Consider a CES production function $f(x) = \left(\sum_{j=1}^{n} x_j^{\rho}\right)^{\frac{1}{\rho}}$. This has a marginal product which can be simplified as follows:

$$MP_i = x_i^{\rho-1}\left(\sum_{j=1}^{n} x_j^{\rho}\right)^{\frac{1}{\rho}-1} = \frac{\left(\sum_{j=1}^{n} x_j^{\rho}\right)^{\frac{1}{\rho}}}{\sum_{j=1}^{n} x_j^{\rho}}\frac{x_i^{\rho}}{x_i}$$

$$= \frac{\left(\sum_{j=1}^{n} x_j^{\rho}\right)^{\frac{1}{\rho}}}{x_i}\frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}} = \frac{f(x)}{x_i}\frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}} = AP_i\frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}}$$

. Notice that $\frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}} \leq 1$, thus, $MP_i \leq AP_i$. As x_i increases, it's instantaneous productivity is always less than it's average productivity has been in the past. Another form of the marginal product, is a bit more elegant:

$$MP_i = \left(\sum_{j=1}^{n}\left(\frac{x_j}{x_i}\right)^{p}\right)^{\frac{1-p}{p}} = \left(1 + \sum_{j\neq i}\left(\frac{x_j}{x_i}\right)^{p}\right)^{\frac{1-p}{p}}$$

Notice that MP_i is decreasing in x_i for $0 < p < 1$ since x_i appears only the the denominator of each of the terms in the sum and the sum is raised to a positive power. For $p < 0$, MP_i is decreasing as well, since each term $\left(\frac{x_j}{x_i}\right)^{p}$ is increasing in x_i in this range, but the entire function is decreasing in the summation term $1 + \sum_{j\neq i}\left(\frac{x_j}{x_i}\right)^{p}$ since it is raised to the power $\frac{1-p}{p}$ which is negative for $p < 0$. Thus, the CES function has diminishing marginal product for $p \in (-\infty, 1)$. When $p > 1$, this is not the case.

Returning to this form of the marginal product: $MP_i = AP_i\frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}}$ and noting output elasticity is $\frac{MP_i}{AP_i}$, we have output elasticity equal to:

$$E_i = \frac{x_i^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}}$$

This is always less than one. Note further that:

$$\sum_{i=1}^{n} E_i = \frac{\sum_{i=1}^{n} x_j^{\rho}}{\sum_{j=1}^{n} x_j^{\rho}} = 1$$

This is the elasticity of scale. This is expected since the CES functions being analyzed here are all homogeneous of degree one and have globally constant returns to scale.