III: CONSISTENCY AND SIMULTANEOUS-EQUATION BIAS

In this note, we shall reveal some of the properties of the least-squares estimators which follow from the classical assumptions. We shall also consider the likelihood that these assumptions will be fulfilled in practice, as well as some consequences of their violation.

We assume that the disturbance term ε is a random variable with

$$E(\varepsilon_t) = 0, \quad \text{and} \quad V(\varepsilon_t) = \sigma^2 \quad \text{for all } t. \tag{1}$$

We have avoided making statistical assumptions about x since we are unwilling to assume that its assembled values will manifest the sort of the regularities which are inherent in a statistical distribution. Therefore, we cannot express the assumption that ε is independent of x in terms of a joint distribution of these quantities; and, in particular, we should not assert that $C(x, \varepsilon) = 0$. However, if we are prepared to regard the x_t as predetermined values which have no effect on the ε_t, then we can say that

$$E(x_t \varepsilon_t) = x_t E(\varepsilon_t) = 0, \quad \text{for all } t. \tag{2}$$

In place of an assumption attributing a finite variance to x, we may assert that

$$\lim(T \to \infty) \frac{1}{T} \sum_{t=1}^{T} x_t^2 = m_{xx} < \infty. \tag{3}$$

For the random sequence $\{x_t \varepsilon_t\}$, we assert that

$$\text{plim}(T \to \infty) \frac{1}{T} \sum_{t=1}^{T} x_t \varepsilon_t = 0. \tag{4}$$

To see the effect of these assumptions, let us substitute the expression

$$y_t - \bar{y} = \beta(x_t - \bar{x}) + \varepsilon_t - \bar{\varepsilon} \tag{5}$$

in the expression for $\hat{\beta}$ found under (45). By rearranging the result, we have

$$\hat{\beta} = \beta + \frac{\sum (x_t - \bar{x})\varepsilon_t}{\sum (x_t - \bar{x})^2}. \tag{6}$$

The numerator of the second term on the RHS is obtained with the help of the identity

$$\begin{aligned} \sum (x_t - \bar{x})(\varepsilon_t - \bar{\varepsilon}) &= \sum (x_t \varepsilon_t - \bar{x}\varepsilon_t - x_t \bar{\varepsilon} + \bar{x}\bar{\varepsilon}) \\ &= \sum (x_t - \bar{x})\varepsilon_t. \end{aligned} \tag{7}$$

From the assumption under (2), it follows that

$$E\big\{(x_t - \bar{x})\varepsilon_t\big\} = (x_t - \bar{x})E(\varepsilon_t) = 0 \quad \text{for all } t. \tag{8}$$

Therefore

$$
\begin{aligned}
E(\hat{\beta}) &= \beta + \frac{\sum(x_t - \bar{x})E(\varepsilon_t)}{\sum(x_t - \bar{x})^2} \\
&= \beta;
\end{aligned} \tag{9}
$$

and $\hat{\beta}$ is seen to be an unbiased estimator of β.

The consistency of the estimator follows, likewise, from the assumptions under (3) and (4). Thus

$$
\begin{aligned}
\text{plim}(\hat{\beta}) &= \beta + \frac{\text{plim}\Big\{T^{-1}\sum(x_t - \bar{x})\varepsilon_t\Big\}}{\text{plim}\Big\{T^{-1}\sum(x_t - \bar{x})^2\Big\}} \\
&= \beta;
\end{aligned} \tag{10}
$$

and $\hat{\beta}$ is seen to be a consistent estimator of β.

The consistency of $\hat{\beta}$ depends crucially upon the assumption that the disturbance term is independent of, or uncorrelated with, the explanatory variable or regressor x. In many econometric contexts, we should be particularly wary of this assumption. For, as we have suggested earlier, the disturbance term is liable to be compounded from the variables which have been omitted from the equation which explains y in terms of x. In a time-dependent context, these variables are liable to be correlated amongst themselves; and there may be scant justification for assuming that they are not likewise correlated with x.

There are other reasons of a more subtle nature for why the assumption of the independence of ε and x may be violated. The following example illustrates one of the classical problems of econometrics.

Example. In elementary macroeconomic theory, a simple model of the economy is postulated which comprises two equations:

$$
y = c + i, \tag{11}
$$

$$
c = \alpha + \beta y + \varepsilon. \tag{12}
$$

Here y stands for the gross product of the economy, which is also the income of consumers, i stands for investment and c stands for consumption. An additional identity $s = y - c$ or $s = i$, where s is savings, is also entailed. The disturbance term ε, which is omitted from the usual presentation in economics textbooks, is assumed to be independent of the variable i.

On substituting the consumption function of (12) into the income identity of (11) and rearranging the result, we find that

$$
y = \frac{1}{1-\beta}\big(\alpha + i + \varepsilon\big), \tag{13}
$$

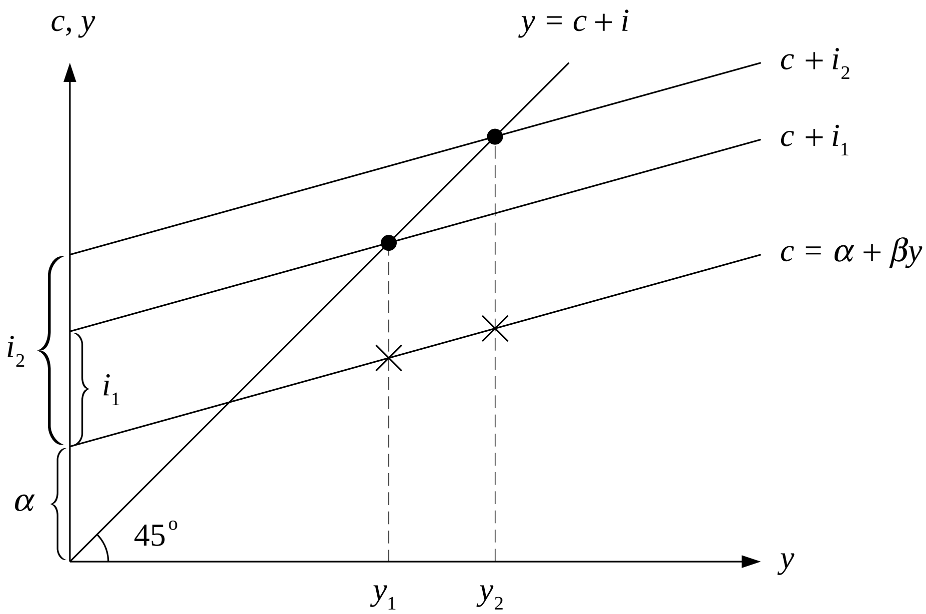

Figure 1. If the only source of variation in y is the variation in i, then the observations on y and c will delineate the consumption function.

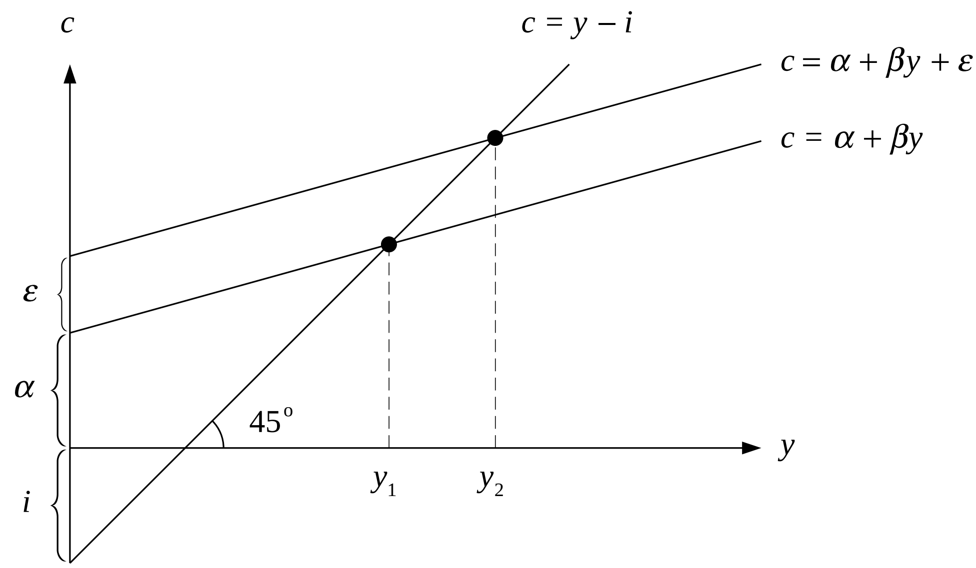

Figure 2. If the only source of variation in y are the disturbances to c, then the observations on y and c will line along a 45° line.

from which

$$y_t - \bar{y} = \frac{1}{1-\beta}\big(i_t - \bar{i} + \varepsilon_t - \bar{\varepsilon}\big). \tag{14}$$

The ordinary least-squares estimator of the parameter β, which is called the marginal propensity to consume, gives rise to the following equation:

$$\hat{\beta} = \beta + \frac{\sum(y_t - \bar{y})\varepsilon_t}{\sum(y_t - \bar{y})^2}. \tag{15}$$

Equation (13), which shows that y is dependent on ε, suggests that $\hat{\beta}$ cannot be a consistent estimator of β.

To determine the probability limit of the estimator, we must assess the separate probability limits of the numerator and the denominator of the term on the RHS of (15).

The following results are available:

$$\begin{aligned} &\lim \frac{1}{T}\sum_{t=1}^{T}(i_t - \bar{i})^2 = m_{ii} = V(i),\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2 = \frac{m_{ii} + \sigma^2}{(1-\beta)^2} = V(y),\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})\varepsilon_t = \frac{\sigma^2}{1-\beta} = C(y,\varepsilon). \end{aligned} \tag{16}$$

The results indicate that

$$\begin{aligned} \text{plim}\ \hat{\beta} &= \beta + \frac{\sigma^2(1-\beta)}{m_{ii} + \sigma^2}\\ &= \frac{\beta m_{ii} + \sigma^2}{m_{ii} + \sigma^2}; \end{aligned} \tag{17}$$

and it can be seen that the limiting value of $\hat{\beta}$ has an upward bias which increases as the ratio σ^2/m_{ii} increases.

On the assumption that the model is valid, it is easy to understand why the parameter of the regression of y on c exceeds the value of the marginal propensity to consume. We can do so by considering the extreme cases.

Imagine, first, that $\sigma^2 = V(\varepsilon) = 0$. Then the only source of variation in y and c is the variation in i. In that case, the parameter of the regression of c on y will coincide with β. This is illustrated in Figure 1. Now imagine, instead, that i is constant and that the only variations in c and y are due ε which is disturbs consumption. Then the expected value of consumption is provided by the equation $c = y - i$ in which the coefficient associated with

y is unity. Figure 2 illustrates this case. Assuming now that both $m_{ii} > 0$ and $\sigma^2 > 0$, it follows that the value of the regression parameter must lie somewhere in the interval $[\beta, 1]$.

Although it may be inappropriate for estimating the structural parameter β, the direct regression of c on y does provide the conditional expectation $E(c|y)$; and this endows it with a validity which it retains even if the Keynesian model of (11) and (12) is misspecified.

In fact, the simple Keynesian model of (11) and (12) is more an epigram than a serious scientific theory. Common sense dictates that we should give more credence to the estimate of the conditional expectation $E(c|y)$ than to a putative estimate of the marginal propensity to consume devised within the context of a doubtful model.