

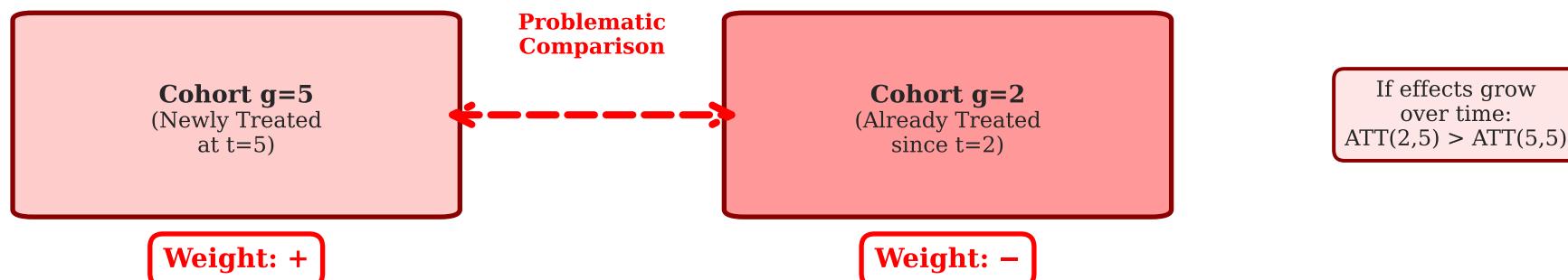
## Valid Comparisons: Treated vs Never-Treated or Not-Yet-Treated



*Under parallel trends, these comparisons identify  $ATT(g,t)$*

*All weights are positive → TWFE is a weighted average of valid effects*

## Problematic Comparisons: Newly Treated vs Already-Treated (Negative Weights)



*Comparison confounds treatment effect with duration of treatment*

*Negative weight on already-treated unit → TWFE can be biased or wrong sign*

*Problem severity increases with heterogeneity and staggered timing*