Time Series: Dynamic Models and Autocorrelation

Modeling Temporal Dependence in Economic Data

Jake Anderson

May 16, 2026

Outline

Cross-Section vs Time Series

Cross-sectional data: observations are independent draws from a population.

- Randomly sample 500 workers $\implies$ one worker's wage tells you nothing about another's

Cross-Section vs Time Series

Cross-sectional data: observations are independent draws from a population.

- Randomly sample 500 workers $\implies$ one worker's wage tells you nothing about another's

Time series data: observations are ordered in time and *dependent on each other*.

- This quarter's GDP is strongly related to last quarter's GDP
- Today's stock return depends on yesterday's

Cross-Section vs Time Series

Cross-sectional data: observations are independent draws from a population.

- Randomly sample 500 workers $\Longrightarrow$ one worker's wage tells you nothing about another's

Time series data: observations are ordered in time and *dependent on each other*.

- This quarter's GDP is strongly related to last quarter's GDP
- Today's stock return depends on yesterday's

$\Longrightarrow$ The i.i.d. assumption fails. We need new tools that account for **temporal dependence**.

Consequences of Temporal Dependence

When observations are correlated over time:

1. Standard OLS **standard errors are wrong** (usually too small)
2. The **order** of observations contains information we should exploit
3. Past values of Y can help **predict** future values

Consequences of Temporal Dependence

When observations are correlated over time:

1. Standard OLS **standard errors are wrong** (usually too small)
2. The **order** of observations contains information we should exploit
3. Past values of Y can help **predict** future values

This lecture covers:

- How to characterize temporal dependence (stationarity, ACF)
- Models that exploit it (AR, ARDL)
- How to detect it in regression residuals (Breusch-Godfrey)
- How to forecast with it

Covariance Stationarity: Definition

A time series $\{Y_t\}$ is **covariance stationary** if:

1. **Constant mean:** $E(Y_t) = \mu$ for all t
2. **Constant variance:** $\text{Var}(Y_t) = \sigma^2$ for all t
3. **Covariance depends only on distance:** $\text{Cov}(Y_t, Y_{t-k})$ depends only on k, not on t

Covariance Stationarity: Definition

A time series $\{Y_t\}$ is **covariance stationary** if:

1. **Constant mean:** $E(Y_t) = \mu$ for all t
2. **Constant variance:** $\text{Var}(Y_t) = \sigma^2$ for all t
3. **Covariance depends only on distance:** $\text{Cov}(Y_t, Y_{t-k})$ depends only on k, not on t

Why does this matter?

Covariance Stationarity: Definition

A time series $\{Y_t\}$ is **covariance stationary** if:

1. **Constant mean:** $E(Y_t) = \mu$ for all t
2. **Constant variance:** $\text{Var}(Y_t) = \sigma^2$ for all t
3. **Covariance depends only on distance:** $\text{Cov}(Y_t, Y_{t-k})$ depends only on k, not on t

Why does this matter?

If the mean or variance drifts over time, a regression estimated on one period will not apply to another. The coefficients themselves would be unstable.

Covariance Stationarity: Definition

A time series $\{Y_t\}$ is **covariance stationary** if:

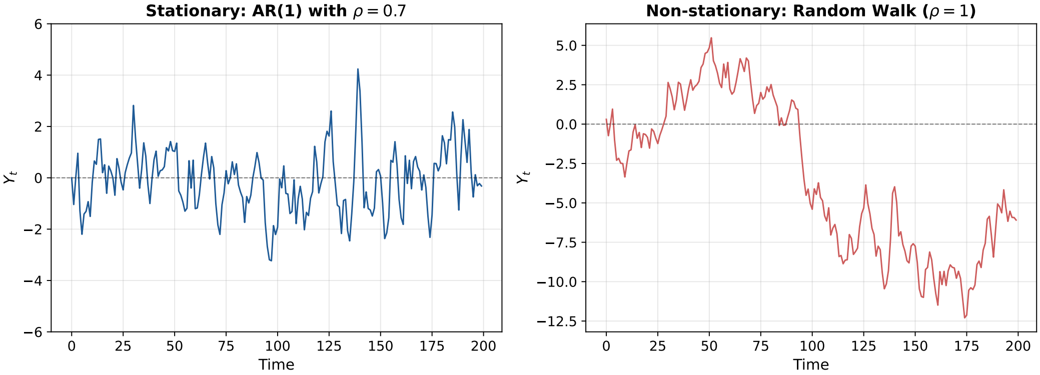
1. **Constant mean:** $E(Y_t) = \mu$ for all t
2. **Constant variance:** $\text{Var}(Y_t) = \sigma^2$ for all t
3. **Covariance depends only on distance:** $\text{Cov}(Y_t, Y_{t-k})$ depends only on k, not on t

Why does this matter?

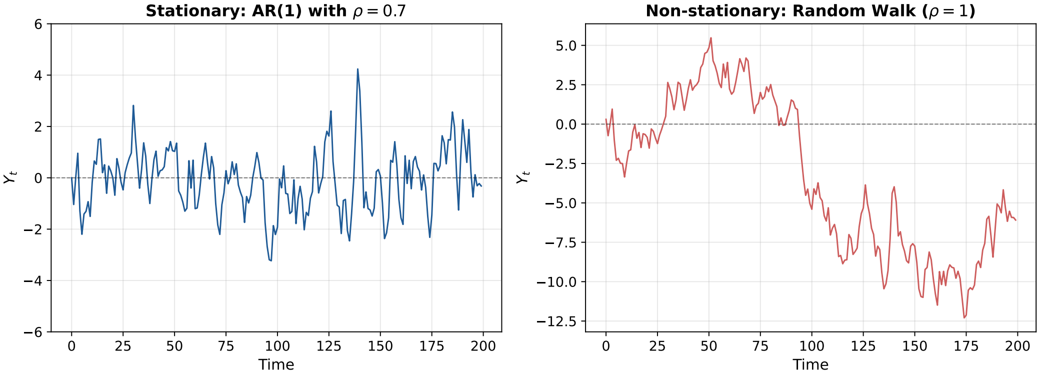
If the mean or variance drifts over time, a regression estimated on one period will not apply to another. The coefficients themselves would be unstable.

$\Longrightarrow$ Stationarity ensures that estimated statistical properties are meaningful and stable.

Stationary vs Non-stationary: Visual Comparison



Stationary vs Non-stationary: Visual Comparison



Left: mean-reverting; always pulled back toward zero. **Right:** wanders without bound; no tendency to return.

The Random Walk

The simplest non-stationary process:

$$Y_t = Y_{t-1} + v_t, \qquad v_t \sim (0, \sigma_v^2)$$

The Random Walk

The simplest non-stationary process:

$$Y_t = Y_{t-1} + v_t, \qquad v_t \sim (0, \sigma_v^2)$$

Unraveling: $Y_t = Y_0 + v_1 + v_2 + \cdots + v_t$

The Random Walk

The simplest non-stationary process:

$$Y_t = Y_{t-1} + v_t, \qquad v_t \sim (0, \sigma_v^2)$$

Unraveling: $Y_t = Y_0 + v_1 + v_2 + \cdots + v_t$

Properties:

- $E(Y_t) = Y_0$ (constant, so mean is OK)
- $\mathrm{Var}(Y_t) = t \cdot \sigma_v^2$ (grows over time $\implies$ non-stationary)

The Random Walk

The simplest non-stationary process:

$$Y_t = Y_{t-1} + v_t, \qquad v_t \sim (0, \sigma_v^2)$$

Unraveling: $Y_t = Y_0 + v_1 + v_2 + \cdots + v_t$

Properties:

- $E(Y_t) = Y_0$ (constant, so mean is OK)
- $\text{Var}(Y_t) = t \cdot \sigma_v^2$ (grows over time $\implies$ non-stationary)

$\implies$ Shocks accumulate forever. The series has no tendency to revert to any level.

The Random Walk

The simplest non-stationary process:

$$Y_t = Y_{t-1} + v_t, \qquad v_t \sim (0, \sigma_v^2)$$

Unraveling: $Y_t = Y_0 + v_1 + v_2 + \cdots + v_t$

Properties:

- $E(Y_t) = Y_0$ (constant, so mean is OK)
- $\text{Var}(Y_t) = t \cdot \sigma_v^2$ (grows over time $\implies$ non-stationary)

$\implies$ Shocks accumulate forever. The series has no tendency to revert to any level.

Example: Stock prices are approximately random walks. That is why financial economists work with *returns* (which are approximately stationary) rather than price levels.

Measuring Temporal Dependence: The ACF

The **autocorrelation function** measures how correlated a series is with its own past:

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)}$$

Measuring Temporal Dependence: The ACF

The **autocorrelation function** measures how correlated a series is with its own past:

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)}$$

- $\rho_0 = 1$ always (correlation with itself)
- $\rho_1 =$ correlation between consecutive observations
- $\rho_k =$ correlation between observations k periods apart

Measuring Temporal Dependence: The ACF

The **autocorrelation function** measures how correlated a series is with its own past:

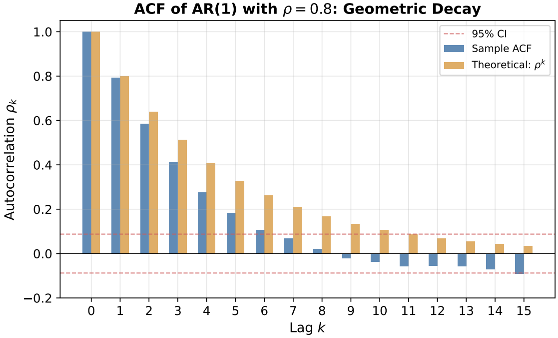
$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\text{Var}(Y_t)}$$

- $\rho_0 = 1$ always (correlation with itself)
- $\rho_1 =$ correlation between consecutive observations
- $\rho_k =$ correlation between observations k periods apart

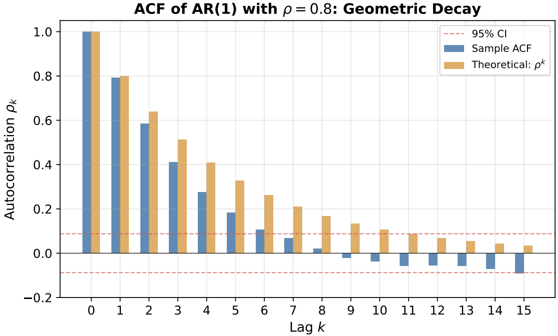
The **sample autocorrelation**:

$$r_k = \frac{\sum_{t=k+1}^{T}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2}$$

Reading an ACF Plot



Reading an ACF Plot



How to read it:

- Bars beyond the dashed lines = statistically significant autocorrelation
- For an AR(1), the ACF decays **geometrically**: $\rho_k = \rho^k$
- If residuals from a model show significant ACF spikes $\implies$ the model is missing dynamics

Autoregressive Models: AR(p)

An **AR(p) model** says today's value depends on its own past p values:

$$Y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots + \theta_p Y_{t-p} + v_t$$

where v_t is **white noise**: zero mean, constant variance, no serial correlation.

Autoregressive Models: AR(p)

An **AR(p) model** says today's value depends on its own past p values:

$$Y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots + \theta_p Y_{t-p} + v_t$$

where v_t is **white noise**: zero mean, constant variance, no serial correlation.

AR(1): $Y_t = \delta + \theta_1 Y_{t-1} + v_t$

- $|\theta_1| < 1$: stationary; shocks fade over time
- $\theta_1 = 1$: random walk (non-stationary)
- $|\theta_1| > 1$: explosive (non-stationary)

Autoregressive Models: AR(p)

An **AR(p) model** says today's value depends on its own past p values:

$$Y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \cdots + \theta_p Y_{t-p} + v_t$$

where v_t is **white noise**: zero mean, constant variance, no serial correlation.

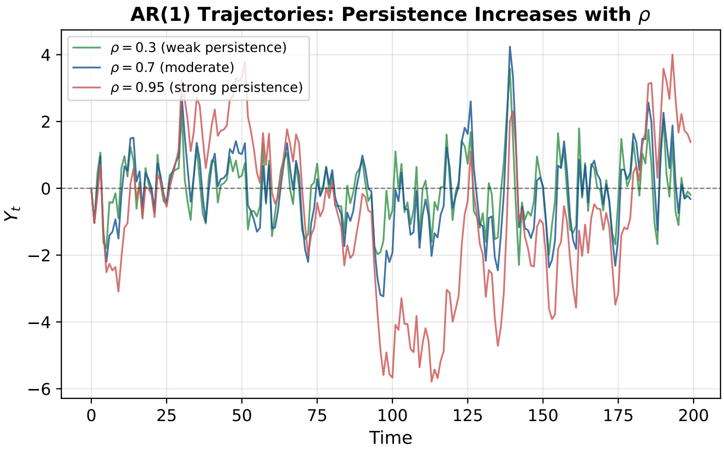
AR(1): $Y_t = \delta + \theta_1 Y_{t-1} + v_t$

- $|\theta_1| < 1$: stationary; shocks fade over time
- $\theta_1 = 1$: random walk (non-stationary)
- $|\theta_1| > 1$: explosive (non-stationary)

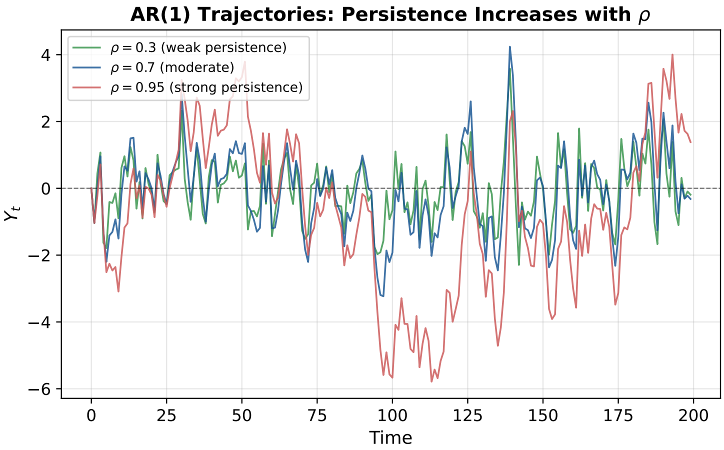
AR(2): $Y_t = \delta + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + v_t$

- Allows richer dynamics: oscillations, humps

How Persistence Changes with ρ



How Persistence Changes with ρ



Same shocks, different ρ. Higher $\rho \implies$ slower mean reversion, longer "memory."

Model Selection: AIC and BIC

How many lags? Use **information criteria**:

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2K}{T}, \qquad \text{BIC} = \ln(\hat{\sigma}^2) + \frac{K\ln(T)}{T}$$

Model Selection: AIC and BIC

How many lags? Use **information criteria**:

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2K}{T}, \qquad \text{BIC} = \ln(\hat{\sigma}^2) + \frac{K\ln(T)}{T}$$

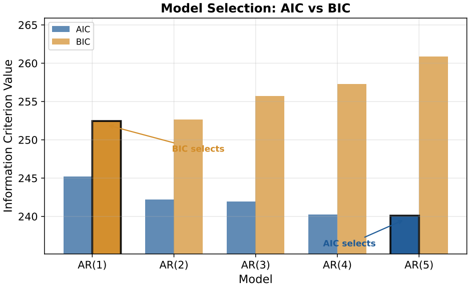
- Both penalize complexity (K = number of parameters); BIC penalizes more heavily
- Choose the model with the **lowest** criterion value

Model Selection: AIC and BIC

How many lags? Use **information criteria**:

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2K}{T}, \qquad \text{BIC} = \ln(\hat{\sigma}^2) + \frac{K\ln(T)}{T}$$

- Both penalize complexity ($K =$ number of parameters); BIC penalizes more heavily
- Choose the model with the **lowest** criterion value

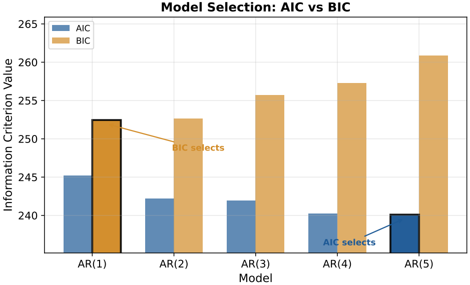


Model Selection: AIC and BIC

How many lags? Use **information criteria**:

$$\text{AIC} = \ln(\hat{\sigma}^2) + \frac{2K}{T}, \qquad \text{BIC} = \ln(\hat{\sigma}^2) + \frac{K\ln(T)}{T}$$

- Both penalize complexity (K = number of parameters); BIC penalizes more heavily
- Choose the model with the **lowest** criterion value



When AIC and BIC disagree, BIC is typically preferred for consistent selection.

Autoregressive Distributed Lag: ARDL(p, q)

An **ARDL** model includes lags of both Y and an explanatory variable X:

$$Y_t = \delta + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_q X_{t-q} + v_t$$

Autoregressive Distributed Lag: ARDL(p, q)

An **ARDL** model includes lags of both Y and an explanatory variable X:

$$Y_t = \delta + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_q X_{t-q} + v_t$$

Example: ARDL(1,1) Phillips Curve

$$\mathrm{INF}_t = \delta + \theta_1 \,\mathrm{INF}_{t-1} + \delta_0 \,\mathrm{DU}_t + \delta_1 \,\mathrm{DU}_{t-1} + v_t$$

Autoregressive Distributed Lag: ARDL(p, q)

An **ARDL** model includes lags of both Y and an explanatory variable X:

$$Y_t = \delta + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_q X_{t-q} + v_t$$

Example: ARDL(1,1) Phillips Curve

$$\text{INF}_t = \delta + \theta_1 \, \text{INF}_{t-1} + \delta_0 \, \text{DU}_t + \delta_1 \, \text{DU}_{t-1} + v_t$$

where INF = inflation, DU = change in unemployment.

Autoregressive Distributed Lag: ARDL(p, q)

An **ARDL** model includes lags of both Y and an explanatory variable X:

$$Y_t = \delta + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + \delta_0 X_t + \delta_1 X_{t-1} + \cdots + \delta_q X_{t-q} + v_t$$

Example: ARDL(1,1) Phillips Curve

$$\text{INF}_t = \delta + \theta_1 \, \text{INF}_{t-1} + \delta_0 \, \text{DU}_t + \delta_1 \, \text{DU}_{t-1} + v_t$$

where INF = inflation, DU = change in unemployment.

$\implies$ Current inflation depends on: (1) its own recent history, and (2) current *and* past unemployment changes.

Multipliers: How Effects Unfold Over Time

A one-unit change in X does not have an instantaneous-only effect. It unfolds:

Multipliers: How Effects Unfold Over Time

A one-unit change in X does not have an instantaneous-only effect. It unfolds:

Multiplier	**Formula**	**Interpretation**
Impact	δ_0	Immediate effect this period
Interim (1 period)	$\delta_0 + \delta_1$	Cumulative effect after 1 period
Long-run	$\dfrac{\delta_0 + \delta_1}{1 - \theta_1}$	Total effect after all dynamics play out

Multipliers: How Effects Unfold Over Time

A one-unit change in X does not have an instantaneous-only effect. It unfolds:

Multiplier	**Formula**	**Interpretation**
Impact	δ_0	Immediate effect this period
Interim (1 period)	$\delta_0 + \delta_1$	Cumulative effect after 1 period
Long-run	$\dfrac{\delta_0 + \delta_1}{1 - \theta_1}$	Total effect after all dynamics play out

The lagged dependent variable (θ_1) **propagates** the initial shock forward in time.

Multipliers: How Effects Unfold Over Time

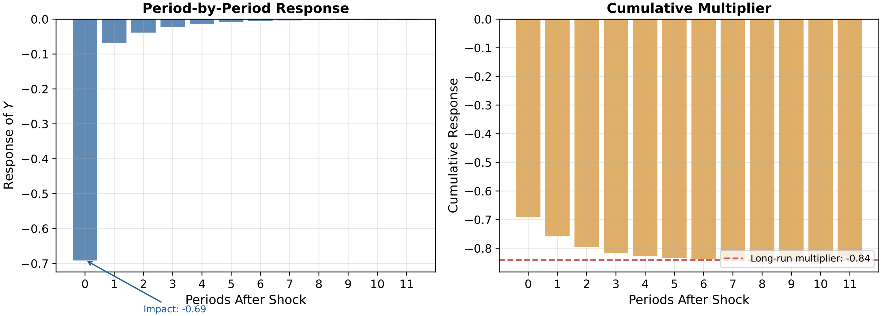
A one-unit change in X does not have an instantaneous-only effect. It unfolds:

Multiplier	**Formula**	**Interpretation**
Impact	δ_0	Immediate effect this period
Interim (1 period)	$\delta_0 + \delta_1$	Cumulative effect after 1 period
Long-run	$\frac{\delta_0 + \delta_1}{1 - \theta_1}$	Total effect after all dynamics play out

The lagged dependent variable (θ_1) **propagates** the initial shock forward in time.
$\Longrightarrow$ The long-run effect is larger (in absolute value) than the impact effect whenever $|\theta_1| > 0$.

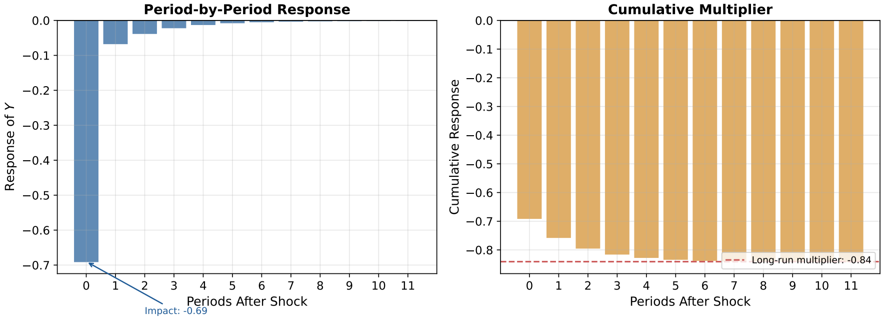
Impulse Response: Visualizing the Multipliers

ARDL(1,1) with $\hat{\theta}_1 = 0.56$, $\hat{\delta}_0 = -0.69$, $\hat{\delta}_1 = 0.32$:



Impulse Response: Visualizing the Multipliers

ARDL(1,1) with $\hat{\theta}_1 = 0.56$, $\hat{\delta}_0 = -0.69$, $\hat{\delta}_1 = 0.32$:



- Impact multiplier: -0.69
- Interim: $-0.69 + 0.32 = -0.37$
- Long-run: $\frac{-0.69+0.32}{1-0.56} = -0.84$

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\mathsf{INF}}_t = 1.2 + 0.56\,\mathsf{INF}_{t-1} - 0.69\,\mathsf{DU}_t + 0.32\,\mathsf{DU}_{t-1}$$

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\text{INF}}_t = 1.2 + 0.56\,\text{INF}_{t-1} - 0.69\,\text{DU}_t + 0.32\,\text{DU}_{t-1}$$

(a) What is the impact of a 1-unit increase in DU_t on inflation?

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\mathsf{INF}}_t = 1.2 + 0.56\,\mathsf{INF}_{t-1} - 0.69\,\mathsf{DU}_t + 0.32\,\mathsf{DU}_{t-1}$$

(a) What is the impact of a 1-unit increase in DU_t on inflation?

Solution: Impact multiplier $= \delta_0 = -0.69$. Inflation falls by 0.69 points immediately.

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\text{INF}}_t = 1.2 + 0.56\,\text{INF}_{t-1} - 0.69\,\text{DU}_t + 0.32\,\text{DU}_{t-1}$$

(a) What is the impact of a 1-unit increase in DU_t on inflation?

Solution: Impact multiplier $= \delta_0 = -0.69$. Inflation falls by 0.69 points immediately.

(b) What is the long-run effect?

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\text{INF}}_t = 1.2 + 0.56\,\text{INF}_{t-1} - 0.69\,\text{DU}_t + 0.32\,\text{DU}_{t-1}$$

(a) What is the impact of a 1-unit increase in DU_t on inflation?

Solution: Impact multiplier $= \delta_0 = -0.69$. Inflation falls by 0.69 points immediately.

(b) What is the long-run effect?

Solution: $\dfrac{\delta_0 + \delta_1}{1 - \theta_1} = \dfrac{-0.69 + 0.32}{1 - 0.56} = \dfrac{-0.37}{0.44} = -0.84$

Multiplier Practice Problem

An ARDL(1,1) model for quarterly inflation gives:

$$\widehat{\text{INF}}_t = 1.2 + 0.56\,\text{INF}_{t-1} - 0.69\,\text{DU}_t + 0.32\,\text{DU}_{t-1}$$

(a) What is the impact of a 1-unit increase in DU_t on inflation?

Solution: Impact multiplier $= \delta_0 = -0.69$. Inflation falls by 0.69 points immediately.

(b) What is the long-run effect?

Solution: $\dfrac{\delta_0 + \delta_1}{1 - \theta_1} = \dfrac{-0.69 + 0.32}{1 - 0.56} = \dfrac{-0.37}{0.44} = -0.84$

$\implies$ The long-run effect (-0.84) exceeds the impact (-0.69) because the AR term propagates the shock.

Serial Correlation in Regression Residuals

Suppose we estimate:

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

and the residuals $\hat{e}_t$ are correlated over time: $\text{Cov}(\hat{e}_t, \hat{e}_{t-1}) \neq 0$.

Serial Correlation in Regression Residuals

Suppose we estimate:

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

and the residuals $\hat{e}_t$ are correlated over time: $\mathrm{Cov}(\hat{e}_t, \hat{e}_{t-1}) \neq 0$.

Consequences:

1. OLS coefficient estimates are still **unbiased** (if the model is correct)
2. But OLS **standard errors are wrong** (typically too small)
3. $\implies$ t-statistics are inflated, confidence intervals are too narrow
4. We reject H_0 too often $\implies$ spurious "significance"

Serial Correlation in Regression Residuals

Suppose we estimate:

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

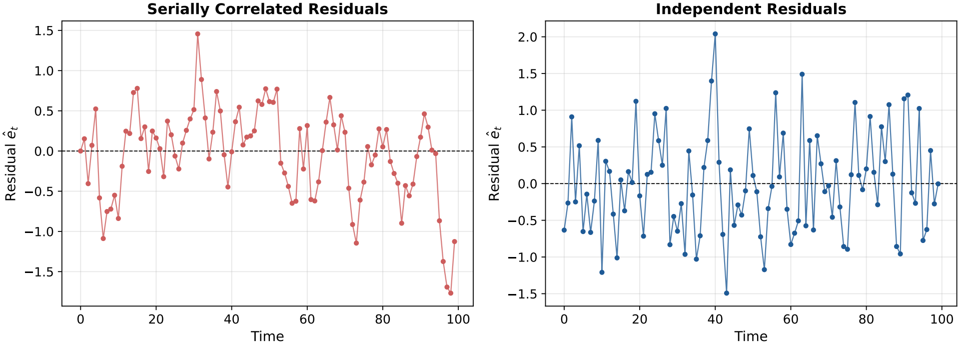
and the residuals $\hat{e}_t$ are correlated over time: $\text{Cov}(\hat{e}_t, \hat{e}_{t-1}) \neq 0$.

Consequences:

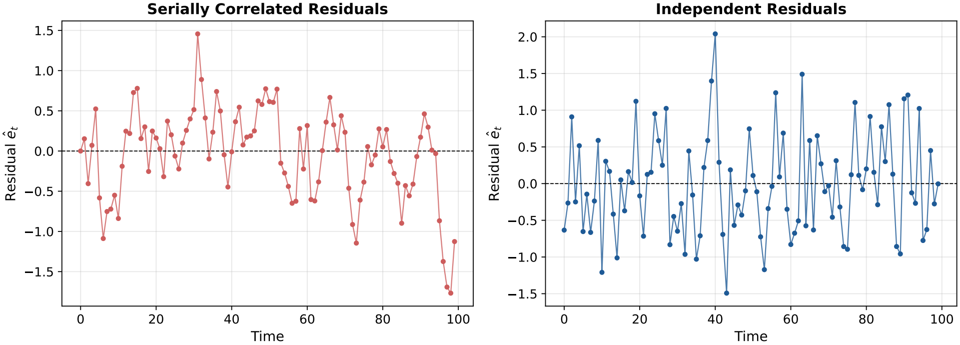
1. OLS coefficient estimates are still **unbiased** (if the model is correct)
2. But OLS **standard errors are wrong** (typically too small)
3. $\implies$ t-statistics are inflated, confidence intervals are too narrow
4. We reject H_0 too often $\implies$ spurious "significance"

$\implies$ The coefficients are fine, but our **inference** (tests, CIs) is unreliable.

What Serially Correlated Residuals Look Like



What Serially Correlated Residuals Look Like



Left: runs of positive then negative residuals $\Longrightarrow$ positive autocorrelation. **Right:** residuals bounce randomly around zero $\Longrightarrow$ no serial correlation.

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

Procedure:

1. Estimate the original model by OLS; get residuals $\hat{e}_t$

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

Procedure:

1. Estimate the original model by OLS; get residuals $\hat{e}_t$
2. Run the **auxiliary regression**:

$$\hat{e}_t = \gamma_0 + \gamma_1(\text{all original regressors}) + \rho_1 \hat{e}_{t-1} + \cdots + \rho_q \hat{e}_{t-q} + v_t$$

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

Procedure:

1. Estimate the original model by OLS; get residuals $\hat{e}_t$
2. Run the **auxiliary regression**:

$$\hat{e}_t = \gamma_0 + \gamma_1(\text{all original regressors}) + \rho_1 \hat{e}_{t-1} + \cdots + \rho_q \hat{e}_{t-q} + v_t$$

3. Compute $LM = T \times R^2$ from the auxiliary regression

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

Procedure:

1. Estimate the original model by OLS; get residuals $\hat{e}_t$
2. Run the **auxiliary regression**:

$$\hat{e}_t = \gamma_0 + \gamma_1(\text{all original regressors}) + \rho_1 \hat{e}_{t-1} + \cdots + \rho_q \hat{e}_{t-q} + v_t$$

3. Compute $LM = T \times R^2$ from the auxiliary regression
4. Under H_0: $\rho_1 = \rho_2 = \cdots = \rho_q = 0, \quad LM \sim \chi^2_{(q)}$

The Breusch-Godfrey (BG) Test

Tests whether residuals are autocorrelated up to order q.

Procedure:

1. Estimate the original model by OLS; get residuals $\hat{e}_t$
2. Run the **auxiliary regression**:

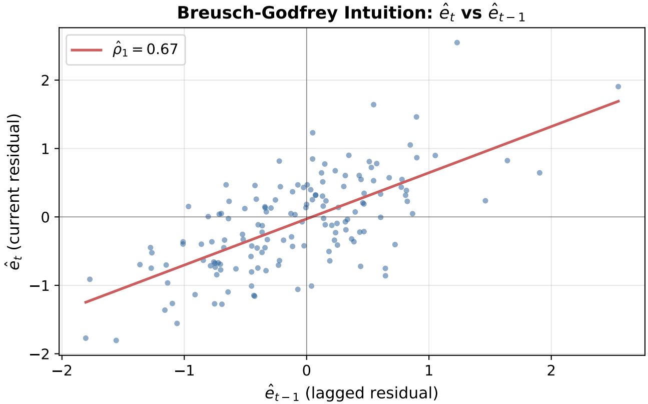
$$\hat{e}_t = \gamma_0 + \gamma_1(\text{all original regressors}) + \rho_1\hat{e}_{t-1} + \cdots + \rho_q\hat{e}_{t-q} + v_t$$

3. Compute $LM = T \times R^2$ from the auxiliary regression
4. Under H_0: $\rho_1 = \rho_2 = \cdots = \rho_q = 0, \quad LM \sim \chi^2_{(q)}$

Important: the auxiliary regression *includes all original regressors*. This is what makes BG valid even with lagged dependent variables (unlike Durbin-Watson).

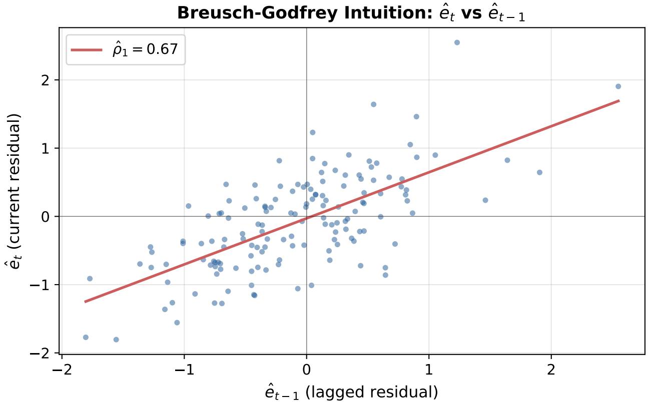
BG Test: The Intuition

If residuals are serially correlated, $\hat{e}_t$ and $\hat{e}_{t-1}$ should be correlated:



BG Test: The Intuition

If residuals are serially correlated, $\hat{e}_t$ and $\hat{e}_{t-1}$ should be correlated:



The BG auxiliary regression checks whether $\hat{\rho}_1$ (the slope in this scatter) is significantly different from zero, *after controlling for the original regressors*.

Forecasting with AR Models

One-step-ahead forecast from an AR(2):

$$\hat{Y}_{T+1} = \hat{\delta} + \hat{\theta}_1 Y_T + \hat{\theta}_2 Y_{T-1}$$

Plug in the most recent observed values and estimated coefficients.

Forecasting with AR Models

One-step-ahead forecast from an AR(2):

$$\hat{Y}_{T+1} = \hat{\delta} + \hat{\theta}_1 Y_T + \hat{\theta}_2 Y_{T-1}$$

Plug in the most recent observed values and estimated coefficients.

Multi-step forecasts: iterate forward, using forecasts as inputs:

$$\hat{Y}_{T+2} = \hat{\delta} + \hat{\theta}_1 \hat{Y}_{T+1} + \hat{\theta}_2 Y_T$$

Forecasting with AR Models

One-step-ahead forecast from an AR(2):

$$\hat{Y}_{T+1} = \hat{\delta} + \hat{\theta}_1 Y_T + \hat{\theta}_2 Y_{T-1}$$

Plug in the most recent observed values and estimated coefficients.

Multi-step forecasts: iterate forward, using forecasts as inputs:

$$\hat{Y}_{T+2} = \hat{\delta} + \hat{\theta}_1 \hat{Y}_{T+1} + \hat{\theta}_2 Y_T$$

$\Longrightarrow$ Each forecast builds on previous forecasts, compounding uncertainty.

Forecast Practice Problem

An AR(2) model for quarterly inflation gives:

$$\hat{\delta} = 0.4523, \quad \hat{\theta}_1 = 0.6234, \quad \hat{\theta}_2 = 0.2145$$

with $\text{INF}_T = 2.5$ and $\text{INF}_{T-1} = 3.0$.

Forecast Practice Problem

An AR(2) model for quarterly inflation gives:

$$\hat{\delta} = 0.4523, \quad \hat{\theta}_1 = 0.6234, \quad \hat{\theta}_2 = 0.2145$$

with $\mathsf{INF}_T = 2.5$ and $\mathsf{INF}_{T-1} = 3.0$.

Compute the one-step-ahead forecast:

Forecast Practice Problem

An AR(2) model for quarterly inflation gives:

$$\hat{\delta} = 0.4523, \quad \hat{\theta}_1 = 0.6234, \quad \hat{\theta}_2 = 0.2145$$

with $\text{INF}_T = 2.5$ and $\text{INF}_{T-1} = 3.0$.

Compute the one-step-ahead forecast:

Solution:

$$\begin{aligned}\widehat{\text{INF}}_{T+1} &= 0.4523 + 0.6234(2.5) + 0.2145(3.0) \\ &= 0.4523 + 1.5585 + 0.6435 \\ &= 2.65\end{aligned}$$

Forecast Intervals

As we forecast further ahead, uncertainty grows:

$$\hat{Y}_{T+h} \pm t_c \cdot \hat{\sigma}_h$$

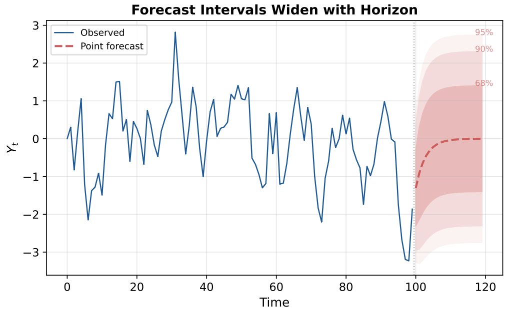
where $\hat{\sigma}_h$ **increases** with the horizon h.

Forecast Intervals

As we forecast further ahead, uncertainty grows:

$$\hat{Y}_{T+h} \pm t_c \cdot \hat{\sigma}_h$$

where $\hat{\sigma}_h$ **increases** with the horizon h.

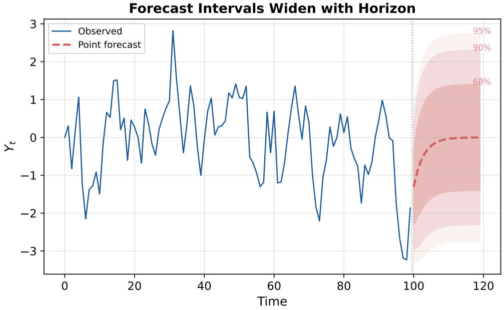


Forecast Intervals

As we forecast further ahead, uncertainty grows:

$$\hat{Y}_{T+h} \pm t_c \cdot \hat{\sigma}_h$$

where $\hat{\sigma}_h$ **increases** with the horizon h.



$\Longrightarrow$ Short-horizon forecasts are relatively precise. Long-horizon forecasts revert toward the unconditional mean with wide bands.

In-Sample vs Out-of-Sample Evaluation

In-sample: fit the model on all available data, check residuals.

- Measures how well the model describes the past
- Always improves with more parameters (overfitting risk)

In-Sample vs Out-of-Sample Evaluation

In-sample: fit the model on all available data, check residuals.

- Measures how well the model describes the past
- Always improves with more parameters (overfitting risk)

Out-of-sample: hold back the last P observations, forecast them, compare to actuals.

- Measures true predictive ability on unseen data
- Penalizes overfitting naturally

In-Sample vs Out-of-Sample Evaluation

In-sample: fit the model on all available data, check residuals.

- Measures how well the model describes the past
- Always improves with more parameters (overfitting risk)

Out-of-sample: hold back the last P observations, forecast them, compare to actuals.

- Measures true predictive ability on unseen data
- Penalizes overfitting naturally

Common metrics:

- **RMSE** $= \sqrt{\frac{1}{P}\sum_{h=1}^{P}(\hat{Y}_{T+h} - Y_{T+h})^2}$
- **MAE** $= \frac{1}{P}\sum_{h=1}^{P}|\hat{Y}_{T+h} - Y_{T+h}|$

In-Sample vs Out-of-Sample Evaluation

In-sample: fit the model on all available data, check residuals.

- Measures how well the model describes the past
- Always improves with more parameters (overfitting risk)

Out-of-sample: hold back the last P observations, forecast them, compare to actuals.

- Measures true predictive ability on unseen data
- Penalizes overfitting naturally

Common metrics:

- **RMSE** $= \sqrt{\frac{1}{P}\sum_{h=1}^{P}(\hat{Y}_{T+h} - Y_{T+h})^2}$
- **MAE** $= \frac{1}{P}\sum_{h=1}^{P}|\hat{Y}_{T+h} - Y_{T+h}|$

$\Longrightarrow$ Always evaluate forecasts out-of-sample when the goal is prediction.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.
3. **ACF:** measures autocorrelation at each lag; geometric decay signals an AR process.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.
3. **ACF:** measures autocorrelation at each lag; geometric decay signals an AR process.
4. **AR(p):** Y_t depends on its own p lags. Stationarity requires $|\theta_1| < 1$ (for AR(1)).

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.
3. **ACF:** measures autocorrelation at each lag; geometric decay signals an AR process.
4. **AR(p):** Y_t depends on its own p lags. Stationarity requires $|\theta_1| < 1$ (for AR(1)).
5. **ARDL(p,q):** includes lags of both Y and X. Distinguish impact, interim, and long-run multipliers.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.
3. **ACF:** measures autocorrelation at each lag; geometric decay signals an AR process.
4. **AR(p):** Y_t depends on its own p lags. Stationarity requires $|\theta_1| < 1$ (for AR(1)).
5. **ARDL(p,q):** includes lags of both Y and X. Distinguish impact, interim, and long-run multipliers.
6. **Serial correlation:** does not bias coefficients, but invalidates standard errors. Detect with Breusch-Godfrey.

Summary

1. **Time series $\neq$ cross-section:** observations are temporally dependent, not i.i.d.
2. **Stationarity:** constant mean, variance, and autocovariance structure. Required for standard regression to work.
3. **ACF:** measures autocorrelation at each lag; geometric decay signals an AR process.
4. **AR(p):** Y_t depends on its own p lags. Stationarity requires $|\theta_1| < 1$ (for AR(1)).
5. **ARDL(p,q):** includes lags of both Y and X. Distinguish impact, interim, and long-run multipliers.
6. **Serial correlation:** does not bias coefficients, but invalidates standard errors. Detect with Breusch-Godfrey.
7. **Forecasting:** iterate AR/ARDL forward; confidence bands widen with horizon. Evaluate out-of-sample.

Thank you!

jakeanderson@g.ucla.edu