ECONOMETRIC TECHNIQUES FOR REVENUE FORECASTING USING EVIEWS

Session 2: Classical Linear Regression

Patrick Grady

Global Economics Ltd.

Introduction to Econometrics

- What is Econometrics?
  - Application of statistical methods to economic phenomena.
  - Provides statistical tests of economic theory.
  - Provides statistical tools for economic forecasting.
  - Key feature is stochastic disturbance term.
  - Develops estimators for parameters of economic relationships.

Criteria for Estimators

- Computational Cost
- Least squares
- Highest R^2
- Unbiasedness
- Efficiency
- Mean Square Error (MSE)
- Assmptotic Properties
- Maximum Likelihood

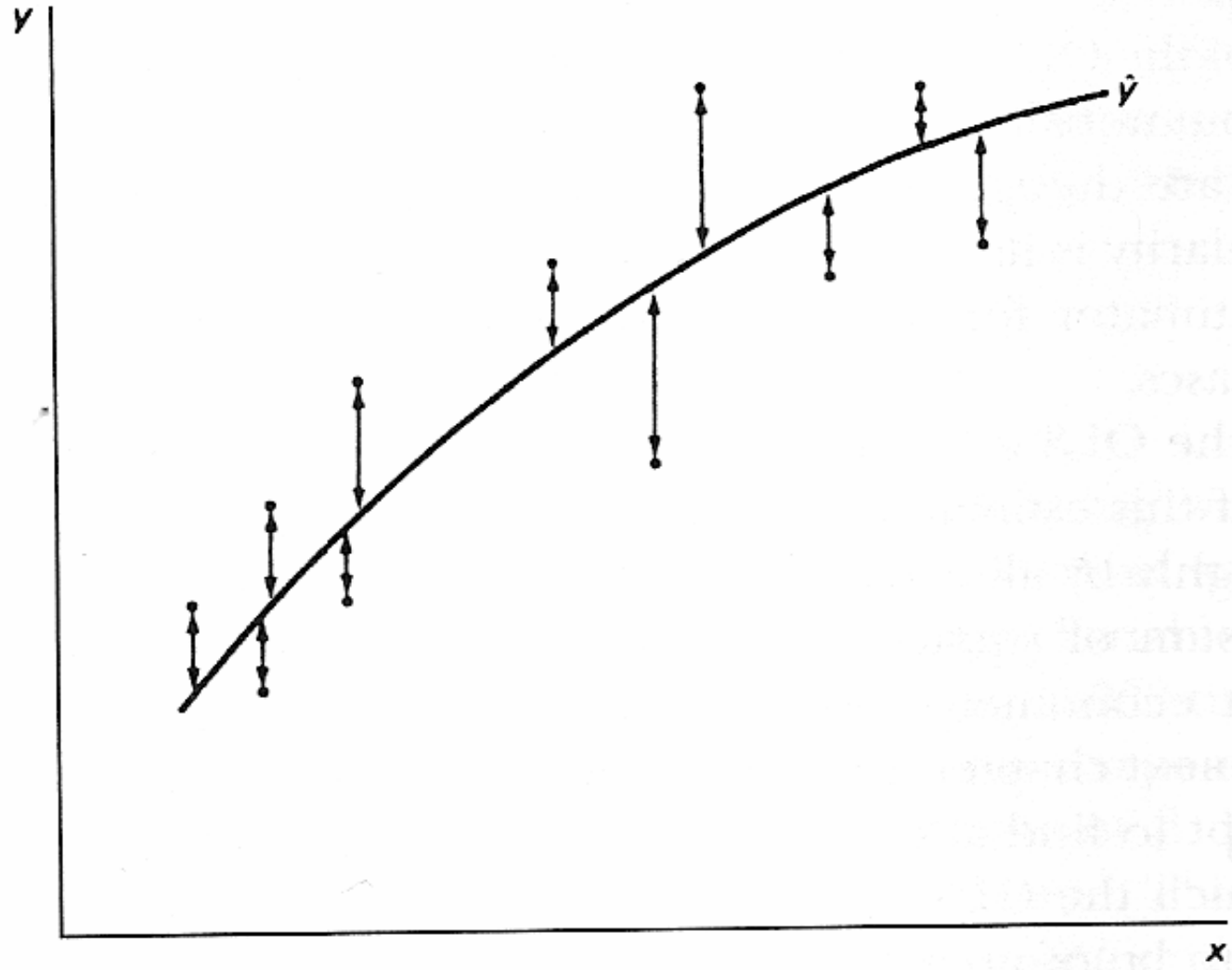

Figure 2.1 Minimizing the sum of squared residuals

Source: Peter Kennedy, *A Guide to Econometrics* (2003).

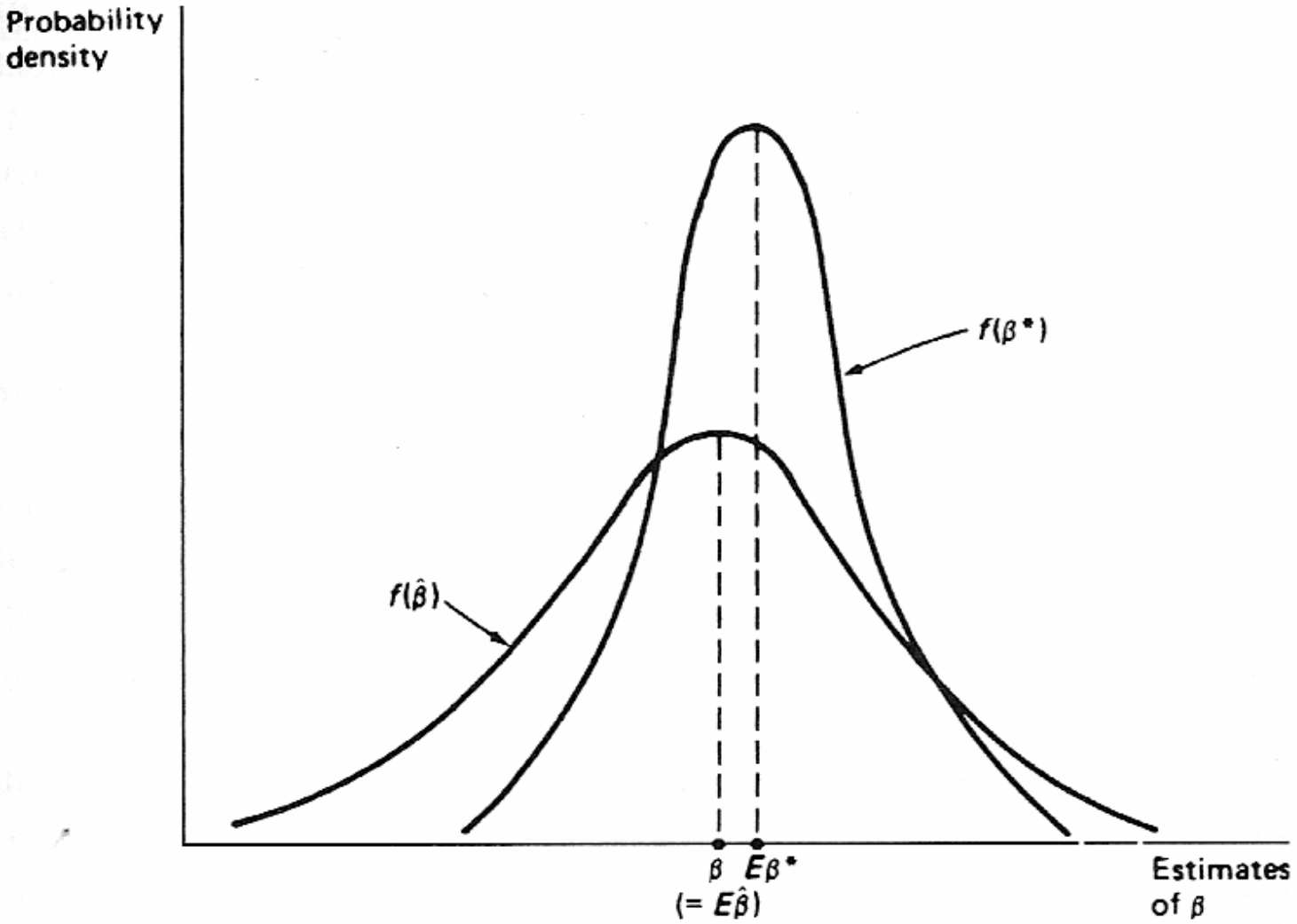

Figure 2.4 MSE trades off bias and variance

Source: Peter Kennedy, *A Guide to Econometrics* (2003).

Table 3.1 The assumptions of the CLR model

Assumption	*Mathematical expression*		*Violations*
	Bivariate	*Multivariate*	
(1) Dependent variable a linear function of a specific set of independent variables, plus a disturbance	$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$, $t = 1, \ldots, T$	$Y = X\beta + \varepsilon$	Wrong regressors Nonlinearity Changing parameters
(2) Expected value of disturbance term is zero	$E\varepsilon_t = 0$, for all t	$E\varepsilon = 0$	Biased intercept
(3) Disturbances have uniform variance and are uncorrelated	$E\varepsilon_t\varepsilon_r = 0$, $t \neq r$ $= \sigma^2$, $t = r$	$E\varepsilon\varepsilon' = \sigma^2 I$	Heteroskedasticity Autocorrelated errors
(4) Observations on independent variables can be considered fixed in repeated samples	x_t fixed in repeated samples	X fixed in repeated samples	Errors in variables Autoregression Simultaneous equations
(5) No exact linear relationships between independent variables and more observations than independent variables	$\sum_{t=1}^{T}(x_t - \bar{x})^2 \neq 0$	Rank of $X = K \leq T$	Perfect multicollinearity

The mathematical terminology is explained in the technical notes to this section. The notation is as follows: Y is a vector of observations on the dependent variable; X is a matrix of observations on the independent variables; ε is a vector of disturbances; σ^2 is the variance of the disturbances; I is the identity matrix; K is the number of independent variables; T is the number of observations.

Source: Peter Kennedy, *A Guide to Econometrics* (2003), p.50.

Interval estimation and hypothesis testing

- Concept of confidence interval
  - dependent variable
  - Coefficients
- Hypothesis testing
  - T tests
  - Chow tests
  - More sophisticated tests on constraints

Specification

- In general, should be based on economic theory.
- But for revenues is simpler and should be based on tax law taking behaviour into account if necessary.
- Alternative specifications should be subjected to statistical tests and comparisons made.
- Sensitivity tests should be carried out.

Some Types of Mispecification Tests

- Omitted Variable Tests
- RESET Tests
- Tests for functional form
- Tests for structural change
- Tests for outliers
- Tests for non-spherical errors
- Tests for exogeneity