PEMILIHAN MODEL “TERBAIK”

I. Uji Model Secara Keseluruhan

$$Y_i = \boxed{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p} + \varepsilon_i$$

- **Apakah Model dapat menjelaskan keragaman Y**
- **Hipotesis Statistik:**

H_0: Model tdk dpt menjelaskan keragaman Y

$\sigma^2_{regresi} = \sigma^2_\varepsilon$ **atau** $\sigma^2_{regresi} / \sigma^2_\varepsilon = 1$

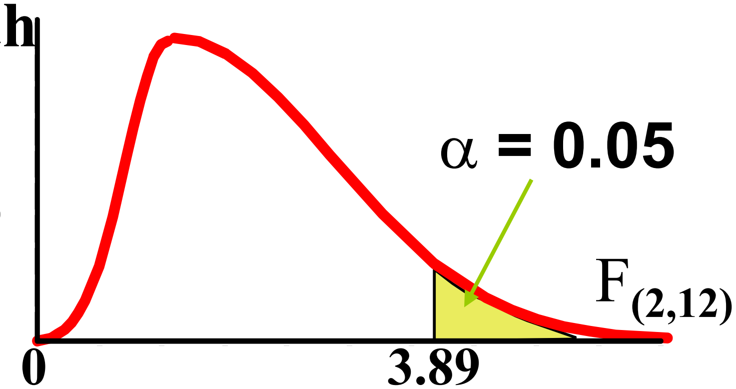
$\beta_1 = \beta_2 = \ldots = \beta_p = 0$ **(tidak dapat menjelaskan)**

H_1: Model dpt menjelaskan keragaman Y

$\sigma^2_{regresi} > \sigma^2_\varepsilon$ **atau** $\sigma^2_{regresi} / \sigma^2_\varepsilon > 1$

Minimal ada $\beta_i \neq 0$ (ada peubah bebas yg mempengaruhi Y)

- **Statistik uji-F = KTR/KTS** $\sim F_{(p,\ n-1-p)}$



R^2 sering digunakan secara “informal” sbg ukuran “*goodness-of-fit*”, dan utk membandingkan validitas hasil regresi dari berbagai spesifikasi model

Beberapa Masalah Penggunaan R^2

1. Hasil uji statistik berdasarkan asumsi awal bahwa Model tsb benar
2. R^2 sensitif thd jumlah peubah bebas dlm model
3. Interpretasi & Penggunaan R^2 jadi sulit, jika suatu model diformulasikan mempunyai intersep 0.

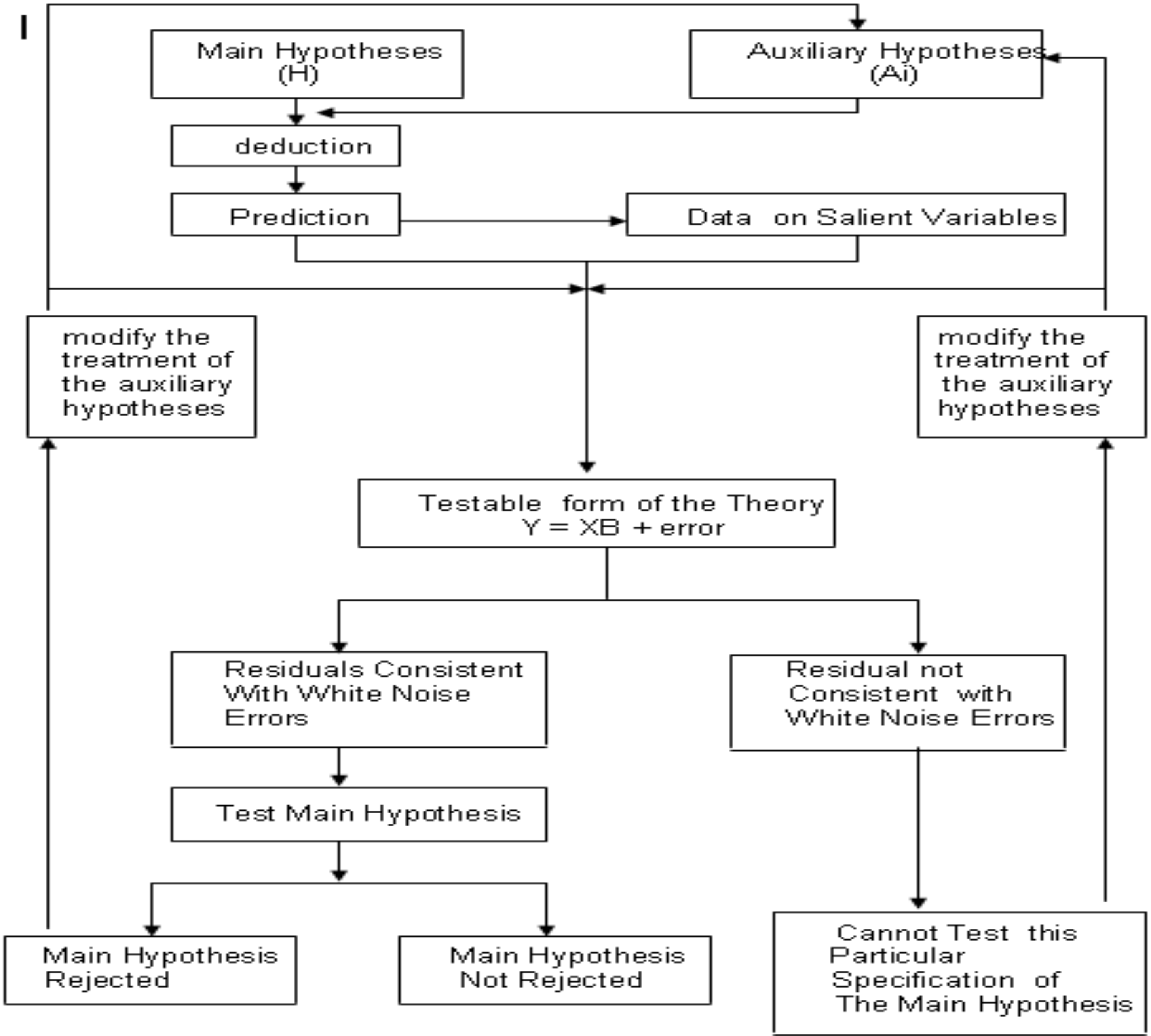
   → R^2 bisa diluar selang (0 s/d1)

Jika jumlah peubah bebas berbeda, sebaiknya menggunakan

R^2 terkoreksi:

$$\overline{R}^2 = 1 - \frac{\hat{V}ar(\in)}{\hat{V}ar(Y)} = 1 - \frac{\sum e_i^{\ 2}}{\sum (Y_i - \overline{Y})^2} \left(\frac{N-1}{N-k} \right)$$

$$\overline{R}^2 = 1 - \left[1 - R^2 \right] \frac{N-1}{N-k}$$



Gambar 1. Tahapan Studi Empiris (untuk menguji hipotesis, perlu diperiksa dulu apakah modelnya sudah "terspesifikasi dengan benar")