3 : CHAPTER

Latent Variables and Simultaneous Equations

In the previous chapter, we introduced, through a brief example, the idea of a signal buried in noise; and we indicated a method of extracting the signal which required some prior knowledge of the statistical parameters of the error process. The example was expressed in the terminology of communications engineering; and, in this field, it is usually possible to obtain information on the error process by experimental means. The problem of signal extraction has an analogy in the so-called errors-in-variables problem of econometrics.

There are good reasons for expecting economic variables to be measured with error. There are bound to be numerous recording errors in the process of compiling economic indices from individual measurements of prices and quantities. Moreover, econometricians are often constrained to use alternative or proxy measurements in place of those which they would ideally employ. The difference between a proxy measurement and an ideal measurement is akin to an error of observation.

A major problem in dealing with measurement errors in econometrics is the difficulty in determining the parameters of the error processes. There is usually no way of telling the extent to which recording errors and errors of compilation afflict the official economic statistics.

Reflection on the question of the accuracy of economic statistics raises philosophical problems. An economic index, such as the level of investment, is often the empirical counterpart of an abstract construct which would have little meaning apart from the context of an economic model. Only rarely do the facts which are gathered in a statistical enquiry match perfectly with the categories of economic analysis. In order to bring theory and reality together, the data gatherers must arbitrate on matters which cannot be expressed in terms of the theory; and therefore economic indices are unavoidably imprecise. The idea, borrowed from the physical sciences, that there is an exact quantity underlying each erroneous measurement is difficult to sustain.

Thus there are both practical and philosophical reasons for why the notion of errors in variables has few direct applications in econometrics; and, in many textbooks, the topic is accorded only minor importance. Nevertheless the so-called errors-in-variables model does have a important role to play in classical econometric theory. The mathematical structure of the errors-in-variables model is identical to that of a structural equation within a simultaneous-equation econometric model. The simultaneous-equation model has occupied a position of prime importance within econometric theory.

In this chapter, we shall begin by developing the error-in-variables model in its own right, as if it were directly applicable to a practical problem. Then we shall apply our results to the problem of estimating a simultaneous-equation system which we shall exemplify with a model of supply and demand. Finally, we shall apply the errors-in-variables model to the problem of estimating a dynamic model of the consumption function which entails the notion of unobservable or latent variables.

The Errors-in-Variables Model

Imagine that the variables ξ_1, ξ_2 have an exact linear relationship

$$\xi_1\beta_1 + \xi_2\beta_2 = \alpha. \tag{1}$$

Imagine also that, instead of observing these variables, we observe

$$y_1 = \xi_1 + \eta_1 \quad \text{and} \quad y_2 = \xi_2 + \eta_2, \tag{2}$$

where η_1 and η_2 are errors of observation which are distributed independently of each other and of the true values ξ_1 and ξ_2. We shall assume that

$$E(\eta_i) = 0, \quad V(\eta_i) = \omega_{ii} \quad \text{and} \quad C(\eta_i, \eta_j) = \omega_{ij}, \tag{3}$$

where $i, j = 1, 2$.

The equations of (1) and (2) may be combined to give

$$(y_1 - \eta_1)\beta_1 + (y_2 - \eta_2)\beta_2 = \alpha. \tag{4}$$

The object is to find expressions for the parameters α, β_1 and β_2 which are in terms of the variances and covariances of the observations y_1, y_2 and of the errors which afflict them.

We shall begin the search for these estimators by resorting to the method of moments. The approach is similar to one which we have applied to the simple regression model. Later, we shall develop a least-squares estimator. A maximum-likelihood estimator is also available.

Multiplying (4) by y_1 and taking expectations gives

$$\bigl\{E(y_1^2) - E(y_1\eta_1)\bigr\}\beta_1 + \bigl\{E(y_1y_2) - E(y_1\eta_2)\bigr\}\beta_2 = E(y_1)\alpha. \tag{5}$$

From the assumption that the error η_j and the true value ξ_i are statistically independent, whether or not the subscripts i and j agree, it follows that

$$(6)\qquad E(y_i\eta_j) = E\big\{(\xi_i + \eta_i)\eta_j\big\} = E(\eta_i\eta_j) = \omega_{ij}.$$

Therefore (5) can be written as

$$(7)\qquad \big\{E(y_1^2) - \omega_{11}\big\}\beta_1 + \big\{E(y_1y_2) - \omega_{12}\big\}\beta_2 = E(y_1)\alpha.$$

Taking expectations in equation (1) gives

$$(8)\qquad E(y_1)\beta_1 + E(y_2)\beta_2 = \alpha,$$

and, on multiplying both sides of this by $E(y_1)$, we get

$$(9)\qquad \big\{E(y_1)\big\}^2\beta_1 + E(y_1)E(y_2)\beta_2 = E(y_1)\alpha.$$

On taking (9) from (7) we get

$$(10)\qquad \big\{V(y_1) - \omega_{11}\big\}\beta_1 + \big\{C(y_1, y_2) - \omega_{12}\big\}\beta_2 = 0,$$

where we have used

$$(11)\qquad \begin{aligned} V(y_1) &= E(y_1^2) - \big\{E(y_1)\big\}^2 \quad \text{and} \\ C(y_1, y_2) &= E(y_1y_2) - E(y_1)E(y_2). \end{aligned}$$

By premultiplying equation (4) by y_2 and taking expectations, and by performing the same set of manipulations as before, we can get

$$(12)\qquad \big\{C(y_2, y_1) - \omega_{21}\big\}\beta_1 + \big\{V(y_2) - \omega_{22}\big\}\beta_2 = 0.$$

Putting (10) and (12) together gives a system of homogeneous equations:

$$(13)\qquad \left\{\begin{bmatrix} V(y_1) & C(y_1, y_2) \\ C(y_2, y_1) & V(y_2) \end{bmatrix} - \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}\right\}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This pair of equations cannot be solved uniquely for both β_1 and β_2. In other words, the vector $\beta' = [\beta_1, \beta_2]$ is determined only up to a factor of proportionality. Therefore an arbitrary normalisation must be imposed. One possibility is to set $\beta_1^2 + \beta_2^2 = 1$. Another is to set $\beta_1 = -1$ or $\beta_2 = -1$ which is to give one or other of y_1 and y_2 the role of the dependent variable.

Once values for β_1 and β_2 have been obtained, the value of α is given by equation (8).

The foregoing solution depends upon our knowing the precise values of the moments within equation (13). When the moments of y_1 and y_2 are unknown, they may be estimated from a sample of observations $(y_1, y_2)_t; t = 1, \ldots, T$. The estimates are

$$\begin{aligned} s_{11} &= \frac{1}{T}\sum(y_{1t} - \bar{y}_1)^2, \\ s_{22} &= \frac{1}{T}\sum(y_{2t} - \bar{y}_2)^2, \\ s_{21} &= \frac{1}{T}\sum(y_{2t} - \bar{y}_2)(y_{1t} - \bar{y}_1). \end{aligned} \tag{14}$$

The errors are not directly observable; and there is, as yet, no indication of how their moments might be estimated. For the present, we shall assume that these are given in prior knowledge.

When the unknown moments of y_1 and y_2 are replaced by their empirical counterparts, the system will almost certainly become algebraically inconsistent; which means that it can have no solution. To render the system solvable, we must interpolate an additional element λ so as form

$$\left\{ \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} - \lambda \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \right\} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{15}$$

The factor λ should be given the value closest to unity which will reconcile the two equations. This value will converge to unity as the empirical moments converge to the true values.

We shall refer to equation (15) as the errors-in-variables estimator.

To see how λ may be determined, let us assume, for the sake of simplicity, that the two errors η_1, η_2 are uncorrelated, so that $\omega_{12} = \omega_{21} = 0$, and that they have equal variance, so that $\omega_{11} = \omega_{22}$. Then the value of the common variance need not be specified, since it may be absorbed in the value of λ. The resulting equation system is

$$\left\{ \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{16}$$

The requirement that the equations should be mutually consistent is equivalent to the condition that

$$\begin{aligned} 0 &= \mathrm{Det} \begin{bmatrix} s_{11} - \lambda & s_{12} \\ s_{21} & s_{22} - \lambda \end{bmatrix} \\ &= \lambda^2 - \lambda(s_{11} + s_{22}) + (s_{11}s_{22} - s_{12}s_{21}). \end{aligned} \tag{17}$$

Therefore λ is found as the solution to a quadratic equation.

Once the estimates for β_1 and β_2 have been determined, the estimate for α may be obtained from the empirical counterpart of equation (8):

$$\bar{y}_1\hat{\beta}_1 + \bar{y}_2\hat{\beta}_2 = \hat{\alpha}. \tag{18}$$

Least-Squares Estimation of The Errors-in-Variables Model

The set of estimating equations under (15) may be obtained from minimising the criterion function

$$\sum_{t=1}^{T}\Big\{\omega_{22}(y_{1t}-\xi_{1t})^2 - 2\omega_{12}(y_{1t}-\xi_{1t})(y_{2t}-\xi_{2t}) + \omega_{11}(y_{2t}-\xi_{2t})^2\Big\}, \tag{19}$$

which is a weighted sum of squares of the errors of observation, subject to the condition

$$\xi_1\beta_1 + \xi_2\beta_2 = \alpha, \tag{20}$$

which defines the relationship between the underlying variables.

Here the elements which weight the terms of the quadratic function are provided by the inverse of the dispersion matrix of the errors:

$$\begin{bmatrix}\omega_{11} & \omega_{12}\\ \omega_{21} & \omega_{22}\end{bmatrix}^{-1} = \frac{1}{\omega_{11}\omega_{22}-\omega_{21}\omega_{12}}\begin{bmatrix}\omega_{22} & -\omega_{12}\\ -\omega_{21} & \omega_{11}\end{bmatrix}. \tag{21}$$

In the appendix, we shall demonstrate that the equation under (16) comes from minimising the function

$$\sum_{t=1}^{T}\Big\{(y_{1t}-\xi_{1t})^2 + (y_{2t}-\xi_{2t})^2\Big\} \tag{22}$$

subject to the constraint of (20). This is just the sum of squares of the perpendicular distances of the observations $(y_1, y_2)_t$ from the interpolated regression line defined by the constraint. The resulting least-squares procedure is described as an orthogonal regression to distinguish it from the ordinary regression which was the subject of the previous chapter. In the case of ordinary least-squares regression, the criterion of estimation is to minimise the sum of squares of the distances in the direction of the y-axis.

It should be clear that a wide variety of regression procedures can be generated by varying the direction of the minimisation. In fact, it is the matrix under (21) which determines the direction.

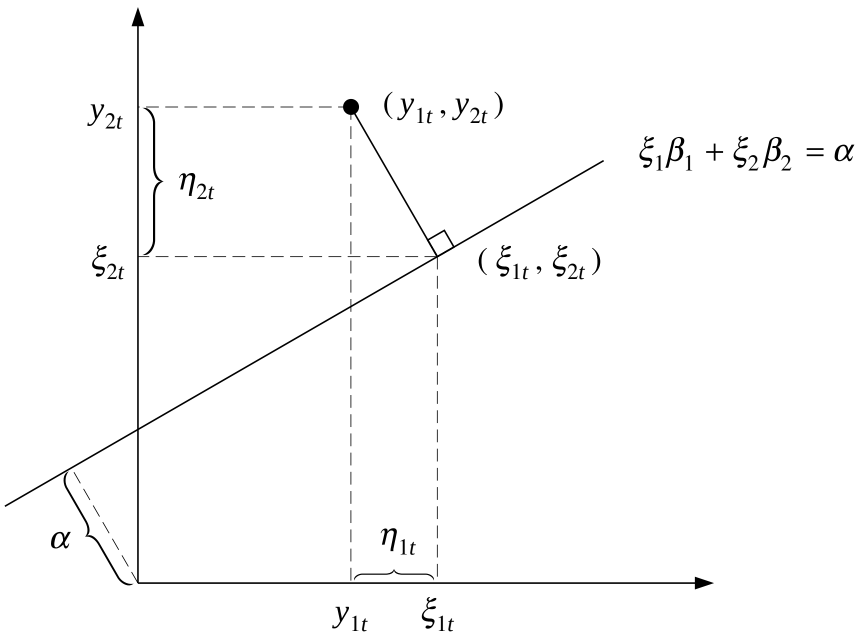

Figure 1. The perpendicular projection of the point (y_{1t}, y_{2t}) onto the regression line $\xi_1\beta_1 + \xi_2\beta_2 = \alpha$.

Ordinary Least-Squares Regression as a Limiting Case.

Imagine that the variance of the error η_1 is tending to zero. In that case, the covariance of η_1 and η_2 must also be tending to zero. With a change of notation and with a particular normalisation of the parameter vector, the limiting form of equation (15) can be written as

$$\left\{\begin{bmatrix} s_{xx} & s_{xy} \\ s_{yx} & s_{yy} \end{bmatrix} - \lambda \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right\} \begin{bmatrix} \beta \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{23}$$

where

$$\begin{aligned} s_{xx} &= \frac{1}{T}\sum(x_t - \bar{x})^2, \\ s_{yy} &= \frac{1}{T}\sum(y_t - \bar{y})^2, \\ s_{xy} &= \frac{1}{T}\sum(x_t - \bar{x})(y_t - \bar{y}). \end{aligned} \tag{24}$$

On solving the first equation $s_{xx}\beta - s_{xy} = 0$, we find that

$$\hat{\beta} = \frac{\sum(x_t - \bar{x})(y_t - \bar{y})}{\sum(x_t - \bar{x})^2}, \tag{25}$$

which is nothing but the ordinary least-squares estimator of the regression parameter in the equation $E(y|x) - x\beta = \alpha$.

In solving the second equation $s_{yy} - s_{yx}\beta = \lambda\sigma^2$, we are faced with two unknowns, λ and σ^2. If we set $\lambda = 1$, then the solution for σ^2 is

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{T}\sum(y_t - \bar{y})^2 - \frac{1}{T}\sum(y_t - \bar{y})(x_t - \bar{x})\hat{\beta} \\ &= \frac{1}{T}\sum(y_t - \bar{y})^2 - \frac{1}{T}\sum(x_t - \bar{x})^2\hat{\beta}^2.\end{aligned} \tag{26}$$

It is straightforward to demonstrate that this formula is equivalent to the formula

$$\hat{\sigma}^2 = \frac{1}{T}\sum(y_t - \hat{\alpha} - x_t\hat{\beta})^2, \qquad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}, \tag{27}$$

which is to be found under (1.46)

The Classical Simultaneous-Equations Model

Consider the system

$$y_1 = y_2\gamma_{21} + \varepsilon_1 : \qquad \text{The Demand Equation,} \tag{28}$$

$$y_2 = y_1\gamma_{12} + x\beta + \varepsilon_2 : \qquad \text{The Supply Equation.} \tag{29}$$

Here

y_1 represents the quantity of popcorn consumed and produced
y_2 represents the price of popcorn, and
x represents the cost of maize.

These variables, which are deviations from mean values, have expected values of zero. The effect of taking deviations is to simplify the algebra; for the intercept terms are thereby eliminated from the equations.

Another feature to take note of is the use of indices. The subscripts on the parameter γ_{21}, for example, indicate a mapping from y_2, which is the dependent variable of the second equation, to y_1, which is the dependent variable of the first equation. We shall assume that the disturbances ε_1 and ε_2 are independent of the variable x, which is described as an exogenous variable to indicate that it is generated in a context which lies outside the model.

The notion which lies behind this model is that the consumers of popcorn, whose behaviour is represented by the demand equation, respond to the price of popcorn, whereas the producers, whose behaviour is represented by the supply equation, set the price in view of the demand for their product and in view of their costs of production. The market is in a state of equilibrium where the quantity produced is equal to the quantity consumed.

Although the cost of maize is not the only cost of production, we shall assume, for the moment, that it is the only one which varies. The other costs, which are fixed, will have an effect which is subsumed in an intercept term which has been eliminated. The factors, other than price, which determine the demand for popcorn are likewise assumed to be constant and are subsumed in another intercept.

There are markets where output is ostensibly determined by supply factors and where the price adjusts to ensure that all of the output is sold. Some markets for agricultural produce are examples. In such cases, we might wish to place y_2 on the LHS of equation (28) and y_1 on the LHS of equation (29). However, there is no need to adapt the equations; for, in a situation of equilibrium, where both sides of the market are reconciled, it cannot be said that either is peculiarly responsible for the price of the item or for the quantity transacted.

The economist Alfred Marshall, who may be credited with formulating much of modern microeconomic theory, likened the supply and demand equations of a market in equilibrium to the blades of scissors. It is no more appropriate to ask which of the equations determines the price and which of them determines the quantity than it is to ask which of blades is cutting a sheet of paper.

It follows that, given a state of equilibrium, the choice of dependent variables in equations (28) and (29) is arbitrary. Nevertheless, the choice should reflect our understanding of how the two parties might behave in the process of achieving the equilibrium.

Now let us consider using a method of moments in estimating the parameters of the system. This entails finding expressions for the parameters which are in terms of the moments of the observable variables. Once these expressions have been found, we may consider replacing the theoretical moments by the empirical counterparts to derive the estimating equations.

To find an expression for γ_{21}, we multiply the demand equation (28) by x and we take expectations. This gives

$$E(xy_1) = E(xy_2)\gamma_{21}, \tag{30}$$

from which we see that

$$\gamma_{21} = \frac{E(xy_1)}{E(xy_2)}. \tag{31}$$

When we attempt to apply the same method to the supply equation (29), we find that there is not sufficient information to determine the two remaining parameters. Multiplying the equation by x and taking expectations leads to

$$E(xy_2) = E(xy_1)\gamma_{12} + E(x^2)\beta. \tag{32}$$

If we seek another equation by multiplying equation (29) by y_1 and by taking expectations, then we shall introduce another unknown quantity which is the nonzero moment $E(y_1\varepsilon_2)$:

$$E(y_1y_2) = E(y_1^2)\gamma_{12} + E(y_1x)\beta + E(y_1\varepsilon_2). \tag{33}$$

We have an equal lack of success in attempting to form an estimating equation by multiplying the equation (29) by y_2 and taking expectations.

In view of its role in generating estimating equations, the exogenous variable x is apt to be described as an instrumental variable. The problem of the supply equation is the impossibility of estimating two parameters γ_{12} and β when there is only one instrumental variable. The two parameters are said to be unidentified. A necessary condition for the identification of the parameters of any equation is that their number should not exceed the number of the available instrumental variables.

Example. An attempt to estimate equation (28) by the ordinary method of regression would lead to a biased estimator. The method is inappropriate because the disturbance term ε_1 is correlated with the variable y_2 which is the putative regressor. This correlation is evident in the fact that we can trace a connection running from ε_1 to y_1, within equation (28) and thence from y_1 to y_2 through equation (29). To find an expression for the covariance of y_2 and ε_1, we may substitute equation (28) into equation (29) to give

$$y_2 = (y_2\gamma_{21} + \varepsilon_1)\gamma_{12} + x\beta + \varepsilon_2. \tag{34}$$

Rearranging this gives

$$y_2 = \frac{x\beta}{1-\gamma_{21}\gamma_{12}} + \frac{\varepsilon_1\gamma_{12} + \varepsilon_2}{1-\gamma_{21}\gamma_{12}}. \tag{35}$$

Therefore

$$C(y_2, \varepsilon_1) = \frac{V(\varepsilon_1)\gamma_{12} + C(\varepsilon_1, \varepsilon_2)}{1-\gamma_{21}\gamma_{12}}. \tag{36}$$

Now let us consider some circumstances which would enable us to estimate both the supply equation and the demand equation. Consider the system

$$y_1 = y_2\gamma_{21} + x_1\beta_{11} + \varepsilon_1 : \qquad \text{The Demand Equation,} \tag{37}$$

$$y_2 = y_1\gamma_{12} + x_2\beta_{22} + \varepsilon_2 : \qquad \text{The Supply Equation.} \tag{38}$$

Compared with equation (28), the revised demand equation incorporates an extra variable x_1 which represents the price of candy floss. If candy floss and popcorn are attractive to the same people, then one may expect the demand for popcorn to fall if the price of candy floss is reduced. Given the additional instrumental variable, we can now estimate the parameters of both equations, which have an identical structure.

To derive estimating equations for the parameters of the demand equation, we multiply the latter in turn by x_1 and x_2 and we take expectations. The results are

$$E(x_1y_1) = E(x_1y_2)\gamma_{21} + E(x_1^2)\beta_{11}, \tag{39}$$

$$E(x_2y_1) = E(x_2y_2)\gamma_{21} + E(x_2x_1)\beta_{11}. \tag{40}$$

These equations serve simultaneously to determine both γ_{21} and β_{11}. Their empirical counterparts, which are derived by replacing the theoretical moments by the corresponding sample moments serve as estimating equations for the parameters. We may use exactly the same device in estimating the supply equation.

We must avoid the false impression that new variables may be introduced at will. The presence, in the demand equation, of the price of candy floss can be justified only if the latter has an active effect on the level of demand. That is to say, x_1 must vary within the sample of observations if it is to assist in identifying the parameters of the model. If this price is constant, then its effect will be subsumed, as before, in the intercept term of the demand equation. It is also required that the price of candy floss should not enter the supply equation for popcorn, which seems plausible.

Now let us consider a third possibility which puts a different construction on the problem of estimation. Consider the system

$$y_1 = y_2\gamma_{21} + \varepsilon_1 : \qquad \text{The Demand Equation,} \tag{41}$$

$$y_2 = y_1\gamma_{12} + x_1\beta_{12} + x_2\beta_{22} + \varepsilon_2 : \qquad \text{The Supply Equation.} \tag{42}$$

Here

x_1 represents the cost of maize, and
x_2 represents the cost of pink sugar.

The price of candy floss no longer enters the demand equation; and we might imagine that the makers of candy floss no longer occupy their stalls on the seaside promenade. There are now two instrumental variables which can serve to identify the demand equation. Thus

$$E(x_1y_1) = E(x_1y_2)\gamma_{21}, \tag{43}$$

$$E(x_2y_1) = E(x_2y_2)\gamma_{21}. \tag{44}$$

The parameter γ_{21} is said to be overidentified.

In practice, when we replace the theoretical moments by their empirical counterparts, the estimates which are generated by the two equations are liable to differ. Since both of the estimates are valid, we should attempt, in the interests of statistical efficiency, to combine them.

In order to resolve the conflict between the two estimates of γ_{21} we shall resort to a procedure which involves the errors-in-variables estimator. We begin by deriving the so-called reduced-form equations. Substituting equation (42) into equation (41) gives

$$y_1 = (y_1\gamma_{12} + x_1\beta_{12} + x_2\beta_{22})\gamma_{21} + (\varepsilon_1 + \varepsilon_2\gamma_{21}). \tag{45}$$

On rearranging this we get

$$\begin{aligned} y_1 &= \frac{(x_1\beta_{12} + x_2\beta_{22})\gamma_{21}}{1 - \gamma_{12}\gamma_{21}} + \frac{\varepsilon_1 + \varepsilon_2\gamma_{21}}{1 - \gamma_{12}\gamma_{21}} \\ &= x_1\pi_{11} + x_2\pi_{21} + \eta_1 \\ &= \xi_1 + \eta_1, \end{aligned} \tag{46}$$

which is the so-called reduced-form equation for y_1. Substituting equation (41) into equation (42) gives

$$y_2 = y_2\gamma_{21}\gamma_{12} + x_1\beta_{12} + x_2\beta_{22} + (\varepsilon_2 + \varepsilon_1\gamma_{12}). \tag{47}$$

On rearranging this we get

$$\begin{aligned} y_2 &= \frac{x_1\beta_{12} + x_2\beta_{22}}{1 - \gamma_{21}\gamma_{12}} + \frac{\varepsilon_2 + \varepsilon_1\gamma_{12}}{1 - \gamma_{21}\gamma_{12}} \\ &= x_1\pi_{12} + x_2\pi_{22} + \eta_2 \\ &= \xi_2 + \eta_2, \end{aligned} \tag{48}$$

which is the reduced-form equation for y_2. On comparing equations (46) and (48), it can be seen that

$$y_1 - \eta_1 = (y_2 - \eta_2)\gamma_{21}. \tag{49}$$

This is the equation of an errors-in-variables model wherein one of the parameters has been normalised with a value of -1.

We can use the errors-in-variables estimator for γ_{21} provided that we can find values for the variances and covariances for the errors η_1 and η_2 which are, in fact, the disturbances of the reduced-form regression equations. Let

$$\begin{aligned} h_{1t} &= y_{1t} - x_{1t}\hat{\pi}_{11} - x_{2t}\hat{\pi}_{21} \quad \text{and} \\ h_{2t} &= y_{2t} - x_{1t}\hat{\pi}_{12} - x_{2t}\hat{\pi}_{22} \end{aligned} \tag{50}$$

be the residuals from using the ordinary method of regression in fitting the reduced-from equations to a sample of T observations. Then the moments of the reduced-form disturbances may be estimated as follows:

$$
\begin{aligned}
\hat{\omega}_{11} &= \frac{1}{T}\sum_{t=1}^{T} h_{1t}^2,\\
\hat{\omega}_{22} &= \frac{1}{T}\sum_{t=1}^{T} h_{2t}^2,\\
\hat{\omega}_{12} &= \frac{1}{T}\sum_{t=1}^{T} h_{1t}h_{2t}.
\end{aligned} \tag{51}
$$

Using equation (15) as a model, we can now construct an estimating equation for γ_{21} in the form of

$$
\left\{\begin{bmatrix} s_{11} & s_{12}\\ s_{21} & s_{22}\end{bmatrix} - \lambda\begin{bmatrix}\hat{\omega}_{11} & \hat{\omega}_{12}\\ \hat{\omega}_{21} & \hat{\omega}_{22}\end{bmatrix}\right\}\begin{bmatrix}-1\\ \gamma_{21}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}. \tag{52}
$$

This gives

$$
\hat{\gamma}_{21} = \frac{s_{11}-\lambda\hat{\omega}_{11}}{s_{12}-\lambda\hat{\omega}_{12}} = \frac{s_{21}-\lambda\hat{\omega}_{21}}{s_{22}-\lambda\hat{\omega}_{22}}. \tag{53}
$$

The value of λ which guarantees the equality above, is found by solving the determinental equation

$$
0 = \mathrm{Det}\begin{bmatrix} s_{11}-\lambda\hat{\omega}_{11} & s_{12}-\lambda\hat{\omega}_{12}\\ s_{21}-\lambda\hat{\omega}_{21} & s_{22}-\lambda\hat{\omega}_{22}\end{bmatrix}, \tag{54}
$$

which is a quadratic equation. The root which is closest to unity is taken. As the various empirical moments tend to their true values, so λ will tend to unity.

There are other ways of estimating the parameter which become virtually equivalent to the errors-in-variables method when the sample size is large. One possibility is to use a system which is modelled on equation (23):

$$
\left\{\begin{bmatrix} s_{11}-\hat{\omega}_{11} & s_{12}-\hat{\omega}_{12}\\ s_{21}-\hat{\omega}_{21} & s_{22}-\hat{\omega}_{22}\end{bmatrix} - \mu\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right\}\begin{bmatrix}-1\\ \gamma_{21}\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}. \tag{55}
$$

In comparison with equation (52), it can be seen that λ has been set to unity. Since λ is no longer available for the purpose of rendering the equations algebraically consistent, a new factor μ has been introduced to perform this task.

Whereas λ will tend to unity with the convergence of the moments to the true values, the value of μ will to zero.

The solution of equation (55) is

$$\hat{\gamma}_{21} = \frac{s_{21} - \hat{\omega}_{21}}{s_{22} - \hat{\omega}_{22}}. \tag{56}$$

This is, in fact, the so-called two-stage least-squares estimator of γ_{21}; and it differs from the ordinary least-squares estimator by virtue of the adjustments which are made to the moments s_{21} and s_{22}.

Two-Stage Least Squares and Limited-Information Maximum Likelihood.

The estimator of the demand equation which we have derived under the guise of the errors-in-variables model was originally derived as the limited-information maximum-likelihood (LIML) estimator by Anderson and Rubin in 1949, when they were members of the Cowles Commission for Research in Economics. The Commission consisted of a group of statisticians and economists whose research was funded by the American industrialist Alfred Cowles. It is arguable the era of modern econometrics began with the work of the Commission.

The derivation of the LIML estimator was a *tour de force.* Its complexity was due in part to the the fact that the likelihood function of a full simultaneous-equation model was taken as a starting point. An alternative derivation, which was no less complicated, was provided shortly afterward by Hood and Koopmans. The inaccessibility of both these derivations deterred econometricians from using the estimator. It was not until the alternative two-stage least-squares estimator was invented independently by Theil and Basmann in the late 1950's that the techniques of simultaneous-equation estimation began to be applied.

The affinity of the 2SLS and the LIML estimators is not evident from a comparison of the original derivations. Nor might it be clear to someone familiar with the 2SLS estimator that it corresponds to what is presented under (55) and (56). Therefore we shall give a version of the familiar derivation before showing how the equivalence may be demonstrated.

The point of departure for the original derivation of the 2SLS estimator is the recognition that, in a structural equation such as (41), the disturbance term is liable to be correlated with some of the variables on the RHS. We have already demonstrated, in an example, the correlation between y_2 and ε_1 within equation (28), which is equation (41) in a different context.

The question arising is how we might purge the variable y_2 of the component which is correlated with ε_1. An effective way, if it were available, would be to replace y_2 by the predicted value $\xi_2 = x_1\pi_{12} + x_2\pi_{22}$ which comes from the reduced-form equation. In fact, by substituting the reduced-form expression for y_2 given by (48) into the equation (41), we obtain

$$\begin{aligned} y_1 &= y_2\gamma_{21} + \varepsilon_1 \\ &= (x_1\pi_{12} + x_2\pi_{22})\gamma_{21} + (\varepsilon_1 + \eta_2\gamma_{21}) \\ &= \xi_2\gamma_{21} + \zeta_1. \end{aligned} \tag{57}$$

The composite disturbance term $\zeta_1 = \varepsilon_1 + \eta_2\gamma_{21}$ is clearly uncorrelated with ξ_2 since ε_1 and η_2 are uncorrelated with x_1 and x_2. Therefore a consistent estimator of γ_{21} would be obtained from the regression of y_1 on ξ_2.

In fact, we cannot put the unknown value of ξ_2 in place of y_2, and we have to make do with its estimate $\hat{y}_2 = x_1\hat{\pi}_{12} + x_2\hat{\pi}_{22}$ which can be expected to converge to ξ_2 as the sample size increases. The resulting estimator of γ_{21} is

$$\begin{aligned} \hat{\gamma}_{21} &= \frac{\sum \hat{y}_{2t}y_{1t}}{\sum \hat{y}_{2t}^2} \\ &= \frac{\sum \hat{y}_{2t}\hat{y}_{1t}}{\sum \hat{y}_{2t}^2}. \end{aligned} \tag{58}$$

The second equality depends upon the result that $\sum \hat{y}_{2t}y_{1t} = \sum \hat{y}_{2t}\hat{y}_{1t}$. The latter is due to the fact that the reduced-from disturbance h_1 within $y_1 = \hat{y}_1 + h_1$ is uncorrelated with the reduced-form regressors x_1 and x_2 and hence with $\hat{y}_2 = x_1\hat{\pi}_{12} + x_2\hat{\pi}_{22}$.

The equivalence between the expression for the 2SLS estimator under (58) and the expression under (56), follows from the identities

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} y_{2t}^2 &= \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{2t}^2 + \frac{1}{T}\sum_{t=1}^{T} h_{2t}^2 \quad \text{and} \\ \frac{1}{T}\sum_{t=1}^{T} y_{1t}y_{2t} &= \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{1t}\hat{y}_{2t} + \frac{1}{T}\sum_{t=1}^{T} h_{1t}h_{2t}. \end{aligned} \tag{59}$$

Using the definitions of (14) and (51), and remembering that the variables are in devaition from, we see that these can be rewritten as

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{2t}^2 &= s_{22} - \hat{\omega}_{22} \quad \text{and} \\ \frac{1}{T}\sum_{t=1}^{T} \hat{y}_{1t}\hat{y}_{2t} &= s_{12} - \hat{\omega}_{12}; \end{aligned} \tag{60}$$

and the equivalence of (56) and (58) follows immediately.

The Permanent Income Hypothesis and the Dynamic Consumption Function

In his famous book *The General Theory of Employment, Interest and Money*, Keynes seemed to suggest that it was a universal rule that, as they grow richer, people consume a declining proportion of their income. The empirical studies of consumer behaviour which were made during the war in the United States seemed to confirm this notion. These studies were based mainly on data of family budgets collected in cross-section household surveys, virtually, at a point in time.

The notion of a declining propensity to consume seemed to be supported also by such time-series evidence as was available from the interwar period.

A weak propensity to consume implied that, if prosperity were to be maintained in postwar years, then the high levels of government expenditure which had characterised the war years would have to be maintained. Moreover, most analysts agreed that these levels of expenditure had been responsible for lifting the economy out of its prewar depression.

In fact, in the immediate postwar years, consumption was buoyant and showed signs of keeping pace with income. Moreover, studies of economic growth began to reveal that the ratio of consumption to income tends to fluctuate from year to year around a value which is remarkably stable in the long term.

Economic theorists were faced with the problem of how to reconcile the essential features of the Keynesian model with the statistical facts which were emerging. Their response, in the main, was to introduce a temporal dimension into the Keynesian model which, hitherto, had been expressed in the terms of static equilibrium analysis.

The idea of the dynamic consumption function is that changes in income have a delayed and a cumulative effect upon consumption. If income rises rapidly, then it will outstrip consumption and the ratio of consumption to income will diminish. If income falls rapidly, then the reverse will happen. Only if income is stationary or changing at a constant rate will the proportions of consumption and income be stable. The idea was expressed in similar ways by several authors including Milton Friedman.

The permanent income hypothesis of Friedman suggests that households plan their expenditure in view of their permanent or habitual income. Thus any unforeseen or transitory fluctuations in income will not immediately affect the expenditure plans which may fail for other reasons.

We might imagine that planned expenditure c^p is a simple function of permanent income y^p:

$$c^p = \alpha + y^p\beta. \tag{61}$$

Let y and c be actual consumption and actual income, so that

$$c = c^p + \eta^c \quad \text{and} \quad y = y^p + \eta^y, \tag{62}$$

where η^c and η^y are the transitory components described as unplanned consumption and windfall income respectively. We can regard the transitory components as unobserved random variables which are mutually uncorrelated as well as being uncorrelated with the corresponding permanent quantities. Thus

$$\begin{aligned} &C(\eta^c, \eta^y) = 0, \\ &E(\eta^c) = 0, \quad C(\eta^c, c^p) = 0, \\ &E(\eta^y) = 0, \quad C(\eta^y, y^p) = 0. \end{aligned} \tag{63}$$

On substituting for $c^p = c - \eta^c$ and $y^p = y - \eta^y$ in equation (61), we get

$$c - \eta^c = \alpha + (y - \eta^y)\beta, \tag{64}$$

or equivalently,

$$\begin{aligned} c &= \alpha + y\beta + (\eta^c - \eta^y\beta) \\ &= \alpha + y\beta + \varepsilon. \end{aligned} \tag{65}$$

Equation (64) suggests that we have a errors-in-variables model. However, if we take equation (65) and if we ignore the structure of the disturbance term $\varepsilon = \eta^c - \eta^y\beta$, then we might imagine that we have to deal with a simple regression equation. Therefore, let us examine the consequences using an ordinary regression estimator. The estimator is

$$\begin{aligned} \hat{\beta} &= \frac{\sum(y_t - \bar{y})(c_t - \bar{c})}{\sum(y_t - \bar{y})^2} \\ &= \frac{\sum(y_t - \bar{y})\big\{(y_t - \bar{y})\beta + (\varepsilon_t - \bar{\varepsilon})\big\}}{\sum(y_t - \bar{y})^2} \\ &= \beta + \frac{\sum(y_t - \bar{y})\varepsilon_t}{\sum(y_t - \bar{y})^2}. \end{aligned} \tag{66}$$

In favourable circumstances, we could expect the second term on the RHS of the final expression to vanish when the expectation of $\hat{\beta}$ is taken. This will not happen in the present case because the disturbance $\varepsilon = \eta^c - \eta^y\beta$ is statistically related to the regressor $y = y^p + \eta^y$.

To highlight the problem, let us examine the asymptotic properties of the estimator. Consider the numerator term

$$\sum_{t=1}^{T}(y_t - \bar{y})\varepsilon_t = \sum_{t=1}^{T}\big\{(y_t^p - \bar{y}^p) + (\eta_t^y - \bar{\eta}^y)\big\}(\eta_t^c + \eta_t^y\beta). \tag{67}$$

Expanding this gives

$$\begin{aligned}\sum_{t=1}^{T}(y_t-\bar{y})\varepsilon_t &= \sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^c+\beta\sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^y\\ &\quad+\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)\eta_t^c+\beta\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)^2.\end{aligned}\tag{68}$$

On dividing the terms on the RHS by T and taking limits, we find that

$$\begin{aligned}&\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^c=C(y^p,\eta^c)=0,\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^y=C(y^p,\eta^y)=0,\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)\eta_t^c=C(\eta^y,\eta^c)=0,\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)^2=V(\eta^y)\neq 0.\end{aligned}\tag{69}$$

The expression in the denominator is

$$\begin{aligned}\sum_{t=1}^{T}(y_t-\bar{y})^2 &= \sum_{t=1}^{T}\left\{(y_t^p-\bar{y}^p)+(\eta^y-\bar{\eta}^y)\right\}^2\\ &=\sum_{t=1}^{T}(y_t^p-\bar{y}^p)^2+2\sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^y+\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)^2.\end{aligned}\tag{70}$$

Dividing by T and taking limits gives

$$\begin{aligned}&\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t^p-\bar{y}^p)^2=V(y^p),\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(\eta_t^y-\bar{\eta}^y)^2=V(\eta^y),\\ &\text{plim}\frac{1}{T}\sum_{t=1}^{T}(y_t^p-\bar{y}^p)\eta_t^y=C(y^p,\eta^y)=0.\end{aligned}\tag{71}$$

Putting these results together gives

$$
\begin{aligned}
\text{plim}(\hat{\beta}) &= \beta + \frac{\text{plim}\Big\{T^{-1}\sum(y_t-\bar{y})\varepsilon_t\Big\}}{\text{plim}\Big\{T^{-1}\sum(y_t-\bar{y})^2\Big\}} \\
&= \beta - \frac{\beta V(\eta^y)}{V(y^p)+V(\eta^y)} = \frac{\beta V(y^p)}{V(y^p)+V(\eta^y)}.
\end{aligned} \tag{72}
$$

The ordinary regression estimate is $\hat{\beta}$ is inconsistent, and it tends to underestimate the propensity to consume to an extent which depends upon the variance of the transitory income component η^y.

To overcome these problems, Friedman resorted to a subsidiary hypothesis concerning the determination of the value of habitual or permanent income y^p. He proposed that

$$
\begin{aligned}
y_t^p &= (1-\lambda)y_t + \lambda y_{t-1}^p \\
&= y_{t-1}^p + (1-\lambda)(y_t - y_{t-1}^p); \quad \text{with} \quad 0 \leq \lambda < 1.
\end{aligned} \tag{73}
$$

This implies that, in each period, the value of permanent income is modified in to take account of the value of the income actually received. In fact, if the actual income is constant over a long period, then permanent income will adjust to this value gradually. The result can be understood readily by examining the second expression which shows that the value of permanent income for the current period is formed by adding a fraction of the discrepancy $y_t - y_{t-1}^p$ to its value in the previous period. If actual income is constant over a long period, then the discrepancy, which is akin to a prediction error, will gradually disappear.

Substituting the expression $y_{t-1}^p = (1-\lambda)y_{t-1} + \lambda y_{t-2}^p$ into equation (73) gives

$$
y_t^p = (1-\lambda)y_t + \lambda(1-\lambda)y_{t-1} + \lambda^2 y_{t-2}^p. \tag{74}
$$

Substituting, in turn, for $y_{t-2}^p = (1-\lambda)y_{t-2} + \lambda y_{t-3}^p$ gives

$$
y_t^p = (1-\lambda)\big\{y_t + \lambda y_{t-1} + \lambda^2 y_{t-2}\big\} + \lambda^3 y_{t-3}^p; \tag{75}
$$

and, by continuing this process indefinitely, one obtains

$$
y_t^p = (1-\lambda)\sum_{i=0}^{\infty}\lambda^i y_{t-i}. \tag{76}
$$

The sequence of the coefficients $\{(1-\lambda)\lambda^i; i = 1, 2\ldots\}$, which are to be found on the RHS of equation (76), define a so-called transfer function which describes how the effects of a change of actual income are distributed over time. The initial impact on permanent income of a change Δy in received income is $(1-\lambda)\Delta y$. If the new level of received income is maintained, then, in the next period, the accumulated change in permanent income will be $\Delta y(1-\lambda)(1+\lambda)$ and, after two periods, it will be $\Delta y(1-\lambda)(1+\lambda+\lambda^2)$. If income were maintained indefinitely at the new level, then permanent income would increase, ultimately, by a multiple of Δy which is given by

$$(1-\lambda)\sum_{i=0}^{\infty}\lambda^i = (1-\lambda)(1+\lambda+\lambda^2+\cdots). \tag{77}$$

The value of this multiple is described as the gain of the transfer function. In the present case, the gain is clearly unity.

On substituting the expression from (76) into the consumption relationship

$$c_t = \alpha + y_t^p\beta + \eta_t^c, \tag{78}$$

we get

$$\begin{aligned} c_t &= \alpha + \beta(1-\lambda)\sum_{i=0}^{\infty}\lambda^i y_{t-i} + \eta_t^c \\ &= \alpha + \gamma\sum_{i=0}^{\infty}\lambda^i y_{t-i} + \eta_t^c. \end{aligned} \tag{79}$$

This is a regression model comprising an infinite number of lagged values of the explanatory variable y. In other respects it fulfils the standard assumptions.

To make the model amenable to estimation, we may write it as

$$c_t = \alpha + \gamma z_t + \lambda^t\delta + \eta_t^c, \tag{80}$$

where

$$z_t = y_t + \lambda y_{t-1} + \lambda^2 y_{t-2} + \cdots + \lambda^{t-1}y_1. \tag{81}$$

comprises elements of the sample, and

$$\begin{aligned} \lambda^t\delta &= \lambda^t\gamma\big(y_0 + \lambda y_{-1} + \lambda^2 y_{-2} + \cdots\big) \\ &= \gamma\big(\lambda^t y_0 + \lambda^{t+1}y_{-1} + \lambda^{t+2}y_{-2} + \cdots\big) \end{aligned} \tag{82}$$

comprises only presample values.

The parameters of equation (80) can be estimated by a combination of an ordinary least-squares regression procedure and a search procedure. Given a specific value of $\lambda \in [0,1)$, one may form the regressors λ^t and $z_t = z_t(\lambda)$ for $t = 1, \ldots, T$. One can proceed to estimate α, $\gamma = \beta(1-\lambda)$ and the nuisance parameter δ by ordinary least squares. Then another value for λ can be selected and the exercise repeated. At each stage, the value of the sum of squares of the residuals will be recorded. In this way we can locate, to whatever degree of accuracy is desired, the values of α, γ and δ which minimise the residual sum of squares unconditionally. These are the definitive estimates.

We shall now describe an alternative method of estimation. Consider again the equation under (78). On taking λc_{t-1} from both sides, we get

(83) $$c_t - \lambda c_{t-1} = (1-\lambda)\alpha + (y_t^p - \lambda y_{t-1}^p)\beta + (\eta_t^c - \lambda\eta_{t-1}^c).$$

From equation (73), which defines permanent income, we get

(84) $$y_t^p - \lambda y_{t-1}^p = (1-\lambda)y_t.$$

On substituting the latter into equation (83) and on defining $\gamma = \beta(1-\lambda)$ and $\theta = (1-\lambda)\alpha$, we get

(85) $$c_t - \lambda c_{t-1} = \theta + \gamma y_t + (\eta_t^c - \lambda\eta_{t-1}^c).$$

This can be rearranged to give

(86) $$c_t - \eta_t^c = \lambda(c_{t-1} - \eta_{t-1}^c) + \gamma y_t + \theta.$$

Equation (86) describes an errors-in-variables model with three variables c_t, c_{t-1} and y_t which are associated respectively with the errors η_t^c, η_{t-1}^c and 0. The estimating equations for λ and δ are

(87) $$\left\{\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} - \mu \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right\} \begin{bmatrix} -1 \\ \lambda \\ \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

where

(88) $$\begin{aligned} s_{11} &= \frac{1}{T-1}\sum(c_t - \bar{c}_0)^2, \\ s_{12} &= \frac{1}{T-1}\sum(c_t - \bar{c}_0)(c_{t-1} - \bar{c}_{-1}), \\ s_{13} &= \frac{1}{T-1}\sum(c_t - \bar{c}_0)(y_t - \bar{y}), \\ s_{22} &= \frac{1}{T-1}\sum(c_{t-1} - \bar{c}_{-1})^2, \\ s_{23} &= \frac{1}{T-1}\sum(c_{t-1} - \bar{c}_{-1})(y_t - \bar{y}), \\ s_{33} &= \frac{1}{T-1}\sum(y_t - \bar{y})^2, \end{aligned}$$

where the summations run from $t = 2, \cdots, T$ and where $\bar{c}_0 = (T-1)^{-1}\sum c_t$ and $\bar{c}_{-1} = (T-1)^{-1}\sum c_{t-1}$. For estimating θ, we use the equation

$$\hat{\theta} = \bar{c}_0 - \lambda\bar{c}_{-1} + \gamma\bar{y}. \tag{89}$$

Adaptive Expectations and Partial Adjustment

The permanent income hypothesis of Friedman may be summarised in two equations:

$$c_t = \alpha + y_t^p\beta + \eta_t^c, \tag{90}$$

$$y_t^p = (1-\lambda)y_t + \lambda y_{t-1}^p. \tag{91}$$

The second equation depicts an adaptive expectations mechanism. It shows how the current value of permanent income is derived from the previous value y_{t-1}^p in a manner which takes account of the amount of income which is actually received. We have shown that, by a simple process of substituting successively for the lagged values of permanent income, the current value of permanent income may be expressed as a geometrically weighted average of all past values of received income:

$$y_t^p = (1-\lambda)\sum_{i=0}^{\infty}\lambda^i y_{t-i}. \tag{92}$$

In contrast to the effects of received income, the effects of disturbances upon the stream of consumption are transitory. They are forgotten after one period. Thus the consumer navigates on a strict course in the face of the disturbances; and it is as if he were applying a firm hand to the tiller of a small boat in a choppy sea.

An alternative model for the process of consumption is provided by the partial adjustment hypothesis. This is also expressed in two equations:

$$c_t = (1-\lambda)c_t^* + \lambda c_{t-1} + \varepsilon_t, \tag{93}$$

$$c_t^* = \alpha + \beta y_t. \tag{94}$$

Here is c^* is a latent variable which represents desired consumption.

The notion underlying this model it that a consumer faces costs in realising his desires. Therefore his adaptation to a new level of income may be a sluggish one; and his tendency will be to adhere to established habits of consumption. Such habits are influenced by the disturbances; and past disturbances will have the same lingering effect on consumption as past receipts of income.

By substituting (94) into (93), we derive the following equation:

$$\begin{aligned} c_t &= (1-\lambda)\{\alpha + \beta y_t\} + \lambda c_{t-1} + \varepsilon_t \\ &= \theta + \gamma y_t + \lambda c_{t-1} + \varepsilon_t. \end{aligned} \tag{95}$$

By a process of back substitution, analogous to that which serves to derive equation (92) from equation (91), this may be reexpressed as

$$\begin{aligned} c_t &= (1-\lambda)\sum_{i=0}^{\infty}\lambda^i\{\alpha + \beta y_{t-i}\} + \sum_{i=0}^{\infty}\lambda^i\varepsilon_{t-i} \\ &= \alpha + \beta y_t^p + \zeta_t, \end{aligned} \tag{96}$$

where $\zeta_t = \sum \lambda^i \varepsilon_{t-i}$ is a geometrically weighted average of all past disturbances. In the face of the disturbances, the consumer navigates a course which could be compared with that of a big boat in a choppy sea which is steered by a loose hand.

The advantage which the partial-adjustment model has over the adaptive-expectations model is its ease of estimation. The fact that the current disturbance ε_t is uncorrelated with the explanatory variable c_{t-1} justifies us in treating equation (95) as if it were an ordinary regression equation. Therefore the method of ordinary least squares may be used to estimate the parameters θ, γ and λ. However, there is likely to be a strong correlation between c and the past values of ε; and, whilst this does not affect their consistency—assuming that the model is correct—it means that the estimates will not be unbiased.

Little reflection is needed to understand that ease of estimation is a poor criterion for adopting a model. We should accept whichever model proves to be more consistent with the data.

Even on first appearances, it seems that the adaptive-expectations model is too restrictive. The notion that the sequence $\{\zeta_t\}$, which summarised the effect of the disturbances, should be formed in the same way as the sequence $\{y_t^p\}$ of permanent income, and that it should entail the same parameters, is doubtful. In the face of such doubts, a model of the form

$$\begin{aligned} c_t &= (1-\lambda)\sum_{i=0}^{\infty}\lambda^i\{\alpha + \beta y_{t-i}\} + \sum_{i=0}^{\infty}\phi^i\varepsilon_{t-i} \\ &= \alpha + \beta y_t^p + \zeta_t, \end{aligned} \tag{97}$$

with gives separate parameters to the two processes, is called for. Only if the parameters λ and ϕ were to show themselves, after estimation, to have similar values, should we be prepared to adopt the more restrictive model of equation (96) which constrains them to be equal.

The advantage of the model of (97), which is shared with the simpler permanent-income model of (90) and (91), is that a misspecification of the nature of the disturbance process will not vitiate the estimation parameters α, λ and β of the systematic part. Whilst the efficiency of the estimates might be prejudiced by such a misspecification, their consistency will not be affected.