Econometric Analysis - Dynamic Relationship Between Inflation and Unemployment

JAKUB PODZIEWSKI, MARCH 2024

1.0 Introduction

This report examines the relationship between inflation, as measured by the Consumer Price Index (CPI), and unemployment rate to understand how inflation trends may influence labour market conditions. Using monthly data from 1947 to 2024, we apply econometric methods, including autoregressive (AR) and autoregressive distributed lag (ADL) models, to assess the persistence of inflation shocks and the dynamic effects on unemployment.

Key analyses include out-of-sample forecast evaluation, cointegration testing for long-run relationships, and a GARCH model to detect volatility clustering in CPI rates. Together, these insights aim to shed light on the complex interaction between inflation and unemployment, offering relevant considerations for economic policy. The analysis reveals that inflation shows persistence, with past CPI values significantly impacting current rates. The ADL(4,4) model indicates that while unemployment has a delayed influence on inflation, its direct effect is limited. Out-of-sample forecasts highlight weaker predictive power, suggesting model limitations for future projections. Cointegration tests show a weak long-term relationship between CPI and unemployment, but its suitability is questionable given differing economic influences. Finally, GARCH results confirm volatility clustering in inflation, with periods of high volatility persisting longer during economic shocks. These findings underscore inflation's lasting impact and the complexity of its relationship with unemployment, which is essential for policy considerations.

1.1 Descriptive analysis

1.1.1 Data Description

The primary outcome variable in this analysis is the monthly unemployment rate, which reflects the percentage of unemployed individuals in the labor force, serving as an indicator of economic health and labor market conditions. The primary explanatory variable is the monthly Consumer Price Index (CPI) rate, measured as the month-over-month percentage change. These metric capture inflation trends, which are expected to influence unemployment rates through the economic climate.

The primary research question of this report is as follows: Does the monthly inflation rate, as indicated by the CPI, impact the unemployment rate? The question examines if monthly inflation, measured by CPI changes, affects the unemployment rate. It aims to determine whether inflation trends influence labor market conditions, shedding light on possible economic linkages between price levels and employment.

1.1.1 Data Description

The time series plot of the CPI rate illustrates inflation trends from 1947 to 2024 (Figure 1.). There are notable fluctuations, with visible peaks and dips corresponding to economic cycles. Before beginning the analysis, significant outliers in the CPI and unemployment rate were identified using a Z-score threshold of 3. These outliers were set to 'NaN' values and then handled through linear interpolation to ensure a more stable trend without abrupt spikes.

The resulting CPI time series appears consistent, with no remaining significant outliers, allowing for a smoother analysis of inflation's impact on unemployment.

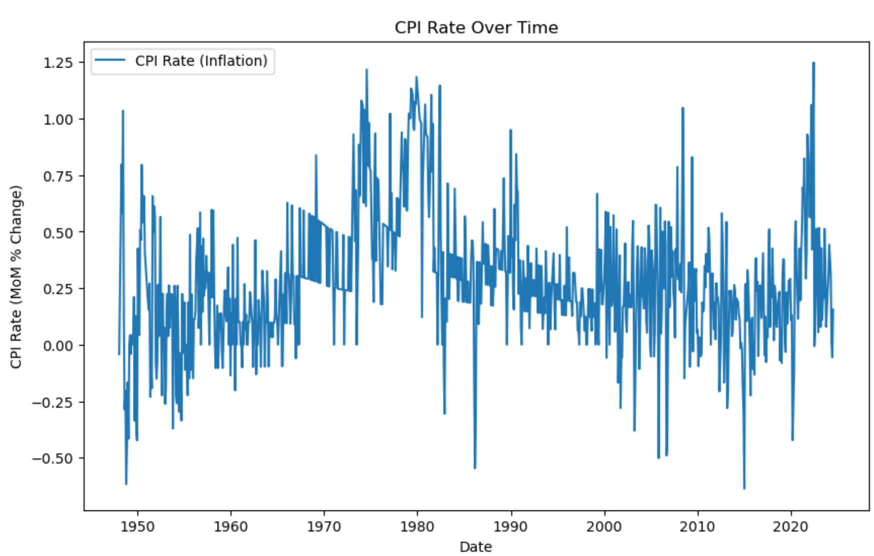

Figure 1

1.2 Autoregression Analysis of a Time Series

1.2.1 Estimating an Autoregression Model

The Bayesian Information Criterion (BIC) was calculated for AR models with lag lengths from 1 to 4. The BIC values for each model were as follows:

BIC(1): 4.8229
BIC(2): -40.8318
BIC(3): -73.4340
BIC(4): -99.1625

The AR(4) model had the lowest BIC, indicating it as the optimal lag length. Thus, an AR(4) model was selected for the final analysis.

```
                            AutoReg Model Results
==============================================================================
Dep. Variable:               CPI_RATE   No. Observations:                  918
Model:                     AutoReg(4)   Log Likelihood                  66.626
Method:               Conditional MLE   S.D. of innovations              0.225
Date:                Sat, 02 Nov 2024   AIC                           -123.252
Time:                        02:16:03   BIC                            -99.163
Sample:                             4   HQIC                          -114.056
                                  918
===============================================================================
                  coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------
CPI_RATE.L1     0.5267      0.032     16.216      0.000       0.463       0.590
CPI_RATE.L2     0.0933      0.037      2.542      0.011       0.021       0.165
CPI_RATE.L3     0.1038      0.037      2.827      0.005       0.032       0.176
CPI_RATE.L4     0.1765      0.032      5.439      0.000       0.113       0.240
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0550           -0.0000j            1.0550           -0.0000
AR.2            0.1389           -1.6664j            1.6722           -0.2368
AR.3            0.1389           +1.6664j            1.6722            0.2368
...
-----------------------------------------------------------------------------
```

Each lag coefficient is statistically significant, suggesting that past CPI values up to four months ago have a meaningful impact on the current CPI rate, indicating persistence in inflationary shocks. The coefficient on Y_{t-1}in the AR(1) model was significant, indicating that last month's inflation rate influences the current rate. In the AR(4) model, all lag terms up to four months were significant, suggesting that inflation shocks persist over multiple months. This persistence highlights the gradual adjustment of inflation rates in response to economic changes.

To test for stationarity, the Augmented Dickey-Fuller (ADF) test was conducted. The results were as follows:

ADF Statistic: -3.5516

p-value: 0.0068

Critical Values: {'1%': -3.4376, '5%': -2.8647, '10%': -2.5685}

Since the ADF statistic is lower than the 1% critical value, we reject the null hypothesis of a unit root. This indicates that the CPI rate series is stationary, which aligns with economic theory that inflation, while cyclical, tends to fluctuate around a stable mean rather than drift over time (Ascari & Sbordone, 2014). Because the series is stationary, no transformation (such as differencing) is necessary, and the AR(4) model can be reliably used for further analysis.

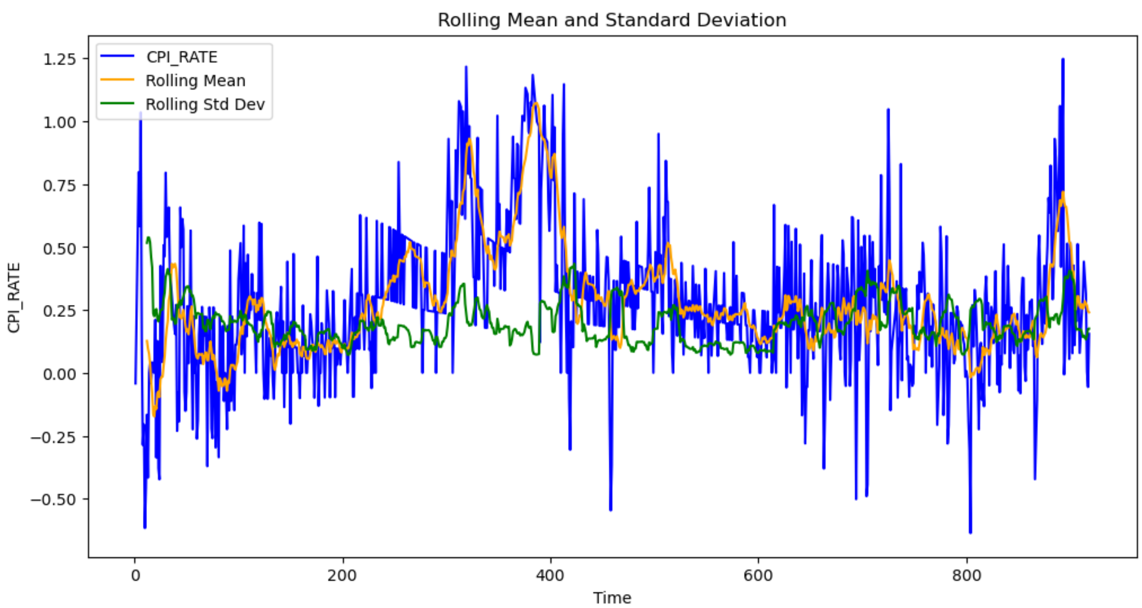

Figure 2.

The positive and significant coefficients in the AR(4) model suggest that a shock to the CPI rate has a lasting impact, diminishing gradually over the following months. Using a rolling window of 12 months (Figure 2.), the rolling standard deviation remained constant over time, while the rolling mean showed a noticeable spike, indicating a temporary disturbance rather than a permanent shift in inflation dynamics. This observation reinforces the persistence of inflationary shocks but suggests that the system eventually reverts to a stable mean after such events.

1.2.2 Estimating an Autoregression Distributed Lag Model

The ADL(1,1) model was estimated with the CPI rate as the dependent variable Y_t and an additional explanatory variable X_t (e.g., unemployment rate). The results for the ADL(1,1) model are as follows:

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0730	0.027	2.712	0.007	0.020	0.126
y.L1	0.6233	0.026	24.151	0.000	0.573	0.674
X_lag1	0.0057	0.004	1.291	0.197	-0.003	0.014

The coefficient of X_{t-1} is positive but not statistically significant at the 5% level ($p = 0.197$), indicating that the lagged value of X does not have a significant effect on Y in this ADL(1,1) model.

The lowest BIC value was observed at lag length 4, so the ADL(4,4) model was chosen as the optimal specification.

The ADL(4,4) model results are as follows:

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0431	0.026	1.650	0.099	-0.008	0.094
y.L1	0.4820	0.033	14.745	0.000	0.418	0.546
y.L2	0.0693	0.036	1.910	0.056	-0.002	0.140
y.L3	0.0773	0.036	2.143	0.032	0.007	0.148
y.L4	0.1414	0.032	4.358	0.000	0.078	0.205
X_lag4	0.0037	0.004	0.868	0.385	-0.005	0.012

Roots

The coefficients for Y_{t-1} through Y_{t-4} were significant, showing that past values of the CPI rate have a cumulative and prolonged effect on current values. However, Y_{t-4} was not significant ($p = 0.385$), indicating that the lagged value of unemployment rate does not strongly influence the CPI rate at this lag length.

The Granger causality test was conducted to determine if the lags of X_t are jointly significant predictors of Y_t. Only at lag 4 was the test statistically significant ($p = 0.0104$), suggesting that the unemployment rate lags have some predictive power over the CPI rate at the 4th lag length.

The ADL(4,4) model indicates that the CPI rate (inflation) is influenced by its own past values up to four months back, showing persistence in inflation shocks. However, the lags of the unemployment rate do not have a statistically significant direct impact on CPI in this model, except as indicated by the Granger causality test at lag 4. This result implies that while the unemployment rate may have some delayed influence on inflation, its impact is not immediate or strong within this model. This could suggest that factors affecting inflation are primarily internal (Blanchard & Quah, 1989) or that other macroeconomic variables should be considered for a more comprehensive understanding of inflation dynamics.

1.2.3 Out-Of-Sample Forecast Performance

Using the ADL(1,1) model, a pseudo out-of-sample forecasting approach was applied by excluding the final 25% of the sample as the test set. The model was then evaluated by comparing the within-sample Standard Error of Regression (SER) and the out-of-sample Root Mean Squared Forecast Error (RMSFE).

Within-Sample SER: 0.2224

Out-Of-Sample RMSFE: 0.2920

The RMSFE is larger than the within-sample SER, indicating that the model performs less accurately when predicting out-of-sample data compared to within-sample. This discrepancy suggests that the model's forecast errors are larger out-of-sample, implying that it struggles to generalize well beyond the data it was trained on. Since the RMSFE exceeds the SER, it indicates that the ADL(1,1) model may not be fully capable of accurately predicting future values outside the training sample. This outcome could be due to the inherent variability in the test set, model limitations, or unaccounted-for factors affecting the series (Diebold & Mariano, 1995). In practice, this means that while the model may capture within-sample dynamics, its out-of-sample predictive power is relatively weaker, limiting its reliability for forecasting purposes.

1.3 Dynamic causal effects

To estimate the dynamic multipliers for the effect of X_t on Y_t, a distributed lag model with r=3r = 3r=3 was used, including X_t, X_{t-1}, and X_{t-2} as predictors. An AR(1) error structure was assumed, and the Cochrane-Orcutt method was applied (without iteration) to correct for autocorrelation.

Results of the GLS regression (Figure 3.):

const	0.0798	0.026	3.012	0.003	0.028	0.132
X_adj	-0.0394	0.034	-1.174	0.241	-0.105	0.026
X_lag1_adj	0.0874	0.035	2.489	0.013	0.018	0.156
X_lag2_adj	-0.0363	0.034	-1.084	0.279	-0.102	0.029

The coefficient for X_{t-1}, 1 is positive and statistically significant ($p = 0.013$), suggesting that the lagged effect of X on Y at a one-period lag is the strongest among the included terms. Both X_t, and X_{t-2} have non-significant effects, implying that the immediate and two-period lagged effects of X on Y are weaker and statistically insignificant.

In this context, it is essential to discuss the assumption of strict exogeneity. Strict exogeneity would imply that X_t and its lags are uncorrelated with the error term across all time periods (Wooldridge, 2010). However, in economic settings, this assumption may often be violated (Stock & Watson, 2015). For example, if there is feedback between y_t and X_t, where economic agents base current decisions on anticipated future changes, then strict exogeneity could be compromised. This can be looked at in terms of policy decisions. If policymakers adjust unemployment-related policies based on expected future inflation, this could create a feedback loop (Sims, 1980) where X_t is influenced by anticipated future values of Y, violating strict exogeneity. In this scenario, we can only assume weak exogeneity, meaning X_t may affect Y_t but is not influenced by future values of Y_t.

The significant one-period lagged effect implies that changes in X have a delayed but noticeable impact on Y one period later, possibly due to the time it takes for economic adjustments to take effect. The lack of significant effects at the contemporaneous and two-period lag levels indicates that this effect does not persist over longer periods

1.4 Multiperiod Forecasting

Using the ADL(4,4) model developed, two forecasting methods—iterative and direct—were applied to project the values for the next ten periods.

In iterative forecasting, each period's forecasted value is used as an input for the following period's prediction. The iterative forecasts for each period up to 10 periods are as follows:

```
Iterative forecasts for each period up to 10:
Forecast for period 1: 0.6438940684869734
Forecast for period 2: 0.9742853510259678
Forecast for period 3: 1.2376939640147266
Forecast for period 4: 1.4815386494497025
Forecast for period 5: 1.6175285622769646
Forecast for period 6: 1.7058627854608908
Forecast for period 7: 1.7665282571790932
Forecast for period 8: 1.7977701885929511
Forecast for period 9: 1.8164323411463004
Forecast for period 10: 1.8275743586728377
```

The final forecasted value at the end of the ten-period horizon is **1.8276**.

Direct forecasting estimates each period independently based on the model parameters, without using the preceding forecasted values. This approach minimizes error accumulation but may not fully capture sequential dependencies.

```
Direct forecast for period 1: 0.7597756225694081
Direct forecast for period 2: 0.8123719550266331
Direct forecast for period 3: 0.7575013636679808
Direct forecast for period 4: 0.827096576709365
Direct forecast for period 5: 0.827096576709365
Direct forecast for period 6: 0.827096576709365
Direct forecast for period 7: 0.827096576709365
Direct forecast for period 8: 0.827096576709365
Direct forecast for period 9: 0.827096576709365
Direct forecast for period 10: 0.827096576709365
```

The direct method projects a stable value of **0.8271** starting from Period 5, which may indicate a steady-state expectation without iterative error propagation.

The iterative method's reliance on previous forecasts introduces a potential for cumulative error (Box, Jenkins, & Reinsel, 2008), which could lead to upward or downward drift over multiple periods. In contrast, direct forecasts avoid this issue by treating each period independently, leading to more stable predictions. Iterative forecasting may capture more of the sequential dependencies in the data, but with possible error compounding. Direct forecasting, while potentially less dynamic, might offer greater stability, especially over longer horizons where cumulative error could distort the trend. In this case, the iterative forecast predicts a continued increase up to **1.8276** by Period 10, reflecting recent trends and the lagged effects modelled in the ADL(4,4) structure.

1.5 Cointegration

To test for cointegration between Y_t and X_t, we applied the two-stage Engle-Granger method:

1. First Stage (Cointegration Regression): We regressed Y_t on X_t to estimate the long-run equilibrium relationship. The regression results are as follows:

   Intercept (const): 0.2073, significant at the 1% level.

   Coefficient of UNRATE: 0.0128, significant at the 5% level (p = 0.028).

These results indicate a statistically significant relationship between the unemployment rate and CPI rate in the long run, with a low R2R^2R2 value (0.5%), suggesting that unemployment explains only a small portion of the variance in CPI.

2. Second Stage (ADF Test on Residuals): After obtaining the residuals from the first-stage regression, we conducted the Augmented Dickey-Fuller (ADF) test to check for stationarity. The results are as follows:

   ADF Statistic: -3.7236

   p-value: 0.0038

   Critical Values: {'1%': -3.4376, '5%': -2.8647, '10%': -2.5685}

Since the ADF statistic is below the 1% critical value, we reject the null hypothesis of a unit root, indicating that the residuals are stationary. This suggests that Y_t and X_t are likely cointegrated.

Testing for cointegration in this context might not be entirely appropriate. Cointegration is typically expected in relationships where two or more series move together in the long run due to shared economic forces, such as GDP and consumption (Engle & Granger, 1987). Here, while CPI and unemployment may have some economic interaction, they are influenced by a range of distinct factors, so a long-term equilibrium relationship may not be plausible.

Despite the questionable suitability, the test results suggest cointegration, implying that CPI and unemployment rates have a stable long-run relationship, with deviations that revert to equilibrium. This might reflect an indirect relationship influenced by broader economic cycles. However, given the low R2R^2R2, this relationship is weak and should be interpreted cautiously. In this setting, strict exogeneity may be violated if CPI affects future unemployment expectations (Hamilton, 1994), as policymakers or economic agents may adjust their actions in response to anticipated inflation.

1.6 Volatility Clustering Analysis

An ARMA(1,1)-GARCH(1,1) model was estimated on the CPI rate data to assess whether the series displays volatility clustering (periods of high volatility followed by more high volatility, and low volatility followed by low volatility). Figure 3 shows the conditional volatility of the CPI rate over time, illustrating these clusters.

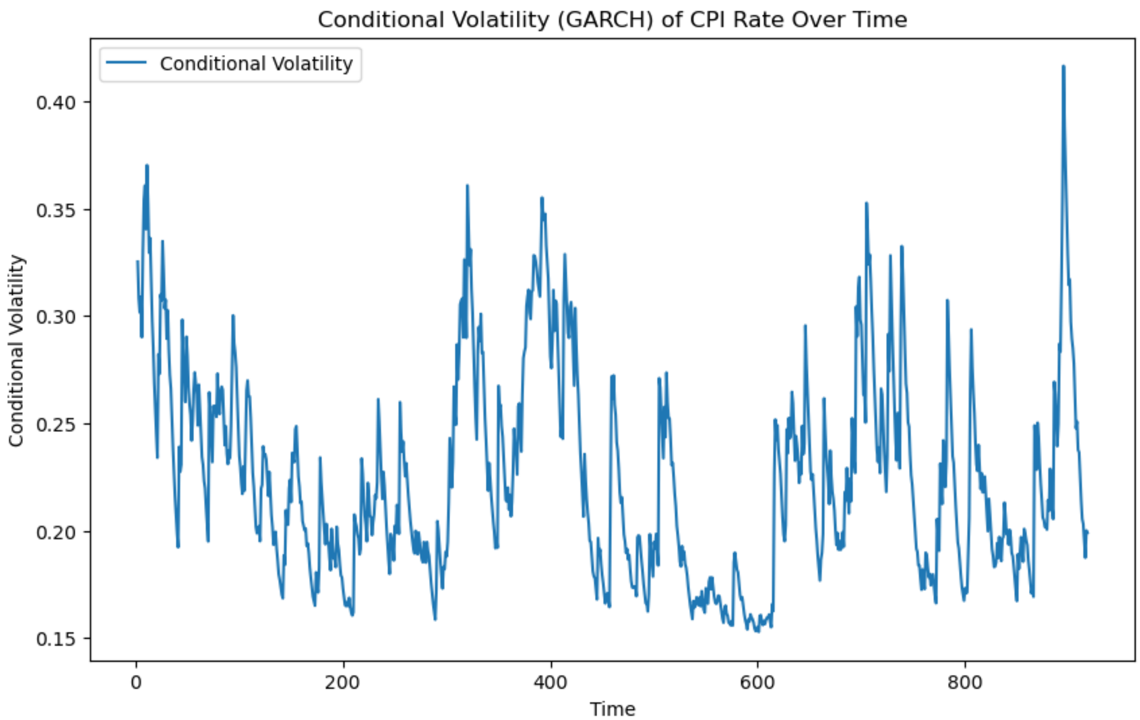

Figure 3.

The results are summarized below:

Mean Model

	coef	std err	t	P>|t|	95.0% Conf. Int.
Const	0.1094	1.056e-02	10.362	3.685e-25	[8.872e-02, 0.130]
CPI_RATE[1]	0.5567	4.201e-02	13.252	4.410e-40	[0.474, 0.639]

Volatility Model

	coef	std err	t	P>|t|	95.0% Conf. Int.
omega	2.7944e-03	2.169e-03	1.288	0.198	[-1.457e-03,7.046e-03]
alpha[1]	0.0951	4.081e-02	2.330	1.982e-02	[1.509e-02, 0.175]
beta[1]	0.8513	7.644e-02	11.138	8.217e-29	[0.702, 1.001]

Mean Model:

The constant term and the AR(1) term (CPI_RATE[1]) are both statistically significant. The AR(1) term suggests a strong dependence on past values, indicating that the CPI rate follows an autoregressive process with high persistence.

Volatility Model (GARCH):

The significant coefficients for α_1 and β_1] indicate the presence of conditional heteroskedasticity in the series. The α_1 coefficient (0.0951) captures the impact of past shocks on current volatility, while the $\beta_1$coefficient (0.8513) represents the persistence of volatility over time. The large and significant $\beta_1$coefficient suggests that periods of high volatility in the CPI rate are likely to be long-lived. This means that once volatility rises, it tends to persist before gradually declining, which is characteristic of volatility clustering.

The presence of significant $\alpha_1$and $\beta_1$coefficients in the GARCH model implies that the CPI rate exhibits conditional heteroskedasticity, with volatility clustering. Economically, this suggests that inflation rates may experience extended periods of instability during economic turbulence or shocks, followed by periods of relative calm (Bollerslev, 1986). The high persistence β_1 indicates that inflation volatility tends to remain elevated once it increases, reflecting a tendency for economic uncertainty to sustain inflation variability over time (Poon & Granger, 2003).

1.7 Conclusion

This report analyzed the relationship between monthly CPI (inflation) and unemployment rates. We found that inflation exhibits persistence, with past values significantly influencing current rates. The ADL(4,4) model indicated that unemployment affects inflation over multiple periods, but out-of-sample forecasts showed limited predictive power. Although a weak cointegration relationship was detected, it may not be entirely appropriate in this context. Volatility analysis confirmed clustering in inflation, suggesting that periods of high volatility tend to persist. These findings underscore the complex, lasting impact of inflation on economic conditions, relevant for policy considerations.

1.8 Reference List

Ascari, G., & Sbordone, A. M. (2014). The macroeconomics of trend inflation. *Journal of Economic Literature*, 52(3), 679-739.

Blanchard, O. J., & Quah, D. (1989). "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, 79(4), 655–673.

Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, 31(3), 307–327.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control* (4th ed.). Wiley.

Diebold, F. X., & Mariano, R. S. (1995). "Comparing predictive accuracy." *Journal of Business & Economic Statistics*, 13(3), 253–263.

Engle, R. F., & Granger, C. W. J. (1987). "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, 55(2), 251–276.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Poon, S.-H., & Granger, C. W. J. (2003). "Forecasting Volatility in Financial Markets: A Review." *Journal of Economic Literature*, 41(2), 478–539.

Sims, C. A. (1980). "Macroeconomics and Reality." *Econometrica*, 48(1), 1–48.

Stock, J. H., & Watson, M. W. (2015). *Introduction to Econometrics* (3rd ed.). Pearson.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.