Under our assumption, y_{i1} is always observed, but y_{i2} may be missing. Let r_{i2} be a missing data indicator for y_{i2}, with $r_{i2} = 1$ if y_{i2} is observed and 0 if otherwise. Consider the following estimate of $\boldsymbol{\mu}$:

$$\widehat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} y_{i1}, \quad \widehat{\mu}_2 = \frac{1}{n_2}\sum_{i=1}^{n} r_{i2}y_{i2}, \quad n_2 = \sum_{i=1}^{n} r_{i2} \qquad (4.70)$$

It is readily seen that the above is simply a fancy way to express the sample mean of y_{i2} based on observed data. If y_{i2} is not subject to missingness, $\widehat{\boldsymbol{\mu}}$ above yields the GEE estimate of $\boldsymbol{\mu}$ in the complete data case. Otherwise, if missing y_{i2} follows MCAR, r_{i2} is independent of y_{i1} as well as y_{i2}. Thus, we have

$$\pi_{i2} = \Pr(r_i = 1 \mid y_{i1}, y_{i2}) = \Pr(r_i = 1) \qquad (4.71)$$

Intuitively, $\widehat{\mu}_2$ should be consistent. This is indeed the case; it follows from LLN, Slutsky's theorem and (4.71) that

$$\widehat{\mu}_2 = \frac{\frac{1}{n}\sum_{i=1}^{n} r_{i2}y_{i2}}{\frac{1}{n}\sum_{i=1}^{n} r_{i2}} \rightarrow_p \frac{1}{E(r_{i2})} E(r_{i2}y_{i2}) = \mu_2 \qquad (4.72)$$

Example 3 (*One-sample repeated measures ANOVA under MAR*). In Example 2, if the missingness depends on y_{i1}, (4.71) no longer holds and $\widehat{\mu}_2$ in general will not be consistent. For example, if y_{i1} and y_{i2} are positively correlated with higher values of y_{i1} more likely leading to missing y_{i2}, $\widehat{\mu}_2$ will be downwardly biased. In treatment studies, this type of response-dependent missingness may happen if a patient feels that he/she has responded to the treatment and decides not to undergo additional treatment. To construct a consistent estimate, we may somehow "recover" the missing response y_{i2} for these study dropouts and include them in the estimation of μ_2. Of course, such a recovery process is not possible in most studies. However, with a probability model for the missingness data indicator r_{i2}, we can "statistically recover" such missing y_{i2}.

Since the missingness of y_{i2} only depends on y_{i1} as guaranteed by the MAR model, we have:

$$\pi_{i2} = \Pr(r_{i2} = 1 \mid y_{i1}, y_{i2}) = \Pr(r_{i2} = 1 \mid y_{i1}) \qquad (4.73)$$

For each subject i, we observe y_{i2} with probability π_{i2}. Thus, this subject represents a subgroup of $\frac{1}{\pi_{i2}}$ subjects with response y_{i1} at time 1 who are unobserved at time 2. By augmenting each observed subject with the weight