Functional Forms

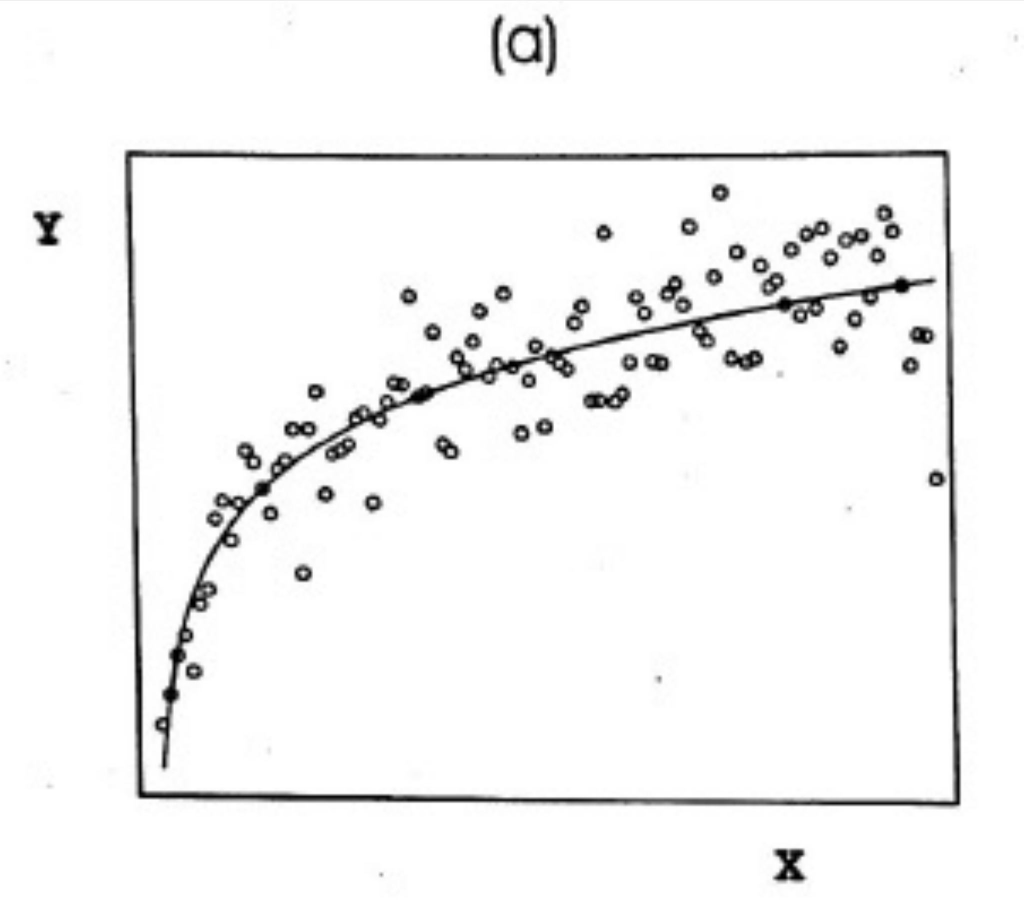
- Difference between “linearity in the parameters” (**MLR.1**) and “linear functional form.”

- **MLR.4**: $E(u \mid x) = 0$
  - Connection to functional form

- Exploration versus theory in specifying different functional forms
  - Pre-regression scatterplots
  - Post-regression “partial residual” plots

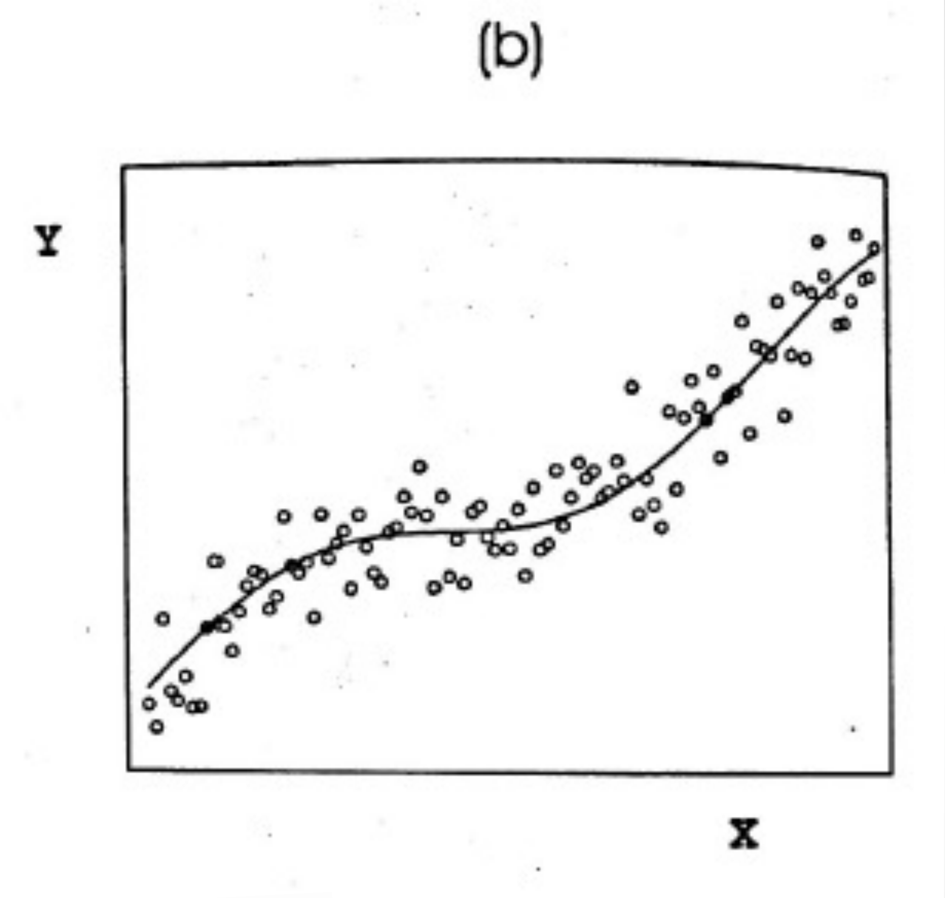
- Transforming nonlinearity
  - Determine whether relationship between x and y is *monotonic* or n*on-monotonic*.
  - Determine whether relationship is *simple* or *not simple*.

Functional Forms



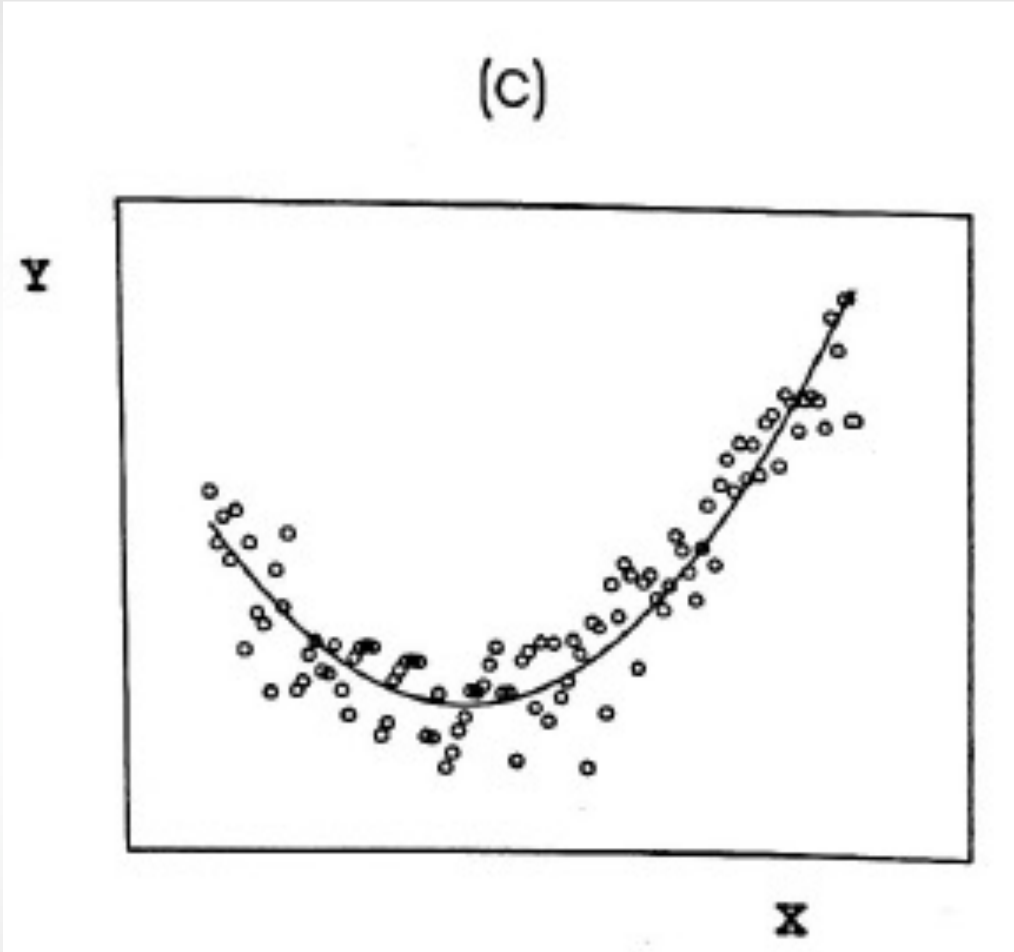
(*a*) A simple monotone relationship between Y and X;

Functional Forms



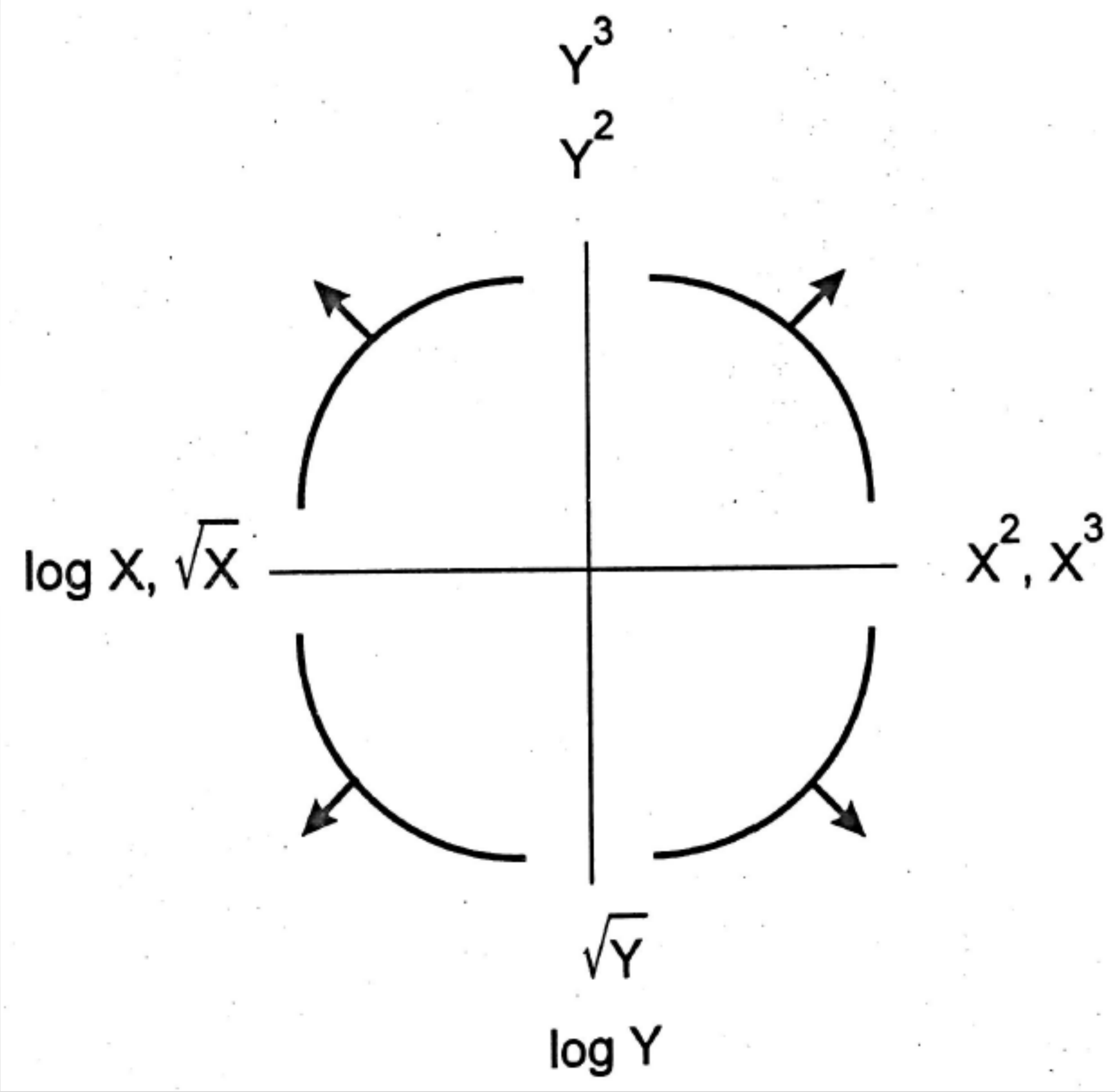
(*b*) a monotone relationship that is not simple;

Functional Forms



(*c*) a relationship that is simple but not monotone.

Functional Forms



Two Common Functional Forms

- **Quadratic operationalization** of x: parabolic relationship between x and y; concave (upside-down u) or convex (u-shaped).
  - Marginal effects
  - Concave or convex?
  - Min or max of function ("inflection" point)
  - Graphing

- **Log transformation** of x: "diminishing marginal returns"
  - Marginal effects
  - Graphing

- Other less common forms: exponential, cubic

Partial Residual Plots

- Partial residual for a given x variable: $\widehat{u\_p}_i = \hat{u}_i + \hat{\beta} x_i$
  - Add the linear component of the partial relationship between a given x and y to the OLS residual.
  - Plot partial residual against the given x; slope of that plot will be $\hat{\beta}$.
  - Use plot to detect any *unmodeled* nonlinear functional form that may be occurring.

- If we have: $y_i = \beta_0 + \beta_1 x1_i + \beta_2 x2_i + \beta_3 x3_i + u_i$

- Calculate three partial residuals:
  - For x1: $\widehat{u\_p}_i = \widehat{u}_i + \hat{\beta}_1 x1_i$
  - For x2: $\widehat{u\_p}_i = \widehat{u}_i + \hat{\beta}_2 x2_i$
  - For x3: $\widehat{u\_p}_i = \widehat{u}_i + \hat{\beta}_3 x3_i$

Transforming Skewness, Log Tranformations

Wooldridge, pp. 191-192

$$\widehat{\log(price)} = \underset{(0.19)}{9.23} - \underset{(.066)}{.718} \log(nox) + \underset{(.019)}{.306}\, rooms \quad [6.7]$$

$$n = 506,\ R^2 = .514.$$

Thus, when *nox* increases by 1%, *price* falls by .718%, holding only *rooms* fixed. When *rooms* increases by one, *price* increases by approximately 100(.306) = 30.6%.

The estimate that one more room increases price by about 30.6% turns out to be somewhat inaccurate for this application. The approximation error occurs because, as the change in $\log(y)$ becomes larger and larger, the approximation $\%\Delta y \approx 100\cdot\Delta\log(y)$

More accurate way:

$$\%\Delta\hat{y} = 100\cdot[\exp(\hat{\beta}_2) - 1]. \quad [6.9]$$

Applied to the housing price example with x_2 = rooms and $\hat{\beta}_2 = .306$, $\%\Delta\widehat{price} = 100[\exp(.306) - 1] = 35.8\%$, which is notably larger than the approximate percentage change, 30.6%, obtained directly from (6.7). {Incidentally, this is not an unbiased