieved an **R-squared value of 0.939,** indicating that approximately 93.9% of the variation in real GDP is explained by the standardized housing prices.
2. **Coefficient:** The coefficient for the standardized housing prices is 1720.58. This implies that a one-unit increase in standardized housing prices corresponds to an increase of approximately 1720.58 units in real GDP.
3. **Statistical Significance:** The p-value associated with the coefficient is < 0.001, confirming the relationship between housing prices and GDP is statistically significant at 1% level.

To provide a unit-free comparison and enhance numerical stability in the regression model, housing prices were standardized before analysis. The coefficient estimates can be understood in terms of a change in house prices of one standard deviation thanks to standardization. This transformation guarantees that the results are not affected by the magnitude of the variables and enhances the comparability of the results.

```
Performing OLS Regression with Standardized Data...
                            OLS Regression Results
==============================================================================
Dep. Variable:               Real_GDP   R-squared:                       0.939
Model:                            OLS   Adj. R-squared:                  0.938
Method:                 Least Squares   F-statistic:                     760.6
Date:                Sun, 15 Dec 2024   Prob (F-statistic):           1.68e-31
Time:                        23:25:37   Log-Likelihood:                -381.90
No. Observations:                  51   AIC:                             767.8
Df Residuals:                      49   BIC:                             771.7
Df Model:                           1
Covariance Type:            nonrobust
====================================================================================================
                                      coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------------------------
const                            2.007e+04     61.774    324.859      0.000    1.99e+04    2.02e+04
FHFA_House_Price_Index_Standardized 1720.5814  62.388     27.579      0.000    1595.207    1845.956
==============================================================================
Omnibus:                        0.061   Durbin-Watson:                   0.535
Prob(Omnibus):                  0.970   Jarque-Bera (JB):                0.094
Skew:                           0.065   Prob(JB):                        0.954
Kurtosis:                       2.835   Cond. No.                         1.01
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Figure 1: OLS Regression Summary using Standardized Housing Prices.

Residual Diagnostics:

To ensure the validity of the OLS model, residual diagnostics were performed:

1. **Normality of Residuals:**
   - The Jarque-Bera Test resulted in a p-value of 0.954, indicating that the residuals are normally distributed.

```
Jarque-Bera Test for Normality:
JB Statistic: 0.09417503079664892, p-value: 0.9540039039052796
```

Figure 2: Normality Test of Residuals from OLS Regression.

- A histogram of residuals confirms this visually.

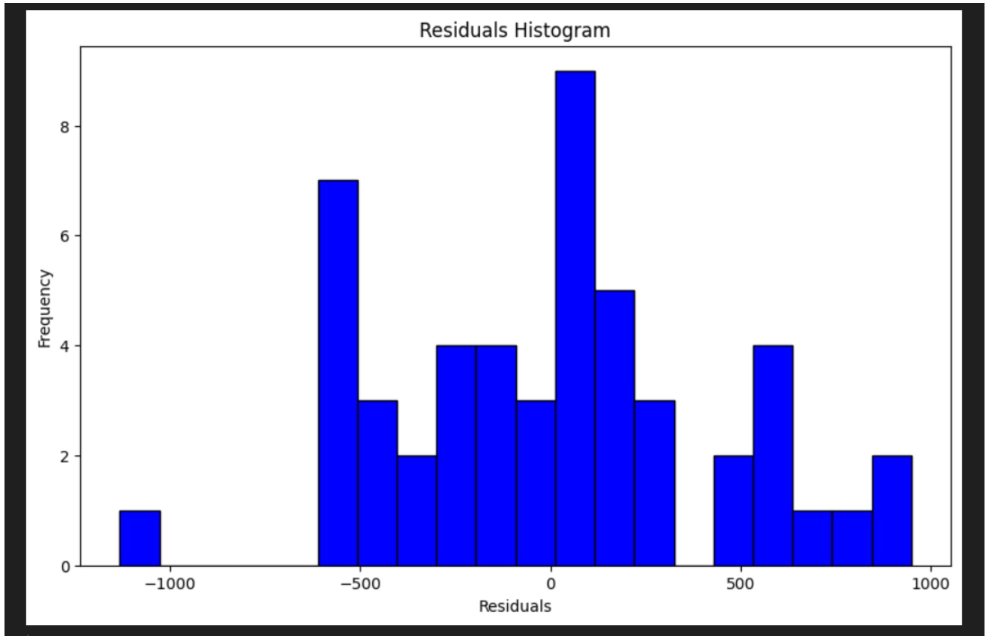

Figure 3: Histogram of Residuals from OLS Regression.

2. **Autocorrelation in Residuals:**
   - The **Durbin-Watson** statistic was 0.535, suggesting the presence of positive autocorrelation.
   - This was further confirmed by Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plot, which show significant lags in the residuals.

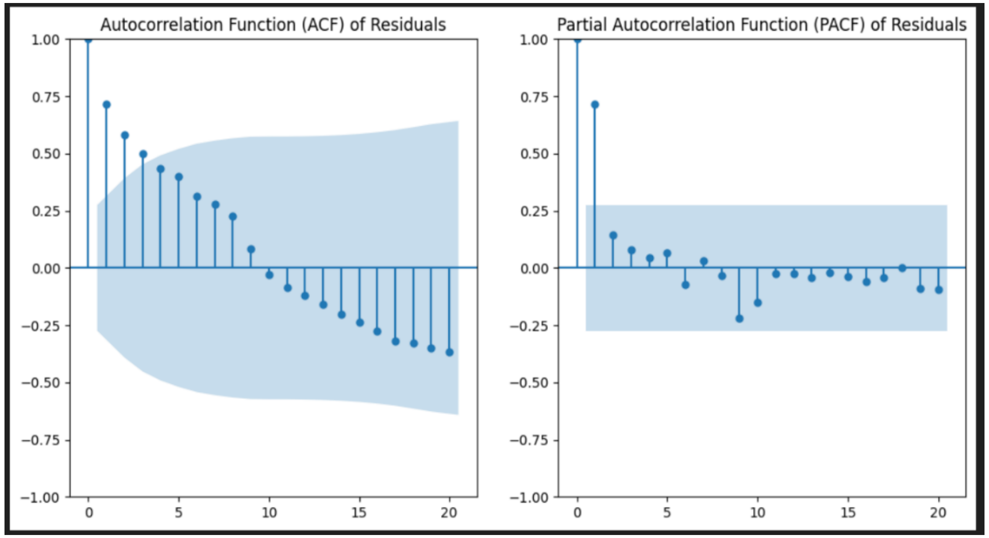

Figure 4: ACF and PACF Plots of OLS Residuals.

Given the evidence of autocorrelation, further modelling using the ARIMA approach was implemented to correct for this issue.

3. **Breusch-Pagan Test for Heteroscedasticity:**
   To ensure that the assumption of homoscedasticity (constant variance of residuals) is satisfied in the OLS model, the Breusch-Pagan Test was performed.
   - Null Hypothesis (H_0): Residuals Exhibit constant variance (no heteroscedasticity).

- Alternative Hypothesis (H_1): Residuals exhibit heteroscedasticity (variance changes over time).

Test Results:

```
Breusch-Pagan Test for Heteroscedasticity:
LM Statistic: 1.5382680313283665
p-value: 0.21487578942817626
```

Figure 5: Breusch-Pagan Test for Heteroscedasticity.

- **LM Statistic:** 1.538
- **P-value:** 0.215

Given the p-value is greater than the significance level of 0.05 we fail to reject the null hypothesis. This indicates that the residuals exhibit no significant heteroscedasticity, and the assumption of constant variance holds.

ARIMA Results:

To address the autocorrelation observed in the residuals of the OLS regression, an **ARIMA(1,0,0)** model was applied. The ARIMA model captures the first-order autoregressive structure in the residuals, ensuring that the residuals become **white noise**.

The ARIMA model is specified as below:

$$\epsilon_t = \emptyset_t \, \epsilon_{t-1} + \eta_t$$

Where:

- ϵ_t: Residual value at time *t*.
- $\emptyset_t$: Coefficient for the first lag of residuals.
- η_t: White noise error term with a mean of zero and constant variance.

Key Results:

The ARIMA(1,0,0) model results are summarized as follows:

- AR(1) Coefficient $(\emptyset_1)$: 0.7258 (statistically significant at the 1% level).
- Log-Likelihood: -362.996.
- AIC: 731.992.
- BIC: 737.788.

```
Fitting ARIMA(1,0,0) model to residuals...
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                   51
Model:                 ARIMA(1, 0, 0)   Log Likelihood                -362.996
Date:                Sun, 15 Dec 2024   AIC                            731.992
Time:                        23:25:38   BIC                            737.788
Sample:                    03-31-2012   HQIC                           734.207
                         - 09-30-2024
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -1.2818    183.254     -0.007      0.994    -360.452     357.889
ar.L1          0.7258      0.106      6.816      0.000       0.517       0.934
sigma2      8.808e+04    1.3e+04      6.798      0.000    6.27e+04    1.13e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.80   Jarque-Bera (JB):               405.98
Prob(Q):                              0.37   Prob(JB):                         0.00
Heteroskedasticity (H):               3.47   Skew:                            -2.09
Prob(H) (two-sided):                  0.01   Kurtosis:                        16.18
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Figure 6: ARIMA(1,0,0) Model Summary.

The significant AR(1) coefficient indicates that the residuals from the OLS Regression exhibit strong first-order autocorrelation. After applying the ARIMA model, the residuals were evaluated for white noise.

Residuals Diagnostics:

The ACF and PCAF plots of the ARIMA residuals confirm that the autocorrelation has been adequately addressed, as there are no significant spikes outside the confidence bands.

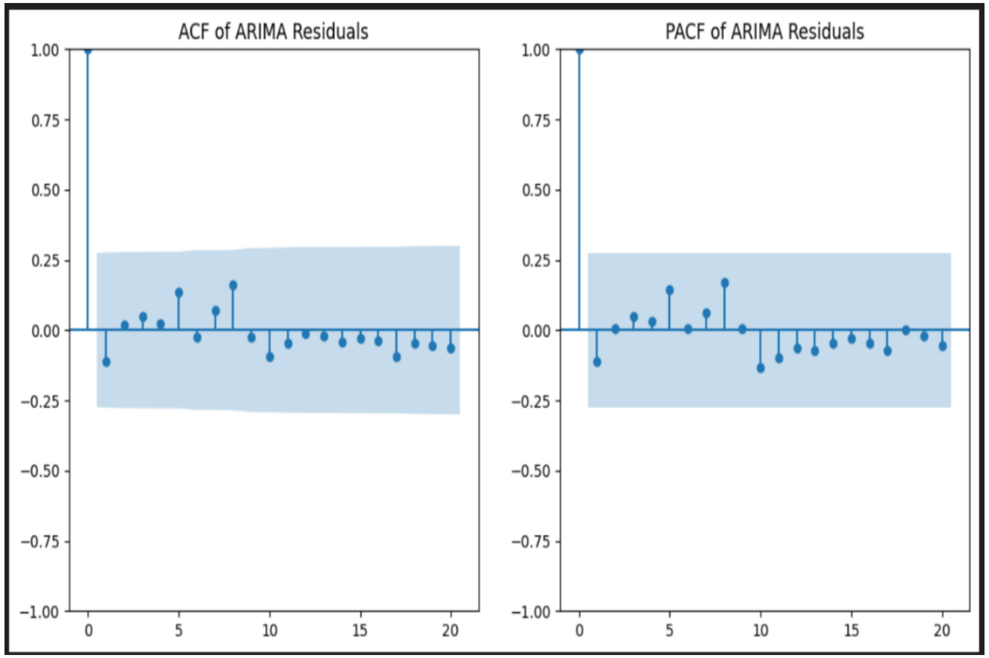

Figure 7: ACF and PACF of ARIMA Residuals.

The plots confirm the absence of significant autocorrelation, indicating that the model successfully addressed linear dependencies in the data. The residuals can now be considered white noise, validating the robustness of the ARIMA(1,0,0) model.

GARCH Results:

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model was applied to the residuals obtained from the ARIMA(1,0,0) model to analyze the presence of volatility clustering. The phenomenon known as volatility clustering occurs when high volatility periods are followed by more high volatility and low volatility periods are followed by more low volatility.

The GARCH(1,1) model specification is as follows:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Where:

- σ_t^2: Conditional variance (volatility) at time *t.*
- ϵ_{t-1}^2: Lagged squared residuals (past shocks).
- σ_{t-1}^2: Lagged conditional variance (past volatility).
- $\omega, \alpha_1, \beta_1$: GARCH model parameters.

Key Results:

The GARCH(1,1) model results are summarized as follows:

- Omega (ω): 9089.02 (constant term)
- Alpha (α): 0.5292 (impact of past shocks)
- Beta (β): 0.4708 (impact of past volatility)

```
Fitting GARCH(1,1) model to ARIMA residuals...

GARCH(1,1) Model Results:
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                   None   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:               -353.048
Distribution:                  Normal   AIC:                           714.097
Method:            Maximum Likelihood   BIC:                           721.824
                                        No. Observations:                   51
Date:                Sun, Dec 15 2024   Df Residuals:                       50
Time:                        23:25:38   Df Model:                            1
                                 Mean Model
===========================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
---------------------------------------------------------------------------
mu           -16.7648     57.235     -0.293      0.770 [-1.289e+02, 95.414]
                               Volatility Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
omega       9089.0164   5253.650      1.730  8.362e-02 [-1.208e+03,1.939e+04]
alpha[1]       0.5292      0.151      3.514  4.407e-04      [  0.234,  0.824]
beta[1]        0.4708      0.140      3.366  7.637e-04      [  0.197,  0.745]
=============================================================================

Covariance estimator: robust
```

Figure 8: GARCH(1,1) Model Summary Output.

The significant values of α_1 and β_1 confirm the presence of **volatility clustering,** indicating that volatility depends on both past shocks and past volatility.

The GARCH(1,1) model's results show that volatility clustering exists, especially during times of economic shocks like COVID-19 Pandemic in 2020. Potential threats to economic stability in the near future are indicated by the anticipated rise in the conditional volatility. These results highlight the necessity for proactive steps to reduce economic uncertainty for policymakers, such as monetary policy implementation to control credit flows or housing market stabilization. These insights into volatility can also be used by financial analysts and investors to modify investment portfolios and risk management plans in anticipation of times of increased economic volatility.

Conditional Volatility Analysis:

The conditional volatility plot reveals a sharp spike around the year 2020, which aligns with the economic disruptions caused by COVID-19 pandemic. The fact that the volatility declines after 2020 but still high shows how persistent it is over time.

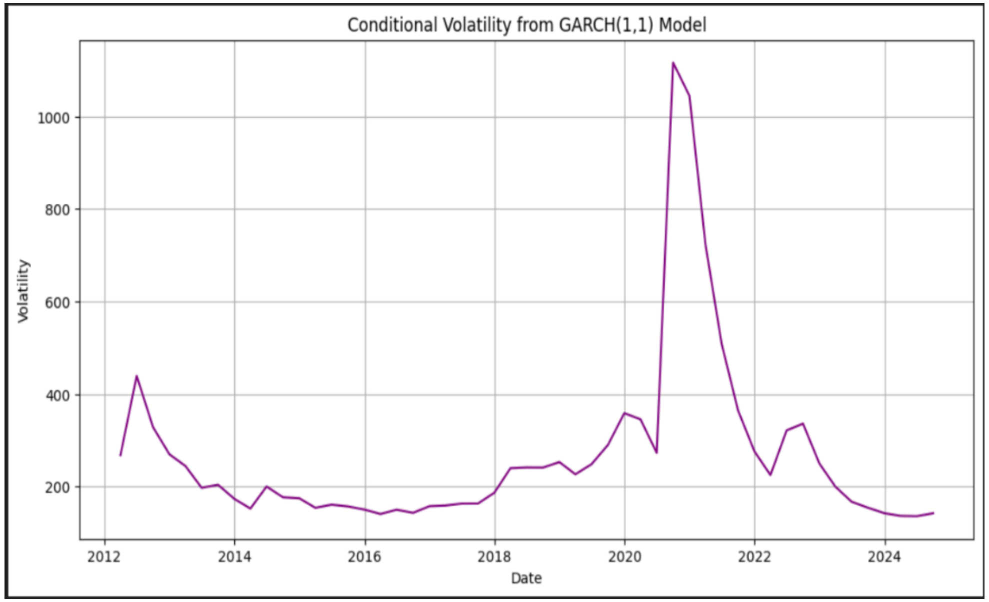

Figure 9: Conditional Volatility from GARCH(1,1) Model.

Forecasting Results:

Using the ARIMA and GARCH models, respectively, we provide Real GDP and Conditional Volatility Projections for the upcoming eight quarters in this section. Policymakers and analysts are better equipped and get ready for future economic conditions thanks to these forecasts, which provide insightful information about future economic growth and possible volatility tendencies.

1. **GDP Forecasting:**

To predict future Real GDP, the ARIMA(1,0,0) model was applied. The ARIMA model captures the autoregressive structure in GDP trends and provides reliable short-term forecasts based on historical patterns.

The key findings from the ARIMA-based GDP forecast are as follows:

- The forecast suggest a continuation of the upward trend in GDP.
- Real GDP is expected to stabilize around 23400 units by the end forecast period.
- No significant shocks or disruptions are observed, indicating steady economic growth under current conditions.

This forecast aligns with economic expectations, given historical data trends and the observed relationship between housing prices and GDP.

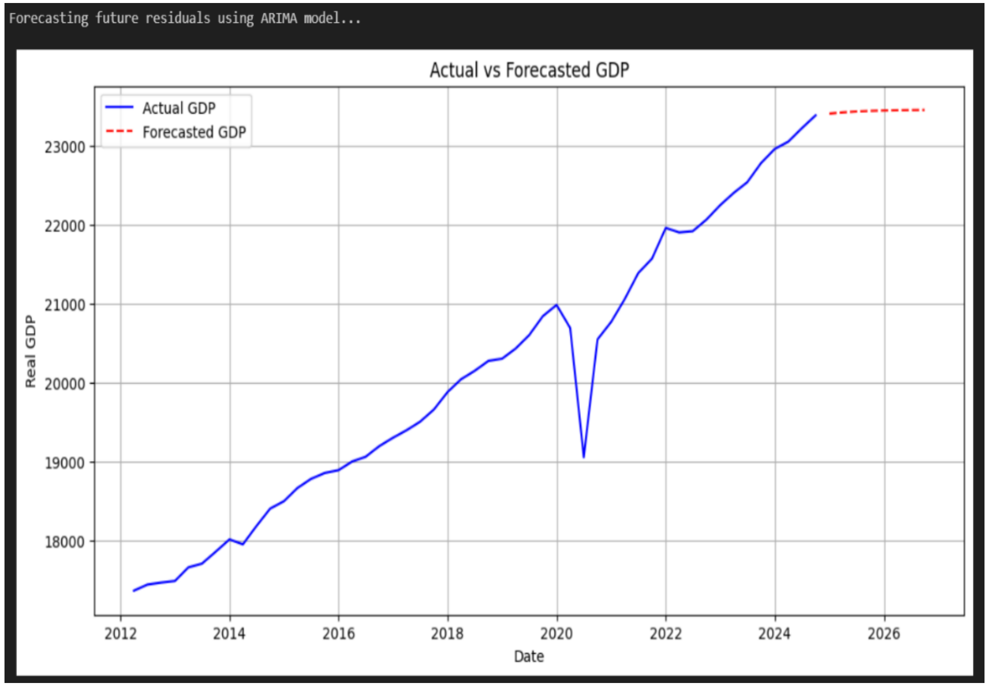

Figure 10: Actual vs Forecasted GDP for the Next 8 Quarters.

Forecasted GDP Table:

Table 1: Forecasted Real GDP Values for the Next 8 Quarters.

Period	Forecasted GDP
2024-12-31	23408.44
2025-03-31	23423.85
2025-06-30	23434.68
2025-09-30	23442.19
2025-12-31	23447.29
2026-03-31	23456.63
2026-06-30	23452.71
2026-09-30	23453.57

The GDP forecasts assume that historical patterns and relationships will continue throughout the projected period. External events like abrupt policy changes, economic shocks, global crisis could cause projections to deviate from the ARIMA model, even if it is a dependable model for short term predictions.

Interpretation:

The GDP forecasts show a steady increasing trend, pointing to a rebound and long-term expansion. Businesses, investors, and politicians might use this data as a baseline when making plans for upcoming economic activity. The lack of abrupt changes indicates that, under the current circumstances, the macroeconomic climate should continue to be positive.

2. **Volatility Forecasting:**

The GARCH(1,1) model was used to forecast the conditional volatility of the ARIMA residuals for the next eight quarters. This helps capture the persistence of volatility and assess the risk levels associated with future economic conditions.

The key findings from the volatility forecasts are as follows:

- Volatility begins at 151.06 and gradually rises to 294.01 over the forecast horizon.

- The steady increase in the volatility suggests persistent uncertainty and potential economic risks, although no extreme shocks are predicted.
- A gradual upward trend in volatility reflects that while economic growth remains steady, small disruptions or fluctuations may occur.

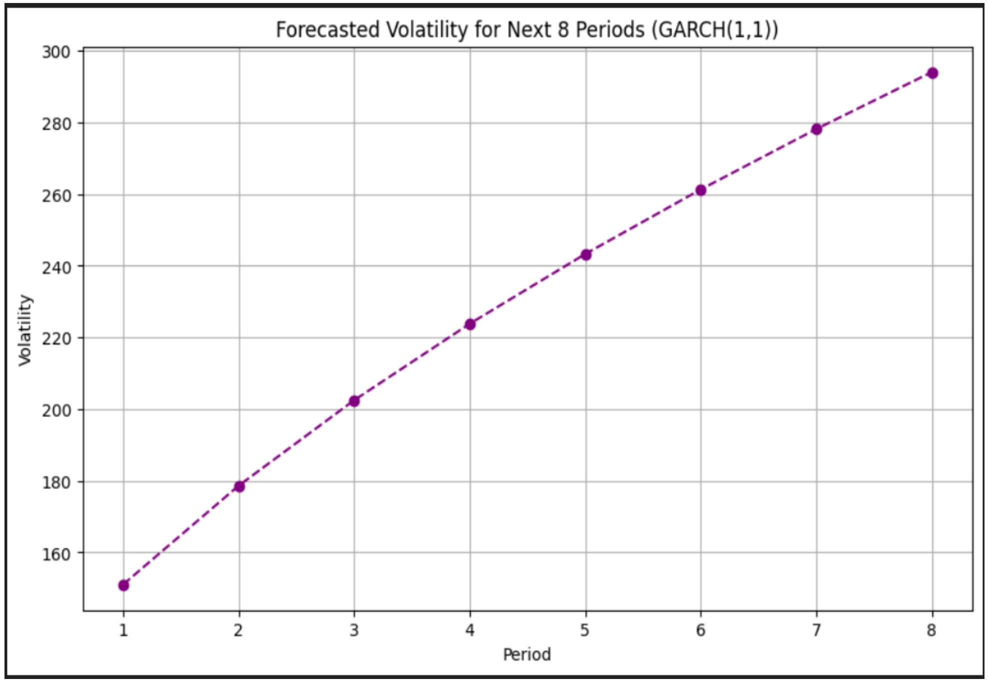

Figure 11: Forecasted Conditional Volatility for the Next 8 Quarters.

Forecasted Volatility Table:

The table below summarizes the forecasted volatility values:

Table 2: Forecasted Conditional Volatility for the Next 8 Quarters.

Period	Forecasted Volatility
2024-12-31	151.06
2025-03-31	178.62
2025-06-30	202.47
2025-09-30	223.80
2025-12-31	243.26
2026-03-31	261.27
2026-06-30	278.12
2026-09-30	294.01

Interpretation:

Over the next two years, there maybe minor economic uncertainty, as indicated by the conditional volatility's progressive increase. Policymakers and businesses should exercise caution and keep an eye on external economic developments that could lead to increased volatility, even though volatility is still moderate.

A comprehensive overview of the macroeconomic environment for the future is given by the combined insights from GDP and volatility predictions, which highlight the significance of home prices as an economic indicator.

Conclusion:

The study used time series econometric models, including OLS Regression, ARIMA, and GARCH to analyze the relationship between housing prices and the real GDP in the United States from 2012 to 2024.

The key findings are follows:

1. **OLS Regression Results**:
    - A strong positive relationship was found between housing prices (standardized FHFA House Price Index) and real GDP.
    - The OLS model achieved an R-Squared value of 0.939, indicating that approximately 93.9% of the variation in real GDP is explained by housing prices.
2. **ARIMA Test Results**:
    - The residuals from the OLS model exhibited significant autocorrelation, which was addressed using as ARIMA(1,0,0) model.
    - After applying ARIMA, the residuals were confirmed to be white noise, validating the model's reliability.
3. **GARCH Results**:
    - The GARCH(1,1) model captured volatility clustering in the residuals, with significant coefficients for both past shocks and past volatility.
    - The analysis revealed a spike in volatility around 2020, aligning with economic disruptions caused by the COVID-19 pandemic.
4. **Forecasting Results**:
    - **GDP Forecasts**: ARIMA forecasts showed a steady upward trend in GDP, with values stabilizing around 23400 units by 2026.
    - **Volatility Forecasts**: GARCH forecasts indicated a gradual increase in conditional volatility, reflecting persistent economic uncertainties over the next eight quarters.

Implications:

The findings of this study provide important insights for policymakers, economists, and market participants:

- **Housing Prices** serve as a significant indicator of economic growth and stability. Monitoring housing market trends can help forecast GDP movements.
- The presence of **volatility clustering** underscores the importance of managing risks during periods of economic uncertainty.
- Forecasts of GDP and volatility offer a basis for informed decision-making in policy planning and investment strategies.

Future Research:

This study can be extended in the following ways:

- Including additional macroeconomic factors such as interest rates, unemployment rates, or inflation to improve the robustness of the model.
- Applying more advanced models like Vector Autoregression (VAR) or Wavelet Analysis for a multi-variable analysis of housing markets and GDP.
- Exploring regional housing price indices to capture local economic dynamics.

Final Thoughts:

One of the basic pillars of macroeconomic analysis is still the correlation between GDP and housing prices. This research emphasizes the crucial role that housing prices play in predicting economic development and identifying times of increased uncertainty by integrating linear regression, time series model, and volatility analysis.

References:

1. Main Research Paper:

   [1] T. Klarl, "The nexus between housing and GDP re-visited: A wavelet coherence view on housing and GDP for the U.S.," *Economics Bulletin,* vol. 36, no. 2, pp. 704-720, Apr. 2016.

2. Textbook for Methodologies:

   [2] A. Hamilton, *Time Series Analysis,* Princeton University Press, 1994.

3. ARIMA and GARCH Methods:

   [3] R. Engle, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, vol. 50, no. 40, pp. 987-1007, 1982.

4. Data Source for Real GDP:

   [4] Federal Reserve Economic Data (FRED), *Real Gross Domestic Product,* [online]. Available: https://fred.stlouisfed.org.

5. Data Source for Housing Prices:

   [5] Federal Housing Finance Agency (FHFA), *House Price Index (HPI)*, [online]. Available: https://www.fhfa.gov.

6. OLS and Time Series Techniques:

   [6] D. Gujarati and D. Porter, *Basic Econometrics*, 5h ed., McGraw-Hill, 2009.

7. The coding and analysis implementations were supported using Python guidance provided by OpenAI's ChatGPT.

   [7] OpenAI, ChatGPT: A Conversational AI Model, [online]. Available: https://chat.openai.com.

Appendix:

Python Codes

```
import pandas_datareader.data as web
from datetime import datetime
import pandas as pd

# Define the date range
start_date = "2012-01-01"
end_date = "2024-12-31"

# Fetch FHFA House Price Index (Housing Prices) from FRED
print("Fetching FHFA House Price Index...")
fhfa_data = web.DataReader("HPIPONM226S", "fred", start_date,
end_date)

# Fetch Real GDP data from FRED
print("Fetching Real GDP data...")
gdp_data = web.DataReader("GDPC1", "fred", start_date, end_date)

# Resample data to quarterly frequency if needed (both are quarterly
by default)
fhfa_data = fhfa_data.resample('QE').mean()
gdp_data = gdp_data.resample('QE').mean()

# Combine the datasets
combined_data = pd.concat([fhfa_data, gdp_data], axis=1)
combined_data.columns = ["FHFA_House_Price_Index",
"Real_GDP"]
combined_data.dropna(inplace=True)

print("Combined Dataset:\n", combined_data.head())

# Save the dataset to a CSV file
combined_data.to_csv("housing_gdp_data.csv")
print("Data saved to 'housing_gdp_data.csv'")


import matplotlib.pyplot as plt

# Load the dataset (if not already in memory)
data = pd.read_csv("housing_gdp_data.csv", parse_dates=["DATE"],
index_col="DATE")

# Step 1: Summary statistics
print("Summary Statistics:\n", data.describe())

# Step 2: Plot Housing Prices and Real GDP
plt.figure(figsize=(12, 6))
plt.plot(data.index, data["FHFA_House_Price_Index"], label="FHFA
House Price Index", color="blue")
plt.plot(data.index, data["Real_GDP"], label="Real GDP",
color="green")
plt.title("Housing Prices and Real GDP (2012-2024)")
plt.xlabel("Date")
plt.ylabel("Value")
plt.legend()
plt.grid()
```

```
plt.show()


import statsmodels.api as sm

# Define the dependent (Y) and independent (X) variables
Y = data["Real_GDP"]
X = data["FHFA_House_Price_Index"]
X = sm.add_constant(X)  # Add a constant for the intercept

# Perform OLS Regression
print("Performing OLS Regression...")
model = sm.OLS(Y, X).fit()

# Display the summary results
print(model.summary())


# Standardize the independent variable
data['FHFA_House_Price_Index_Standardized'] =
(data['FHFA_House_Price_Index'] -
data['FHFA_House_Price_Index'].mean()) /
data['FHFA_House_Price_Index'].std()

# Define the dependent (Y) and independent (X) variables
Y = data["Real_GDP"]
X = data["FHFA_House_Price_Index_Standardized"]
X = sm.add_constant(X)  # Add a constant for the intercept

# Perform OLS Regression
print("Performing OLS Regression with Standardized Data...")
model_standardized = sm.OLS(Y, X).fit()

# Display the summary results
print(model_standardized.summary())


import statsmodels.stats.api as sms

# Extract residuals
residuals = model_standardized.resid

# 1. Normality Test: Jarque-Bera
jb_test = sms.jarque_bera(residuals)
print("\nJarque-Bera Test for Normality:")
print(f"JB Statistic: {jb_test[0]}, p-value: {jb_test[1]}")

import matplotlib.pyplot as plt

# Plot Residuals Histogram
plt.figure(figsize=(10, 6))
plt.hist(residuals, bins=20, color='blue', edgecolor='black')
plt.title("Residuals Histogram")
plt.xlabel("Residuals")
plt.ylabel("Frequency")
plt.show()

from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
```

```
# Plot ACF and PACF for Residuals
fig, ax = plt.subplots(1, 2, figsize=(12, 6))

# ACF Plot
plot_acf(residuals, lags=20, ax=ax[0])
ax[0].set_title("Autocorrelation Function (ACF) of Residuals")

# PACF Plot
plot_pacf(residuals, lags=20, ax=ax[1])
ax[1].set_title("Partial Autocorrelation Function (PACF) of
Residuals")

plt.show()

from statsmodels.stats.diagnostic import het_breuschpagan

# Perform Breusch-Pagan test
bp_test = het_breuschpagan(residuals, X)

# Display results
print("\nBreusch-Pagan Test for Heteroscedasticity:")
print(f"LM Statistic: {bp_test[0]}")
print(f"p-value: {bp_test[1]}")

from statsmodels.tsa.arima.model import ARIMA

# Fit ARIMA model to the residuals
print("Fitting ARIMA(1,0,0) model to residuals...")
arima_model = ARIMA(residuals, order=(1, 0, 0)).fit()

# Display ARIMA model summary
print(arima_model.summary())

# Extract ARIMA residuals
arima_residuals = arima_model.resid

# Plot ACF and PACF for ARIMA residuals
fig, ax = plt.subplots(1, 2, figsize=(12, 6))

plot_acf(arima_residuals, lags=20, ax=ax[0])
ax[0].set_title("ACF of ARIMA Residuals")

plot_pacf(arima_residuals, lags=20, ax=ax[1])
ax[1].set_title("PACF of ARIMA Residuals")

plt.show()

from arch import arch_model

# Fit GARCH(1,1) model on ARIMA residuals
print("Fitting GARCH(1,1) model to ARIMA residuals...")
garch_model = arch_model(arima_residuals, vol='Garch', p=1, q=1,
rescale=False)
garch_results = garch_model.fit(disp='off')

# Display GARCH model summary
print("\nGARCH(1,1) Model Results:")
print(garch_results.summary())
```

```
# Plot conditional volatility
plt.figure(figsize=(12, 6))
plt.plot(garch_results.conditional_volatility, color='purple')
plt.title("Conditional Volatility from GARCH(1,1) Model")
plt.xlabel("Date")
plt.ylabel("Volatility")
plt.grid()
plt.show()

# Forecast future volatility
forecast_horizon = 8
garch_forecast = garch_results.forecast(horizon=forecast_horizon)

# Extract forecasted volatility
forecast_volatility = garch_forecast.variance.iloc[-1] ** 0.5
print("\nForecasted Volatility for Next Periods:")
print(forecast_volatility)

# Forecast future volatility
forecast_horizon = 8
garch_forecast = garch_results.forecast(horizon=forecast_horizon)

# Extract forecasted volatility
forecast_volatility = garch_forecast.variance.iloc[-1] ** 0.5

# Create a plot for forecasted volatility
import matplotlib.pyplot as plt
forecast_periods = range(1, forecast_horizon + 1)
plt.figure(figsize=(10, 6))
plt.plot(forecast_periods, forecast_volatility, marker='o', linestyle='-
-', color='purple')
plt.title("Forecasted Volatility for Next 8 Periods (GARCH(1,1))")
plt.xlabel("Period")
plt.ylabel("Volatility")
plt.grid()
plt.show()

# Forecast GDP for the next 8 quarters (adjust as needed)
forecast_periods = 8

# Forecasting future residuals
print("Forecasting future residuals using ARIMA model...")
forecast_residuals =
arima_model.forecast(steps=forecast_periods)

# Add ARIMA residual forecasts back to the predicted GDP (adjusting
for mean)
future_index = pd.date_range(data.index[-1],
periods=forecast_periods + 1, freq='QE')[1:]
future_gdp = data['Real_GDP'].iloc[-1] +
forecast_residuals.cumsum()

# Combine actual and forecasted GDP values
forecast_df = pd.DataFrame({
  'Forecasted_GDP': future_gdp
}, index=future_index)
```

```python
# Plot Actual and Forecasted GDP
plt.figure(figsize=(12, 6))
plt.plot(data.index, data['Real_GDP'], label="Actual GDP", color='blue')
plt.plot(forecast_df.index, forecast_df['Forecasted_GDP'], label="Forecasted GDP", linestyle='--', color='red')
plt.title("Actual vs Forecasted GDP")
plt.xlabel("Date")
plt.ylabel("Real GDP")
plt.legend()
plt.grid()
plt.show()

# Display forecasted GDP values
print("Forecasted GDP values:\n", forecast_df)
```