Interpreting univariate OLS coefficients

Precise interpretation of a simple univariate, cross-sectional regression is not easy. Here is some help. There are many good (precise) ways to do it, some that are not perfect and some that are not good.

The question

You have a sample on 10,000 people in a country, aged 15-45. You are interested in the relationship between earning (USD /per year) and age. You run a simple linear regression estimated with OLS.

$$y_i^E = \hat{\alpha} + \hat{\beta} \times age_i$$

Interpret the coefficients, where $\hat{\alpha} = 7000, \hat{\beta} = 400$.

Good answers

$\hat{\alpha}$

- **For people aged zero (when age=0), earnings is $7000, on average**
- For people aged zero (when age=0), the expected earning is $7000 (you may or may not add on average, "expected" includes it)
- For people aged zero (when age=0), earning is $7000, on average
- The constant cannot be interpreted in this context (because newborns make no money)
- Intercept of the regression line, in this case has no realistic meaning, no earnings at age = 0

$\hat{\beta}$

- **People who are one year older, earn $400 more, on average**
- People who are one year older tend to earn $400 more / are expected to earn $400 more / tend to have higher earnings by $400) (you may or may not add on average, "expected" kind of includes it)
- One additional year of age is associated with $400 higher earning, on average
- One year age difference is associated with and an average of $400 extra earnings
- One additional year in age corresponds to an average of $400 extra earnings
- Earnings of people who are one year older, are (tend to be / is expected to be) on average $400 higher in the data
- Comparing two people, the one who is one year older, is expected to (tend to have) have $400 higher earning

Partially ok

$\hat{\alpha}$

- Newborn/Zero aged people earn $7000 (missed: on average)
- Average values of earning without considering the age is $7000 (we consider it, but it's zero)
- The person earns 7000usd/year at least, no matter what is his age (true but only because beta is positive)
- 7000 is the minimum income that has to be given irrespective of the age (true but only because beta is positive)

$\hat{\beta}$

- People who are one year older will have $400 higher earnings, on average ("will have": the data is about the past, we don't know what the future brings. Yes, will can mean "likely" but should be avoided)
- One additional year in age corresponds to $400 higher earning (missed: on average)
- One extra year means (implies) 400 more (suggests causality, and missed on average)
- any extra age adds up 400 to earnings (suggests causality, and missed on average)

Not good

$\hat{\alpha}$

- The intercept is 7000 (not interpretation)
- Average earning is 7000 if age=15 (not at the minimum age in the sample)
- Average earning is 7000 (no, 7000 is average earnings at age zero)

$\hat{\beta}$

- for every unit change in age the change in earning on average is 400USD (it's about cross-section differences between people, not changes)
- One year increase will get $400 increase in wage (no time series or causality, no increase!)
- Each year in the age increases earnings by $400 (no time series or causality, no increase!) The slope is 400 (not interpretation)