Glossary

ARCH model. An autoregressive conditional heteroskedasticity (ARCH) model is a regression model in which the conditional variance is modeled as an autoregressive (AR) process. The ARCH(m) model is

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + \epsilon_t$$
$$E(\epsilon_t^2|\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots) = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \cdots + \alpha_m\epsilon_{t-m}^2$$

where ϵ_t is a white-noise error term. The equation for y_t represents the conditional mean of the process, and the equation for $E(\epsilon_t^2|\epsilon_{t-1}^2, \epsilon_{t-2}^2, \ldots)$ specifies the conditional variance as an autoregressive function of its past realizations. Although the conditional variance changes over time, the unconditional variance is time invariant because y_t is a stationary process. Modeling the conditional variance as an AR process raises the implied unconditional variance, making this model particularly appealing to researchers modeling fat-tailed data, such as financial data.

ARFIMA model. An autoregressive fractionally integrated moving-average (ARFIMA) model is a time-series model suitable for use with long-memory processes. ARFIMA models generalize autoregressive integrated moving-average (ARIMA) models by allowing the differencing parameter to be a real number in $(-0.5, 0.5)$ instead of requiring it to be an integer.

ARIMA model. An autoregressive integrated moving-average (ARIMA) model is a time-series model suitable for use with integrated processes. In an ARIMA(p, d, q) model, the data is differenced d times to obtain a stationary series, and then an ARMA(p, q) model is fit to this differenced data. ARIMA models that include exogenous explanatory variables are known as ARMAX models.

ARMA model. An autoregressive moving-average (ARMA) model is a time-series model in which the current period's realization is the sum of an autoregressive (AR) process and a moving-average (MA) process. An ARMA(p, q) model includes p AR terms and q MA terms. ARMA models with just a few lags are often able to fit data, as well as pure AR or MA models with many more lags.

ARMAX model. An ARMAX model is a time-series model in which the current period's realization is an ARMA process plus a linear function of a set of a exogenous variables. Equivalently, an ARMAX model is a linear regression model in which the error term is specified to follow an ARMA process.

autocorrelation function. The autocorrelation function (ACF) expresses the correlation between periods t and $t-k$ of a time series as function of the time t and the lag k. For a stationary time series, the ACF does not depend on t and is symmetric about $k = 0$, meaning that the correlation between periods t and $t-k$ is equal to the correlation between periods t and $t+k$.

autoregressive process. An autoregressive process is a time-series model in which the current value of a variable is a linear function of its own past values and a white-noise error term. A first-order autoregressive process, denoted as an AR(1) process, is $y_t = \rho y_{t-1} + \epsilon_t$. An AR($p$) model contains p lagged values of the dependent variable.

band-pass filter. Time-series filters are designed to pass or block stochastic cycles at specified frequencies. Band-pass filters, such as those implemented in `tsfilter bk` and `tsfilter cf`, pass through stochastic cycles in the specified range of frequencies and block all other stochastic cycles.

Cholesky ordering. Cholesky ordering is a method used to orthogonalize the error term in a VAR or VECM to impose a recursive structure on the dynamic model, so that the resulting impulse–response functions can be given a causal interpretation. The method is so named because it uses the Cholesky decomposition of the error covariance matrix.

Cochrane–Orcutt estimator. This estimation is a linear regression estimator that can be used when the error term exhibits first-order autocorrelation. An initial estimate of the autocorrelation parameter ρ is obtained from OLS residuals, and then OLS is performed on the transformed data $\widetilde{y}_t = y_t - \rho y_{t-1}$ and $\widetilde{\mathbf{x}}_t = \mathbf{x}_t - \rho \mathbf{x}_{t-1}$.

cointegrating vector. A cointegrating vector specifies a stationary linear combination of nonstationary variables. Specifically, if each of the variables $x_1, x_2, \dots, x_k$ is integrated of order one and there exists a set of parameters $\beta_1, \beta_2, \dots, \beta_k$ such that $z_t = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$ is a stationary process, the variables $x_1, x_2, \dots, x_k$ are said to be cointegrated, and the vector $\boldsymbol{\beta}$ is known as a cointegrating vector.

conditional variance. Although the conditional variance is simply the variance of a conditional distribution, in time-series analysis the conditional variance is often modeled as an autoregressive process, giving rise to ARCH models.

correlogram. A correlogram is a table or graph showing the sample autocorrelations or partial autocorrelations of a time series.

covariance stationarity. A process is covariance stationary if the mean of the process is finite and independent of t, the unconditional variance of the process is finite and independent of t, and the covariance between periods t and $t - s$ is finite and depends on $t - s$ but not on t or s themselves. Covariance-stationary processes are also known as weakly stationary processes.

cross-correlation function. The cross-correlation function expresses the correlation between one series at time t and another series at time $t - k$ as a function of the time t and lag k. If both series are stationary, the function does not depend on t. The function is not symmetric about $k = 0$: $\rho_{12}(k) \neq \rho_{12}(-k)$.

cyclical component. A cyclical component is a part of a time series that is a periodic function of time. Deterministic functions of time are deterministic cyclical components, and random functions of time are stochastic cyclical components. For example, fixed seasonal effects are deterministic cyclical components and random seasonal effects are stochastic seasonal components.

Random coefficients on time inside of periodic functions form an especially useful class of stochastic cyclical components; see [TS] **ucm**.

deterministic trend. A deterministic trend is a deterministic function of time that specifies the long-run tendency of a time series.

difference operator. The difference operator Δ denotes the change in the value of a variable from period $t - 1$ to period t. Formally, $\Delta y_t = y_t - y_{t-1}$, and $\Delta^2 y_t = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$.

drift. Drift is the constant term in a unit-root process. In

$$y_t = \alpha + y_{t-1} + \epsilon_t$$

α is the drift when ϵ_t is a stationary, zero-mean process.

dynamic forecast. A dynamic forecast is one in which the current period's forecast is calculated using forecasted values for prior periods.

dynamic-multiplier function. A dynamic-multiplier function measures the effect of a shock to an exogenous variable on an endogenous variable. The kth dynamic-multiplier function of variable i on variable j measures the effect on variable j in period $t + k$ in response to a one-unit shock to variable i in period t, holding everything else constant.

endogenous variable. An endogenous variable is a regressor that is correlated with the unobservable error term. Equivalently, an endogenous variable is one whose values are determined by the equilibrium or outcome of a structural model.

exogenous variable. An exogenous variable is one that is correlated with none of the unobservable error terms in the model. Equivalently, an exogenous variable is one whose values change independently of the other variables in a structural model.

exponential smoothing. Exponential smoothing is a method of smoothing a time series in which the smoothed value at period t is equal to a fraction α of the series value at time t plus a fraction $1-\alpha$ of the previous period's smoothed value. The fraction α is known as the smoothing parameter.

forecast-error variance decomposition. Forecast-error variance decompositions measure the fraction of the error in forecasting variable i after h periods that is attributable to the orthogonalized shocks to variable j.

forward operator. The forward operator F denotes the value of a variable at time $t+1$. Formally, $Fy_t = y_{t+1}$, and $F^2y_t = Fy_{t+1} = y_{t+2}$.

frequency-domain analysis. Frequency-domain analysis is analysis of time-series data by considering its frequency properties. The spectral density and distribution functions are key components of frequency-domain analysis, so it is often called spectral analysis. In Stata, the `cumsp` and `pergram` commands are used to analyze the sample spectral distribution and density functions, respectively. `psdensity` estimates the spectral density or the spectral distribution function after estimating the parameters of a parametric model using `arfima`, `arima`, or `ucm`.

gain (of a linear filter). The gain of a linear filter scales the spectral density of the unfiltered series into the spectral density of the filtered series for each frequency. Specifically, at each frequency, multiplying the spectral density of the unfiltered series by the square of the gain of a linear filter yields the spectral density of the filtered series. If the gain at a particular frequency is 1, the filtered and unfiltered spectral densities are the same at that frequency and the corresponding stochastic cycles are passed through perfectly. If the gain at a particular frequency is 0, the filter removes all the corresponding stochastic cycles from the unfiltered series.

GARCH model. A generalized autoregressive conditional heteroskedasticity (GARCH) model is a regression model in which the conditional variance is modeled as an ARMA process. The GARCH(m,k) model is

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + \epsilon_t$$
$$\sigma_t^2 = \gamma_0 + \gamma_1\epsilon_{t-1}^2 + \cdots + \gamma_m\epsilon_{t-m}^2 + \delta_1\sigma_{t-1}^2 + \cdots + \delta_k\sigma_{t-k}^2$$

where the equation for y_t represents the conditional mean of the process and σ_t represents the conditional variance. See [TS] **arch** or Hamilton (1994, chap. 21) for details on how the conditional variance equation can be viewed as an ARMA process. GARCH models are often used because the ARMA specification often allows the conditional variance to be modeled with fewer parameters than are required by a pure ARCH model. Many extensions to the basic GARCH model exist; see [TS] **arch** for those that are implemented in Stata. See also *ARCH model*.

generalized least-squares estimator. A generalized least-squares (GLS) estimator is used to estimate the parameters of a regression function when the error term is heteroskedastic or autocorrelated. In the linear case, GLS is sometimes described as "OLS on transformed data" because the GLS estimator can be implemented by applying an appropriate transformation to the dataset and then using OLS.

Granger causality. The variable *x* is said to Granger-cause variable *y* if, given the past values of *y*, past values of *x* are useful for predicting *y*.

high-pass filter. Time-series filters are designed to pass or block stochastic cycles at specified frequencies. High-pass filters, such as those implemented in `tsfilter bw` and `tsfilter hp`, pass through stochastic cycles above the cutoff frequency and block all other stochastic cycles.

Holt–Winters smoothing. A set of methods for smoothing time-series data that assume that the value of a time series at time t can be approximated as the sum of a mean term that drifts over time, as well as a time trend whose strength also drifts over time. Variations of the basic method allow for seasonal patterns in data, as well.

impulse–response function. An impulse–response function (IRF) measures the effect of a shock to an endogenous variable on itself or another endogenous variable. The kth impulse–response function of variable i on variable j measures the effect on variable j in period $t+k$ in response to a one-unit shock to variable i in period t, holding everything else constant.

independent and identically distributed. A series of observations is independently and identically distributed (i.i.d.) if each observation is an independent realization from the same underlying distribution. In some contexts, the definition is relaxed to mean only that the observations are independent and have identical means and variances; see Davidson and MacKinnon (1993, 42).

integrated process. A nonstationary process is integrated of order d, written $\mathrm{I}(d)$, if the process must be differenced d times to produce a stationary series. An $\mathrm{I}(1)$ process y_t is one in which Δy_t is stationary.

Kalman filter. The Kalman filter is a recursive procedure for predicting the state vector in a state-space model.

lag operator. The lag operator L denotes the value of a variable at time $t-1$. Formally, $Ly_t = y_{t-1}$, and $L^2 y_t = Ly_{t-1} = y_{t-2}$.

linear filter. A linear filter is a sequence of weights used to compute a weighted average of a time series at each time period. More formally, a linear filter $\alpha(L)$ is

$$\alpha(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + \cdots = \sum_{\tau=0}^{\infty} \alpha_\tau L^\tau$$

where L is the lag operator. Applying the linear filter $\alpha(L)$ to the time series x_t yields a sequence of weighted averages of x_t:

$$\alpha(L)x_t = \sum_{\tau=0}^{\infty} \alpha_\tau L^\tau x_{t-\tau}$$

long-memory process. A long-memory process is a stationary process whose autocorrelations decay at a slower rate than a short-memory process. ARFIMA models are typically used to represent long-memory processes, and ARMA models are typically used to represent short-memory processes.

moving-average process. A moving-average process is a time-series process in which the current value of a variable is modeled as a weighted average of current and past realizations of a white-noise process and, optionally, a time-invariant constant. By convention, the weight on the current realization of the white-noise process is equal to one, and the weights on the past realizations are known as the moving-average (MA) coefficients. A first-order moving-average process, denoted as an MA(1) process, is $y_t = \theta\epsilon_{t-1} + \epsilon_t$.

multivariate GARCH models. Multivariate GARCH models are multivariate time-series models in which the conditional covariance matrix of the errors depends on its own past and its past shocks. The acute trade-off between parsimony and flexibility has given rise to a plethora of models; see [TS] **mgarch**.

Newey–West covariance matrix. The Newey–West covariance matrix is a member of the class of heteroskedasticity- and autocorrelation-consistent (HAC) covariance matrix estimators used with time-series data that produces covariance estimates that are robust to both arbitrary heteroskedasticity and autocorrelation up to a prespecified lag.

orthogonalized impulse–response function. An orthogonalized impulse–response function (OIRF) measures the effect of an orthogonalized shock to an endogenous variable on itself or another endogenous variable. An orthogonalized shock is one that affects one variable at time t but no other variables. See [TS] **irf create** for a discussion of the difference between IRFs and OIRFs.

partial autocorrelation function. The partial autocorrelation function (PACF) expresses the correlation between periods t and $t-k$ of a time series as a function of the time t and lag k, after controlling for the effects of intervening lags. For a stationary time series, the PACF does not depend on t. The PACF is not symmetric about $k=0$: the partial autocorrelation between y_t and y_{t-k} is not equal to the partial autocorrelation between y_t and y_{t+k}.

periodogram. A periodogram is a graph of the spectral density function of a time series as a function of frequency. The `pergram` command first standardizes the amplitude of the density by the sample variance of the time series, and then plots the logarithm of that standardized density. Peaks in the periodogram represent cyclical behavior in the data.

phase function. The phase function of a linear filter specifies how the filter changes the relative importance of the random components at different frequencies in the frequency domain.

portmanteau statistic. The portmanteau, or Q, statistic is used to test for white noise and is calculated using the first m autocorrelations of the series, where m is chosen by the user. Under the null hypothesis that the series is a white-noise process, the portmanteau statistic has a χ^2 distribution with m degrees of freedom.

Prais–Winsten estimator. A Prais–Winsten estimator is a linear regression estimator that is used when the error term exhibits first-order autocorrelation; see also *Cochrane–Orcutt estimator*. Here the first observation in the dataset is transformed as $\widetilde{y}_1 = \sqrt{1-\rho^2}\, y_1$ and $\widetilde{\mathbf{x}}_1 = \sqrt{1-\rho^2}\, \mathbf{x}_1$, so that the first observation is not lost. The Prais–Winsten estimator is a generalized least-squares estimator.

priming values. Priming values are the initial, preestimation values used to begin a recursive process.

random walk. A random walk is a time-series process in which the current period's realization is equal to the previous period's realization plus a white-noise error term: $y_t = y_{t-1} + \epsilon_t$. A *random walk with drift* also contains a nonzero time-invariant constant: $y_t = \delta + y_{t-1} + \epsilon_t$. The constant term δ is known as the drift parameter. An important property of random-walk processes is that the best predictor of the value at time $t+1$ is the value at time t plus the value of the drift parameter.

recursive regression analysis. A recursive regression analysis involves performing a regression at time t by using all available observations from some starting time t_0 through time t, performing another regression at time $t+1$ by using all observations from time t_0 through time $t+1$, and so on. Unlike a rolling regression analysis, the first period used for all regressions is held fixed.

rolling regression analysis. A rolling, or moving window, regression analysis involves performing regressions for each period by using the most recent m periods' data, where m is known as the window size. At time t the regression is fit using observations for times $t-19$ through time t; at time $t+1$ the regression is fit using the observations for time $t-18$ through $t+1$; and so on.

seasonal difference operator. The period-s seasonal difference operator Δ_s denotes the difference in the value of a variable at time t and time $t-s$. Formally, $\Delta_s y_t = y_t - y_{t-s}$, and $\Delta_s^2 y_t = \Delta_s(y_t - y_{t-s}) = (y_t - y_{t-s}) - (y_{t-s} - y_{t-2s}) = y_t - 2y_{t-s} + y_{t-2s}$.

serial correlation. Serial correlation refers to regression errors that are correlated over time. If a regression model does not contained lagged dependent variables as regressors, the OLS estimates are consistent in the presence of mild serial correlation, but the covariance matrix is incorrect. When the model includes lagged dependent variables and the residuals are serially correlated, the

OLS estimates are biased and inconsistent. See, for example, Davidson and MacKinnon (1993, chap. 10) for more information.

serial correlation tests. Because OLS estimates are at least inefficient and potentially biased in the presence of serial correlation, econometricians have developed many tests to detect it. Popular ones include the Durbin–Watson (1950, 1951, 1971) test, the Breusch–Pagan (1980) test, and Durbin's (1970) alternative test. See [R] **regress postestimation time series**.

smoothing. Smoothing a time series refers to the process of extracting an overall trend in the data. The motivation behind smoothing is the belief that a time series exhibits a trend component as well as an irregular component and that the analyst is interested only in the trend component. Some smoothers also account for seasonal or other cyclical patterns.

spectral analysis. See *frequency-domain analysis*.

spectral density function. The spectral density function is the derivative of the spectral distribution function. Intuitively, the spectral density function $f(\omega)$ indicates the amount of variance in a time series that is attributable to sinusoidal components with frequency ω. See also *spectral distribution function*. The spectral density function is sometimes called the *spectrum*.

spectral distribution function. The (normalized) spectral distribution function $F(\omega)$ of a process describes the proportion of variance that can be explained by sinusoids with frequencies in the range $(0, \omega)$, where $0 \leq \omega \leq \pi$. The spectral distribution and density functions used in frequency-domain analysis are closely related to the autocorrelation function used in time-domain analysis; see Chatfield (2004, chap. 6) and Wei (2006, chap. 12).

spectrum. See *spectral density function*.

state-space model. A state-space model describes the relationship between an observed time series and an unobservable state vector that represents the "state" of the world. The measurement equation expresses the observed series as a function of the state vector, and the transition equation describes how the unobserved state vector evolves over time. By defining the parameters of the measurement and transition equations appropriately, one can write a wide variety of time-series models in the state-space form.

steady-state equilibrium. The steady-state equilibrium is the predicted value of a variable in a dynamic model, ignoring the effects of past shocks, or, equivalently, the value of a variable, assuming that the effects of past shocks have fully died out and no longer affect the variable of interest.

stochastic trend. A stochastic trend is a nonstationary random process. Unit-root process and random coefficients on time are two common stochastic trends. See [TS] **ucm** for examples and discussions of more commonly applied stochastic trends.

strict stationarity. A process is strictly stationary if the joint distribution of $y_1, \ldots, y_k$ is the same as the joint distribution of $y_{1+\tau}, \ldots, y_{k+\tau}$ for all k and τ. Intuitively, shifting the origin of the series by τ units has no effect on the joint distributions.

structural model. In time-series analysis, a structural model is one that describes the relationship among a set of variables, based on underlying theoretical considerations. Structural models may contain both endogenous and exogenous variables.

SVAR. A structural vector autoregressive (SVAR) model is a type of VAR in which short- or long-run constraints are placed on the resulting impulse–response functions. The constraints are usually motivated by economic theory and therefore allow causal interpretations of the IRFs to be made.

time-domain analysis. Time-domain analysis is analysis of data viewed as a sequence of observations observed over time. The autocorrelation function, linear regression, ARCH models, and ARIMA models are common tools used in time-domain analysis.

trend. The trend specifies the long-run behavior in a time series. The trend can be deterministic or stochastic. Many economic, biological, health, and social time series have long-run tendencies to increase or decrease. Before the 1980s, most time-series analysis specified the long-run tendencies as deterministic functions of time. Since the 1980s, the stochastic trends implied by unit-root processes have become a standard part of the toolkit.

unit-root process. A unit-root process is one that is integrated of order one, meaning that the process is nonstationary but that first-differencing the process produces a stationary series. The simplest example of a unit-root process is the random walk. See Hamilton (1994, chap. 15) for a discussion of when general ARMA processes may contain a unit root.

unit-root tests. Whether a process has a unit root has both important statistical and economic ramifications, so a variety of tests have been developed to test for them. Among the earliest tests proposed is the one by Dickey and Fuller (1979), though most researchers now use an improved variant called the augmented Dickey–Fuller test instead of the original version. Other common unit-root tests implemented in Stata include the DF–GLS test of Elliott, Rothenberg, and Stock (1996) and the Phillips–Perron (1988) test. See [TS] **dfuller**, [TS] **dfgls**, and [TS] **pperron**.

Variants of unit-root tests suitable for panel data have also been developed; see [XT] **xtunitroot**.

VAR. A vector autoregressive (VAR) model is a multivariate regression technique in which each dependent variable is regressed on lags of itself and on lags of all the other dependent variables in the model. Occasionally, exogenous variables are also included in the model.

VECM. A vector error-correction model (VECM) is a type of VAR that is used with variables that are cointegrated. Although first-differencing variables that are integrated of order one makes them stationary, fitting a VAR to such first-differenced variables results in misspecification error if the variables are cointegrated. See *The multivariate VECM specification* in [TS] **vec intro** for more on this point.

white noise. A variable u_t represents a white-noise process if the mean of u_t is zero, the variance of u_t is σ^2, and the covariance between u_t and u_s is zero for all $s \neq t$. Gaussian white noise refers to white noise in which u_t is normally distributed.

Yule–Walker equations. The Yule–Walker equations are a set of difference equations that describe the relationship among the autocovariances and autocorrelations of an autoregressive moving-average (ARMA) process.

References

Breusch, T. S., and A. R. Pagan. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47: 239–253.

Chatfield, C. 2004. *The Analysis of Time Series: An Introduction*. 6th ed. Boca Raton, FL: Chapman & Hall/CRC.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Dickey, D. A., and W. A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74: 427–431.

Durbin, J. 1970. Testing for serial correlation in least-squares regressions when some of the regressors are lagged dependent variables. *Econometrica* 38: 410–421.

Durbin, J., and G. S. Watson. 1950. Testing for serial correlation in least squares regression. I. *Biometrika* 37: 409–428.

———. 1951. Testing for serial correlation in least squares regression. II. *Biometrika* 38: 159–177.

———. 1971. Testing for serial correlation in least squares regression. III. *Biometrika* 58: 1–19.

Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64: 813–836.

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

Phillips, P. C. B., and P. Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75: 335–346.

Wei, W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods.* 2nd ed. Boston: Pearson.