**Seven Sins of Regression Analysis**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Problem** | | **Does it create bias?** | **Does it affect SE's?** | **Solution** |
| 1 | Multicollinearity | NO | YES | Depends upon whether the issue is measurement or causal. If it is measurement, omit a variable. If it is causal, you need good research design. |
| 2 | Omitted Variable Bias | YES | YES | Research design, Fixed Effects, Instrumental Variables |
| 3 | Measurement Error DV | NO | YES | Better measurement |
| Measurement Error IV | YES - attenuation | YES | Better measurement |
| 4 | Misspecification | YES | YES | Visualize the data, add non-linear terms to the model (X2, X3, etc) use log transformations, outlier analysis |
| Outliers | Maybe | If bias exists |
| 5 | Heterogeneity | YES | YES | Fixed effects if it is unobservable omitted variables that cause the observed differences in the data (intelligence, natural ability, efficiency in task completion). |
| 6 | Selection (non-random into or out of a group) | YES | YES | Control non-random selection through research design (randomization to control and treatment groups). Model the selection process (Heckman Selection Model - not covered here, just know it exists). |
| Random attrition | NO | YES |
| 7 | Simultaneity | YES | YES | Panel data, Structural Equation Models (SEM) (not covered here - just know they exist) |