

Proxy

$$\ln(\text{wage}) = \beta_1 + \beta_2 \ln(S_i) + \beta_3 \ln(\text{exp}) + \epsilon_i$$

Ability_i is omitted, correlated with S.
OLS biased (inconsistent)

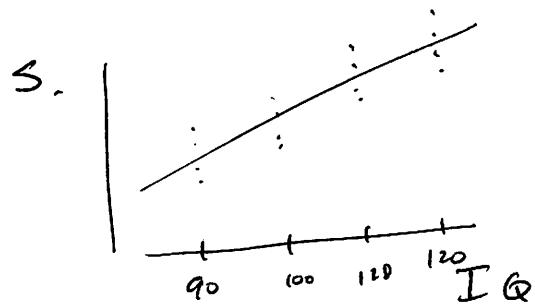
(1) one solution: control variable.
Something correlated with Ability

We can "control" for Ability, if
we have another variable that is
correlated with it

Conditionally mean independent.

Once we know IQ, knowing the person's
schooling tells us nothing new about
his Ability

Education is randomly assigned to
people of given IQ



$$E(\text{Ability}_i | E_{\text{edu}}, \text{IQ}) = E[\text{Ability}_i | \text{IQ}]$$

~~Ability is randomly distributed across IQ.~~

$$\ln(\omega) = \beta_1 + \beta_2 \ln(s) + \beta_3 IQ + \epsilon.$$

$$E(\ln(\omega) | s, \text{Ability}) = \beta_1 + \beta_2 \ln(s)$$

$$E(\text{Ability} | s, IQ) = E(\text{Ability} | IQ)$$

Mac view:

$$\ln(\omega) = \beta_1 + \beta_2 s + \beta_3 \exp + \beta_4 \text{Age} + \beta_5 IQ + \epsilon.$$

$$E[A | s, IQ, \exp, \text{Age}] = E[A | IQ, \exp, \text{Age}]$$

if True then you can
assert causal relations between
lnw and s, Exp, Age. if IQ included as
a "control".

OMB

Omitting a variable that

(1) has an effect &

AND (2) is correlated with other regressors.

$$\ln(w) = \beta_1 + \beta_2 S + \beta_3 \text{Exp} + \epsilon$$

Ability \rightarrow affects wage
 \rightarrow is correlated with S .

So N:

$$\ln(w) = \beta_1 + \beta_2 S + \beta_3 \text{Exp} + \beta_4 \text{Ability} + \epsilon.$$

However: Ability unobserved.
 omitting it from model MAKES CS
 Biased for β_2 (and possibly β_3 if Exp
 correlated with Ability.)

CBS.

$$\hat{\beta}_2 \xrightarrow{P} \beta_2 + \beta_4 \frac{\text{Cov}(S, \text{Ability})}{\text{Var}(S)}$$

NOTE: Bias is \pm dependent on sign of
 β_4 and $\text{Cov}(S, \text{Ab})$.

Ability inc. wage $\beta_4 > 0$

and
 Ability is pos. corr. Then

$$E(\hat{\beta}_2) > \beta_2 \text{ and } \text{Plim}(\hat{\beta}_2) > \beta_2$$

Irrelevant Vars

$$\ln(w) = \beta_1 + \beta_2 S + \beta_3 \# \text{children} + \epsilon.$$

$$\text{Var}(\hat{\beta}_2) = \sigma^2 \left[\frac{1}{(1 - r_{S,\#C}^2) \sum (S_i - \bar{S}_c)^2} \right]$$

$r_{S,\#C}^2$ is ~~a~~ from $\frac{r_{23}}{r_{23}}$ is sample corr.
~~corr~~: Between S & $\# \text{children}$

$$S = \pi_1 + \pi_2 \# \text{children} + \nu.$$

~~Also,~~

$$\beta_3 = 0 \quad \text{But correlated with } S.$$

$$\text{Var}(\hat{\beta}_S) \uparrow$$

Recall: ^{more} independent variation in S
 makes CSE of β_S more precise.
 Adding irrelevant vars to model
 that are correlated with other Reg.
 Increases SE, reduces size of t, increases
 size of intervals

Cornish Model

$$y = \beta_1 + \beta_2 t_2 + \epsilon \quad y = \beta_1 + \beta_2 T_2 + \beta_3 T_3 + \epsilon$$

Est. β_1, β_2

\checkmark OLS Blue. Tests valid.	\times OLS Biased. More precise. Tests NOT Valid.
\checkmark OLS Unbiased, NOT Blue Overspec. Tests Valid. lose Power.	\checkmark OLS Blue Tests Valid.

$\beta_1, \beta_2, \beta_3$