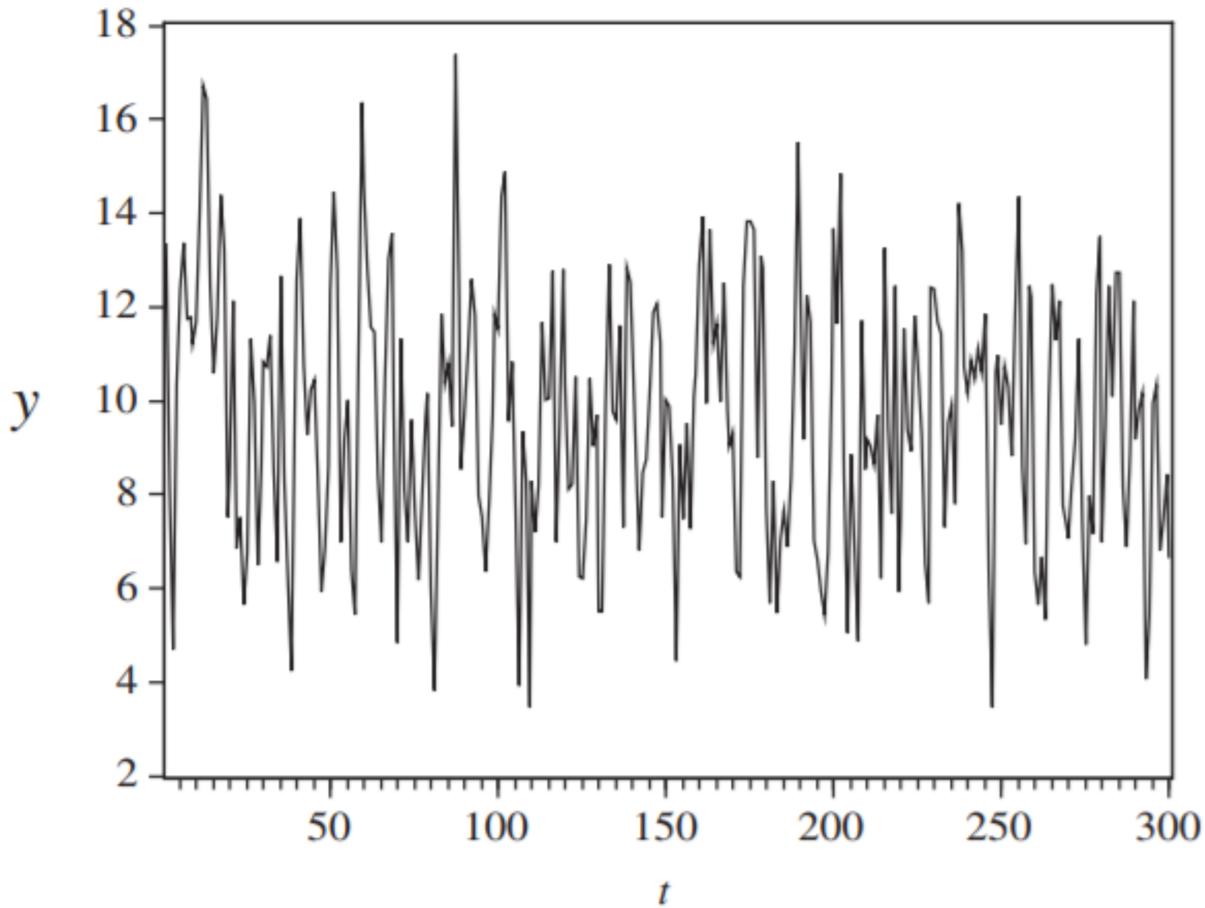
Stationary and Weak Dependence

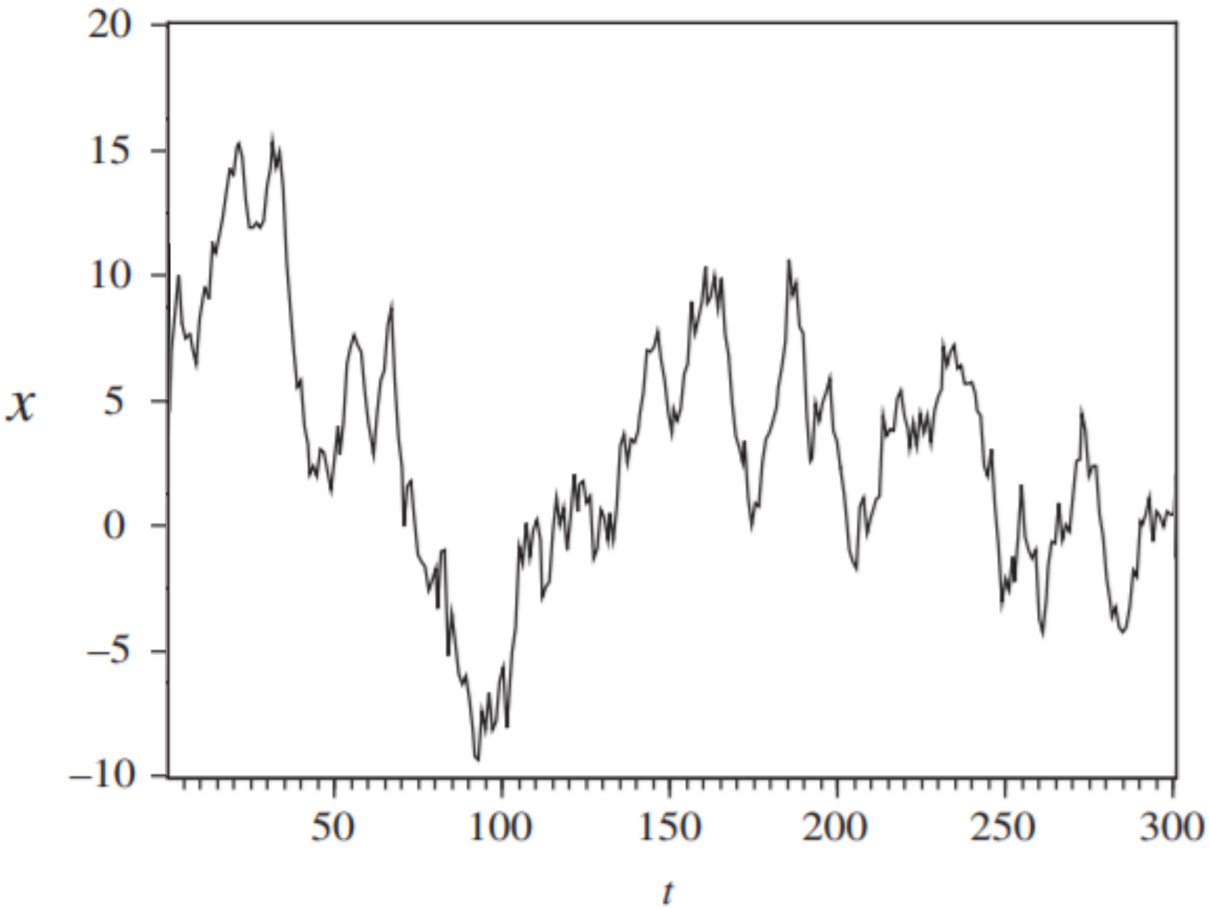
FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK
Gabelli School of Business

Stationary and Weak Dependence

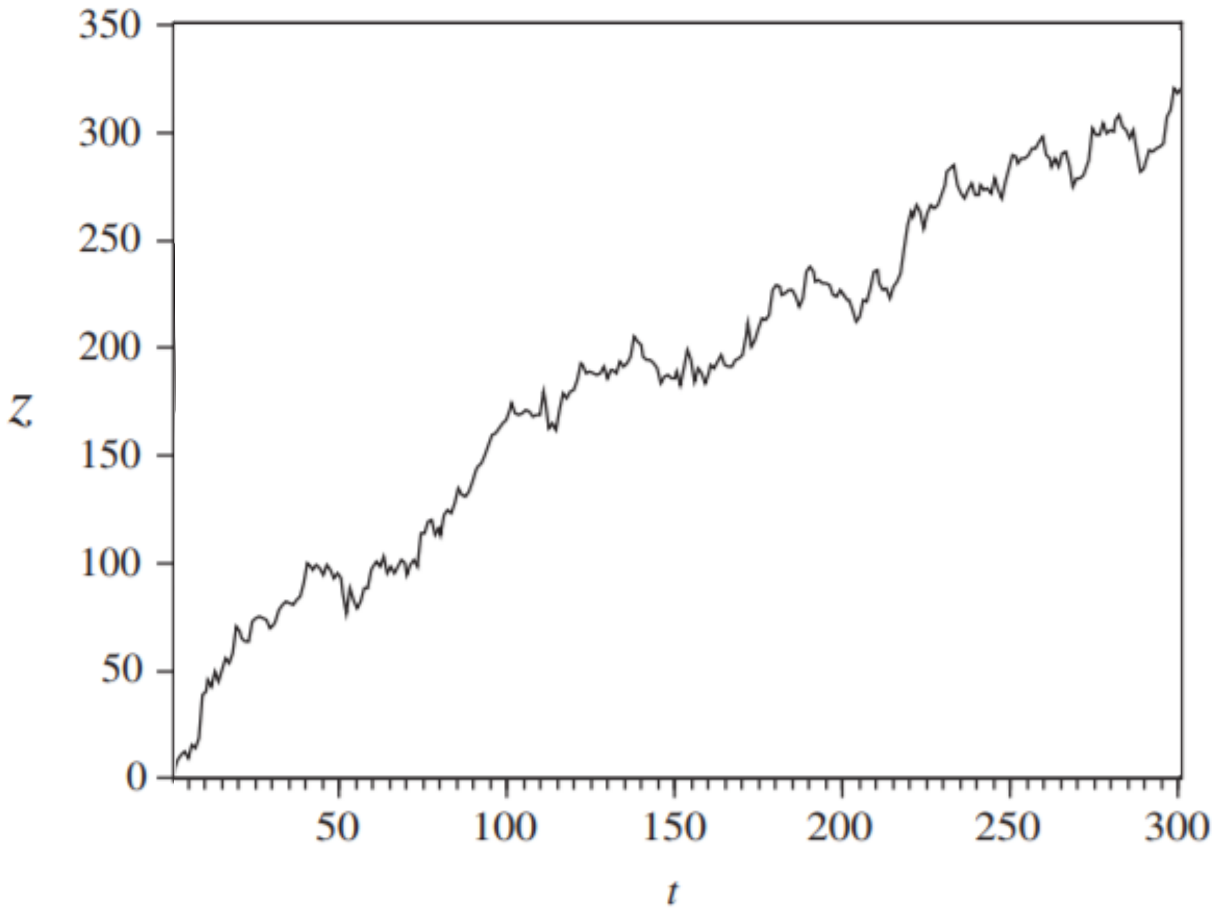
- A critical assumption that is maintained throughout this chapter is that the variables in our equations are **stationary**.
- Stationary variables have means and variances that do not change over time and autocorrelations that depend only on how far apart the observations are in time.
- In addition to assuming that the variables are stationary, in this chapter we also assume they are weakly dependent.
- Weak dependence implies that, as $s \rightarrow \infty$ (observations get further and further apart in time), they become almost independent.



(a)



(b)



(c)

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Forecasting

Forecasting, Part I

- The forecasting of values of economic variables is a major activity for many institutions including firms, banks, governments, and individuals
- In this section, we consider forecasting using two different models: an AR model and an ARDL model
- Our focus is on **short-term forecasting**, typically up to three periods into the future

Forecasting, Part II

TABLE 9.2 **Spreadsheet of Observations for AR(2) Model**

t	Quarter	U_t	U_{t-1}	U_{t-2}
1	1948Q1	3.7	•	•
2	1948Q2	3.7	3.7	•
3	1948Q3	3.8	3.7	3.7
4	1948Q4	3.8	3.8	3.7
5	1949Q1	4.7	3.8	3.8
271	2015Q3	5.2	5.4	5.6
272	2015Q4	5.0	5.2	5.4
273	2016Q1	4.9	5.0	5.2

Forecasting, Part III

- Using the observations in Table 9.2 to find OLS estimates of the model in equation 9.27 yields:

$$\widehat{U_t} = 0.2885 + 1.6128U_{t-1} - 0.6621U_{t-2} \;\; \hat{\sigma} = 0.294$$
$$(se) \quad (0.0666) \quad (0.0457) \qquad (0.0456)$$

- These standard errors and the estimate $\hat{\sigma} = 0.2947$ will be valid with the conditional homoskedasticity assumption:

$$var(e_t|U_{t-1},\, U_{t-2}) \;= \sigma^2$$

Forecasting, Part IV

- Having estimated the AR(2) model, we are now in a position to use it for forecasting
- The unemployment rates for the two most recent quarters are $\widehat{U}_{2016Q1}$= 4.9 and $\widehat{U}_{2015Q4}$ = 5
- The forecast for U_{2016Q2} = 0.28852 + 1.61282 × 4.9 − 0.66209 × 5 = 4.8809
- Two quarters ahead it is $\widehat{U}_{2016Q3}$ = 0.28852 + 1.61282 × 4.8809 − 0.66209 × 4.9 = 4.9163

Forecasting, Part V

- Three quarters ahead it is $\widehat{U}_{2016Q4}$ = 0.28852 + 1.61282 × 4.9163 − 0.66209 × 4.8809 = 4.986
- The forecast unemployment rates for 2016Q2, 2016Q3, and 2016Q4 are approximately 4.88%, 4.92%, and 4.99%, respectively

Example: Forecasting Unemployment With an AR(2) Model

- Consider an AR(2) model for real GDP growth:

$$U_t = \delta + \theta_1 U_{t-1} + \theta_2 U_{t-2} + e_t$$

- The expressions for forecasts for the remainder of 2016 are:

$$\hat{U}_{2016Q2} = E(U_{2016Q2} | I_{2016Q1})$$
$$= \delta + \theta_1 U_{2016Q1} + \theta_2 U_{2015Q4}$$

$$\hat{U}_{2016Q3} = E(U_{2016Q3} | I_{2016Q1})$$
$$= \delta + \theta_1 U_{2016Q2} + \theta_2 U_{2016Q1}$$

$$\hat{U}_{2016Q3} = E(U_{2016Q3} | I_{2016Q1}) = \delta + \vartheta_1 U_{2016Q2} + \vartheta_2 U_{2016Q1}$$

Forecast Intervals and Standard Errors, Part I

- We are interested in interval forecasts that give a likely range in which a future value could fall and indicate the reliability of a point forecast
- The forecast error for one quarter ahead is:

$$f_1 = e_{T+1}$$

- We will be using:

$$\hat{y}_{T+2} = \delta + \theta_1 \hat{y}_{T+1} + \theta_2 y_T + \delta_1 \hat{x}_{T+1} + \delta_2 x_T$$

to forecast:

$$y_{T+2} = \delta + \theta_1 \hat{y}_{T+1} + \theta_2 y_T + \delta_1 \hat{x}_{T+1} + \delta_2 x_T + e_{T+2}$$

Forecast Intervals and Standard Errors, Part II

- The two-period ahead forecast error is:

$$f_2 = \theta_1(y_{T+1} - \hat{y}_{T+1}) + e_{T+2}$$
$$= \theta_1(y_{T+1} - \hat{y}_{T+1}) + e_{T+2} f_1 + e_{T+2} = \theta_1 e_{T+1} + e_{T+2}$$

- For three periods ahead, the error can be shown to be:

$$f_3 = \theta_1 f_2 + \theta_2 f_2 + e_{T+3} = (\theta_1^2 + \theta_2)e_{T+1} + \theta_1 e_{T+2} + e_{T+3} 1$$

- Expressing the forecast errors in terms of the e_t's is convenient for deriving expressions for the forecast error variances

Forecast Intervals and Standard Errors, Part III

- Because $E(e_t | I_{T-1}) = 0$

- And $\text{var}(e_t | y_{t-1}, y_{t-2}, x_{t-1}, x_{t-2}) = \sigma^2$

- We can show:

$$\sigma_{f1}^2 = \text{var}(f_1 | I_T) = \sigma^2$$
$$\sigma_{f2}^2 = \text{var}(f_2 | I_T) = \sigma^2(1 + \theta_1^2)$$
$$\sigma_{f3}^2 = \text{var}(f_3 | I_T) = \sigma^2 \left((\theta_1^2 + \vartheta_2)^2 + \theta_1^2 + 1 \right)$$

TABLE 9.3 Forecasts and Forecast Intervals for Unemployment from AR(2) Model

Quarter	Forecast $\hat{U}_{T+j}$	Standard Error of Forecast Error $(\hat{\sigma}_{f,j})$	Forecast Interval $\left(\hat{U}_{T+j} \pm 1.9689 \times \hat{\sigma}_{f,j}\right)$
2016Q2 ($j = 1$)	4.881	0.2947	(4.301, 5.461)
2016Q3 ($j = 2$)	4.916	0.5593	(3.815, 6.017)
2016Q4 ($j = 3$)	4.986	0.7996	(3.412, 6.560)

TABLE 9.4 Forecasts and Forecast Intervals for Unemployment from ARDL(2, 1) Model

Quarter	Forecast $\hat{U}_{T+j}$	Standard Error of Forecast Error $(\hat{\sigma}_{uj})$	Forecast Interval $\left(\hat{U}_{T+j} \pm 1.9689 \times \hat{\sigma}_{uj}\right)$
2016Q2 ($j = 1$)	4.950	0.2919	(4.375, 5.525)
2016Q3 ($j = 2$)	5.058	0.5343	(4.006, 6.110)
2016Q4 ($j = 3$)	5.184	0.7430	(3.721, 6.647)

Assumptions for Forecasting

These assumptions ensure an ARDL model can be estimated consistently and used for forecasting

- F1: the time series y and x are stationary and weakly dependent
- F2: the conditional expectation E($y_t|I_{t-1}$) is a linear function of a finite number of lags of y and x
- F3: The errors are conditionally homoskedastic: var($e_t|\boldsymbol{Z_t}$) = σ^2

Selecting Lag Lengths

- A critical assumption to ensure that we had the best forecast in a minimum mean-squared-error sense was that no lags beyond those included in the model contained extra information that could improve the forecast.
- There are four ways to decide on p and q.
  1. Extend the lag lengths for y and x as long as their estimated coefficients are significantly different from zero.

Selecting Lag Lengths (cont.)

2. Choose p and q to minimize either the AIC or the SC variable selection criterion.
3. Evaluate the out-of-sample forecasting performance of each (p, q) combination using a hold-out sample.
4. Check for serial correlation in the error term because E($e_t | I_{t-1}$) = 0 implies that the lag lengths p and q are sufficient and the errors are not serially correlated.

Testing for Granger Causality

- Granger causality refers to the ability of lags of one variable to contribute to the forecast of another variable
- Testing for Granger causality is equivalent to testing
- $H_0: \delta_1 = 0, \delta_2 = 0, \ldots, \delta_q = 0 \; and \; H_1: at \; least \; one \; \delta_i \neq 0$
- Rejection of H_0 implies x Granger causes y
- Note that, if x Granger causes y, it does not necessarily imply a direct causal relationship between x and y

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Testing for Serially Correlated Errors

Testing for Serially Correlated Errors

- Consider again the ARDL(p, q) model.
- For the absence of serial correlation, we require the conditional covariance between any two different errors to be zero.
- One way of assessing whether sufficient lags have been included to get the best forecast is to test for serially correlated errors.
- Not using the best model for forecasting is not the only implication of serially correlated errors.

Checking the Correlogram of the Least Squares Residuals

- We can use the correlogram of the least squares residuals to check for serially correlated errors
- The *k*th order autocorrelation for the residuals can be written as:

$$r_k = \frac{\sum_{t=k+1}^{T} \hat{e}_t \hat{e}_{t-k}}{\sum_{t=1}^{T} \hat{e}_t^2}$$

- Ideally, for the correlogram to suggest no serial correlation, we like to have $|r_k| < 2 \,/\, \sqrt{T}$ for k = 1, 2,…

Lagrange Multiplier Test

- An advantage of the Lagrange multiplier test is that it readily generalizes to a **joint** test of correlations at more than one lag
- Consider the ARDL(1,1) model:

$$y_t = \delta + \theta_1 y_{t-1} + \delta_1 x_{t-1} + e_t$$

- The null hypothesis for the test is that the errors e_t are uncorrelated
- To express this null hypothesis in terms of restrictions on one or more parameters, we can introduce a model for an alternative hypothesis

Testing for AR(1) Errors

- Consider an alternative hypothesis that the errors are correlated through the AR(1) process: $e_t = \rho e_{t-1} + v_t$
- Substituting for e_t in the original equation yields $y_t = \delta + \theta_1 y_{t-1} + \delta_1 x_{t-1} + \rho e_{t-1} + v_t$
- Now, if ρ = 0, then e_t = v_t, and because v_t is not serially correlated, e_t will not be serially correlated
- The hypotheses H_0:ρ = 0 and H_1:ρ ≠ 0

Testing for MA(1) Errors

- Another useful class of models is what is known as moving-average models

$$\phi$$

- Following the previous strategy we get:

$$\phi$$

$$y_t = \delta + \theta_1 y_{t-1} + \delta_1 x_{t-1} + \phi v_{t-1} + v_t$$

- Notice that ϕ = 0 implies e_t = v_t, and so we can test for autocorrelation through the hypotheses H_0:ϕ = 0 and H_1:$\phi \neq 0$

Finite Distributed Lags, Part I

- The finite distributed lag model where we are interested in the impact of current and past values of a variable *x* on current and future values of a variable *y* can be written as:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t$$

- It is called a finite distributed lag because the impact of *x* on *y* cuts off after *q* lags

$$E(y_t|x_t, x_{t-1}, \ldots) = \alpha + \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q}$$

$$E(y_t|x_t, x_{t-1}, \ldots) = E(y_t|x_t)$$

Finite Distributed Lags, Part II

- Once *q* lags of *x* have been included in the equation, further lags of *x* will not have an impact on *y*
- Given this assumption, a lag-coefficient β_s can be interpreted as the change in E($y_t|x_t$) when x_{t-s} changes by one unit, but *x* is held constant in other periods
- In terms of derivatives (9.56), $\frac{\partial E(y_t|x_t)}{\partial x_{t-s}} = \frac{\partial E(y_{t+s}|x_t)}{\partial x_t} = \beta_s$

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Time-Series Regressions for Policy Analysis

FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK
Gabelli School of Business

Time-Series Regressions for Policy Analysis

- Models for policy analysis differ in a number of ways. The individual coefficients are of interest because they might have a causal interpretation
- In the following four sections, we are concerned with three main issues that add to our time-series regression results from earlier chapters:
    1. Interpretation of coefficients of lagged variables in finite and infinite distributed lag models
    2. Estimation and inference for coefficients when the errors are autocorrelated
    3. The assumptions necessary for interpretation and estimation

Finite Distributed Lags, Part I

- The finite distributed lag model where we are interested in the impact of current and past values of a variable *x* on current and future values of a variable *y* can be written as:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t$$

- It is called a finite distributed lag because the impact of *x* on *y* cuts off after *q* lags

$$E(y_t|x_t, x_{t-1}, \dots) = \alpha + \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q}$$

$$E(y_t|x_t, x_{t-1}, \dots) = E(y_t|x_t)$$

Finite Distributed Lags, Part II

- Once *q* lags of *x* have been included in the equation, further lags of *x* will not have an impact on *y*
- Given this assumption, a lag-coefficient β_s can be interpreted as the change in E($y_t | x_t$) when x_{t-s} changes by one unit, but *x* is held constant in other periods
- In terms of derivatives (9.56), $\frac{\partial E(y_t|x_t)}{\partial x_{t-s}} = \frac{\partial E(y_{t+s}|x_t)}{\partial x_t} = \beta_s$

Finite Distributed Lags, Part III

- The effect of a one-unit change in x_t is **distributed** over the current and next q periods, from which we get the term "distributed lag model"
  - It is called a **finite distributed lag model of order *q*** It is assumed that, after a finite number of periods q, changes in x no longer have an impact on y
  - The coefficient β_s is called a **distributed-lag weight** or an ***s*-period delay multiplier**
  - The coefficient β_0 ($s = 0$) is called the **impact multiplier**

Assumptions for Finite Distributed Lag Mode

- FDL1: the time series *y* and *x* are stationary and weakly dependent
- FDL2: the finite distributed lag model describing how *y* responds to current and past values of *x* can be written as $y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots \beta_q x_{t-q} + e_t$
- FDL3: the error term is exogenous with respect to the current and all past values of *x*
- FDL4: the error term is not autocorrelated
- FDL5: the error term is homoskedastic

HAC Standard Errors

- HAC (heteroskedasticity and autocorrelation consistent) standard errors, or known as Newey–West standard errors
- Different software packages may yield different HAC standard errors because there are a large number of possibilities
- The analysis in this section extends to the finite distributed lag model with *q* lags and indeed to any time-series regression involving stationary variables

Estimation With AR(1) Errors

- Consider the simple regression model:

$$y_t = \alpha + \beta_0 x_t + e_t$$

- This model can be extended to include extra lags from an FDL model and other variables; the AR(1) error model is given by equation 9.67:

$$e_t = \rho e_{t-1} + v_t |\rho| < 1$$

- Assume that the v_t are uncorrelated random errors with zero mean and constant variances:

$$E(v_t | x_t, x_{t-1}, \dots) = 0 \quad \mathrm{var}(v_t | x_t) = \sigma_v^2 \quad \mathrm{cov}(v_t, v_s | x_t, x_s) = 0 \quad for\ t \neq s$$

Nonlinear Least Squares Estimation

- Consider the equation:

$$y_t = \alpha(1-\rho) + \rho y_{t-1} + \beta_0 x_t - \rho\beta_0 x_{t-1} + v_t$$

- We have transformed the original model, with the autocorrelated error term e_t into a new model that has an error term v_t that is uncorrelated over time
- The advantage of doing so is that we can now proceed to find estimates for (α, β_0, ρ) that minimize the sum of squares of uncorrelated errors

Generalized Least Squares Estimation

- To introduce an alternative estimator for (α, β_0, ρ) in the AR(1) error model

$$y_t - \rho y_{t-1} = \alpha(1-\rho) + \beta_0(x_t - \rho x_{t-1}) + v_t$$

- Defining

$$y_t^* = y_t - \rho y_{t-1}, \alpha^* = \alpha(1-\rho) \; and \; x_t^* = x_t - \rho x_{t-1}$$

$$y_t^* = \alpha^* + \beta_0 x_t^* + v_t \; \ldots, T$$

- ρ is not known and must be estimated

Generalized Least Squares Estimation (cont.)

- The steps for obtaining the feasible generalized least squares estimator for α and β_0 using this estimator for ρ are as follows:
  1. Find least-squares estimates a and b0 from the equation
  $$y_t = \alpha\beta_0 x_t + e_t$$
  2. Compute the least squares residuals
  $$\hat{e}_t = y_t - a + \beta_0 x_t + e_t$$
  3. Estimate ρ by applying least squares to the equation
  $$\hat{e}_t = \rho\hat{e}_{t-1} + \hat{v}_t$$
  4. Compute values of the transformed variables
  $$y_t^* = y_t - \rho y_{t-1} \text{ and } x_t^* = x_t - \rho x_{t-1}$$
  5. Apply least squares to the transformed equation
  $$y_t^* = \alpha^* + \beta_0 x_t^* + v_t$$

Assumptions and Properties

- To solve the problem when FDL4 and FDL5 are violated:
  1. Use the HAC estimator for variances and covariances and the corresponding HAC standard errors
  2. Assume a specific model for the autocorrelated errors and use an estimator that is the minimum variance for that model
- Modeling of more general forms of autocorrelated errors with more than one lag requires e_t to be uncorrelated with x values further than one period into the future

Infinite Distributed Lags

- One way of avoiding the need to specify a value for *q* is to consider an IDL model where *y* depends on lags of *x* that go back into the indefinite past

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \cdots + e_t$$

- For it to be feasible, the β_s coefficients must eventually (but not necessarily immediately) decline in magnitude, becoming negligible at long lags

$$\beta_s = \frac{\partial y_t}{\partial x_{t-s}} = s \textit{ period delay multiplier}, \sum_{j=0}^{s} \beta_j = s \textit{ period interim multiplier},$$

$$\sum_{j=0}^{\infty} \beta_j = \textit{total multiplier}$$

Geometrically Declining Lags

- An obvious disadvantage of the IDL model is its infinite number of parameters
- By imposing the restrictions, derived in Section 9.11, we have been able to reduce the infinite number of parameters to just three

$$y_t = \delta + \theta y_{t-1} + \beta_0 x_t + v_t$$

- The delay multipliers can be calculated from the restrictions $\beta_s = \lambda^s \beta_0$

Testing for Consistency in the ARDL Representation of an IDL Model

The test is based on whether or not an estimate of the error e_{t-1} adds explanatory power to the regression:

$$y_t = \delta + \lambda y_{t-1} + \beta_0 x_t + (\rho - \lambda)e_{t-1} + u_t$$

1. Compute the least squares residuals from equation 9.74 under the assumption that H_0 holds

Testing for Consistency in the ARDL Representation of an IDL Model (cont.)

2. Using the least squares estimate $\hat{\lambda}$ from step 1, and starting with $\widehat{e_1} = 0$, compute recursively $\hat{e}_t = \lambda\hat{e}_{t-1} + u_t, t = 2,3,\ldots,T$
3. Find the R^2 from a least squares regression of $\hat{u}_t$ on y_{t-1}, x_t and $\hat{e}_{t-1}$
4. When H_0 is true, and assuming that u_t is homoskedastic, (T − 1) × R^2 has a $\chi^2_{(1)}$ distribution in large samples

Deriving Multipliers From an ARDL Representation

- An alternative strategy is to begin with an ARDL representation whose lags have been chosen using conventional model selection criteria and to derive the restrictions on the IDL model implied by the chosen ARDL model.
- Specifically, we first estimate the finite number of θs and δs from an ARDL model.
- Our task for the general case is made much easier if we can master some heavy machinery known as the lag operator.

The Error Term

- The question we need to ask is whether the error term will be such that the least squares estimator is consistent

$$e_t = (^{1-\theta_1 L-\theta_2 L^2)-1} v_t$$
$$(1-\theta_1 L-\theta_2 L^2)e_t = v_t$$
$$e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} = v_t$$
$$e_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + v_t$$

- In the general ARDL(p, q) model, this equation becomes:

$$e_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_p e_{t-p} + v_t$$

Assumptions for the Infinite Distributed Lag Model

- IDL1: the time series *y* and *x* are stationary and weakly dependent
- IDL2: the infinite distributed lag model describing how *y* responds to current and past values of *x* can be written as:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + e_t \text{ with } \beta_s \longrightarrow 0 \; as \; s \longrightarrow \infty$$

- IDL3: corresponding to equation 9.95 is an ARDL(p, q) model

$$y_t = \delta + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + v_t$$

Assumptions for the Infinite Distributed Lag Model (cont.)

- IDL4: the errors e_t are strictly exogenous

  $E(e_t|X) = 0$, where X includes all current, past, and future values of *x*
- IDL5: the errors e_t follow the AR(p) process
  - Where:

  1. v_t is exogenous with respect to current and past values of *x* and past values of *y*
  2. v_t is homoskedastic, var($v_t|x_t$) =σ_v^2

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business

Key Words

- AR(1) error
- ARDL(p, q) model
- Autocorrelation
- Autoregressive distributed lags
- Autoregressive error
- Autoregressive model
- Correlogram
- Delay multiplier
- Distributed lag weight
- Dynamic models
- Exogeneity
- Finite distributed lag
- Forecast error
- Forecast intervals
- Forecasting
- Generalized least squares
- Geometrically declining lag
- Granger causality
- HAC standard errors
- Impact multiplier
- Infinite distributed lag
- Interim multiplier
- Lag length
- Lag operator
- Lagged dependent variable
- LM test
- Moving average
- Multiplier analysis
- Nonlinear least squares
- Sample autocorrelations
- Serial correlation
- Standard error of forecast error
- Stationarity
- Total multiplier
- $T \times R^2$ form of LM test
- Weak dependence

Copyright

FORDHAM

THE JESUIT UNIVERSITY OF NEW YORK

Gabelli School of Business