**Test Exercise 4**

1. The endogeneity of the variable di in the model can arise due to potential omitted variable bias or reverse causality. In the context of this application, if di is correlated with unobservable factors that affect weight change (e.g., motivation, health conditions), then the OLS estimator for β might not be consistent. If di is endogenous and positively correlated with weight change (i.e., people who followed the diet had greater weight change), OLS would overestimate the true effect of the diet. On the other hand, if di is negatively correlated with weight change, OLS would underestimate the true effect.
2. In the context of instrumental variables, the two important conditions for variables Z to be useful as instruments are:

* Relevance: This condition requires that the instrument (Z) is correlated with the endogenous variable (di). In the context of the advertising variable, it means that the door-to-door advertising (zi) should be correlated with whether individuals followed the diet (di).
* Exogeneity: This condition requires that the instrument (Z) is not directly correlated with the error term (εi) in the main equation. In this case, it means that the door-to-door advertising (zi) should not affect weight change (εi) directly.

1. These assumptions can potentially be tested statistically given the available variables. To test relevance, you could run a regression of di on zi and other relevant control variables. If the coefficient on zi is statistically significant, it suggests that the instrument is relevant. To test exogeneity, you could regress weight change (εi) on zi and other relevant control variables. If the coefficient on zi is statistically insignificant, it suggests that the instrument is exogenous.
2. The 2SLS estimator for β can be written in terms of simple sample averages as follows:

β̂(2SLS) = (∆1 - ∆0) / (¯d1 - ¯d0)

Where:

* ∆1 = Average weight change over individuals with zi = 1
* ∆0 = Average weight change over individuals with zi = 0
* ¯d1 = Proportion of people with zi = 1 taking the diet
* ¯d0 = Proportion of people with zi = 0 taking the diet

This estimator utilizes the differences in average weight changes between those exposed to the advertising (zi = 1) and those not exposed (zi = 0) to estimate the causal effect of the diet, while controlling for the proportions of people who took the diet in each group.