Least Squares Estimates for the CER model

The mean μ_i:

$$\hat{\mu}_i = \frac{1}{T}\sum_{t=1}^{T} r_{it} \ ,$$

where the "hat" denotes the "least squares" estimate, *T* denotes the sample size (i.e., number of observations), and the r_{it}'s are realized c.c. return values.

The variance σ_i^2:

$$\hat{\sigma}_i^2 = \frac{1}{T-1}\sum_{t=1}^{T} (r_{it} - \hat{\mu}_i)^2 \ ,$$

where the "hat" denotes a "least squares" estimate, *T* denotes the sample size (i.e., number of observations), and the r_{it}'s are realized c.c. return values.

The covariance σ_{ij}:

$$\hat{\sigma}_{ij} = \frac{1}{T-1}\sum_{t=1}^{T} (r_{it} - \hat{\mu}_i)(r_{jt} - \hat{\mu}_j) \ ,$$

where the "hat" denotes a "least squares" estimate, *T* denotes the sample size (i.e., number of observations), and the r_{it}'s and r_{jt}'s are realized c.c. return values. Note: the formula used in Excel for the "COVAR" function divides by 1/T instead of (1/(T-1)). The Excel estimator is biased, but consistent. However, you can use either the formula above or the Excel function in any homework assignment.

The correlation ρ_{ij}:

$$\hat{\rho}_{ij} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i \hat{\sigma}_j} \ ,$$

where the "hat" denotes a "least squares" estimate.

The estimated standard errors of the corresponding "least squares" estimat*ors* are given as follows:

$$S\hat{E}(\hat{\mu}_i) = \frac{\hat{\sigma}_i}{\sqrt{T}} \ ,$$

$$SE(\hat{\sigma}_i^2) = \frac{\hat{\sigma}_i^2}{\sqrt{(T-1)/2}},$$

$$SE(\hat{\sigma}_i) = \frac{\hat{\sigma}_i}{\sqrt{2(T-1)}},$$

$$SE(\hat{\rho}_{ij}) = \frac{(1-\hat{\rho}_{ij}^2)}{\sqrt{T-1}},$$

where the “least squares” estimators are on the left hand sides and “least squares” estimates are on the right hand sides. (It should be noted that the formulas for the variance, standard deviation, and correlation are approximate, but valid if the sample size T is reasonably large.)

Finally, in terms of calculating 95% confidence intervals, the following result is useful:

$$\frac{\hat{\theta}_i - \theta_i}{S\hat{E}(\hat{\theta}_i)} \sim t_{T-1},$$

where $\hat{\theta}_i$ is the “least squares” estimat*or* of interest (e.g., $\hat{\theta}_i = \hat{\mu}_i$). Since the estimator is a random variables, the left-hand side expression is a random variable. Because an estimated standard error is used in the expression, the left-hand side expression has a t-distribution instead of a standard Normal distribution. For a sample size of around 60, the above result implies

95% confidence interval = $\left[\hat{\theta}_i - 2 \cdot S\hat{E}(\hat{\theta}), \quad \hat{\theta}_i + 2 \cdot S\hat{E}(\hat{\theta})\right]$,

where 2 is the 95% critical value for a t-distribution with around 60 degrees of freedom.