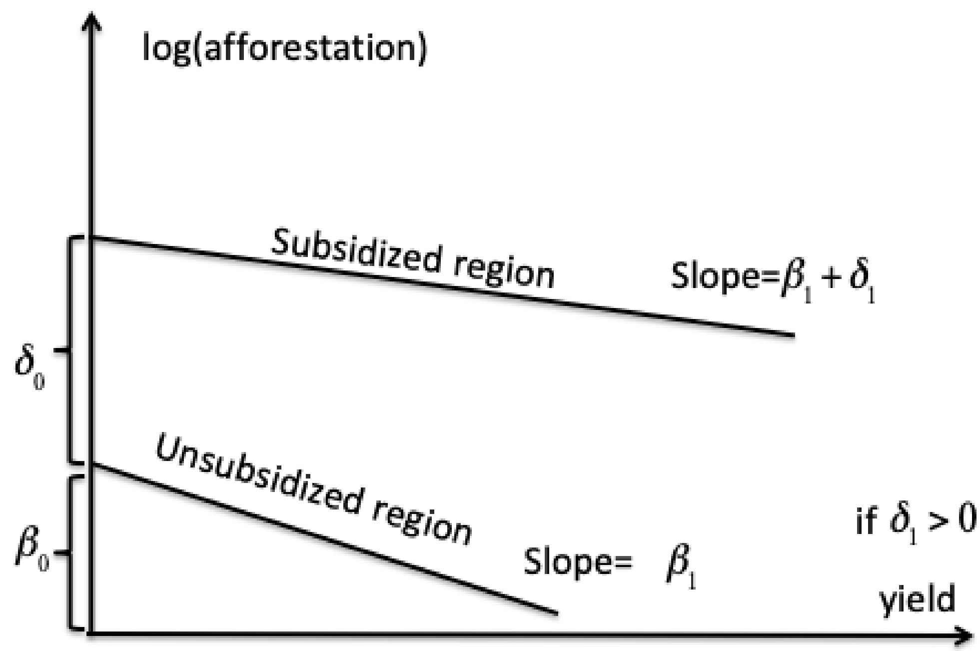
AREC422 Notes

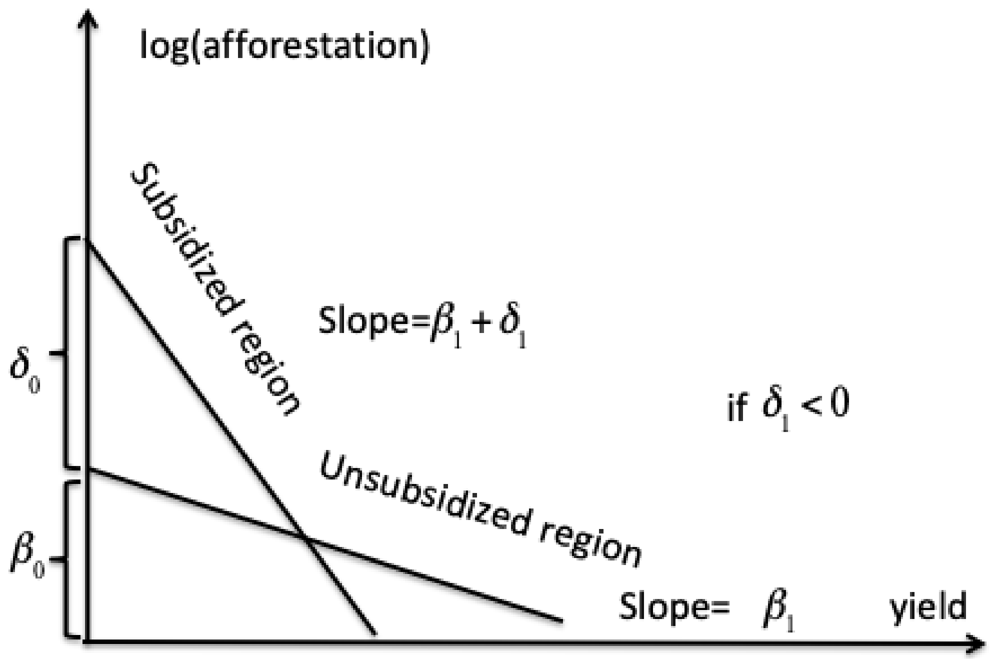
Youpei Yan

April 3, 2019

e.g.2: $log(afforestation) = \beta_0 + \delta_0 subsidy + \beta_1 yield + \delta_1 subsidy \cdot yield + u$ with only $subsidy$ the dummy.

If $subsidy = 0$, $log(afforestation) = \beta_0 + \beta_1 yield + u$
If $subsidy = 1$, $log(afforestation) = (\beta_0 + \delta_0) + (\beta_1 + \delta_1) yield + u$





* Testing between groups: Chow Test

The change in the structure of the regression can be due to the intercept, the slopes, or both.

When we have two groups of observations:
group 1: $y_i = \alpha_1 + \alpha_2 X_i + u_{1i}$, with n_1 observations.
group 2: $y_i = \beta_1 + \beta_2 X_i + u_{2i}$, with n_2 observations.

It is possible that X have different impacts to y: $\alpha_2 \neq \beta_2, \alpha_1 \neq \beta_1$. Or, it is possible that there is no difference between the two groups:
$y_i = \lambda + \lambda X_i + u_i$ with $n = n_1 + n_2$ observations.
i.e., we can fit n points into one model: $\alpha_2 = \beta_2 = \lambda_2, \alpha_1 = \beta_1 = \lambda_1$

e.g.: $pollution = \lambda_0 + \lambda_1 production + \lambda_2 abatement + \lambda_3 resi_d ensity + u$, $n = n_1 + n_2$. There are n_1 firms inside an inspection zone, n_2 firms outside the zone.

How to test if they perform the same?

Unrestricted (UR) model (group of equations):
$pollution = \alpha_0 + \alpha_1 production + \alpha_2 abatement + \alpha_3 resi_d ensity + u$, with n_1 obs.
$pollution = \beta_0 + \beta_1 production + \beta_2 abatement + \beta_3 resi_d ensity + u$, with n_2 obs.

$H_0: \alpha_0 = \beta_0, \alpha_1 = \beta_1, \alpha_2 = \beta_2, \alpha_3 = \beta_3$ vs. $H_1: H_0$ is not true.
Apparently, this is an F-test. (Named Chow Test in this case)

$\#$ Restrictions $= q = k + 1 = 4 = df_r - df_{ur}$

$df_{ur} = (n_1 + n_2) - 2(k+1) = n - 2(k+1)$

Recall the SSR-form of the F-value:

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/df_{ur}}$$

$$F = \frac{[SSR_r - (SSR_1 + SSR_2)]/(k+1)}{(SSR_1 + SSR_2)/[n - 2(k+1)]}$$

What if we have m categories?
UR model:
$y_1 = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + ...\beta_{k1}x_k + u$, with n_1 obs.
$y_2 = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + ...\beta_{k2}x_k + u$, with n_2 obs.
...
$y_m = \beta_{0m} + \beta_{1m}x_1 + \beta_{2m}x_2 + ...\beta_{km}x_k + u$, with n_m obs.
R model:
$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + u$, with $n = n_1 + n_2 + ... + n_m$ obs.

Restrictions $= q = (m-1)(k+1)$ with slopes and intercept restrictions.
Restrictions $= q = (m-1)k$ with just slopes restrictions.

$df_{ur} = (n_1 + n_2 + ... + n_k) - m(k+1)$

$$F_{q,[n-m(k+1)]} = \frac{[SSR_r - (SSR_1 + SSR_2 ... + SSR_m)]/q}{(SSR_1 + SSR_2 ... + SSR_m)/[n - m(k+1)]}$$